"Development of molecular approaches for the control of *Odoiporus longicollis* (Oliver), a major pest of bananas and plantains in Asia."

A Thesis

Submitted to the

SAVITRIBAI PHULE PUNE UNIVERSITY

For the degree of

Doctor of Philosophy

In

Biochemistry

By
Pallavi Shankar

Under the guidance of

Dr. Lalitha Sunil Kumar

Division of Biochemical Sciences

CSIR-NATIONAL CHEMICAL LABORATORY

Pune-411 008

India

September, 2014

राष्ट्रीय रासायनिक प्रयोगशाला
(वैज्ञानिक तथा औद्योगिक अनुसंधान परिषद)
डॉ. होमी भाभा रोड, पुणे - 411 008. भारत
NATIONAL CHEMICAL LABORATORY
(Council of Scientific & Industrial Research)
Dr. Homi Bhabha Road, Pune - 411008. India

CERTIFICATE

This is to certify that the work described in the thesis **"Development of molecular approaches for the control of *Odoiporus longicollis* (Oliver), a major pest of bananas and plantains in Asia."** by Pallavi Shankar, submitted for the degree of **Doctor of Philosophy in Biochemistry** was carried out under my supervision at the Biochemical Sciences Division of the National Chemical Laboratory, Pune, India. Such materials which have been obtained by other sources have been duly acknowledged in this thesis. To the best of my knowledge, the present work or any part thereof has not been submitted to any other University for the award of any other degree or diploma.

Dr Lalitha Sunil Kumar
(Research Guide)

Dr Sushma M Gaikawad
(Research Co-guide)

Communications Channels ☎ +91 20 25902000 +91 20 25893300 +91 20 25893400 Fax +91 20 25902601 (Director) +91 20 25902660 (Admin.) +91 20 25902639 (Business Development) URL : www.ncl-india.org

DECLARATION BY THE CANDIDATE

I hereby declare that the thesis entitled **"Development of molecular approaches for the control of *Odoiporus longicollis* (Oliver), a major pest of bananas and plantains in Asia."** submitted to Savitribai Phule Pune University for the degree of **Doctor of Philosophy in Biochemistry**, was carried out by me under the guidance of **Dr. Lalitha Sunil Kumar** and has not formed the basis for the award of any degree, diploma, associate-ship, fellowship, titles in this or any other University or other institute of higher learning. I further declare that the material obtained from other sources has been duly acknowledged in the thesis.

Date: Signature of the Candidate

Place: **Pallavi Shankar**

Dedicated to my

Beloved

Husband Raj

Contents

Acknowledgments

At the end of my thesis, it is a pleasant task to express my thanks to all those who contributed in many ways to the success of this study and all my friends who provided me the moral support to complete my study.

First and foremost I offer my sincerest gratitude to my supervisor, ***Dr Lalitha S. Kumar****, who has supported me throughout my PhD with patience and I salute her efforts and valuable guidance. One simply could not wish for a better or friendlier supervisor and mentor. Her valuable insight has been invaluable on both an academic and a personal level, for which I am extremely grateful to her.*

I am also extremely indebted to my co-guide ***Dr. Sushama M. Gaikawad****, for providing necessary infrastructure and resources to accomplish my research work. My sincere thanks to her for the valuable advice and, constructive criticism. I am also grateful to* ***Dr. Meenakshi V. Rele*** *for her valuable suggestions during the initial course of investigation. My sincere thanks to* ***Dr CG Suresh*** *and* ***Mr Priyabrata Panigrahi*** *(NCL) for helping me in the homology modeling studies, I am also thankful to* ***Dr. B. M. Khan, Dr Urmil Mehta****,* ***Dr. Absar Ahamad****,* ***Dr. R.V. Gadre*** *and* ***Dr. Rishikishore Vishwakarma*** *for allowing me to use some of the facilities in their laboratories. My thanks are due to* ***Dr Varsha Pardeshi*** *for her valuable suggestions during RAPD and ISSR work. I am also thankful to* ***Ms Rasika Bhagawat*** *for her help in the genetic diversity analysis.*

I am also extremely indebted to ***Dr. Vidya Gupta,*** *the present Chair of Biochemical Sciences Division,* ***Dr. S. Sivaram****, Ex-Director, National Chemical Laboratory, Pune (India) and present Director,* ***Dr. Sourav Pal****, for providing the necessary infrastructure and resources to accomplish my research work.*

Most of the results described in this thesis would not have been obtained without a close collaboration with few banana research centers. I owe a great deal of appreciation and gratitude to ***Dr. B. Padmanaban****, Senior entomologist, NRCB Trichy, for help in collection of the insect samples from Trichy, Narayangaon and Wayanad. I thank* ***Dr. N. M. Patil****, entomologist, Banana Research Centre, Jalgaon, for his help in collection of samples from Jalgaon. I would like to thank* ***Dr. D. N. Kalita****, Programme co-ordinator, Krishi Vigyan Kendra Kamrup Assam, for providing the insect samples from this region.*

*My thanks are due to **Ms Saroj Devi**, Sarpanch, Saidpur (Hajipur, Vaishali), for her help in sample collection. I want to thank **Dr. Vishwas M. Kulkarni**, Scientist, BARC, Mumbai for help in DNA sequencing of the clones.*

I take this opportunity to sincerely acknowledge the Council of Scientific and Industrial Research (CSIR), Government of India, New Delhi, for providing financial assistance in the form of Junior Research Fellowship which allowed me to enroll for this PhD programme.

I extend my thanks to Ms Padmavati Sahare who helped me with my accommodation when I initially shifted to Pune. I feel highly indebted to the Students Academic Office for providing family quarters till the completion of my PhD.

I am indebted to student trainees who worked with me for their M.Sc. projects, in particular to Ms Sohini, Ms Namita and Ms Nishansala who helped me a lot. My special appreciation goes to Sana for her friendship and encouragement. This work would not have been possible without the cooperation and help of my friends - Poonam, Madhurima, Sonali, Sayali, Avinash, Priyabrata, Ravi, and Shadab, I am also thankful to my lab members - Trupti, Sagar, Ashwini, Shweta and Anirban and all friends from FB for their best wishes.

I would like to express my gratitude to scientists and staff of Biochemical Sciences Division, NCL, for their help during the course of my stay in the Division. Special thanks to Mr Trehan, Mr Giri, Mr Ramakant, Mr Jagtap, Mrs Indira Mohandas and Sheetal.

This thesis would not have been possible without the support and patience of my husband (Rajkumar) and words fail me to express my feelings for him. He has stood by me in my toughest of times and was my constant source of encouragement and motivation. I am indebted to my brother (Bhaskar Vikram Shankar) and sister (Dr Padmini Shankar) for their moral support and encouragement.

Last but not the least, I would like to pay high regards to my Mummy and Daddy for their constant love and faith in me. Their sincere encouragement and inspiration throughout my research work has enabled me to reach this stage in my life. I owe everything to them.

Pallavi Shankar

ABBREVIATIONS

Arg	arginine
Asp	aspartic acid
Cys	Cysteine
DNA	deoxyribonucleic acid
DNSA	Di-nitro Salicylic Acid
FPLC	fast protein liquid chromatography
Gly	glycine
His	histidine
Hour	h
HPA	human pancreatic α-amylase
HSA	human salivary amylase
LB	Luria Bertani Broth
Lys	Lysine
MALDI-TOF	matrix-assisted laser desorption/ionization-Time of flight
Minutes	min
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PDB	protein data bank
PPA	porcine pancreatic amylase
Pro	Proline
RBI	ragi bifunctional inhibitor
rpm	revolutions per minute
SDS	sodium dodecyl sulphate
Tm ^{o}C	melting temperature
TMA	*Tenebrio moliter* amylase
Trp	tryptophan
Tyr	tyrosine
WMAI	wheat monomeric alpha amylase inhibitor
WDAI	wheat dimeric alpha amylase inhibitor

ABSTRACT

Banana is the fourth most important crop in the developing world and India is the world's largest producer of bananas and plantains. In India, eight pests commonly infest the banana crop; of these, the banana rhizome weevil *i.e. Cosmopolites sordidus* (Germar) and the banana pseudostem weevil *i.e. Odoiporus longicollis* (Oliver) are the major pests. *O. longicollis* (Oliver) is a monophagous pest of banana and the larvae as well as adults cause severe damage affecting the production of bananas. The loss varies from 10% to almost 90% depending on the stage of plant growth at which pest infestation occurs and also the efficiency of the management or cultivation practice that is followed.

Use of conventional breeding to transfer resistance is restricted due to parthenocarpy, non-seeded nature and male/female sterility of the banana plants. Therefore, there is an increasing interest in developing banana cultivars resistant to banana stem weevil using biotechnological approaches. A broad integrated pest management strategy has the potential to provide the best strategy for controlling this pest. The components of a successful IPM programme would include identifying host plant resistance, cultural and biological control, study of the population structure of the pest and identifying useful genes whose products would have a detrimental effect on the growth and development of the pest, so that a biotechnological approach can be used to develop banana transgenics carrying such genes.

In this thesis, an attempt has been made to study the following two important components which could contribute to the IPM for the control of this pest:

(i) Study of the population structure of the pest *i.e.* assessment of the genetic variation between and within populations and how this variation is partitioned geographically.

(ii) Identification of useful genes whose products would deter the growth and development of the pest so that these genes could be used in the future for developing banana transgenics with improved tolerance/resistance to *O. longicollis* (Oliver).

Genetic diversity analysis of thirty adult beetles of *Odoiporus longicollis* (Oliver) representing six populations collected from infested banana stems from banana fields in Assam (Kamrup), Bihar (Vaishali), Kerala (Wayanad), Maharashtra (Jalgaon,

Narayangaon) and Tamilnadu (Trichy) was done using both nuclear and mitochondrial markers including RAPDs, ISSRs, AFLPs, rDNA (ITS1 and ITS2) and COI-tRNA^{Leu}-COII markers. The data generated using each of these markers has been analyzed phylogenetically and statistically to determine the utility of each of these marker systems in genetic diversity analysis. The RAPDs and ISSRs based UPGMA dendrograms did not reveal any phylogeographic clustering of the populations. The AFLP based dendrogram showed a strong correlation between geographic and genetic distance, suggesting that AFLPs are more useful than RAPDs and ISSRs in the present study.

The phylogenetic analysis of the six weevil populations based on the ITS1 and ITS2 regions reveals that there is gene flow between these populations and there is no phylogeographical distribution of these six populations. Among the six populations, the Assam population is the oldest as **(i)** it shows the highest genetic diversity based on the ITS1 and ITS2 sequence analysis **(ii)** it is separated from the other five populations by a longer branch length in the phylogenetic trees derived by AFLPs, ITS1 and ITS2 data. This observation supports the initial migration of this pest into India via the north-east, from the centre of origin of bananas in south-east Asia which is considered to be the primary centre of diversification and earliest domestication of this crop

Consensus secondary structures of the ITS1 and ITS2 regions of *O. longicollis* (Oliver) are also presented. The consensus secondary structures of the ITS2 region conform to the pan-eukaryotic model. In the ITS1 secondary structure phylogenetic tree, the individuals grouped according to their secondary structures such that each group showed specific structural characteristics. The phylogenetic trees based on the secondary structures of ITS1 and ITS2 of the thirty *O. longicollis* (Oliver) individuals were congruent, though the secondary structure of these two regions bears no resemblance to each other, suggesting that secondary structures of these two regions are important `markers' in phylogenetic analysis. The present study of the secondary structure of the ITS1 and ITS2 regions of *O. longicollis* (Oliver) and its use in assessing the phylogenetic relationships is the first detailed report for insects, in particular family Curculionidae.

Sequence analysis of the mitochondrial COI-tRNA^{Leu}-COII region of the thirty individuals reveals AT bias which is typical of insect mitochondrial DNA. The nucleotide composition of the partial COI and COII genes is AT rich as observed in insect

mitochondrial genes. There was no phylogeographic distribution of the populations. The Fu and Li's D and F tests were non-significant for this mitochondrial region. No *Wolbachia* infection was detected in any of the populations. The genetic differentiation amongst the populations was highly significant ($p<0.001$; $\chi^2 = 123.333$; $df=75$), suggesting restricted gene flow between the populations. This result did not correlate with that obtained with nuclear markers *i.e*. ITS1 and ITS2, suggesting a male-biased gene flow between the populations.

Alpha-amylases and proteinases are important digestive enzymes which play a central role in the digestive metabolism in those insects that live on seeds and plant parts. The larvae and adults of banana pseudostem weevil feed voraciously on the stem parts and complete their development within the stem. Hence they depend to a large extent on their proteinases and alpha-amylases for their survival. The purified monomeric and dimeric alpha-amylase inhibitor fractions from a local variety of wheat seeds completely inhibit the alpha amylase activity from this pest and do not inhibit porcine pancreatic amylase and human salivary amylase. These characteristics make these inhibitors attractive candidates for genetic engineering to develop banana transgenics with improved tolerance towards banana stem weevil. In this thesis, cloning of the partial gene encoding the amylase (the cloned segment includes all the conserved regions and the catalytic domain) and cloning of the genes encoding the monomeric and dimeric α-amylase inhibitors from a local variety of wheat has been described. The molecular basis of the specificity of inhibition between the α-amylase and the inhibitors has been studied by homology modeling. Such a study is the first step towards developing banana transgenics using such genes.

CHAPTER 1

The agricultural significance of *Odoiporus longicollis* (Oliver)

INTRODUCTION

Banana is the fourth most important crop in the developing world and India is the world's largest producer of bananas and plantains accounting for almost one-fifth of the world's production (Padmanaban *et al.* 2001a, b; Mustaffa & Sathiamoorthy 2004; Gailce Leo Justin *et al.* 2008). In India, about eight pests commonly infest the banana crop; but of these, the banana rhizome weevil *i.e. Cosmopolites sordidus* (Germar) and the banana pseudostem weevil *i.e. O. longicollis* (Oliver) are the major pests (Valmayor *et al.* 1994; Sripriya *et al.* 2000; Padmanaban & Kandasamy 2003). *O. longicollis* (Oliver) is a monophagous pest of banana which belongs to the order Coleoptera and family Curculionidea. The larvae as well as adults cause severe damage affecting the production of bananas (Ostmark 1974; Visalakshi *et al.* 1989; Gailce-Leo Justin *et al.* 2008). The larvae feed voraciously on the tissues of the succulent sheath and may reach the true stem. Infestation by this pest can occur at different development stages of the plant growth. Infestation at the pre-flowering stage results in non-emergence of the flower. Due to the tunneling activity of the larvae, the stem rots and falls down due to loss of tensile strength (Padmanaban & Sathiamoorthy 2001). The banana plants are also damaged indirectly by the tunneling activity of the larvae because the damaged areas provide easy access to rot promoting organisms (Gold *et al.* 2001). The loss varies from 10% to almost 90% depending on the stage of plant growth at which pest infestation occurs and the efficiency of the management or cultivation practice that is followed (Padmanaban & Sathiamoorthy 2001).

Geographical distribution of *O. longicollis* (Oliver)

Worldwide distribution of *O. longicollis* (Oliver) includes diverse geographical locations including Africa and Asia (Bhutan, China, Guizhau, Hong Kong, Indonesia, Java, Sulawesi, Sumatra, Japan, Laos, Malaysia, Sabah, Myanmar (Burma), Nepal, Nicobar Islands, Phillippines, Sri Lanka, Taiwan and Vietnam) (Valmayor *et al.* 1994). Life cycle, control measures and status of the pest in different countries has been investigated by the following researchers: Java (Froggatt 1928); Kwangtung, Hong Kong

(Hoffman 1933); Sri Lanka (Jepson 1935; Dias 1936); Taiwan (Kung 1955); China, Malaysia, Indonesia and Thailand (Valmayor *et al.* 1994); Vietnam (Kiem 1995); China (Luo *et al.* 1985; Zhou & Wu 1986); Kathmandu valley, Nepal (Lefroy 1909; Froggatt 1928; Singh 1966; Shukla & Kumar 1970; Isahaque 1978); China, Japan, Andaman Isaland, Myanmar (Shukla & Kumar 1969 1970); Java, Sri Lanka (Alonso-Zarazaga *et al.* 1999).

Distribution of the pest in India

O. longicollis (Oliver) is a serious pest in Nepal and India (Simmmonds 1966; Singh 1966; Shanmugavelu 1992; Thapa 1993; Shrestha 1994; Sripriya 2000). The pest is widespread in different geographical regions of India such as: Tamil Nadu (Padmanaban *et al.* 2001); Kerala (Visalakshi *et al.* 1989); Manipur (Prasad & Singh 1989; Mathew *et al.* 1997); Assam (Isahaque 1978); Gorakhpur -U.P. (Shukla & Tripathi 1978); Eastern Uttar Pradesh (Shukla & Kumar 1969 1970); Cachar, Sibsagar and Dalfa Assam; Kolkata, Murshidabad, Behrampore, Siliguri, Darjeeling, Kalimpong and Pashok; Purnea West Bengal; Andaman Islands (Lefroy 1909); Hooghly, Howrah, Burdwan, 24-Parganas, Midnapur, Nadia, Murshidabad; Kalyani-West Bengal (Dutt & Maiti 1972a); Imphal, Manipur (Ram & Pathak 1987); Gujarat (Patel & Jagadale 2003); Bihar (Tiwary 1971); Jammu and Kashmir (Azam *et al.* 2010).

Life cycle of *O. longicollis* (Oliver)

All life stages are present on the plant throughout the year (**Fig. 1**). The adult weevils are black-colored (red colored morphs) and measure 23 - 39 mm. The life cycle of the pest is shown in **Fig. 2**. The pest has a pre-oviposition period of 15 - 30 days. The adult weevils mate throughout the day and night and lay an average number of one egg per day for 9 days after a single mating. The female lay yellowish white, 3.14 x 1.1 mm sized (**Fig. 1**). These elliptical eggs are deposited on the outer epidermal layer of the leaf sheath of the pseudostem through an ovipositional slit cut by the rostrum. The incubation period may range between 3-8 days. The larvae are fleshy, yellowish white and apodous.

The larva pass through 5 instars in the process of development and the fifth instar larva enters a nonfeeding pre-pupal stage, forms a cocoon and the exarate pupa remains inside the cocoon. The developmental rates vary and are highly dependent on climatic factors with the duration of the life stages longer in the winter season than in the summer. Under laboratory conditions this whole process from egg to adult is completed in 44 days (Padmanabhan & Sathiamoorthy 2001). Two population peaks observed by several authors are April-May and September-October (Zhou & Wu 1986). The activity is high from February - November (Prasad & Singh 1988).

The density is high from late May to June and from late September to mid October and hence heavy damage of the crop occurs during this period (Luo 1985). The adult weevil is robust, reddish brown and black measuring 1.3–2cm (Singh 1966) (**Fig. 1**). The weevil has a long life span and many adults live for a year. Since the adults are strong fliers they have the ability to spread quickly among the field.

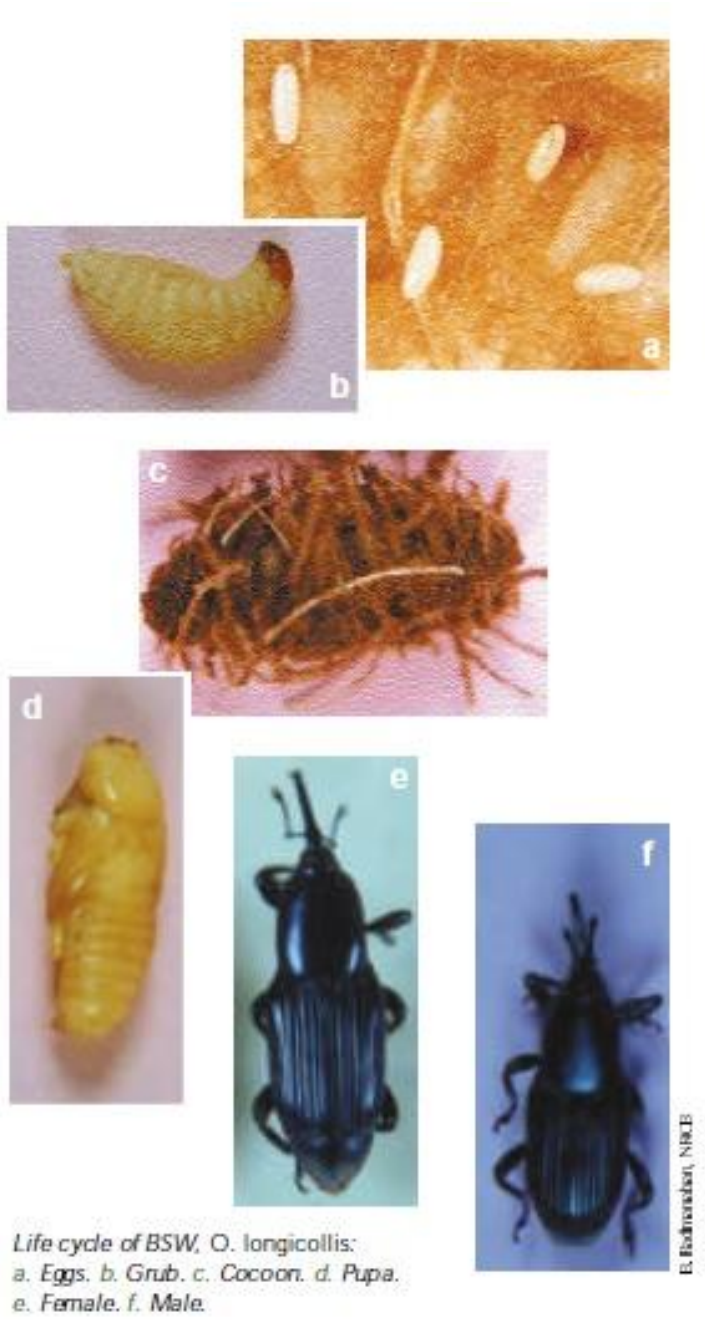


Fig. 1. Different stages in the life cycle of *O. longicollis* (Oliver) (Padmanaban & Sathiamoorthy 2001) **a.** Eggs, **b.** Grub, **c.** Cocoon, **d.** Pupa, **e.** Adult female, **F.** Adult male

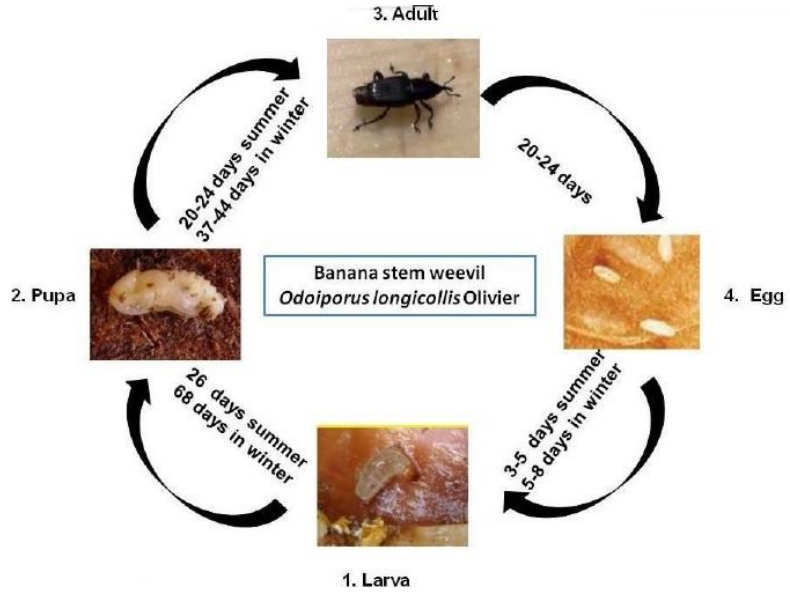


Fig. 2. Life-cycle of *O. longicollis* insects*

***Image taken from the website of Department of Agriculture and Cooperation Ministry of Agriculture Government of India (http://www.agritech.tnau.ac.in/expert_system/banana/crop protection.html#4)

Nature of damage caused by *O. longicollis* (Oliver)

Infestation starts in 5 month old plants and can be recognized by the presence of small pinhead sized holes on the stem and fibrous extrusions from bases of leaf petiole**.** Justin *et al.* (2008) has reviewed bionomics and management of banana pseudostem weevil. *O. longicollis* (Oliver) is a monophagous pest of banana and the larvae as well as adults cause severe damage affecting the production of bananas (Ostmark 1974; Visalakshi *et al.* 1989; Gailce-Leo Justin *et al.* 2008). According to the reports, one of the most important visible symptoms of infestation include exudation of sap from the leaf sheaths, yellowing and whitening of leaves, immature ripening of fruits, reduction in bunch size and toppling of infested pseudostem **(Fig. 3)**.

Fig. 3. Various stages of damage to the host plant (Padmanaban & Sathiamoorthy 2001)

Both adult and larvae can be found in the pseudostem region of plants. The larvae feed voraciously and in the process they reach deep into the pseudostem and cause severe damage to the plants, whereas adults feed under the leaf sheath (Shukla & Kumar 1970; Isahaque 1978). The extensive tunneling activity of larvae may reach the true stem, the length of which ranges between 8-10 cm. These widespread tunnels may reach the fruit peduncle and rhizome. If it reaches the fruit peduncle, larval activity may lead to non-emergence of flower bud which decays inside the pseudostem (Padmanabhan *et al.* 2001). Weakening of the stem due to tunneling activity of larvae often result in breakage by wind because the plant is unable to bear the weight of the bunch (Padmanaban & Sathiamoorthy 2001). Infestation by this pest is a serious problem in several banana growing areas of India. Since the larvae complete a major part of their life-cycle within the stem, they are not easily accessible to sprayed pesticides. Moreover, use of insecticides gives rise to several environmental issues and has also resulted in the emergence of insecticide resistant weevil strains (Shanahan & Goodyer 1974; Wright 1977; Swaine *et al.* 1980). The banana plants are also damaged indirectly by the tunneling

activity of the larvae because the damaged areas provide easy access to rot promoting organisms (Gold 2001). The loss varies from 10% to almost 90% depending on the stage of plant growth at which pest infestation occurs and also the efficiency of the management practice that is followed (Padmanaban & Sathiamoorthy 2001).

Management of *O. longicollis* (Oliver)

Management of this pest is necessary to sustain productivity and obtain higher economic returns. Different control methods to curb damage are available which depends upon the type of banana production systems being practiced. The control practices range from use of chemical insecticides to cultural control strategies depending on the size of the field and availability of resources to the farmers. Presently, a combination of chemical and cultural methods are being practiced to control banana pseudostem weevil (Padmanaban & Sathiamoorthy 2001).

Cultural control

Cultural practices can help in the control of the pest to some extent. Such practices include field sanitation, use of healthy suckers, eradication of infested stems, removal of dried old leaves, early detection of infestation, periodic pruning of suckers and removal as well as destruction of infested pseudostems (Tiwary 1971). The disc-on stump and longitudinal split pseudostem traps can be efficiently used to monitor and reduce the pest population. The disc-on stump is effective due to the higher exudation of the plant fluids (Padmanabhan & Sathiamoorthy 2001). Banana pseudostem (longitudinal) @100/ha effectively traps all banana weevils. Cultural practices such as use of clean planting material, field sanitation, systematic trapping of adult weevils, etc, are labour intensive and not economically feasible for small household farmers who are the main producers of bananas (Sikora *et al.* 1989).

Chemical control

Chemical control methods are preferred due to their ease of application and their compatibility with other methods of control. Control of *O. longicollis* (Oliver) is a difficult and complex task because the pest completes almost all of its life cycle inside the pseudostem. Application of organochlorine insecticides has resulted in the development of insecticide resistant weevil strains and environmental consequences are also a major concern (Padmanaban & Sathiamoorthy 2001). Systemic injection of organophosphorus compound (*e.g.* monocrotophos), swabbing along with surfactants, swabbing with mud slurry containing the candidate insecticide, spraying and fumigation of the spaces between the leaf sheaths in the pseudostem can control damage to some extent but they have their own limitations, are labour intensive and can be discouraging for marginal farmers. Fumigation by Celphos (aluminium phosphide tablets), especially during the vegetative phase is phytotoxic and should not be used (Padmanaban & Sathiamoorthy 2001). Application of pesticides i.e. carbofuran 3g @ 30g/plant at planting and @ 15g/plant at 60^{th} and 90^{th} day after planting and spray application of quinalphos 0.05% or chlorpyriphos 0.03% or carbaryl 0.2% at planting, is also practiced. In case of severe infestation, spraying may be repeated after 3 weeks (Gailce Leo Justin *et al.* 2008). Bananas are the only food crop harvested daily on the same area for years, which greatly restricts the use of insecticides that require a certain period between application and harvest.

Biological control

The work done on biological control of this pest is limited and success reports are fewer. One report from China describes the use of two species of earwigs feeding on larvae and pupae China (Valmayor 1994) There are studies with limited success done using parasitic acarid mite for larvae and adults; ectoparasitic mites, *Uropodia sp.* on adult banana pseudostem weevils; *Metarhizium anisopliae*, an entomopathogenic fungus; fungal pathogens such as *Fusarium solani*, *Mucor heimalis*, *Aspergillus niger* and *Scopulariopsis brevicaulis*; Entomopathogenic nematodes are under study (Padmanaban

& Sathiamoorthy 2001). Among several endophytic entomopathogens *Beauveria bassiana* offers a promising option for the management of the banana weevil. This entomopathogen survives as an endophyte in a wide range of plants, offering substantial protection against tunneling insect pests. Akello *et al.* (2008a, b) determined the effects of endophytic *B. bassiana* on oviposition rate and populations of the banana weevil, and their plant damage. Results of green house trials using *B. bassiana* inoculation in tissue-cultured banana plants showed a mortality of 53.4 -57.7% and a reduction of plant damage by 29.1–62.7% demonstrating for the first time that endophytic *B. bassiana* could be used as an effective biological control agent for banana pseutostem weevils.

Akello *et al.* (2007) demonstrated that that all *B. bassiana* strains were able to colonize banana plant roots, rhizomes and pseudostem bases. Their study also demonstrated that, depending on the strain and inoculation method, *B. bassiana* can form an endophytic relationship with tissue culture banana plants, causing no harmful effects and might provide an alternative method for biological control of *C. sordidus*. Strains of *Beauveria bassiana* virulent against the banana weevil have been identified in East Africa (Nankinga 1994), West Africa (Godonou *et al*. 2000) and Latin America (reviewed by Gold *et al.* 2001). For example, laboratory bioassays using different *B. bassiana* strains originating from insect cadavers and soil samples in Uganda resulted in adult banana weevil mortalities of up to 100% (Nankinga 1994). In laboratory bioassays conducted by Godonou *et al.* (2000), *B. bassiana* strain IMI330194 was identified as a possible control agent for *C. sordidus* based on its virulence and its potential for mass production. *Griesbach (2000)* obtained 200 isolates of endophytes from 64 recently harvested banana plants in Ntungamo district, Uganda. Spore suspensions of 12 isolates (8 *Fusarium*, 3 *Acremonium*, 1 *Geotrichium*) caused 80-100% mortality in weevil eggs, while 74 additional isolates caused 60-79% mortality. Two *Fusarium* strains also produced 30-48% mortality of weevil larvae.

Sex pheromone-based control

It has been reported that the females of *O. longicollis* (Oliver) produce sex pheromones as detected by bio-assay trials (Ravi & Palaniswami 2002; Prasuna *et al.* 2008). But olfactory behavior of this pest shows that it is under the influence of chemicals produced by the weevils themselves and the host plant. This is the reason for the success of the pseudostem trapping method. There has been an indication of the existence of female-produced sex pheromone system in the pest, but the olfactory activity of both male and female is mediated by the male-produced aggregation pheromone (Prasuna *et al.* 2008). This finding indicates the utility of synthetic aggregation pheromone in integrated pest management of this pest as it would attract both male and female weevils for trapping, leading to higher efficiency due to specificity. There are reports of the presence of sex pheromones in other Curculionids (Jacobson 1966; Cross & Mtchell 1966; Cofflet *et al.* 1978). The use of sex attractant such as these can be used to attract the males at low levels and at higher dose to confuse them thereby leading to reduction in number of pests in the long run (Ravi & Palaniswami 2002). This technique can be an alternative eco-friendly technique for the integrated pest management of these weevils. Fifteen compounds consisting of sex attractants and aggregating pheromones for different pest species have been identified (Keller 1964; Hardee *et al.* 1969; Palaniswamy & Underhill 1988; Yadav & Yadava 2000). Sordidin, a commercial aggregation pheromone which was identified by Budenberg (1993) is being used against *C. sordidus* (Jayaraman 1997).

Host plant resistance

Screening *Musa* germplasm for resistance to banana pseudostem weevil can be a potential tool to identify the source of resistance genes that could be used in plant breeding programmes and development of IPM strategies. *O. longicollis (*Oliver) being a monophagous pest exhibits a high degree of host plant preference. Charles *et al.* (1996) screened 212 banana accessions of various genomes of which 27 accessions exhibited tolerance to the pest (Charles *et al.* 1996). In this regard *Musa balbisiana* can be an

important source of resistance for the development of resistant varieties based on several interesting properties. Wild *Musa balbisiana* is a collective group of many 'wild types' with no specific subspecies status and it shows resistance to various biotic and abiotic stresses in being immune to biotic stress of fusarium wilt, leaf spot diseases like Sigatoka (*Mycosphaerella musicola, Musa fijiensis, Musa eumusae*), Cordana and Septoria leaf spot, rust and bacterial diseases like head rot (*Erwinia* spp.). *Musa balbisiana* has been found to be tolerant to pseudostem weevil and rhizome weevil. *Musa balbisiana* is also highly tolerant to severe drought, cold and poor soil conditions as well. At the National Research Centre for Bananas, Trichy, laboratory screening of 119 accessions led to the identification of a high degree of resistance in *Musa balbisiana* clones, such as Bhimkol, Athiakol, Elavazhai and Sawai. In general, plantains are the most preferred host for this pest (Padmanaban & Sathiamoorthy 2001). As bananas are perennial *i.e.* the new plants emerge as suckers from the rhizome, the pest problem gets worse with successive ratoon crops (Biocontrol News and Information, September 2000, 2(3)). The modern commercial bananas are diploid, triploid and tetraploid, and are mainly derived from two diploid wild species *i.e. Musa acuminata* (A genome) and *Musa balbisiana* (B genome). The wild species are seed bearing whereas the cultivars are sterile. The problems of parthenocarpy, non-seeded nature and male/female sterility in addition to the triploid nature of the best cultivars, limits their use as parents (Uma 2006). Wild *Musa balbisiana* is tolerant to both the banana weevils *i.e. C. sordidus* and *Odoiporus longicollis* (Mesquita *et al.* 1984; Kiggundu *et al.* 2003; Uma 2006); but a major problem in the use of *M. babisiana* as a donor of the resistance is the integration of Banana Streak Virus (BSV) in its genome (Uma 2006). Some of the hybrids developed using *M. balbisiana* have expressed symptoms of BSV (Uma 2006).

Control using natural enemies

Other unconventional measures include use of natural enemies including arthropods, entomopathogenic nematodes and entomopathogens which have great potential to reduce the population of the weevils in severe infestation. Hussaini (2003) has studied the identification and distribution of entomopathogenic nematodes in India, their

interaction with the soil, virulence, storage, compatibility with pesticides, formulation, mass production and their bioefficacy against *Agrotis ipsilon, A. segetum, Chilo partellus, Holotrichia consanguinea, Spodoptera litura, Plutella xylostella, Opisina arenosella, Pthorimaea operculella, Leucinodes orbonalis, Lasioderma serricorne* and *O. longicollis.*

.

Bt toxins

Though several Bt toxins have been shown to be effective against a number of insect pests (de Maagd *et al.* 2001), to date no *Bt* gene has been identified that is effective against both the banana weevils (Kiggundu *et al.* 2003; Montesdeoca *et al*. 2005).

Development of molecular approaches to aid the present IPM for the control of *O. longicollis* (Oliver) - Scope of this thesis

Most of the research on *O. longicollis* (Oliver) has focused on studying different control methods, management strategies, bionomics and evaluation of *Musa* germplasm (Dutt & Maiti 1972a, b; Reghunath *et al.* 1992; Charles *et al.* 1996; Mathew *et al.* 1997; Padmanaban & Sathiamoorthy 2001; Padmanaban *et al.* 2001a, b; Ravi & Palaniswami 2002; Padmanaban & Kandasamy 2003; Prasuna *et al.* 2008).There are many gaps in the knowledge of *O. longicollis* (Oliver) that require further detailed studies. For example, population dynamics and bionomics of the pest are not well understood and studies on yield loss due to this pest are incomplete. The effect of different banana production systems on pest populations must also be studied and economic threshold levels should be estimated for the pest in these different production systems. Screening trials currently utilize a number of different methodologies to identify resistant clones. Screening methodologies require further refinement and more attention must be paid to identify resistant and tolerant genotypes, as host plant resistance is a potential viable long-term control strategy. The different mechanisms of resistance must be studied in detail in order to gain the necessary knowledge for determining the selection criteria for utilization of resistance sources in breeding strategies, which could thus be made more cost-efficient and less time consuming. Development of molecular approaches which is the focus of the

present thesis, could also contribute to the present IPM of this banana pest.

In the present thesis, an attempt has been made to study the following two important components which could contribute significantly to the IPM for the control of this pest. These include:

1. **Study of the population structure of the pest *i.e*. assessment of the genetic variation between and within populations and how this variation is partitioned geographically.**
2. **Identification of useful genes whose products would deter the growth and development of the pest so that these genes could be used in the future for developing banana transgenics with improved tolerance/resistance to banana pseudostem weevil.**

The need to study genetic diversity of O. longicollis (Oliver) populations

To date, there are no reports on molecular biology studies of this pest and this thesis presents the first molecular studies on this important banana pest. Management of *O. longicollis* (Oliver) is a pre-requisite to sustain productivity and to obtain higher economic returns of bananas and plantains. A very important aspect for developing a successful Integrated Pest Management (IPM) protocol for the control of this pest is the study of the population structure i.e. the genetic variability of the pest needs to be assessed between and within populations. It is also important to determine how this variability is partitioned geographically. Knowledge of the structure of the pest population is important for identifying and characterizing the resistance, since resistance genes are defined in terms of their effect on a pest population. When a resistance gene is deployed in a field, it faces a population of the pest, and therefore, it is extremely important to characterize each resistance gene in terms of its spectrum of resistance relative to the pest population. Quantifying the genetic diversity within and between the pest populations from different hot-spot locations in India, will definitely add knowledge which could contribute to an important aspect of IPM for the control of this banana pest. Such studies could also throw light on pest migration, and dispersal. Knowledge about pest dispersal and movement and their relationship to pest management practices, environmental factors

and ecosystem can be used to predict pest infestations. An understanding of pest dispersal is important for the success of IPM programmes (Aylor & Irwin 1999; Byrne 1999). In the present work, the genetic diversity within and between six populations of the pest collected from different hot-spot locations in India, have been studied using various nuclear and mitochondrial markers.

The need to develop genetic engineering approaches for the control of O. longicollis (Oliver)

O. longicollis (Oliver) completes all its larval stages within the stem and hence is not easily accessible to sprayed pesticides. Though certain wild species show increased tolerance/resistance to this pest, transfer of resistance is restricted due to problems of parthenocarpy, non-seeded nature and male/female sterility in addition to the triploid nature of the best cultivars, Hence, in such a case, identifying insecticidal genes that would have a detrimental effect on the growth and development of the pest and transforming banana cultivars with such genes would be the ideal approach towards developing banana plants with improved resistance to this pest.

Though several Bt toxins have been shown to be effective against a number of insect pests (de Maagd *et al.* 2001), to date no Bt gene has been identified that is effective against both the banana weevils (Kiggundu *et al.* 2003; Montesdeoca *et al.* 2005). NARO and the Centre de Cooperation Internationale en Recherche Agronomique pour le Development in France (CIRAD) are screening newly isolated Bt toxins for banana weevil toxicity (Kiggundu *et al.* 2003).

Proteinaceous inhibitors to the digestive enzymes of the larvae are attractive candidates for genetic engineering, because they have always been an integral part of the plant's defense against its natural pests. All plants and seeds have varying degrees of these inhibitors which are regularly ingested by humans and animals in the diet and do not produce any harmful effects. Though there are no reports of the use of insecticidal genes for the control of *O. longicollis* (Oliver), insecticidal genes such as papaya cystatin and oryzacystatin I have been shown to have a detrimental effect on the related banana pest *C. sordidus* (Germar) (Kiggundu *et al.* 2003). Remarkable progress has been made in

banana transformation (May *et al*. 1995; Sagi *et al.* 1995; Cote *et al.* 1996; Becker *et al.* 2000; Matsumoto *et al.* 2002; Pei *et al.* 2005). The Uganda National Agricultural Research Organization (NARO), the Forestry and Agricultural Biotechnology Institute of the University of Pretoria (FABI), the International Institute of Tropical Agriculture (IITA) and the University of Leeds have identified and expressed oryzacystatin I (OC-I) and papaya cystatin in banana and have shown that this has a detrimental effect on the development of *C. sordidus* weevils (Kiggundu *et al.* 2003). Site-directed mutagenesis of these inhibitors is also being worked out to improve the stability and inhibition of these two inhibitors (Kiggundu *et al.* 2003). Kiggundu *et al.* (2010) have shown that rice cystatin OsCYS1, papaya cystatin (CpCYS1), cathepsin-B and Cathepsin-L like activities were effective against banana corm weevil in bioassay trials, suggesting their use in developing transgenic banana lines resistant to corm weevil. The John Innes Centre and the University of Leeds, UK, have reported transgenic resistance to the nematode *Radopholus similis* conferred on banana by a modified rice cystatin. (Atkinson *et al.* 2004). The effect of soybean kunitz trypsin inhibitor on in-vitro *C. sordidus* larval performance has been studied (Montesdeoca *et al.* 2005).

NARO, IITA and the African Centre for Banana and Plantain in Cameroon (CARBAP) are developing conventional breeding programs using diploid parents showing resistance to banana weevils (Kiggundu *et al.* 2003)

Transgenic banana plants with improved tolerance to fungal diseases have been developed by *Agrobacterium*-mediated transformation using different genes including rice chitinase gene and a β 1, 3 – glucanase gene (Maziah *et al.* 2007a, b), the GbSGT1 gene (Shengjun *et al.* 2007), chitinase and glucanase genes (Sreeramanan *et al.* 2006 a, b) and stibene synthase, superoxide dismutase and endochitinase genes (Vishnevetsky *et al.* 2011). Pei *et al.* (2005) have demonstrated that transgenic bananas expressing the human lysozyme gene exhibited increased resistance to Panama wilt disease.

Why target the digestive enzymes of O. longicollis* (Oliver)*?

α-amylases and proteinases are important digestive enzymes which play a central role in the digestive metabolism in those insects that live on seeds and plant parts. The larvae and adults of banana pseudostem weevil feed voraciously on the stem parts and complete their development within the stem. Hence they depend to a large extent on their proteinases and α-amylases for their survival. Hence, inhibitors to these digestive enzymes would be attractive candidates for the control of this pest. It is also necessary to identify a proteinaceous inhibitor rather than a non-proteinaceous inhibitor to the amylase of this banana pest, because the former would be the product of a single gene and will be relatively easier to manipulate using genetic engineering strategies, whereas the latter would be a product of a complex metabolic pathway and will be more difficult to manipulate using genetic engineering strategies. An ideal inhibitor should specifically inhibit the pest amylase and should not show any inhibitory effect on mammalian amylases.

The purified monomeric and dimeric alpha-amylase inhibitor fractions isolated in this laboratory, from a local variety of wheat seeds, completely inhibits the alpha amylase activity from this pest and do not inhibit porcine pancreatic amylase and human salivary amylase (Sharma *et al.* 2009). These characteristics would make these inhibitors attractive candidates for genetic engineering to develop banana transgenics with improved tolerance towards banana pseudostem weevil. In the present work, the partial gene, encoding the amylase as well as the complete genes encoding the inhibitors have been cloned and the specificity of interaction between the amylase and the wheat inhibitors has been studied by homology modeling using the available 3D structure of *Tenebrio molitor* amylase. These studies will throw light on the enzyme-inhibitor interaction and would also serve as a base to design inhibitors with more desirable characteristics. These studies are important before these inhibitor genes can be used to develop banana transgenics.

REFERENCES

Alonso-Zarazaga MA, Lyal CH and Vinolas A (1999) Vol. 315. *Barcelona: Entomopraxis,* ISBN: 84-605-9994-9.

Akello J, Dubois T, Gold CS, Coyne D, Nakavuma J and Paparu P (2007) *Journal of Invertebrate Pathology*, **96(1)**, 34-42.

Akello J, Dubois T, Coyne D and Kyamanywa S (2008a) *Crop Protection*, ***27*(11)**, 1437-1441.

Akello J, Dubois T, Coyne D and Kyamanywa S (2008b) *Entomologia Experimentalis et Applicata,* **129(2)**, 157-165.

Atkinson HJ, Grimwood S, Johnston K and Green J (2004) *Transgenic Research,* **13(2)**, 135-142.

Aylor DE and Irwin ME (1999) *Agricultural and Forest Meteorology*, **97(4)**, 233-234.

Azam M, Tara JS, Ayri S, Feroz M and Ramamurthy VV (2010) *Munis Entomology and Zoology*, ***5*(2)**, 627-635.

Becker DK, Dugdale B, Smith MK, Harding RM and Dale JL (2000) *Plant Cell Reports*, **19(3)**, 229-234.

Broadway RM and Duffey SS (1986a) *Journal of Insect Physiology*, **34(12)**, 1111-1117.

Broadway RM and Duffey SS (1986b) *Journal of Insect Physiology*, **32(10)**, 827-833.

Boulter D (1993) *Phytochemistry*, **34(6)**, 1453-1466.

Budenberg WJ, Ndiege IO and Karago FW (1993) *Journal of Chemical Ecology,* **19(9)**, 1905-1916.

Byrne DN (1999) *Agricultural and Forest Meteorology*, **97(4)**, 309-316.

Charles JSK, Thomas MJ, Menon R, Premalatha T and Pillai SJ (1996) In: Sym. Tech. Aadvancement in Banana/Plantain Production and Processing–India-International, 20-24 Aug. 1996. Kerala Agricultural University, India.

Coffelt JA, Vick KW, Sower LL and McClellan WT (1978) *Environmental Entomology*, ***7*(5)**, 756-758.

Cote FX, Domergue R, Monmarson S, Schwendiman J, Teisson C and Escalant JV (1996) *Physiologia Plantarum*, **97(2)**, 285-290.

Cross WH and Mitchell HC (1966) *Journal of economic entomology,* **59(6)**, 1503.

De Graaf J (2006) In South Africa (Doctoral dissertation, University of Pretoria).

de Maagd RA, Bosch D and Stiekema W (1999) *Trends in Plant Science*, **4(1)**, 9-13.

de Maagd RA, Bravo A and Crickmore N (2001) *Trends in Genetics*, **17(4)**, 193-199.

de Maagd RA, Bravo A, Berry C, Crickmore N and Schnepf HE (2003) *Annual Review of Genetics*, **37(1)**, 409-433.

Dias SJF (1936) *Administrative report, Division of Agriculture, Ceylon, D*60-66.

Dutt N and Maiti BB (1972a) *Indian Journal of Entomology*, **34(1)**, 20-30.

Dutt N and Maiti BB (1972b) *Indian Journal of Entomology*, **34(4)**, 272-289.

Froggatt JL (1928) *Queensland Agricultural Journal*, **30(6)**, 530-541.

Gailce Leo Justin C, Leelamathi M and Nirmaljohson SB (2008) *Agricultural Reviews*, **29(3)**, 185-192.

Gatehouse AMR, Boulter D and Hilder VA (1992) *CAB International*, pp. 155–181.

Godonou I, Green KR, Oduro KA, Lomer CJ and Afreh-Nuamah K (2000) *Biocontrol Science and Technology*, **10(6)**, 779-788.

Gold CS, Pena JE and Karamura EB (2001) *Integrated Pest Management Reviews*, **6(2)**, 79-155.

Griesbach M (1999) 131 pp. *Ph.D. Thesis, University of Bonn*.

Hoffmann W (1933) *Hong Kong Nature*, **4(1)**, 48-54.

Hussaini SS (2003) Workshop on the Entomopathogenic Nematodes in India held on 22 and 23rd January, 2003. (pp. 27-68). Project Directorate of Biological Control, Indian Council of Agricultural Research.

Isahaque NMM (1978) *Pesticides*, **12(6)**, 22-24.

Jacobson M (1966) *Annual Review of Entomology*, **11(1)**, 403-422.

Jayaraman S, Ndiege IO, Oehlschlager AC, Gonzalez LM, Alpizar D, Falles M and Ahuya P (1997) *Journal of Chemical Ecology*, **23(4)**, 1145-1161.

Jepson FP (1935) *Administration Report Director Agricultural Ceylon*, D132-D147.

Justin C, Leelamathi M and Nirmaljohson SB (2008) *Agricultural Reviews*, **29(3)**, 185-192.

Keller, JC, Mitchel EB, Mckibben G and Davich TB (1964) *Journal of Economic Entomology*, **57**, 609–610.

Kiem NN (1995, August) In: *JIRCAS International Symposium Series (Japan)*, Food and

Agricultutre Organization of United States, 67-73.

Kiggundu A, Gold CS, Labuschagne MT, Vuylsteke D and Louw S (2003) *Euphytica*, **133(3)**, 267-277.

Kiggundu A, Pillay M, Viljoen A, Gold C, Tushemereirwe W and Kunert K (2004) *African Journal of Biotechnology*, **2(12)**, 563-569.

Kiggundu A, Muchwezi J, Van der Vyver C, Viljoen A, Vorster J, Schlüter U and Michaud D (2010) *Archives of Insect Biochemistry and Physiology*, **73(2)**, 87-105.

Kung KS (1955) *Journal of Agriculture and Forestry*, Taichung, **4**, 80-113.

Lefroy HM (1909) Text Book of "Indian Insect Life", 379-392.

Luo LY, Luo QC, Yao X and Liu ZL (1985) *Insect Knowledge (Kunchong Zhishi)*, **22(6)**, 265-267.

Magana C, Beroiz B. Hernández-Crespo P, Montes de Oca M, Carnero A, Ortego F and Castanera P (2007) *Bulletin of Entomological Research*, **97(06)**, 585-590.

Mathew MP, Nair SR and Sivaraman S (1997) *Indian Journal of Entomology*, **59(3)**, 269-273.

Matsumoto K, Vilarinhos AD and Oka S (2002) *Euphytica*, **125(3)**, *317-324.*

May GD, Afza R, Mason HS, Wiecko A and Novak FJ Arntzen CJ (1995) *Nature Biotechnology*, **13(5)**, 486-492.

Maziah M, Sariah M and Sreeramanan S (2007a) *Plant Pathology Journal*, **2(5)**.

Maziah M, Sreeramanan S, Puad A and Sariah M (2007) *Journal of Plant Sciences*, **2(5)**.

Mesquita ALM, Alves EJ and Caldas RC (1984) *Fruits*, **39(4)**.

Montesdeoca M, Lobo MG, Casañas N, Carnero A, Castañera P and Ortego F (2005) *Entomologia Experimentalis et Applicata*,**116(3)**, 227-236.

Mustaffa MM and Sathiamoorthy S (2004) Proceedings of the 2st BAPNET Steering Committee meeting. Jakarta (IDN), 2003/10/06-07. Los Baños (PHL): INIBAP-AP 2004. ISSN: 1729-0805.

Nankinga CM, Latigo WMO, Allard BG and Ogwang J (1994) In *African Crop Science Proceedings*, **1**, 300-302.

Ostmark HE (1974) *Annual Review of Entomology*, **19**, 161-176.

Padmanaban B and Sathiamoorthy S (2001) *Musa Pest Fact Sheet No.* 5.

Padmanaban B, Sundararaju P, Velayudhan KC and Sathiamoorthy S (2001a) *InfoMusa*,

10(1), 26-28.

Padmanaban B, Sundararaju P and Sathiamoorthy S (2001b) *Indian Journal of Entomology*, **63**, 204-206.

Padmanaban B and Kandasamy M (2003) *Indian Journal of Entomology*, **65(3)**, 424-425.

Palaniswamy P and Underhill EW (1988) *Environmental Entomology*, **17(3)**, 432-441.

Parniski PF, Treverrow N, Bedding RA and Sikora RA (1990) In *Second International Nematology Congress* (121-122).

Patel ZP and Jagadale VS (2003) *Insect Environment*, **9(3)**, 120-121.

Pei XW, Chen SK, Wen RM, Ye S, Huang JQ, Zhang YQ and Jia SR (2005) *Journal of Integrative Plant Biology*, **47(8)**, 971-977.

Prasad B and Singh OL (1988) *Bulletin of Entomology*, **29(1)**, 97-99.

Prasad B and Singh OL (1989) *Indian Journal Hill Farming*, **I**, 71-73.

Prasuna AL, Jyothi KN, Prasad AR and Yadav JS Padmanaban B (2008) *Current Science (00113891)*, ***94*(7)**.

Ram S and Pathak KA (1987) *Bulletin of Entomology (New Delhi)*, ***28*(1)**, 12-18.

Ravi G and Palaniswami MS (2002) *Current Science*, **83(7)**, 893-898.

Reghunath P, Visalakshi A, Mathew T, Mohandas N, Beevi S and Remamoni K (1992) *Entomon*, **17(1-2)**, 113-115.

Sagi L, Panis B, Remy S, Schoofs H, De Smet K, Swennen R and Cammue BP (1995) *Nature Biotechnology*, **13(5)**, 481-485.

Shanahan GJ and Goodyer GJ (1974) *Journal of Economic Entomology*, 446-447.

Shanmugavelu KG, Aravindakshan K and Sathiamoorthy S (1992) Banana Taxonomy, Breeding and Production Technology. Metropolitan Book Co. *PVT. Ltd. New Delhi. In: Fullerton RA, Stover RH (eds) Sigatoka Leaf Spot Diseases of Bananas, INIBAP, Montpellier, France*, 252-266.

Sharma P, Shankar PR, Subramaniam G, Kumar A, Tandon A, Suresh CG, Meenakshi V and Kumar LS (2009) *International Journal of Insect Science*, 1, 29–44.

Shengjun L, Ruiming W, Xinwu P, Shikai C, Weimin L and Shirong J (2007) *Web of Science*: 2,**17(10)**.

Shrestha GK, Thapa RB, Baral DR and Pokharel RR (1994) *IAAS, Rampur, Chitwan, Nepal*. pp. 88-98.

Shukla GS and Kumar K (1969) *Science and Culture*, **35**, 481-482.

Shukla GS and Kumar K (1970) *Science and Culture*, **36(9)**, 515-516.

Shukla GS and Tripathi AK (1978) *Entomological News*, **89 (9 and 10)**, 249.

Sikora RA, Bafokuzara ND, Mbwana ASS, Oloo GW and Uronu B and Reddy KS (1989) *FAO Plant Protection Bulletin*, **37(4)**, 151-157.

Simmonds NW (1966) Bananas (2nd ed.). Longmans, Green, London. 512 p.

Singh SS (1966) *Journal of Entomology*, **28**, 410.

Subramaniam S, Maziah M, Sariah M, Puad MP and Xavier R (2006a) *Scientia Horticulturae*, **108(4)**, 378-389.

Sreeramanan S, Maziah M, Rosli NM, Sarria M and Xavier R (2006) *Biotechnology*, **5(2)**, 203-216.

Sripriya C, Padmanaban B and Uma S (2000b) *Indian Journal of Entomology*, **62(4)**, 382-390.

Swaine G, Pinese B and Corcoran RJ (1980) *The Queensland Journal of Agriculture and Animal Science*, **37**, 1.

Thapa RB (1993) *IAAS Res. Rep.* (1985-1991). 66-69.

Tiwary M (1971) Proceeding of 58th Indian Science Congress, III. 772.

Uma S, Saraswathi MS, Durai P and Sathiamoorthy S (2006) *Plant Genetic Resources Newsletter (IPGRI/FAO)*.

Valmayor RV, Davide RG, Stanton JM, Treverrow NL and Rao VN (eds) INIBAP/ASPNET, Los Banos, Philippines, 258.

Visalakshi A, Nair GM, Beevi SN, and Amma AMK (1989) *Entomon*, **14(3-4)**, 367-368.

Vishnevetsky J, White Jr TL, Palmateer AJ, Flaishman M, Cohen Y and Elad Y and Perl A (2011) *Transgenic Research*, **20(1)**, 61-72.

Wright WE (1977) *Animal Production Science*, *17*(86), 499-504.

Yadav VK and Yadava CPS (2000) 3rd Internationational Crop Science Congress, 17-22 August 2000, Hamburg, Germany.

Zhou SF and Wu XZ (1986) *Acta Phytophylactica Sinica*, **13(3)**, 195-199.

CHAPTER 2

Literature survey

SUMMARY

The advances in DNA marker technology has been used effectively in studies varying from genetic diversity analysis to the positional cloning of genes. The high-density molecular maps based on molecular markers have made the mapping and tagging of almost any trait possible. Marker-assisted selection has proved to be a potentialy powerful technique to combat diseases in economically important crop. Additionally, DNA marker technology has been successfully used in fingerprinting genotypes, in determining seed purity, in systematic sampling of germplasm, and in phylogenetic analysis. Since the focus of this thesis is on the development of molecular approaches for the control of *O. longicollis* (Oliver), this introductory chapter gives a comprehensive review of literature pertaining to the theme of this thesis. The literature survey has been divided into two sections: the first section describes the use of various molecular makers in the genetic diversity studies of insect populations and insect systematics; and the second section describes the use of genetic engineering approaches for the control of agriculturally important pests and their crop plants.

SECTION I

Use of molecular markers in the analysis of genetic diversity and systematics in insects

Insects are a major life form on earth and represent 75% of all the animal species. They can be found in almost all ecosystems due to their immense adaptation potential in diverse environments. Due to this ability to survive in such diverse conditions they form a delicate relationship to human life and the environment. Some of the insects are beneficial since they pollinate crops, act as natural enemies of damaging pests, and produce useful products such as honey, silk and wax for humans. On the contrary several insects are major pests of food crops, act as vectors for transmitting deadly diseases, cause damage to urban infrastructure, environment, forest and natural resources.

To understand this diversity and how the genes and genetic make-up of insects contribute to their adaptable life forms, it is important to study genetic diversity of insect populations. Even within a species, insect populations vary in their behaviour pattern and morphology which contributes to their complex interaction with the environment (Dempster & McLean 1999).

Classical genetic principles were initially applied to understand insect diversity and ecology. Initially phenotypic markers such as eye colour, body spots or bands and hairs or spines etc., were used as common tools to study pattern of dispersal, mating behaviour and inheritance of genetic traits in insects (Bartlett *et al.* 1968; Fay & Craig 1969; Bond *et al.* 1970; Bartlett & Butler 1975). However, such kind of phenotypic markers suffer from many practical limitations, the most important being their relatively infrequent and often hard to score traits. Moreover the relatively small number of phenotype markers, makes it difficult to map a trait. Due to the limitations of such morphological and physiological markers, population biologists adopted protein electrophoresis to measure genetic variations post 1960s (Loukas 1979; Loxdale & den Hollander 1989; Smith & Wayne 1996; Symondson & Liddell 1996). In recent years,

different DNA based molecular markers have revolutionized this field and are described in this section.

Genetic diversity

Genetic diversity has been defined as any measure that quantifies the magnitude of genetic variability within and between populations. It is a fundamental source of biodiversity and provides the raw material for evolution by natural selection. Ecological effects of genetic diversity has been seen in evolutionary biology. In agronomy, there have been efforts to increase crop yield by planting genetically diverse varieties within a single field (Wolfe 1985; Smithson & Lenne 1996). By increasing the varietal diversity, greater yield and decreased damage by herbivores and pathogens can be achieved.

Need to quantify genetic diversity

Genetic diversity is most often characterized using data that depict variation in either allelic status or continuously distributed (*i.e.* quantitative) characters, which lead to different measures of genetic diversity. For the study of genetic diversity based on molecular markers, neutral alleles provide the most important source of variation. Molecular markers which represent neutral alleles are microsatellites, AFLPs, direct DNA sequence and protein polymorphisms (Avise 2004). Neutral variation in itself cannot have ecological consequences. In theory quantitative traits can also be neutral. Diversity, if measured by neutral markers is not expected to have any ecological consequences in itself. In some cases, it is correlated with levels of genetic diversity for ecologically relevant traits (Reed & Frankham 2001).

Population genetics

Population genetics studies the underlying processes of genetic variation, defined in samples of individuals from different populations and species. Genetic variation is a huge source of information about the biology of individuals carrying a given variant. It also keeps track of the history and spatial relationships of populations composed by these individuals. This information can be extracted using genetic markers. Two main characteristics of genetic markers are that they are distinguishable and are inherited genetically. The genetic basis of having markers in an individual or in a population is the presence of different alleles at a given genetic locus, which results in genetic variability or genetic diversity or genetic polymorphisms. These genetic differences between individuals can be used for carrying out genetic analysis on them for different purposes like individual identification (genetic fingerprinting), genetic mapping, breeding or to describe their genetic relatedness.

Use of protein polymorphism as molecular markers

Protein markers made a significant contribution at that time when DNA technologies were not so much advanced as it is now (Loxdale & Lushai 1998). The electrophoretic pattern of allozymes on polyacrylamide or starch gels were used to identify different alleles of a given gene. It was possible to determine the extent of heterogeneity at a locus in a population based on the variant allozymes produced by the individuals based on banding pattern. Similarly, isozymes (products of different loci, but with similar function) banding pattern was also used to study genetic variation within and between populations (Steiner & Joslyn 1979; Bartlett 1981; Loxdale *et al.* 1985). Protein polymorphisms were the first markers used for genetic studies in livestock (Manwell & Baker 1970). However, the number of polymorphic loci that can be assayed, and the level of polymorphisms observed at the loci are often low, which greatly limits their application in genetic diversity studies.

Allozymes represent allelic variation in gene products. The electrophoretic technique for the detection of allozymes was developed in the 1950s. Before the advent of this technique very few single-locus genetic markers were available for the study of population biology. These markers were the genes for Mendelian traits like flower or fruit colour and other serological incompatibility reactions. New marker genes were provided by the protein electrophoresis technique. Based on this technique individuals could be identified as homozygotes or heterozygotes at a given locus.

There are two terms related to each other which can be used for analysis based on protein markers. One is "Allozyme" which refers to different allelic forms of nuclear-encoded enzymes. The other is "Isozyme" which is a more general term and refers to the different biochemical forms of an enzyme identified by electrophoresis. Protein markers have been successfully exploited in studying insecticide resistance (Maa & Terriere 1983), pathogen identification (Wilding *et al.* 1993), chromosome mapping (Loukas *et al.* 1979) or detection of prey in insect predators (Solomon *et al.* 1996) to name a few.

Enzyme polymorphisms as a measure of genetic variation in insect populations

Several examples of the use of protein and enzyme polymorphisms as a measure of genetic variation in insect populations are available in literature and are described below: Harris (1966), Johnson (1966) and Lewontin & Hubby (1966) used enzyme polymorphism based on electrophoretic mobility of proteins as an estimate of the level of genetic variation within and between the populations. Protein electrophoresis was used to investigate genetic variation of the *Drepanosiphum platanoidis* aphid populations from Southern Britain on the basis of mobility of isoenzymes. This information was used to investigate the population structure and gene flow for these aphids Harris (1966). Polyacrylamide gel electrophoresis was used to investigate polymorphism in natural populations of *Chrysoperla carnea* complex in France (Nice & Shapiro 2001). Allozyme polymorphisms were studied to estimate variations in populations of Stenotopic peat bog (tryphobiontic) nocturnal moth and *Coenophila subrosea stephens* (Sula & Spitzer 2000). Isoenzyme markers were used for population genetic study of *Aedis aegypti* in Brazil

which included the characterization of the geographic and seasonal structure of the vector population and gene flow among it from different cities (Costa-Ribeiro 2007). Genetic diversity and differentiation of Atlantic and Mediterranean populations of *Pogonus littoralis* was studied using allozyme electrophoresis. Several enzymes studied were aldehyde oxidase, peptidases, phosphoglucomutase, glucose-6-phosphate isomerase, malate dehydrogenase, isocitrate dehydrogenase 1 and 2 (Dhuyvetter 2005). Allozyme polymorphism was used to analyze the genetic diversity and the genetic structure of *Aedes aegypti* populations from different ecological regions (Huber 2008). Protein banding pattern (SDS-PAGE) was used to discriminate between the three Scarab beetles form sugarcane fields in upper Egypt (Ibrahim 2009). Population genetic structure of white pine weevil, *Pissodes strabi* (Peck) was studied using isozyme markers (Lewis 2000). Isozyme variation was studied for the susceptibility of *Aedes aegypti* strains to the dog heart worm (Nayar & Knight 2002). Population dyanamics of *Aedes aegypti* in a dengue endemic urban region was studied using two phenotypic characters *i.e.* insecticide resistance and vector competence in association with the analysis of genotypic markers. Both isoenzyme as well as RAPD markers were used to analyse genetic differentiation of *Aedes aegypti* populations (Ocampo & Wesson 2004). Genetic differentiation at five poplymorphic isoenzyme loci were analysed for the study of population genetic structure of *Aedis aegypti* with relation to human population densities (Paupy *et al.* 2000). Isoenzyme allele frequency was used to determine intra- and inter- specific allele frequency and gene flow in the *Octhebius* complex (Urbanelli 2002). Isoenzyme polymorphism was assessed in single mosquitos for different strains of *Aedes albopictus* from northern Medagascar (Vazeille 2001). Genetic heterozygosity of *Aedes aegypti* in Martinique was analysed based on isoenzyme variation and its relation to infection rate to a dengue virus (Yebakima 2004). Population genetic structure of two beetles *Acanthoscelides obtachus* and *A. obvelatus* from wild and cultivated phaseolus was studied using allozyme electrophoresis (Gonzalez-Rodriguez 2000). Phylogenetic relationships among social parasites and their hosts in the ant tribe Tetramoniini was analysed using allozyme electrophoresis (Carpenter 1993). Pattern of genetic variation among Canadian populations of the aphid *Rhopalosiphum padi* L. was determined using allozyme analysis (Simon & Hebert 1995).

Advantages and disadvantages of protein polymorphisms as a measure of genetic diversity

Advantages

1. This technique is relatively inexpensive and straightforward
2. Large number of samples can be processed
3. Can be used for both fine-and broad-scale genetic variation
4. Most allozyme markers represent co-domonant Mendelian loci.

Disadvantages

1. The protein/enzyme electrophoretic profiles should be polymorphic for the data to have statistical significance. However, some species are monomorphic for most allozymes. It has been seen that not only narrow endemic species, but widespread species can be monomorphic at all or most allozyme loci.
2. Another limitation is that allozymes may differ in metabolic function. There are exceptions to the common assumption that phenotypic differences among allozymes are minimal and selectively neutral, the reason being selection working on allozymes or on traits to which they are genetically linked. These studies suggest that non-coding DNA rather than a gene product can be an ideal genetic marker because of selective processes working on the gene products.

Use of DNA polymorphisms as molecular markers

After the advent of PCR by Karry Mullis (Mullis & Faloona 1986), different techniques were used to detect variation in DNA sequence directly. With the development of DNA-based techniques greater level of polymorphism was obtained with DNA markers than protein markers (Richardson *et al.* 1986), because DNA markers provide more genetic variation as well as different choices of techniques to detect various patterns of inheritance. Protein polymorphism is limited and hence cannot provide such a vast

amount of data. Moreover, mutations in introns or exons of a gene can potentially provide more variation at the DNA level than at the protein level. DNA markers are ubiquitous and numerous and DNA samples are more stable than proteins and are unchanged for detection at all time and tissue of the organism unlike proteins. Due to these properties DNA markers have become popular tools for measuring genetic differences between individuals or within and between related species or populations.

DNA based molecular markers in the study of genetic diversity and systematic in insects

DNA marker technology which is one of the recent advancements in modern molecular biology has found useful applications in molecular ecology research in insects (Hoy 2003). The has been possible due to **(1)** the development of the polymerase chain reaction (PCR) **(2)** the application and availability of evolutionarily conserved sets of PCR primers (Simon 1994) **(3)** the advent of hypervariable microsatellite loci (Jarne & Lagoda 1996, Goldstein & Schlotterer 1999) **(4)** the advancements in DNA sequencing technology and **(5)** the availability of powerful analyses and relatively user friendly computer programs (Luikart & England 1999). The different DNA based molecular markers that have been commonly used in the study of the population structure of various groups of insects are listed in **Table 1**.

DNA-based makers have contributed significantly to rapid increase in molecular studies of genetic relatedness, phylogeny, population dynamics and gene and genome mapping in insects (Loxdale & Lushai 1998; Avise 2000, 2004; Severson *et al.* 2001; Heckel 2003). Many improvements in techniques have led to enhanced reproducibility, power of resolution (ability to reveal more informative polymorphisms from less number of loci) and reduced cost and time in developing and scoring the marker loci. Due to these advances, the application of DNA markers in entomology has undergone a noticeable change by accommodating new technologies for more robust and less expensive genotyping methods.

Basic properties of DNA markers in population biology

The different genetic markers show different rates of changes due to the differential action of fundamental processes, including recombination, mutation and selective constraint on them, which can be quantified at various population levels. These markers can be applied for the study of attributes related to individuals such as inbreeding and/or outbreeding and fitness (Luikart & England 1999), linkage disequilibrium (LD), effective population size (Ne) and natural selection. At the population level, these markers can be used for the estimation of effective number of alleles Ne, understanding metapopulation dynamics and recognizing recent colonizations and introductions (Schierwater 1994; Tishkoff 1996). Allele and/or haplotype frequencies are based on population level statistics which can be changed by genetic drift, founder effect, gene flow and selection. Hence these markers have found application in estimating gene flow and population subdivisions (Burke 1998). The gene genealogies of populations can be studied using sequences of mtDNA and nuclear genetic regions such as microsatellites and single copy nuclear (scn) markers which evolve as determined by the mutation rate, selection and population parameters, such as Ne and changes in Ne and hence have found application in the assessment of intra-specific phylogeography and population history, systematics and biodiversity (Simon 1994; Avise 1994; Burke 1998; Templeton 1998; Crozier 1999). Considering the diverse types of markers available for studies at different levels of population hierarchy, it is important to select the appropriate marker and genetic analysis which is suitable for a particular study. The advancement in development of software related to genetic diversity analysis has made it possible to select the appropriate marker system and analysis for any particular study.

Following are some of the favourable attributes of genetic markers:

1. They can be assayed by polymerase chain reaction (PCR) because PCR uses low quantities of template DNA and provides high specificity.
2. The data generated by PCR primers that amplify homologous regions over a wide taxonomic range can be used in meta-analysis Burke (1998), Johns & Avise (1998), Avise (1998), Avise & Johns (1999).

3. DNA which is used as a template for PCR is easier to extract and store, as compared to proteins.
4. Gene genealogies and frequency data can yield information on demographic trends (Templeton 1998). Molecular markers which give both allele and/or haplotype frequency and sequence data are informative over various range of population parametres (Avise 1994; Hillis 1996; Templeton 1998).
5. The use of multiple markers has two main advantages: (**1**) overcoming variation in biological sampling that can cause different patterns in loci with similar histories; and (**2**) detection of non-concordant characters that can be biologically informative.
6. The markers can be used for screening new taxa with little further development and are amenable to rapid screening.
7. Multilocus approaches (for example, RAPD and AFLP) are technically convenient but imprecise, and have many major technical and/or analytical drawbacks, such as dominance (only one allele identified) (Hurme & Savolainen 1999; Perez 1998; Rabouam 1999). The data generated are of limited comparability among studies. By contrast, single-locus, codominant or haploid organellar markers supply robust data which can be used for more precise analyses.

DNA markers can be classified into three types based on the method used for their detection ie. **(1)** hybridization-based **(2)** PCR-based and **(3)** DNA sequence-based (Gupta 1999; Jones 1997; Joshi 1999; Winter & Kahl 1995).

Table 1. Types of molecular markers which have been used to study genetic diversity of insects

DNA marker	Type of marker	Examples
Microsatellites	Simple Sequence repeats (SSRs) or Short Tandem Repeats (STRs)	Social wasp *Parachartergus colobopterus* (Coudhary 1993); Lepidoptera (Zhang 2004); Hemiptera (Weng 2007); Meglecz 2007
PCR-based markers	RAPD	Hadrys 1992 (mini-review)
	ISSR	Aphids (Abbot 2001)

	AFLP	Drosophila (Luckinbill & Goldenberg 2002)
Nuclear DNA markers	**Nuclear Ribosomal RNA genes**	
	18S rRNA	Homoptera and Hemiptera (several families of insects) (Campbell 1995)
	28S rRNA	Holometabola (several insect taxas) (Whiting 1997)
	5.8S rRNA	Diptera (mosquito family Culicidae) (Miller 1997)
	ITS1	Drosophila (Schlottere 1994)
	ITS2	
	Nuclear protein coding genes	
	a-amylase (*amy*)	Diptera (*D. melanogaster*) and Copleoptera (*Tribolium castaneum*) (Hickey 1987; Inomata 1987; Shibata & Yamazaki 1995; Okuyama 1996)
	Acetylcholine esterase (*achE*)	Diptera (Drosophilila) (Baker 1997)
		Lepidoptera (*Bombyx mori*) (Mange 1999)
	Alcohol dehydrogenase (*adh*)	Diptera (*Drosophila*) (Albalat 1993)
	*EF-1*a	Lepidoptera (*Heliothine* moths) (Cho 1995), Hymenoptera (Aphidiinae) (Belshaw 1997), Hemipetra, Collembola (flea *Isotoma klovstadi*) (Frati 1998), Archaeognatha (crustaceans) (Regier 1997)
	Period (*per*)	Diptera (*Drosophila*) (Costa 1991), Lepidoptera (*Bombyx mori* and 25 other species) (Regier 1998)
	Resistance to dieldrin (*Rdl*)	Coleoptera (*Hypofhenemus hampe*) (Andreev 1998)
	Wingless (*wg*)	Lepidoptera (*Heliconius* butterflies) (Brower 1997) and Diptera (Drosophilila) (Baker 1997)

Organelle DNA markers	**Mitochondrial genes (rDNA and protein coding genes)**	
	rDNA 12S	*Anopheles hilli* (Ballard 1992)
	rDNA 16S	*Aedes Albopictus* (HsuChen *et al*.1984)
	COXI	*Apis mellifera* (Garnery 1992)
	COXII	*Apis domesticus* (Liu & Beckenbach 1992)
	COXII	*Cicindela dorsalis* (Vogler & DeSalle 1993)
	Cytb	*Tetraponera rufoniger* (Jermiin & Crozier 1994)
	ND1	*Cicindela dorsalis* (Vogler & DeSalle 1993)
	ND2	*Drosophila melanogaster* (Nigro 1991)
	ND3	*Aedes albopictus* (Hsuchen & Dubin 1984)
	ND5	*Ceratitis capitata* (Frohlich 1993)
	ND6	*T. rufoniger* (Jermiin & Crozier 1994)

Microsatellites

Microsatellites (also known as STRs or short tandem repeats) are tandem repeat sequences, usually less than five bases long, which differ in size of the repeat unit. These are run on acrylamide gel and stained with silver stain for greater resolution or sequenced. They can also be detected by hubridisation. Expected polymorphism depends on the length characteristics of the loci. The target region is amplified by using specific primer pairs which represent different loci. Resourses such as NCBI-GENBANK (http://www.ncbi.nlm.nhi.gov), EMBL (http://www.embl-heidelberg.de) as well as species specific databases for microsatellites can be used to find microsatellite-enriched sequences efficiently (Scribner & Pearce 2000).

Advantages

1. Specific primers are used to assay by PCR, which gives enormous possibility of sampling.
2. Microsatellites are highly variable and in a single locus as many as 50 alleles can be examined.
3. Co-dominant multilocus allelic pattern is obtained where both alleles at every locus can be identified.
4. They provide sufficient statistical power for individual identification, parentage determination and in inferring relatedness of individuals.
5. They can be successfully used for pedigree or population level studies as well as determining mating structure, estimating effective population size, determining populations differentiation and gene flow.
6. The information obtained can be utilized for taxonomic and evolutionary biology studies for determining speciation and hybridisation (Scribner & Pearce 2000).

Disadvantages

1. Different primers are needed for different loci, which leads to change in PCR conditions also.
2. Because of the requirement of specific primers, the major drawback of microsatellites is that they need to be isolated *de novo* from most species being examined for the first time (so without prior sequence information it cannot be used for new species).

Microsatellites have proved to be extremely useful markers for studies with wide applicability from individual to population level studies. Some of the examples are cited below:

In case of *Drosophila melanogaster* variation between dinucleotide repeat motifs indicated differences in mutation rates between loci (Bachtrog 2000). Schlotterer (2000) has reviewed the various mechanisms of mutations which can give rise to the observed variability in microsatellites. These include recombination, DNA (replication) slippage

and efficiency of mismatch repair. Recombination is not the predominant mechanism generating microsatellite variability because experiments have demonstrated that DNA slippage occurs at very high rates and therefore is an important contributer of variation (Hentschel 1982; Streisinger & Owen 1985; Schlotterer & Tautz 1992). Phylogeny of the social wasp of subfamily Polistinae was determined by using microsatellite markers (Arevalo 2003). Eight microsatellites were used to determine the population genetic structure of the cotton-melon aphid *Aphis gossypii*, collected from wide geographical locations (Carletto 2009). Within- and between-population variation was investigated at eight microsatellite markers in *Diabrotica virgifera virgifera* (WCR) to determine the routes by which WCR was introduced into Europe, and to assess the effect of introduction events on genetic variation (Ciosi 2008). Microsatellite markers suggested that geographical isolation was the cause of population structure and inbreeding in bumblebee populations, *Bombus muscorum* (Hymenoptera: Apidae) (Darvill 2006). Polymorphic microsatellite loci in the coffee berry borer, *Hypothenemus hampei* (Coleoptera, Scolytidae) indicated the presence of low to moderate genetic diversity per locus (Gauthier & Rasplus 2004). A meta-analysis conducted on 32 insect species using microsatellite markers showed strong differences in the abundance of microsatellites among species and individuals within species, and that species which showed similarities in flanking regions of microsatellite often grouped into families (Meglecz 2007). Microsatellites markers have been used to determine population genetic structure of various insect species such as *Coenagrion mercurial* (Watts 2006), lepidopteran insects (Zhang 2004), aphids (Weng 2007), leaf-cutting ant societies (Hughes & Boomsma 2004) and spider mite pest *Tetranychus turkestani* (Bailly 2004).

RAPDs (Random Amplified Polymorphic DNA)

RAPD is a molecular marker technique developed by Williams (1990) which is based on PCR amplification of genome DNA at random locations using ten bases long single primers of arbitrary nucleotide sequence. This technique amplifies small inverted repeats scattered throughout the genome and in the process generates the polymorphic

amplified DNA regions. The polymorphism is detected as amplified DNA segments, which amplify from one parent but not the other, show Mendelian inheritance and can be used to construct genetic maps in a range of species. In this case the primer is short 10 bases long whose sequence is randomly selected as there is no prior sequence information is required. These 10 mer oligonucleotides hybridise to the randomly dispersed complementary sequences in the genome which results in the amplification of multiple fragments. The location and the sequence of these fragments amplified is unknown but they represent fragment length differences between different individuals, if any deletion/insertion has been taken place in the regions amplified. The primer-binding site can also be altered by mutations in some individuals resulting in differences in the amplification pattern. The amplified fragments are separated by electrophoresis and are visualized by ethidium bromide staining on UV transilluminator. The results are documented by Gel-documentation system and approximately 5-10 loci can be detected in one experiment. The RAPD single primer of arbitrary nucleotide sequence detects polymorphism in the form of presence and absence of specific nucleotide sequence and this polymorphism can be used as genetic markers along the DNA, to construct genetic maps as well as various genetic diversity studies. It requires no prior knowledge of the sequence for primer designing. The amplification results in PCR products which when run on agarose gels give a banding pattern which is unique for the particular primer and gDNA of the organism used. The amplified bands need not show any complimentarity on hybridisation. These banding patterns can be scored and used for analysis for the type of study it was intended with the use of a wide variety of softwares available on the internet. The polymorphism detected is dominant and other than genetic maps, these are also suitable for breeding applications in animals and plants, DNA fingerprinting, population genetics, polymorphism-based identification of hybrid cell lines or genetic stocks. The advantages of RAPD include: universal set of primers are available, no prior information about genome sequence is needed, can be used as sequenced tag sits markers and most importantly, determination of genotype can be automated. The disadvantages include problems in reproducibility, as the procedure is sensitive to the reaction condition, DNA quality and temperature of PCR.

Hadrys (1992) have reviewed the use of RAPDs in the area of molecular ecology since its development. The advantages of RAPDs for use in molecular ecology include their application on anonymous genomes and on samples where limited quantity of DNA is available as a starting material, efficiency, simplicity and low cost. RAPD marker technique has been applied to all major groups of insects such as Coleoptera, Lepidoptera, Diptera, Hymenoptera, Homoptera and Aphids. For *e.g*. the root borer (Rhinoceros beetle), *Oryctes agamemnon* Burmeister, an invasive coleopteran pest of date palm trees in south-western Tunisia was introduced accidentally and spread out into areas of oases of Tozeur and Kebili. This introduced pest was studied using RAPD markers for level and distribution of genetic diversity, degree of gene flow between populations and number of introduction events (Abdallah 2012). Genetic differentiation among six Florida populations of *Diaprepes abbreviatus* (L.) was determined using protein as well as RAPD markers. Although no differentiation was observed among populations when protein markers were used, RAPD-PCR data showed significant differentiation among populations (Bas 2000). Genetic differentiation among *Anthonomus grandis* Boheman populations were studied by Kim & Sappington (2004) and Martins (2007). Genetic properties of sweet potato weevil, *Cylas formicarius* from different geographical regions of Japan were determined using RAPD-PCR (Kawamura 2007). The genetic structure of Colorado potato beetle (*Leptinotarsa decemlineata*) populations were examined using RAPD markers. The obtained clustering indicated that the populations were structured which may due to adaptation to pyrethroid insecticides used for controlling the population of this pest (Sidorenko & Berezovska 2002). RAPD markers were used to discriminate between three species of scarab beetles (*Orycetes agamemmon*) commonly found in sugarcane fields in Upper Egypt (Ibrahim 2009).

Several studies using RAPDs have been carried out in Insects of Lepidoptera group which include several pests of economic importance as well as beneficial insects. For example, molecular diversity of *Helicoverpa armigera* Hubner was studied by Asokan (2012) to determine the genetic relationships amongst the intra-specific populations and thereby estimating genetic diversity which led to insight into the genetic structure of the populations. RAPDs were used to study and differentiate five butterfly

species (Sharma 2006). Verovnik & Glogovcan (2007) used RAPD-PCR to generate species specific markers for *Liptidea sinabsis* and *Liptidea reali.*

Chatterjee & Pradeep (2003) used RAPD-PCR to develop markers associated with growth, yield and origin of *Bombax mori,* for help in breeding programmes for crop improvement in silkworm. Silkworm being used for commercial benefits have been studied by several researchers, most important being the work done by Keshva Murthy (2006) (RAPD-based fingerprinting of *Bombyx mori*); genetic diversity of Muga silkworm (Neog 2010) and genetic fingerprinting of *Bombyx mori* (Nagaraja & Nagaraju 2005); Sets of ISSR and RAPD markers were used to study genetic variability between the *E. sorbillans* (a parasitoid of silkworm in India) populations from four different geographic locations in the southern states of India and one from Bengal (Chatterjee 2003).

Ge Mei (2000) studied genetic variation in cotton bollworm (*Helicoverpa armigera*) in relation to population resistance to *Bacillus thuringeinsis* using RAPD-PCR. Dutta (2012) developed RAPD and SCAR markers associated with yield traits for Indian Tasar silkworm (*Antheraea mullita* drury*).* Garner & Slavicek (1996) developed RAPD-PCR markers for distinguishing Asian and North American gypsy moths (*Lymantria disper*). Genetic diversity of *Scirpophaga incertulas,* an important pest of rice was studied using RAPDs and ISSRs (Kumar 2001). Differentiation between the males and females of *Diatraea saccharalis,* the most damaging pest of sugar cane crop, was done using the RAPDs. *G*enetic diversity studies of a *D. saccharalis* laboratory culture population led to the identification of a 700 bp, which was considered as a female sex marker in *D. saccharalis* (Munhoz 2010)**.**

Diptera is a group of insects which includes flies to different mosquitos which are vectors for several diseases. RAPD analysis was used by Buraffi (1995) to assess genetic variability in six wild populations and in five laboratory strains of *Ceratitis capitata* (Med fly). RAPD technique revealed larger amount of genetic variation and improved discrimination within and between populations and strains. RAPD data showed that a major part of intra-specific variability involved the differentiation of central vs. peripheral populations. RAPD-PCR was used to assess the levels of genetic variability and

population differentiation of *Ae. aegypti* in Cordoba, Argentina (Julio 2009), boll weevil (Kim 2004) and *Aedes aegypti* (Santos 2003).

In case of Hymenoptera group of insects, RAPDs have proved to be useful markers for genetic diversity analysis in common pine sawfly (Baumann 2003), stingless bee population (Sriram 2004) and alfalfa leaf-cutting bee (Lu & Rank 1996). Based on haploid males, RAPD markers were used to study genetic variation within and among four French populations and one Finnish out group population of the common pine sawfly, *Diprion pini* (L.), representing a severe European forest pest associated with mass outbreaks (Baumann 2003). RAPD analysis was carried out with insect specific nuclear primers to detect variations in stingless bee population (*Trigona* spp.) from India. Genetic diversity analysis clearly indicated the existence of variations at the genomic level (Sriram 2004). Al-Barrak (2004) studied British species of *Tetramesa*, which are morphologically very similar. They used RAPD and other nuclear (enzyme) markers to distinguish a complex of five cryptic species. With the help of RAPD makers they were able to distinguish these sympatrically evolved species of Tertamesa. Lu & Rank (1996) used RAPD for the study of population genetic variation of alfalfa leaf-cutting bee. Ross (1999) studied hierarchical genetic structure in introduced populations of the fire ant *Solenopsis invicta using* RAPD markers. RAPDs analysis proved the existence of cryptic species in two geographical populations of *Ageniaspis citricola* (Hoy 2000).

Homoptera is also an important group of insects representing aphids. B biotypes and other biotypes of *Bemisia tabaci* in Jordan were surveyed using RAPD-PCR to develop a marker for the identification of *B. tabaci* biotypes. RAPDs were able to distinguish whitefly biotype A, B and BA (Hasan 2006). RAPD-PCR revealed intra-population diversity in six Iberian population of the whitefly *Bemisia tabaci* (Moya 2001). RAPD-PCR analysis was used to determine the DNA polymorphism in populations of aphid *Aphis gossypii* (cotton aphid) collected from different host plants in different seasons (Peng 2008). Similarly RAPD-PCR was used to study host-based genetic differentiation in the aphid *Aphis gossypii* Glover. The populations of the pests were collected from different host plants at eighteen locations in Southern France. The study revealed the first evidence for within species structuring of the populations and it was found that the populations were distributed according to the type of host-plant

(Vanlerberghe 1998). RAPD markers were used to assess the level and distribution of genetic diversity of *Aphis spiraecola* (Patch) population reared in different regions in Tunisia (Maha Mezghani 2012).

Population genetic structure and phylogenetic relationships of several species and species complexes of aphids have been studied. Example include rose aphids (Jalalizand 2012); *Tagosodes orizicolus* (Myrium 2008); *Aphis gossypii* (Masuti & Chavigny 1998); *Diuraphis noxia* (Puterka 1993); *Myzus persicae* complex (Clements 2000); *Aphid* sp (Black IV 2009). Biotypes (B, C, E, F, G, H and I) were ideintified from seven greenbug, *Schizaphis graminum* laboratory maintained colonies, using RAPD primers. Out of 100 primers tested, four diagnostic primers either alone, or in combination, were able to distinguish all biotypes by distinct size differences in amplified fragments. These primers also produced highly reproducible banding patterns for the field populations (Aikhionbare 1997). The two species of aphids (*P. populitransversus* and *P. Obesinymphae*) which are major pests on cruciferous vegetables in Texas are of small size and are very similar morphologically and in feeding beahviour, thus making their identification and distinction difficult. In this study, RAPD and AFLP markers were used to distinguish these two species over a period of time when the two species occurred together, or separately, in cabbage fields. These two techniques proved efficient in discriminating the two species. However, the results suggested that RAPD technique was a better choice (Chen 2009). In another study by Su-fang (2008), RAPD-PCR technique was used to determine genetic differetiation among different taxa of aphids, the results of which can be used to study population differentiation of closely related aphid species as well as identification of aphids. *Planococcus ficus* (Signoret) and *Planococcus citri* (Risso) are important phytophagous components in different agroecosystems. These two species coexist in the same environment and are very difficult to distinguish by morphological features. The study done by Demontis (2007) was aimed to identify genetic markers suitable for distinguishing *P. ficus* from *P. citri*, to assist in the rapid identification of field specimens using RAPD-PCR. The RAPD primers used were able to clearly differentiate the two species and these primers allowed sensitive and reliable PCR identification of both males and females of *P. citri* and *P. ficus* of different geographic origin.

RAPDs have proved useful in genetic diversity studies in the order Hemiptera which includes planthoppers. RAPD was used to test for segregating ratios in two families of the brown planthopper, *Nilaparvata lugens*. Inheritance was found to be of simple Mendelian fashion (Latif 2008). Jun (2000) and Dong (2005) studied population genetic structure of *Nilaparvata lugens* and developed RAPD-based molecular markers from rice cultivars linked to Bph resistance. Figueroa *et al.* (1999) used RAPD markers to differentiate two morphological close species of genus Sitobion (*Sitobion avenae* (Fabricius) and *Sitobion fragariae* (Walker)). Valenzuela *et al.* (2007) developed markers for the identification of aphid species using RAPD-RFLP of COI gene. *De* Oliveira *et al*. (2007) used RAPD-PCR to estimate genetic structuring of leafhopper species *Dalbulus maidis* and showed the high occurrence of gene flow in some populations. Genetic variability of *Aleurodicus disperses* populations from different islands of Canaries was determined using RAPD-PCR (Callejas *et al.* 2005).

The populations of banana weevil, *Cosmopolites sordidus* (Coleoptera: Curculionidae), which is one of the most important insect pests of bananas and plantains have been analysed using RAPD (Magana *et al*. 2007). Their results indicated the existence of local biotypes with limited gene flow and the populations were affected by genetic drift. The RAPD technique was used to establish a relationship between genetic evolution and systematic development of main locust species (Grasshopper, Orthoptera) in Inner Mongolia. The results indicated that the polymorphism existed generally among genera in the same family and among species in same genera. This study was helpful in laying the molecular foundation for research on systematic development and evolution of main locust in Inner Mongolia pastures (ShuJing 2010). RAPDs were applied to analyze the polymorphism between *D. folliculorum* and *D. Brevis* (hair follicle mites, Arachnids) and it was successfully used for genomic DNA polymorphism analysis and species identification of these two species (Yae & Hui 2009). The genetic variation within and between Finnish *Euseius finlandicus* populations (Predatory mite) was investigated by RAPD-PCR and ITS sequence analyses. The study resulted in the identification of amplified bands specific to different strains of species *E. Finlandicus* mites (Yli-Mattila *et al.* 2001). Geographically distinct Brazilian populations of the Southern Green Stink Bug, *Nezara viridula* (L.) (Heteroptera: Pentatomidae) which is a cosmopolitan and

economically important pest to several crops was studied to assess variability, population structure parameters and gene flow among them using RAPD markers (Sosa-Gomez *et al*. 2005). Populations of the bugs from the genus *Orius* (Heteroptera, Anthocoridae) which are used as biological control predators against acari and homopteran pests were studied using RAPD primers to develop a marker to identify predators which are released in the fields. Species are easily distinguishable by these markers, but differentiation of strains from distinct geographic origins was impossible due to the high level of polymorphism observed (Gozlan *et al*. 1997).

ISSRs (Inter-simple sequence repeat polymorphisms)

ISSRs is a PCR based technique developed by Ziekiewicz (1994) which amplifies the regions between inverse-oriented microsatellite loci, using oligo-nucleotides anchored in microsatellites themselves using single primers which are semi-arbitrary based on di- or tri-nucleotide repeat motifs which are about 20 nucleotides long. This technique was initially utilized mostly for the studies of plant population biology (Wolfe & Liston 1999). In recent years it has found application for similar studies for invertebrates and fungi (Kostia *et al.* 2000; Reddy *et al.* 1999). Simple sequence repeats or SSRs or microsatellites are short, hypervariable elements distributed throughout the genome of eukaryotes. ISSR markers amplify the regions between two closely-spaced, inversely oriented SSRs, which results in multilocus and highly polymorphic dominant, di-allelic Mendelian markers when there is divergence in SSR sites or there are chromosomal rearrangements (Wolfe & Liston 1998; Wolfe *et al.* 1998). This technique is similar to RAPD in amplification and data scoring pattern, but it differs in the amplification temperature used which is less for RAPD and higher for ISSR, resulting in more stringency conditions in case of ISSRs (Wolfe & Liston 1998). ISSR markers are highly reproducible due to the following reasons: stringent annealing temperature, long primers and primers are not arbitrary (specific to SSR loci). ISSR-PCR was used for population-level studies in some invertebrate species (aphids, yellow fever mosquito and rotifers) to assess the applicability of these methods for detection of within population variation in invertebrates (Abbot 2001). Lugue *et al.* (2002) studied intra-specific variation between

different species of Noctuides (Lepidoptera) using ISSR markers and genetic markers based on ISSRs were developed. ISSR markers were used for the analysis of polymorphisms in eleven populations of *Antheraea mylitta* (tropical Tasar silkworm) collected from five zones of Dandakaranya forest in Madhya Pradesh, India, which belonged to six ecotypes. From this study nine ISSR markers were identified which could be used to characterize specific populations (Chatterjee *et al.* 2003). ISSR markers were used to determine the genetic variability within and among populations of semi-domestic *Antheraea mylitta* sp from different locations of India (Kar *et al*. 2005). Compound ISSR primers containing CA/GT-repeat motifs were used to estimate the population genetic structure of *Homalodisca coagulate*. ISSRs were able to reveal intra-population diversity of eighteen populations from throughout the United States and a subpopulation from Tahiti (Leon *et al*. 2004). Intra-specific variation was found in French and New Zealand *M. Aethiopoides,* thus explaining the differences in ability to parasitise *S. Lepidus* (Philips *et al.* 2002). ISSR-PCR tool was used as a tool for population discrimination and genetic variation among different populations of *Plutella xylostella* (Diamond back moth), from different geographical areas of the world (Roux *et al.* 2007). Genetic stability of diapauses-induced multivoltine silkworm *Bombax mori* germplasm was assessed using ISSRs. The diapauses were induced in the silkworm for greater ease of conservation as compared to rearing because it requires expenditure, manpower and infrastructure, in addition to the chances of exposure to diseases, pests and predators. The diapauses-induced eggs when reared were compared to the control batches and no significant variation was found (Sarvankumar *et al.* 2010). ISSR markers were used to analyse genetic relationship in mutant silkworm starins of *Bombax mori*. It has proved to be a useful tool for the characterization of a number of silkworm strains maintained at germplasm centers (Velu *et al.* 2008). Khemakhem *et al*. (2005) used ISSR-PCR to study three haplotypes of *Mayetiola hodei* (Diptera) and assessment of genetic variability and phylogenetic relationships within and between the haplotypes. Saha *et al.* (2011) developed few line-specific ISSR markers for lac insect. Leon *et al.* (2004) used ISSR and ITS (1 and 2) markers for the identification of two geographical (Texas and California) populations of *Gonatocerus morrilli* (Hymenoptera) parasitoids of the glassy winged Sharpshooter (Homoptera). The findings were helpful for the development of

glassy-winged Sharpshooter/Pierce's disease biological control programme in California. Paplauskiene *et al*. (2006) used ISSR-PCR for the assessment of genetic diversity of different races and lines of bee (*Apis mellifera*, Hymenoptera). Al-Otaibi (2008) used ISSR method to detect genetic variability within mite-resistant honey bee populations collected from different colonies which represented three races and were tolerant to mites. ISSR-PCR study of the genetic variability of cotton bollworm (*Helicoverpa armigera*, Lepidoptera) revealed that the populations showed high genetic variation and high gene flow which could explain the insecticide resistance pattern, host adaptation and outbreak of the pest (Feng *et al.* 2000). Genetic diversity of the white backed hopper (*Sogatella furcifera*, Hemiptera), which is a serious pest of rice in Asia, was studied using ISSR markers. The populations were collected from different geographical regions of Greater Mekong sub-region in China and the information obtained on the population differentiation provided useful data for the sustainable management of this long-range migratory pest (Liu *et al.* 2010). Subramanian and Mohankumar (2006) determined genetic variability of cotton bollworm (*H. armigera*) populations collected from different host specioes. Genetic diversity and differentiation among populations of Indian silkworm *Samia cynthia* (Lepidoptera) collected from different parts of North-Eastern India using ISSRs revealed that these populations were different ecotypes and *in situ* conservation of these commercial insect species was recommended (Vijayan *et al*. 2006).

AFLP

AFLPs (Amplification Fragment Length Polymorphisms) are the most recent PCR-based fingerprinting technology. This technique developed by Zabeau and Vos (Vos *et al*. 1995) has been patented by Keygene N. V. (Patent no: CA2119557). Basically the technique resembles RFLP (Restriction Fragment Length Polymorphism), the major difference being, PCR amplification is used instead of Southern hybridization for the detection of restriction fragment, which gives an advantage of efficiency and easy detection due to the use of radio-labled or fluorescent labled dyes. There are four main steps involved: isolation of template DNA; restriction enzyme digestion (as in RFLPs); ligation of adaptors to the restricted fragments; PCR amplification using specifically

designed primers which are complementary to the oligonucleotides ligated to the ends of the digested DNA fragments. PCR is carried out in two steps ie first step is pre-amplification and second step is selective amplification where selective primers with 3'extension are used to give specificity to the final amplification. The fragments are separated by polyacrylamide gel electrophoresis and visualised either by silver staining or radioactive/fluorescent reactions. This process gives rise to a number of amplified fragments with different lengths and in this way genetic polymorphisms can be detected. The polymorphisms detected by this method are due to nucleotide changes at restriction enzyme sites (like RFLPs), or the insertion/deletion events resulting in altered fragment sizes between individuals.

Factors affecting reproducibility of this technique includes the quality of genomic DNA ie a high purity, good quality (high molecular weight) gDNA is required, because it is a restriction digestion dependent process and these enzymes require high quality DNA for optimum results. Partial fragments may lead to false positives (Vos *et al*. 1995). Another important factor is template DNA concentration. The unique features of AFLP which makes it a better suitable marker are that this technique can be used for complex genomes and a combination of six-base and four-base restriction enzyme allows the amplification of potentially several hundreds of unique AFLP fragments, out of which generally 50-100 are selected for AFLP reaction due to the use of selective primer combinations and base extensions. These selective primers further reduce the complexity of the genome. These AFLP fragments correspond to specific positions on the chromosome, and can be utilized as a landmark in the genetic and physical mapping of the genome. As compared with RAPD, AFLP is highly reproducible (Mueller & Wolfenbarger 1998). Thus AFLP is a universal DNA fingerprinting system because it is highly reproducible where a large no of loci can be studied simultaneously with any genome irrespective of the origin (human, animal, plant and microbes) (Blears *et al*. 1998; Savelkoul *et al*. 1999; Janssen *et al*. 1997; Rademaker *et al*. 1997; Blears *et al*. 1998; Hill *et al*. 1996; Powell *et al*. 1996; Buntjer *et al*. 1999; Bensch & Akesson 2005; Luckinbill & Goldenberg 2002; Mueller & Wolfenbarger 1999; Mba & Tohme 2005; Meudit & Clarke 2007).

AFLP fingerprinting technique has found application in the following areas:

1. Construction of genomic maps using genetic markers
2. Detection of genomic clones
3. Fingerprinting of cloned DNA fragmentss
4. Construction of high resolution genetic maps in plants
5. Positional cloning of gene of interest
6. Genetic relationships and epidemiological typing of bacteria and higher eukaryotes
7. Monitoring inheritance of traits of agronomic importance in plant and animal breeding
8. Diagnosis of genetically inherited diseases
9. Pedigree analysis
10. Forensic typing
11. Parentage analysis
12. Screening of DNA markers linked to genetic traits
13. Microbial typing

Use of AFLPs in the study of insect populations

The first report of AFLP in insects was in *Drosophila* (Luckinbill & Goldenberg 2002). In this study, amplified fragment length polymorphisms were used for locating Quantitative Trait Loci related to ageing or longivity. Three AFLP markers which assorted with long life phenotype were developed. Later, several researchers used this technique for the phylogeographic studies on various groups of insects.

Genetic diversity within and among different geographic populations of the beet webworm *Loxostege sticticalis* was evaluated using AFLP analysis. Theresults indicated high genetic variability among individuals but little genetic differentiation among geographic populations, which can be explained by the effects of long distance migration of the beet webworm in different geographic locations and consequent gene flow (Jiang *et al*. 2010). Genetic differentiation among populations of the tomato leafminer *Tuta absoluta* (Meyrick) (Lepidoptera: Gelechiidae) was determined using AFLP markers. The

AFLP fingerprints were able to differentiate these populations in two groups, in relation to susceptibility to insecticides, These results were congruent with differences in susceptibility of this insect to insecticides, as previously identified by other authors (Suinaga *et al*. 2004). Genetic variability of *Spodoptera frugiperda* (Lepidoptera: Noctuidae), the fall armyworm, an economically important maize pest in the western hemisphere was studied using AFLP markers over a large geographic area. The results showed that the majority of the genetic variability is within populations and not between populations, indicating minor gene flow and suggesting that *S. frugiperda* in the Western Hemisphere are an inter-breeding population (Clark *et al*. 2007). AFLP fingerprints were used to characterize the population genetic structure and gene flow of the oriental fruit moth, *Grapholita molesta* (Busck) (Lepidoptera: Tortricidae) for populations collected from different geographic regions in South Africa. The results of this study provided important information on the population genetics of the pest for the design and implementation of sustainable pest management strategies, such as the management of insecticide resistance, which is influenced by insect dispersal (Timm *et al*. 2008). AFLP analysis of the pecan nut casebearer *Acrobasis nuxvorella* collected from diverse geographical locations indicated a high degree of genetic structuring in the pecan nut casebearer across its geographical distribution and invasiveness of the pest was determined (Hartfield *et al*. 2012). AFLPs were used to study the genetic structure of the geographic and putative host-associated populations of carob moth, *Ectomyelois ceratonia* (Lepidoptera: Pyralidae), the most important pest of pomegranate in Iran (Zeller 1839). The results of this study suggested that in spite of the effects of geographic barriers, high within-population genetic variation, migration rate and gene flow could provide the opportunity for emerging new phenotypes in pest populations to adapt to difficult conditions when subjected to intensive control methods (Mozaffarian *et al*. 2008). The European corn borer, *Ostrinia nubilalis* (Hübner) (Lepidoptera: Crambidae), is a widely distributed and serious pest of corn in the U.S. Genetic variability studies of this pest using AFLPs revealed a high degree of gene flow indicating the possibility of fast evolution of resistance to Bt in the pest (Krumm *et al*. 2008). The genetic structure of codling moth strains (*Cydia pomonella* L.), one of the most successful insect pest species known today, has been studied using mitochondrial markers as well as AFLPs. AFLP

markers were able to elucidate the genetic structure of codling moth strains and populations from different regions. The results indicated that the high degree of genetic differentiation between the populations was due to limited gene flow among the populations. In addition, several factors such as microclimatic, ecological and selective pressure due to pest control and geographic constraints could favour the splitting of *Cydia pomonella* into many local populations each with its own genetic fingerprint (Thaler *et al*. 2008).

AFLPs were used to map quantitative trait loci for economically important traits in silkworm, *Bombyx mori,* an economically beneficial insect (Mirhoseini *et al*. 2007, 2010). Population genetics of different populations of silkworm was studied (Dalirsefat *et al*. 2009) and AFLP was also used to develop markers associated with yield attributes (Duverney *et al*. 2006). Keyghobadi *et al*. (2009) used AFLP markers for the study of conservation aspects of butterflies. Johan de Graaf (2006) used AFLPs to study the population genetics of banana root corm weevil (*Cosmopolites sordidus*) for future intended use in integrated pest management. This study revealed population differentiation and various control methods were evaluated for their efficacy based on the population dynamics (Johan de Graaf 2006). Bark beetles (Coleoptera: Curculionidae) populations from forests of Arizona were studied using fAFLP (fluorescently labeled amplified fragment length polymorphism) for a better understanding of the population structure of these species to facilitate analysis of their dispersal patterns and improve management strategies. The study revealed that gene flow estimates among the populations were high and they lacked genetic differentiation. These findings suggested that the insects are capable of long distance dispersal and thus exhibit a high degree of gene flow across a broad region (Allander *et al*. 2008). The phylogeography and population structure of dung beetle species of the genus *Trypocopris* (Coleoptera, Geotrupidae) was studied using mitochondrial DNA cytochrome oxidase I sequencing and AFLP techniques. The results of mitochondrial and nuclear data revealed that there was phylogeographical structuring among populations and significant genetic differentiation was found among the populations and the AFLP data was able to clearly separate the geographical populations (Carisio *et al*. 2004). The pollen beetle (*Meligethes aeneus*) is one of the most important insect pests of oilseed rape (*Brassica napus*), and the

only effective control measure for the pest is extensive use of insecticides. AFLP analysis was used to study population dynamics as well as genotyping the pollen beetle pests for insecticide (pyrethroids) resistance. Different genotypes of pollen beetles showing geographical variation in resistance to insecticides were identified (Kazachkova *et al*. 2007). The study showed the existence of a low level of gene flow between pollen beetle populations (Kazachkova *et al.* 2008).

The AFLP technique has been used extensively for investigating the order Diptera. The genetic structure of populations of *Leptopilina boulardi*, a parasitoid of Drosophila was studied for the influence of geographic isolation (including physical distance) and ecology on the populations using AFLP and mitochondrial (Cytochrome Oxidase I) loci. The study led to the observation of high genetic differentiation between populations which indicated genetic structuring in AFLP profiles and gene flow between the populations. The results suggested that climate, geographic distance and physical barriers may all have contributed to the formation of genetically distinct populations of *L. Boulardi* (Seyahooei *et al*. 2011). Tsetse flies (Diptera: Glossinidae) are vectors of trypanosomes that cause sleeping sickness in humans and nagana in livestock across sub-Saharan Africa. A detailed understanding of the epidemiology and ecology of tsetse together with genetic variation within and among populations will help to develop Tsetse control strategies. AFLP markers were developed to analyze genetic variation in different geographical populations and laboratory strains. AFLP markers were able to distinguish genetically similar populations thus providing a useful tool for the genetic analysis of tsetse populations (Lall *et al*. 2010). The Asian rice gall midge (*Orseolia oryzae* Wood Mason) is one of the most important dipteran pests of cereals. AFLP analysis was used to assess the biodiversity of populations collected from different geographical locations. AFLP fingerprints were also able to detect sexual dimorphism in the adult gall midges and to distinguish gall midge from its major parasite *Platygaster oryzae (*Katiyar *et al*. 2000). The population genetic structure in relation to *Bti*-resistance (*Bacillus thuringiensis israelensis*) of the mosquito *Aedes rusticus* was analysed in different geographical regions of French Alps. The study revealed that the *Bti*-treatment had no effect on genetic structure and on within-population genetic diversity and positive selection was detected with AFLP data (Paris & Despers 2010). Yan *et al*. (1999)

compared the AFLP and RFLP markers for the study of population genetics of yellow fever mosquito.

Few of the aphids groups have also been investigated using *AFLP* markers. AFLP methods were employed to detect genetic diversity within grape phylloxera (*Daktulosphaira vitifoliae* Fitch) populations. Genetic variation was detected within all lineages and AFLP technique proved useful for the identification of reproducible banding patterns within clonal lineages. The study also revealed that the host plant did not affect the genetic structure of European phylloxera (Forneck *et al.* 2000, 2007). The geographic populations of woolly apple aphid *Eriosoma lanigerum* (Hausmann) were studied using AFLP to determine genetic structure of the pest populations. The low level of variation found in the study indicated that the possibility of controlling *E. Lanigerum* in the Western Cape using host plant resistance is favourable. This result suggested that AFLP markers could be useful for analysis of other aphid species (Timm *et al.* 2005).

Studies of genetic diversity of various insect groups using AFLP, supports the wide applicabilty of AFLP technique in determination of genetic diversity and studies related to population genetic structure. Examples include : Hemiptera (*Bemissia tabaci*, Cervera *et al*. 2000; Chu *et al*. 2008), Heteroptera (sothern chinch bug, Chandra *et al*. 2011; golden egg bug, Garcia-Gonzalez *et al*. 2005), Isoptera (*Nasutitermes takasogoensis*), Acrinid (grasshoppers, Tatsuta & Butin 2001), Hymemoptera (*Melipona quadrifasciata*, Makert *et al*. 2006; *Trichograma tidae*, Samara *et al.* 2006; *Apis nigrocincta*, Smith *et al*. 2003), aquatic insect species from Trichoptera and Hemiptera (Miller *et al.* 2002), Orthoptera (crickets of the genus *Laupala*, Parsons *et al*. 2001), Odonata (*Calopteryx spenders*, Sadeghi *et al*. 2010).

Nuclear DNA markers

Nuclear Ribosomal RNA genes (18S, 28S, 5.8S, ITS1 and ITS2)

Ribosomal genes have been studied extensively due to their important role in protein assembly. Since they are universal, have highly conserved sequences and

structures and are abundant they have been used for phylogenetic analyses of a wide range of taxa in different organisms (Mindell & Honeycutt 1990; Hamby & Zimmer 1992). The gene organization of rDNA is highly conserved which includes tandemly repeated transcription units separated from each other by intergenic spacers which in turn are arranged in the following order: 18S gene, Inter transcribed spacer 1, 5.8S gene, Inter transcribed spacer 2, and 28S gene. The ribosomal gene cluster has become a useful tool for classifying/diagnosing organisms at various taxonomic levels.

ITS1 and ITS2 genes

Molecular phylogenies of intra-species samples can only be obtained with fast-evolving DNA regions. The low level of evolutionary constraint in the spacers, combined with the rapid tendency of these arrays to be homogenized leads to the rapid spread of a new sequence variant which results in genetically isolated populations to be homogenized for different variants of rDNA. This process leads to high levels of intra-specific and intra-individual variation, which makes these regions highly useful for studies at the inter- and intra-population level (Mukabayire *et al*. 1999; Depaquit *et al*. 2000, 2002; Di Muccio *et al*. 2000; Yahia *et al*. 2004; Moinvaziri *et al*. 2007; Hamasheh *et al*. 2007). Hence, the ITS1 and ITS2 spacer sequences have proved to be useful for both species identification and phylogenetic analysis of closely related species.

The rDNA ITS region is an ideal candidate for such studies as it is present in all eukaryotes; it can easily be amplified with universal primers in the highly conserved flanking regions; and it shows a high rate of divergence, at least in comparison with the coding regions of the rDNA. However, little attention has so far been paid to the question of whether it is still under some selective constraints or whether it evolves with a rate that is close to the neutral rate of the remainder of the genome. Evidence for potential functional constraint comes from in vivo mutational analysis in yeast, where it was shown that the ITS regions are actively involved in the processing of the primary transcript (van der Sande *et al*. 1992). The copies of these rRNA genes evolve in concert, meaning a substitution in one quickly spreads to other copies (Hancock *et al.* 1988, Hillis & Dixon

1991). The rate of evolution of these genes varies along the length of the cluster (Hillis & Dixon 1991).

rRNA spacer sequences with variable regions have proved to be useful for both species identification and phylogenetic analysis of closely related species. The utility of nuclear ribosomal DNA for phylogenetic studies and study of molecular evolution has been described by Hillis and Dixon (1991). *Anopheles* is one of the most studied groups of insects where ITS1 and ITS2 have been used. Collins & Pakewitz (1996) have reviewed the use of rDNA markers for the differentiation of cryptic species by using PCR-mediated diagnostic primers. This is an example of an additional usefulness of ITS regions. But mainly ITS regions have been proved to be very useful in phylogenetic studies, especially in insects. There are several references available on work on *Anopheles* related to phylogenetic analysis. The research done by Dezfouli *et al*. (2002); Chen *et al*. (2006); Beebe (1999); Fritz (1994); Hackett *et al*. (2000); Marinucci *et al*. (1999); Mukabayire *et al*. (1999); Phuc *et al*. (2003); Walton *et al*. (2007); Li & Wilkerson (2005); Sallum *et al*. (2009) and Manonmani *et al*. (2000) are examples of the use of ITS2 for species identification.

Cruickshank (2002) has reviewed the use of various molecular markers including ITS1 and ITS2 for the study of phylogenetics of mites and ticks for use within species or for closely related species. Navajas *et al*. (1992) gave the first ever report of use of ITS2 in several species of tetranychidae mites. To devise IPM strategies using beneficial insects as biological control agents, it is important to have knowledge of population structure and dynamics of the insect under study. In this regard several molecular markers have been used and have been reviewed by MacDonald & Loxdale (2010). Few examples of use of ITS for the study of population structure and understanding the dynamics of beneficial insects include: Parasitic Wasps (Hymenoptera: Braconidae) (Quicke & Bleshaw 1999), Aphid Parasitoids (Hymenoptera: Braconidae: Aphidiinae) (Smith *et al*. 1999; Belshaw *et al*. 2000; Kambhapati *et al*. 2000; Sanchis *et al*. 2000; Bebcock *et al*. 2001; Schmidt *et al*. 2001), Microgastroid Wasps (Dowton & Austin 1998).

The genus *Diabrotica*, which are the most important insect pests of crops in United states, has also been studied extensively using ITS1 and ITS2 markers, by several authors, including Szalanski & Owens (2003), Szalanski *et al*. (1999), Roehrdanz *et al*.

(2003) and Clark *et al*. (2001). Genetic analysis of populations of southern corn rootworms, *Diabrotica undecimpunctata howardi* Barber was undertaken using DNA sequences of the nuclear DNA.

Szalanski *et al*. (2000) studied the population genetics and phylogenetics of the American burying beetle using ITS1 markers. Sequence comparison were done for parasitic ants and their respective host species (Hymenopyera: Fromicidae) by Baur *et al.* (1996). The study was conducted to determine the taxonomic positions of the species *A. atratulus* which led to the conclusion that, this species belonged to Tetramonini which was previously categorised to Solenopsididni based on morphological criteria. ITS1 and ITS2 have been reported to be used even for evolutionary studies of *Drosophila* by Schlotterer *et al*. (1994). Following this work, these regions were used in combination with other molecular markers for the study of inter- and intra-specific molecular phylogeny in mites (McLain *et al*. 1995; Navajas *et al*. 1994, 1998, 1999, 2000; Fenton *et al*. 2000; Murrel *et al*. 2006). ITS1 has similar properties to ITS2 and has been used for various population level studies but not as extensively as ITS2.

The secondary structure of the ITS regions which has been implicated in the processing of the primary transcript, is not directly dependent on primary sequence. In case of ITS1, reports from several species have shown that within this region there is a high propensity to form secondary structure (Furlong & Maden 1983; Michot *et al*. 1983; Gonzalez *et al*. 1990; Torres *et al*. 1990; Yeh & Lee 1990). Studies on the ITS1 sequences of lady bird beetles (Coleoptera, Coccinellidae species) showed that no putatively homologous similarities could be identified in the taxa studied, but some homologous regions were found to be conserved throughout the taxa (Schulenberg *et al*. 2001). ITS1 elements were found to be conserved at the 3'end. Similar findings have been made for drosophilid dipterans (insecta) (Schlotterer *et al.* 1994). Conserved secondary structure motifs primarily found at the 3'and 5'end, contain (as shown in yeasts) elements required for ribosome biogenesis (Henry *et al*. 1994; Van Nues *et al*. 1994; Weaver *et al*. 1997). Functional importance of the 3'end of the ITS1 sequences has also been suggested in the studies on metazoans (Schlotterer *et al.* 1994; Van der Schulenberg *et al*. 1999).

As the secondary structure of the ITS1 and ITS2 regions have been implicated to play an important role in the final processing of the rRNA, their secondary structure has also been used to decipher phylogenetic relationships in different groups of organisms. The few reports on the use of the ITS1 secondary structure include *Drosophila* (Armbuster *et al*. 2000; Armbruster 2001), *Poecilimon chopardi* (Ullrich *et al*. 2010), Digenea, Trematoda (Schulenburg *et al.* 1999), dinoflagellates (Marc & Jorg 2003), bush crickets (Orthoptera) (Ullrich *et al*. 2009), Volvocales (Mai & Coleman 1979), Boringales (Gottschling) and hookworms (Chilton & Gasser 1999).

As compared to the use of secondary structure of ITS1 in phylogenetic studies, several examples on the use of ITS2 secondary structure for this purpose have been reported. Schlotterer *et al*. (1994) studied eight *Drosophila* species and proposed an ITS2 secondary structure. In case of *Anopheles*, Beebe (1999) and Marrelli *et al*. (2005) have studied ITS2 structure for species identification. Severini *et al*. (1996) studied sequence and secondary structure of the rDNA ITS2 to differentiate sibling species of two mosquito species. Several beetles species were also studied using ITS2 secondary structure, *e.g.* meloid beetles (Bologna *et al*. 2007), Timarcha leaf beetles (Gomez-Zurita *et al*. 2000), raspberry beetles (Malloch *et al.* 2001) and pollen beetles (Trizzino *et al.* 2009). The secondary structure obtained was highly conserved and was comparable with the general model of eukaryotic secondary structure. Wolf *et al.* (2005) studied the extent of the conservation in ITS2 structure based on homology modeled secondary structures of more than 20,000 ITS2 covering about 14,000 species. Coleman (2007) studied different aspects of evolution and higher relationships across eukaryotes (including algae, plants, fungi, animals, molluscs and arthropods) using the secondary structure of the ITS2 region.

18S rDNA genes

18S rDNA sequences have been used for higher-level studies. 18S sequence has been successfully used for determining the relationships between various taxa of insects as a standard molecular marker (Caterino *et al*. 2000). Kjer (2004) and Kjer *et al.* (2006)

determined phylogeny of Insecta from 18S rDNA data based on structural and sequence alignment. Most of the studies involving 18S sequence are order level phylogenetic determination. Length variation in SSU/LSU rDNA in the group of sternorrhynchids (Hemiptera), was identified in several studies. Length variations caused difficulties in alignment and could affect phylogenetic results. In this study, the secondary structure model of Hexapoda SSU nrRNA was slightly adjusted and these corrections improved the quality of the data leading to correct phylogeny of the group (Xie *et al*. 2008). The utility of eight DNA sequence markers (5.8S rDNA, 18S rDNA, 28S rDNA, ITS regions, long-wavelength opsin, elongation factor 1-α, cytochrome *b*, and cytochrome oxidase I) were assessed in the reconstruction of phylogenetic relationships at various levels of divergence in gallwasps (Hymenoptera: Cynipidae). The results suggested that 28S rDNA, elongation factor 1-α, and long-wavelength opsin are more useful markers for the resolution of cynipid and other insect within-family-level divergences (circa 50–100). On the other hand mitochondrial loci and ITS regions were most useful for lower-level phylogenetics. In addition to that the 18S rDNA marker is more useful for the resolution of above-family-level relationships (Rokas *et al*. 2002). Phylogeny of the arachnid order Opiliones (Arthropoda) was inferred from a combined analysis of complete 18S rDNA and the D3 region of the 28S rDNA gene sequences and morphology and it helped to resolve the phylogenetic relationships (Giribet *et al*. 1999). Cruickshank (2002) reviewed the role of 18S rRNA in phylogeny of ticks and mites.

5.8S genes

It has been suggested that the sequence regions flanking the ITS2 sequence, *i.e.* the 3' end of the 5.8S gene and the 5' end of the 28S gene can pair with each other and this may be important for the correct processing of these genes and might therefore play an important role in directing the folding of the ITS2 region (Veldman *et al.* 1981; Subrahmanyam *et al.* 1982; Furlong & Maden 1983; Hindenach & Stafford 1983). Since the ITS regions are flanked by three highly conserved coding regions (18S, 5.8S and 28S ribosomal genes), PCR primers for DNA amplification of the region can be easily

designed. Navajas *et al.* (1992) for the very first time used primers in the 5.8S and 28S rDNA genes to amplify the ITS2 region in several species of tetranychid mites and showed that this region was suitable for examining phylogenetic relationships within genera, but high differences between genera prevented correct alignment of sequences. Navajas *et al*. (1999) compared the sequences of ITS1, ITS2 and 5.8S rDNA in several species of phytoseiids and concluded that ITS1 is more variable and difficult to align. ITS1 was also the only region that showed polymorphism within species. 5.8S rDNA was found to be more conserved than either ITS. Fenton *et al.* (2000) used primers in 18S and 28S rDNA to amplify a 1629 bp fragment of the nuclear ribosomal gene cluster spanning both of the internal transcribed spacers and 5.8S to study the coevolution of *Cecidophyopsis* mites and their *Ribes* hosts. The evolutionary origins of the mosquito family *Culicidae* was investigated by analysing the 18S and 5.8S ribosomal gene sequence divergence. Phylogenetic analyses supported the monophyletic position of the four taxas of the mosquitoes and the lack of a spacer in the 5.8S gene was unique to members of the Culicidae (Miller 1997). PCR-based diagnostic assay primers were developed for two of the sibling species of the *Anopheles fluviatilis* complex using species-specific differences in the nucleotide sequences of the second internal transcribed spacer (ITS2) region of the ribosomal DNA (rDNA). The assay thus developed was found to be highly specific and sensitive (Manonmani 2000). *B. bassiana* and *B. Brongniartii* are entomopathogenic fungi useful against several insect pests. These fungal populations collected from temperate, sub-tropical and tropical habitats was analysed using the ITS1-5.8S-ITS2 region (Ghikas 2010).

28S genes

As compared to 18S, the 28S sequence is small (less than 500bp) and it is difficult to compare results for this gene (Wheeler 1989; Carmean *et al*. 1992; Pashley *et al.* 1993; Chalwatzis *et al.* 1996; Whiting *et al.* 1997; Wheeler *et al.* 2001), A fragment of the large subunit of 28S from D3 region has been used in few studies (Holometablous insect orders, hexapod orders) but due to small size of the fragment it gave ambiguous

information (Whiting *et al.* 1997, 2002; Wheeler *et al.* 2001; Hovmoller *et al.* 2002). The very first use of the nearly complete 28S and 18S rRNA gene sequences to classify the arthropods and their kin was reported by Mallatt *et al.* 2003. This study revealed that the 28S gene contained more signal than the 18S gene, and contributed to the improved phylogenetic resolution.

Nuclear protein coding genes

Various nuclear genes have been used for arthropod phylogenetics (**Table 2**) (Brower & DeSalle 1994; Brower & DeSalle 1998; Friedlander *et al*. 1992; Friedlander *et al*. 1994; Graybeal 1994; McHugh 1998). A wide range of evolutionary rates have recently been assessed for phylogenetic utility of nuclear protein-coding genes. These loci can resolve relationships at levels which is too conserved to be easily examined with mitochondrial proteins. Because of the rate variation among codon positions in protein coding genes they may be effective for more recent divergences as well (Brower & DeSalle 1998; Cho *et al.* 1995; Mitchell *et al.* 1997; Reed & Sperling 1999). But the difficulties encountered with nuclear protein coding genes are heterozygosity and problem with amplification due to low copy number, presence of introns, low polymorphism in the coding regions i.e. exons and many single-copy loci may be represented by one or more variants (for examples in *EF-1*a Danforth 1998, Walldorf *et al*. 1985; Zhang & Hewitt 2003).

Table 2. Nuclear protein coding genes used for phylogenetic analysis of insects

Nuclear gene	**Insect**	**Reference**
a-amylase	*Drosophila melanogaster*	Okuyama *et al.* (1996)
	Drosophila melanogaster	Shibata & Yamazaki (1995)
Acetylcholine esterase (*achE*)	*Hawaiian drosophilids*	Baker & DeSalle (1997); DeSalle & Brower (1997)
Actin	*Bombyx mori* and *Helicoverpa armigera*	Mange & Prudhomme (1999)

Alcohol dehydrogenase (*adh*)	*Drosophila* species	Albalat & Gonzalez-Duarte (1993)
soluble guanylate cyclase	Anopheles gambzae complex	Garcia *et al*. (1996)
*EF-1*a	Aphidiinae	Belshaw & Quicke (1997)
	Diopsidae flies	Baker *et al.* (2001)
	heliothine moths	Cho *et al.* (1995)
	Apis mellifera	Danforth & Shuqing (1998)
	Noctuoidea superfamily (e.g. Heliothinae group of species)	Mitchell *et al*. (1997)
	stilleto flies	Yang *et al*. (2000)
	Bicyclus species	Monteiro & Naomi (2000)
	aphids	Moran *et al*. (1999)
	Gallwasps	Rokas *et al.* (2002)
Long wavelength opsin	Gallwasps	Rokas *et al.* (2002)
Period (*per*)	*Drosophila*	Costa *et al.* (1991); Gleason *et al*. (1998)
	Drosophila	Nielsen *et al.* (1994)
Resistance to dieldrin (*Rdl*)	Hypofhenemus hampei (beetles)	Andreev *et al.* (1998)
Wingless (*wg*)	*Hawaiian drosophilids*	Baker & DeSalle (1997)
	Butterflies	Campbel *et al.* (2000)
	Heliconius butterflies	Brower & Egan (1997)
	Diopsidae flies	Baker *et al.* (2001)
white	Diopsidae flies	Baker *et al.* (2001)

Mitochondrial genes (rDNA and protein coding genes)

The most frequently sequenced mitochondrial genes in insects are cytochrome oxidase I (COI), COII, 16S rDNA, 12S rDNA and ND5. These genes are described in the order of their representation for their use in the phylogenetics of insects.

***Mitochondrial cyotochrome oxidase II* (*COII*)**

This is the most widely used and sequenced mitochondrial protein coding gene in insects and has been used in the phylogenetic analysis of several insects (**Table 3**).

Table 3. Use of Mitochondrial COII gene in insects for phylogenetic analysis

Insect	**Reference**
Complete COII	
Sympetrun Strolatum	Liu & Beckenbach (1992)
A. domesticus	Liu & Beckenbach (1992)
S. gregaria	Liu & Beckenbach (1992)
P. americana	Liu & Beckenbach (1992)
Z. angusticollis	Liu & Beckenbach (1992)
o. fasciatus	Liu & Beckenbach (1992)
Adalia bipunctata	Liu & Beckenbach (1992)
Sitophilus granarius	Liu & Beckenbach (1992)
Apis (5 Spp.)	Willis *et al.* (1992)
Excristes roboratus	Liu & Beckenbach (1992)
C. felis	Liu & Beckenbach (1992)
Drosophila pseudoobscura	Liu & Beckenbach (1992)
Drosophila (12 SPP.)	Beckenbach (1993)
Calliphoridae (3 genera)	Sperling *et al*. (1994)
Christoneura (6 spp.)	Sperling & Hickey (1994)
Anopheles gambiae	Beard *et al*. (2007)

Hymenoptera group of insects	Dowton & Austin (1999)
Heterodoxus macropus	Shao *et al*. (2001)
Chryosomelidae spp.	Dunk *et al*. (1999)
Hegeter spp.	Juan *et al*. (1996)
Rhagoletis spp.	Smith & Bush (1997)
Reticulitermes spp.	Austin *et al*. (2002)
Crioceris duodecimpunctata	Stewart & Beckenbach (2003)
Tribolium castanaeum	Friedrich & Muqim (2003)
Rhinotermitidae spp.	Austin *et al*. (2004)
Partial COII	
A. mellifera (10 subspp)	Garneys *et al*. (1992)
A. cerana	Garneys *et al*. (1992)
Greya (15 spp.)	Brown *et al*. (1994)
P. quinquepunctellus	Brown *et al*. (1994)
T. gei	Brown *et al*. (1994)
Heliconius (sensulato, 37 spp.)	Brown (1994)
D. juno	Brown (1994)
D. phaetusa	Brown (1994)
D. iulia	Brown (1994)
P. dido	Brown (1994)
P. telesiphe	Brown (1994)
Cicindela dorsalis (4 subspp.)	Volger & DeSalle (1993)
Cicindela puritama	Volger & DeSalle (1993)
C. dorsalis (2 subspp.)	Volger & DeSalle (1993)
Cicindela (6 spp.)	Volger & DeSalle (1993)

Mitochondrial cyotochrome oxidase I (COI)

The COI gene has been sequenced less frequently than the COII gene and is one of the most conserved genes in terms of amino acid evolution. **Table 4** lists few examples where COI has been used for phylogenetic analysis in various groups of insects.

Table 4. Use of COI gene for the insect phylogeny

Insect	**Reference**
COI complete	
Calliphoridae (3 genera)	Sperling *et al*. (1994)
COI partial	
A. mellifera	Garnery *et al*. (1992)
A. cerana	Garnery *et al*. (1992)
D. melanogaster group	Satta *et al.* (1987); Satta & Takahata (1990); Nigro *et al.* (1991)
Choristoneura (6 spp.)	Sperling & Hickey (1994)
Greya (15 spp.)	Brown *et al.* (1994)
Prodorus quinquepunctellus	Brown *et al*. (1994)
Tetragma gei	Brown *et al.* (1994)
Heliconius (37 spp.)	Brower (1994)
Dione juno	Brower (1994)
Dryadula iulia	Brower (1994)
Dryadula phaetusa	Brower (1994)
Philaethria dido	Brower (1994)
Podotricha telesiphe	Brower (1994)
Chorthippus parallelus	Lunt *et al*. (1998)
Neochlamisus spp.	Funk (1999)

Phylogenetics of all major genera of insects has been studied using this fragment of COI-tRNA^{Leu}-COII. For example, Lepidoptera has been studied by various authors. Beltron *et al.* (2002) used mt DNA encompassing COI and COII to infer phylogenetic gene genealogies and to determine relationships between closely related species of *Heliconius* Butterflies. Bonebrake *et al*. (2011) made phylogenetic estimation and evolutionary assessment of *Chlosyne lacinia.* Species complex of pest insects, the spruce budworms *(Choristoneuru fumiferana* species group) was studied using a combination of polymerase-chain-reaction amplification and automated DNA sequencing to investigate genetic variation. The work proved suitability of these marker systems for assessing population structures, species limits, and phylogenetic relationships among organisms

that have not previously been subjected to DNA analysis (Sperling & Hickey 1994). Genetic variation within two Neotropical butterflies, *Heliconius charithonia* and *Dryas iulia* was studied using mitochondrial DNA and allozyme variation and biogeographic patterns were obtained which explained the phylogeography of the populations (Neil Davis & Berminghan 2002). Phylogeography of *Hyles tithymali* on the Canary Islands was studied usning mtDNA sequence (COI + II, tRNA-leu) and ISSR)-PCR (Hundsdoerfer & Wink 2006).

Mt COI/COII region was used as a molecular diagnostic marker for the estimation of systematic and phylogenetics of *Dioryctria abietella* Species Group (Lepidoptera: Pyralidae) (Roux-Morabito *et al.* 2008). COI/COI gene region was used for the phylogeneic analysis of several Diptera and Hymenoptera species groups (*Apis mellifera* ligustica (Crozier *et al.* 1989); Muscoidea **(**Bernasconi *et al*. 2000); *Apis cerana* and *A. nigrocincta (*Smith *et al.* 2000); *Apis cerana (*Hepburn *et al*. 2001); Themira genera (Laamanen *et al*. 2005); *Chrysotoxum* species (Masetti *et al*. 2006); *Apis mellifera scutellata (*McKern & Szalanski 2007); Sesiidae genera (McKern *et al*. 2008); *Melipona bicolor* (Silvestre *et al*. 2008); *Chrysomya megacephala* (Siew Hwa Tan *et al*. 2009); *Apis cerana* indica (Baskaran 2011). Complete mitochondrial genome of *Cochliomyia hominivorax* (Diptera: Calliphoridae) was sequenced and phylogenetic analysis was done by Lessinger *et al*. (2000).

Beetles have been extensively studied using this approach of mt COI/COII gene region. Boyce *et al.* (1989) studied mt DNA size, structure and heteroplasmy in three species of Curculionidae (Bark beetles). Freidrich & Muquit (2003) described the first ever complete mitochondrial genome sequence from insect order Coleoptera (species of the flour beetle *Tribolium castaneum*). Different beetles species and groups were studied using this region of mt DNA (Iberian *Zabrus (*Sanchez *et al*. 2004); *Hypera postica* (Kuwata *et al*. 2005); *Tomicus destruens* (Horn *et al*. 2006 and Vasconcelos *et al.* 2006); Dendroctonus ponderosae (Mock *et al*. 2001). Complete mitochondrial genome of carabid beetle (*Dasater mirabilissimus* mirabillissim) was analysed and gene arrangement was studied by Wan *et al*. (2012). Szalanski *et al.* (2000) studied *Diabrotica* species for phylogenetic assessment. Host-associated genetic differentiation was studied by Hernandez Vera *et al*. (2010) using a combination of mt DNA COII and EF1-α.

Finn *et al*. (2007) used mt COI/COII to study the Madrean Sky Islands populations of giant water bugs *Abedus herberti* (Hemiptera) to determine population genetic structure and phylogeography. Young *et al*. (2012) used mt genome of Katydis, *Xizicus fascipes* (Orthoptera) for phylogenetics. Shao & Baker (2003) sequenced the mt genome of *Thrips imginis* (Thysanoptera), the plague thrips, to understand the mechanisms of gene rearrangement. Shao *et al*. (2004) studied mt genomes of two major groups of ticks, soft ticks and hard ticks. In this study mitochondrial genome was sequenced and comparative study was done for gene rearrangement. Shu-Jun Wei *et al.* (2010) sequenced mt genome from seven subfamilies of Braconidae and assessed the utility of mt genome data for phylogenetic analysis. Steven *et al.* (2005) studied systematics of three genera of stick insect from New Zealand using COI/COII. Menon *et al*. (2006) used morphological characters and mt DNA COI/COII to describe a new species of *Halys fabricius* (Pentatomidae) and proved usefulness of DNA sequences for Halys taxonomy. Complete mt genome of Olive fly *Bactrocera oleae* was sequenced and phylogenetic analysis was done by Nardi *et al.* (2003).

16S rDNA gene

16S rDNA is similar in pattern of variation as those of 12S rDNA in having a highly conserved 3'segment but studies have suggested that like 3'half of the 12S, the 3'half of the 16S is not very useful for phylogenetic studies of recently diverged populations within species because few sites vary (Uhlenbusch *et al*. 1987, Gutell *et al*. 1992, Pashley & Ke 1982). Few of the examples where 16S gene has been used include *Drosophila* (Desalle 1992a, b), blackflies (Xiong & Kocger 1991, 1993b; Vogler *et al*. 1993a; Vogler & DeSalle 1993). It appears that the conserved region of the 26S is more useful for studies of more distantly related taxa (DaSalle 1992; Cameron 1991, 1993). A region of the 16S gene corresponding roughly to positions 12900–13400 (relative to *Drosophila yakuba;* 103) has been consistently sequenced across most taxa with the exception of Lepidoptera.

12S rDNA gene

The 12S rDNA gene, especially the 3'segment of the gene, has been useful for phylogenetic analysis for distant taxa but not for recently diverged species (Ballard 1994; Thomas *et al*. 1989). This region has been used for phylogenetic analysis of *Anopheles hilli* and *Musca* (Ballard *et al*. 1992) and *Drosophila* (Clary & Wolstenholme 1987; Nigro & Grapputo 1993). 12S rDNA gene is similar in properties to 16S gene and has been used for phylogenetic studies of insects including *Austrosimulium spp.* (Ballard 1994); *Onychophoridae* spp. (Ballard *et al*. 1992); cockroach families (Kambhapati 1996); *Tracheata* spp. (Waegele & Stanjek 1995).

ND5 gene

It has been observed that all mitochondrial protein coding genes behave in a similar fashion when subjected to phylogenetic analysis but the length of the gene is also important before considering a mitochondrial protein coding gene for this purpose (Simon *et al*. 1994). Similar to other mitochondrial protein coding genes, ND5 gene (5^{th} subunit of NADH dehydrogenase complex) has been utilized for the phylogenetic analysis of diverse taxa of insects for example, *Acrythosiphon pisum* (Birkle & Douglas 1999); *Anopheles maculipennis* complex (Walton *et al.* 1999); *Apis mellifera* (Crozier & Crozier 1992); *Sarcophaga* spp. (Bajapi & Tewari 2010); *Aedes albopictus* (Usmani-Brown *et al*. 2009); *Drosophila birchii* (Schiffer *et al*. 2007).

Of these mitochondrial regions, COII has been sequenced in a wide variety of taxa. Similarly, COI has been sequenced in various groups. COIII, NADH dehydrogenase 5 (ND5) and *cyt* b sequences have few representation while ND2, ND4, and ND1 have scarce representation. In general, sequences of larger fragments (more than 1 kb) reveal more information in terms of various levels of variation, giving resolution for a greater range of divergences. Mitochondrial DNA-sequence data have been used successfully for several phylogenetic analyses. It appears that silent sites of mitochondrial protein coding genes evolve at the same rate in all mitochondrial genes (Simon *et al*. 1994). Although

the number of different genes sequenced is highly variable, and there is no explanation for this kind of preference of one gene to another (Caterino *et al.* 2000). Tables 3 and 4 list other mitochondrial genes which have been used for the phylogenetic analysis of insects. It is suggested that ribosomal genes are most likey to be useful at the population level where highly variable sites have not yet experienced multiple substitutions and at deep levels of divergence where the more conserved sites, which are likewise free of homoplasy, supply useful phylogenetic information.

Simon *et al.* (1994) reviewed the utility of mitochondrial gene sequences in molecular phylogenetics. The advantages of using DNA sequence data to estimate phylogenetic relationship is that the processes governing the evolution and inheritance of DNA are already understood. There are several DNA sequence regions which can be utilized for phylogenetic analysis. Although there is no clear cut generalization about which gene to be used, but the majority of genes chosen for study are mitochondrial because they are abundant, easier to manipulate; are clonally inherited; single copy and non-recombining (Brown 1985; Moritzetal 1987; Attardi & Schaltz 1988; Wolstenholme 1992).

Avise *et al.* (1987) has reviewed the role of mitochondrial DNA as a bridge between population genetics and systematics. Animal mt DNA shows properties of maternal inheritance, non recombining mode of inheritance, rapid pace of evolution, and extensive intra-specific polymorphism. It is compact, in the sense that it lacks introns, repetitive DNA, pseudogenes and long spacer sequences between genes. Gene arrangement is mostly stable within a taxonomic class. Rate of evolution at nucleotide sequence level is rapid (1-10 times faster) than single-copy nuclear DNA. Most of the genetic changes are found to be nucleotide substitutions, and there are very few insertions/deletions. Most importantly, due to maternal inheritance, no recombination is observed.

Animal mt genomes are small, circular DNA, 14,000-17,000 bp) long and encode approx 37 genes (13 protein-coding, 22tRNA, and 2 rRNA genes). Evolutionary unique features of mitogenes are length variation (Boyce *et al*. 1989), altered tRNA anticodons or secondary structure (Steinberg & Ledregren 1994; Eddy 200), atypical start codons (*e.g.* Lavrov *et al.* 2004), base compositional bias (Gowri-Shankar & Rattry 2006), codon

usage (Jia & Higgs 2007) and gene rearrangements (Zang & Hewitt 1997; Shao & Barker 2003; Mueller & Boore 2005).

Sheffield *et al.* (2008) has made a comparative analysis of mitochondrial genomes in Coleoptera (arthropoda: insecta). There are four suborders of order Coleoptera (adephaga, Archostemata, Myxophaga and Polyphaga). The mitogenomes of coleopterans suggest that the gene rearrangement in this group follows that of ancestral insects. There were no rearrangements, duplications or deletions of any genes within these mitogenes. This suggests that there have not been significant gene rearrangements during the diversification of coleopterans. This molecular stability is remarkable within such a diverse group because most of the insect taxonomic groups exhibit rearrangements that are diagnostic to the group (Dowton & Austin 1998; Thao *et al.* 2004; Castro *et al.* 2006; Cameron & Whiting 2008). In fact, only Diptera appears to be as conservative with respect to mitogenome structure as coleoptera (Cameron & Cambkin 2007).

This section can be concluded with a note on the importance of comparative studies of both nuclear and mitochondrial derived genes for a complete understanding of population structure and phylogeny. This is due to the varying degree of information obtained from both the systems because they differ in their mutation rates as well as mechanism. If both systems are studied simultaneously, the incongruencies observed in the phylogenies based on these two genes for insect molecular systematics could be complemented for the information lacking in each. A simultaneous study of two genes which differ in their mode of inheritance makes it possible to assess whether the observed genetic differentiation is due to a bias in dispersal of one of the genders, because male and female insects of several species have been reported to differ in their migratory and colonization behaviour (Roderick 1996; Sielezniew *et al*. 2011). For markers such as the mitochondrial markers which are uni-parentally inherited, the male sex does not contribute to the mitochondrial genome of the off-spring, whereas for nuclear genes such as the rDNA genes which are biparentally inherited, both sexes contribute to the diversity in the progeny. Hence a difference in the genetic structure between these two markers is expected when sex biased dispersal occurs. For species in which the males disperse and the females are philopatric, the estimated genetic differentiation between populations is

expected to be higher when estimated using mitochondrial DNA or any other maternal marker, than using a biparental marker.

SECTION 2

Use of genetic engineering techniques in the development of insect resistant crop plants

In modern-day agriculture there are several challenges which are faced by the farmers such as the ever increasing demands on yield and the intensification of farming practice, which has led to increased problem of damage by pests and development of pesticide resistance in insect pests. Hence, an important consideration in Integrated Pest Management (IPM) is development of suitable control measures. Hence protection plays a vital and integral role in agricultural sustainability. The problem has been worsened due to the practice of monoculture (extensive cultivation of certain crops to the exclusion of other plants) leading to drastic increase in the populations of insects that feed upon these crops (Strong 1979). A similar situation is seen in storage of food crops, where accumulation of large quantities of seed attracts several kinds of insects, resulting in significant damage and loss.

An important approach for integrated pest management practice would be the use of crop varieties resistant to insect attack, which will reduce the use of chemical pesticides. In this regard transgenic plant technology can be a useful tool in producing insect resistant crops, by introducing novel resistance genes into a plant species. This technology has been integrated in many crop management programmes. An integrated pest-control programme comprising of a combination of practices including the judicious use of pesticides, crop rotation, field sanitation, and more importantly, use of inherently resistant plant varieties, would provide the best option for improving pest control. Within this last category, the use of genetically engineered insect resistant crops may be included.

The research in this area has focussed on different classes of plant proteins which are shown to be insecticidal towards a range of economically important insect pests from different orders. Genes encoding insecticidal proteins have been isolated from various plant species and transferred to crops by genetic engineering. These genes encode inhibitors of proteases (serine and cysteine) and α-amylases, lectins, and enzymes such as

chitinases and lipoxygenases. This section gives a broad overview of genetically engineered crops expressing insecticidal plant proteins from different plant species, with enhanced resistance to one or more insect pests from different orders of insects. The number of different crop species expressing such genes is very diverse and ever-increasing.

The various strategies used to protect plants against insect pests

Genetic engineering of crops can contribute immensely towards the development of inherently resistant/tolerant crop varieties as it opens a virtually limitless source of germplasm variability which can be used to select insect control genes for introduction into several important crop varieties. As the number of crops subjected to genetic transformation is increasing, identifying genes encoding desirable characteristics, such as pest resistance, becomes increasingly important. To date, two main strategies for producing such plants are being used successfully. One classical approach is to use the entomicidal bacterium *Bacillus thuringiensis* Berliner as a source of resistance genes, and the other is to identify and use the insect resistance genes present in plants themselves i.e. exploit the plants' own defense to the problem.

Natural defense mechanisms of plants

There are three types of plant-based defense mechanisms which occur naturally in plants such as temporal avoidance, physical and chemical defenses. The chemical defense is one of the most important classes of natural defense mechanisms which is a result of the plant’s vast metabolic capability which produces many secondary chemicals which are toxic, anti-nutritional, or aversive to predator species (Sharma & Norris 1991). Examples of such plant-derived chemicals include the pyrethrins from chrysanthemums and alkaloids like nicotine from tobacco. Other classes of plant secondary compounds which have been implicated in protection from insect attack include the terpenoids, steroids, navanoids, phenolics, glucosinolates, cyanogenic glycosides, rotenoids, saponins and nonprotein amino acids (Gatehouse 1991). These secondary compounds are the

products of multi-enzyme pathways and the interaction of many gene products in the process, which makes these systems too complex for use in plant genetic engineering, although efforts are being done to exploit them (Hallahan 1992). Other plant defense mechanisms are based on proteins, the product of a single gene. Because proteins are non-volatile, the only way that they can be effective is that they must be ingested so that they may reach their target site which is the insect digestive system.

One of the earliest examples of such kind of natural defense includes plant protein arcelin-4 which provides resistance to wild accessions of the common bean (*Phaseolus vulgaris*) against the Mexican bean beetle (*Zabrotes subfasciatus*) which replaces the conventional bean storage protein, phaseolin. Arcelin-4 is undigestible by the insect pest and so the larvae starve to death on these resistant beans (Minney 1990). Proteinaceous inhibitors of insect digestive enzymes have a similar effect (Gatehouse 1979) and few lectins of plant origin bind and disrupt cells of the insect gut epithelium (Gatehouse 1984). Similarly ribosome inactivating proteins (RIPs) such as ricin from castor beans have also been shown to be highly toxic to certain insects (Gatehouse 1990). These single gene products are useful in plant protection using genetic engineering techniques as their genes can be transferred form one plant to another thereby providing protection to the transgenic plants. Since they are of plant origin, they have the added advantage of high degree of compatibility with the metabolic system of the transgenic host plants.

Use of Cry proteins in crop protection

The very first insect-resistant transgenic tobacco plant was developed in 1987, with the genes encoding a Cry toxin derived from a soil bacterium *Bacillus thuringiensis* (*Bt*) Berliner (Vaeck 1987). Since then, many novel Cry proteins and new insecticide proteins, such as protease inhibitors, lectins, and α-amylase inhibitors, have been used to develop genetically engineered crop varieties (Malone 2008). The site of action of the *Bt* toxin is the insect mid-gut. In susceptible insects, the *Bt* protein crystals dissolve in the midgut, releasing protoxins that are in turn proteolytically converted into smaller toxic polypeptides. These toxins bind with high affinity to glycoprotein receptors on the midgut epithelium and generate pores in the cell membrane, thereby disturbing the cellular

osmotic balance and causing the cells to swell and lyse through a process called `colloid osmotic lysis'. Binding site assays have indicated that significant differences exist among binding sites in some species such that some sites may bind a single toxin, whereas other sites may bind two or more toxins (Lee 1997; Peferoen 1997; Bhau & Koul 1998).

Transgenic crop cultivar of *Bt* maize and *Bt* cotton were first approved for commercial release in the USA in 1995, and were first planted in 1996 (Hellmich 2008; Naranjo 2008; James 2009). Since then, the number of such crops being grown in different countries worldwide has increased steadily. These *Bt* crops provide an effective alternative tool for controlling target insects (Wu 2008; Wang 2007; Brookes & Barfoot 2010; Choudhary & Gaur 2010; Huang 2010; Hutchison 2010; Tabashnik 2010). **Table 5** lists some important crops which have been transformed with *Bt* gene and being successfully used.

Though Bt toxin is being used for the past 50 years as an exogenous natural pesticide, its use has led to several debates on consumer health and environmental safety. The inhibition of growth of the larvae due to the physiological stress caused by inhibitors is advantageous in tritrophic interactions (plants, pests and predators), because the retarded insects become easy targets for greater parasitism by natural enemies (Heath 1997; Lewis *et al*.1997).

Table 5. Crop varieties transformed with *Bacillus thuringiensis* genes for developing resistance to target pests*

Target pests of crops and order	**Insect toxin genes**	**Crop**
Tobacco budworm (*Helioth virescens*) (Lepidoptera)	*Cry1Ac; Cry1Ab/c; Cry1Ac+Cry2Ab; Cry1F; Vip3A; Vip3A+Cry1Ab*	Cotton, Tomato
Pink bollworm (*Pectinopho gossypiella*) (Lepidoptera)	*Cry1Ac; Cry1Ab/c;* Cry1Ac+CpTI; *Cry1Ac+Cry2Ab; Vip3A; Vip3A+Cry1A*	Cotton, Tomato
Cotton bollworm (*Helicoverpa zea*) (Lepidoptera)	*Cry1Ac; Cry1Ab/c; Cry1Ac+Cry2Ab; Cry1F; Vip3A ; Vip3A+Cry1Ab; Cry1A.105+Cry2Ab +Cry3Bb1+Cry34Ab1 +Cry35Ab1+Cry1Fa2; Vip3Aa20+mCry3A +Cry1Ab*	Cotton, Tomato

Cotton bollworm (*Helicoverpa armigera*) (Lepidoptera)	*Cry1Ac; Cry1Ab/c;* Cry1Ac+CpTI; *Cry1Ac+Cry2Ab; Vip3A+Cry1Ab*	Cotton , Maize, Tomato
Spodoptera spp.(Lepidoptera)	*Cry1A+Cry1F*	Cotton
Beet armyworm (*Spodoptera exigua)* Lepidoptera)	*Cry1F; Vip3A; Vip3Aa20+mCry3A +Cry1Ab*	Cotton, Maize
Fall armyworm (*Spodoptera frugiperda*) (Lepidoptera)	*Vip3A; Cry1F;* Cry1A.105+ Cry2Ab2; *Cry1A.105+Cry2Ab +Cry3Bb1+Cry34Ab1 +Cry35Ab1+Cry1Fa2; Vip3Aa20+mCry3A+Cry1Ab*	Cotton, Maize
Soybean looper (*Pseudoplusia includens*) (Lepidoptera)	*Cry1F*, *Vip3A*	Cotton
Cabbage looper (*Trichoplusia ni*) (Lepidoptera)	*Vip3A*	Cotton
Cotton leaf perforator (*Bucculatrix thurberiella*) (Lepidoptera)	*Vip3A*	Cotton
European corn borer (*Ostrinia nubilalis*) (Lepidoptera)	Cry1Ab; *Cry1Ab+mCry3A; Cry1F;* Cry34Ab1+Cry35Ab1+Cry1Fa2+Cry1; Cry1A.105+Cry2Ab2Fa2; Cry1A.105+ Cry2Ab2; *Vip3Aa20+mCry3A +Cry1Ab; Cry1A.105+Cry2Ab +Cry3Bb1+Cry34Ab1 +Cry35Ab1+Cry1Fa2*	Maize
Corn rootworm (*Diabrotica* spp.) (Coleoptera)	*Cry1Ab+mCry3A;* Cry34Ab1+Cry35Ab1 Cry34Ab1+Cry35Ab1+Cry1Fa2+Cry1Fa Cry3Bb1	Maize
Southwestern corn borer (*Diatraea grandiosella) (*Lepidoptera*)*	*Cry1F; Cry1A.105+Cry2Ab +Cry3Bb1+Cry34Ab1 +Cry35Ab1+Cry1Fa2; Vip3Aa20+mCry3A+Cry1Ab*	Maize
Western bean cutworm (*Striacosta albicosta*) (Lepidoptera)	*Cry1F; Vip3Aa20+mCry3A +Cry1Ab; Cry1A.105+Cry2Ab +Cry3Bb1+Cry34Ab1 +Cry35Ab1+Cry1Fa2*	Maize
Black cutworm (*Agrotis*	*Cry1F; Cry1A.105+Cry2Ab*	Maize

ipsilon) (Lepidoptera)	*+Cry3Bb1+Cry34Ab1 +Cry35Ab1+Cry1Fa2; Vip3Aa20+mCry3A+Cry1Ab*	
Western corn rootworm (*Diabrotica vigifera vigifera*) (Coleoptera)	*mCry3A; Cry1A.105+Cry2Ab +Cry3Bb1+Cry34Ab1 +Cry35Ab1+Cry1Fa2; Vip3Aa20+mCry3A+Cry1Ab*	Maize
Northern corn rootworm (*Diabrotica berberi*) (Coleoptera)	*mCry3A; Cry1A.105+Cry2Ab +Cry3Bb1+Cry34Ab1 +Cry35Ab1+Cry1Fa2; Vip3Aa20+mCry3A+Cry1Ab*	Maize
Mexican corn rootworm (*Diabrotica vigifera zeae*) (Coleoptera)	*mCry3A; Cry1A.105+Cry2Ab +Cry3Bb1+Cry34Ab1 +Cry35Ab1+Cry1Fa2; Vip3Aa20+mCry3A+Cry1Ab*	Maize
Sugarcane borer (*Diatraea saccharalis*) (Coleoptera)	*Cry1A.105+Cry2Ab +Cry3Bb1+Cry34Ab1 +Cry35Ab1+Cry1Fa2; Vip3Aa20+mCry3A+Cry1Ab*	Maize
Armyworm (*Pseudaletia unipunctata*) (Lepidoptera)	*Vip3Aa20+mCry3A+Cry1Ab*	Maize
Southern cornstalk borer (*Diatraea crambidoides*) (Lepidoptera)	*Vip3Aa20+mCry3A+Cry1Ab*	Maize
Common stalk borer (*Papaipema nebris*) (Coleoptera)	*Vip3Aa20+mCry3A+Cry1Ab*	Maize
Colorado potato beetle (*Leptinotarsa decemlineata*) (Coleoptera)	*Cry3A*	Potato
Rice stem borers (*Scirpophaga incertulas*) (Lepidoptera)	*Cry1Ab; Cry1Ab/c*	Rice
Rice stem borers (*Chilo suppressalis*) (Lepidoptera)	*Cry1Ab; Cry1Ab/c*	Rice
Fruit and shoot borer (*Leucinodes orbonalis*) (Lepidoptera)	*Cry1Ac*	Eggplant
Diamondback moth (*Plutella xylostella*) (Lepidoptera)	*Cry1*	Crucifer vegetabl

*Source: Yu *et al.* 2011

Plant derived insecticidal proteins

With the advancement in molecular techniques, a wide range of plants have been genetically transformed and new resistant varieties have been developed with improved resistance to pests, without affecting important traits like palatability, nutritional quality and yield. Efforts in this direction include identification and purification of the potential candidate proteins, bioassay of these selected proteins with the desired insecticidal properties, isolation of the encoding gene(s), and incorporation of these genes into vector constructs which express the protein within the host plant, and in the appropriate plant tissues. Genes encoding protease inhibitors, α-amylase inhibitors, lectins, chitinases, etc have been used as a source of resistance against the insect pests. The mechanism of action of the proteinase inhibitors differs from that of *Bt*. The antimetabolic mode of action of the inhibitors to the digestive enzymes appears to be the hypersecretion of digestive enzymes caused by the presence of the inhibitors, resulting in depletion of the essential amino-acids, in addition to direct inhibition of the digestive enzymes (Broadway & Duffey 1986 a, b; Gatehouse 1992) resulting in physiological stress and thus inhibition in growth. Inhibitors also affect the moulting, water balance and enzyme regulation of the insects (Boulter 1993). The expression of independent insecticidal genes in the same insect, each having its own mode of action on the different biochemical and physiological processes would help to provide a multi-mechanism defense and will help to delay resistance developed by the pests, thus leading to a more durable resistance. These various insecticidal genes have been discussed in the following paragraphs. In addition, few successful applications of pyramiding of genes have also been described (**Table 5**).

There are different kinds of protease inhibitors which have been found to be active against different species of insect (Gatehouse 1979; Stefens 1978; Broadway & Dufey 1986; Burgess 1991; Burgess 1994; Johnston 1991; Johnston 1993; Johnston 1995). Gatehouse & Gatehouse (1998) have reviewed the transgenics which have been developed by the use of various plant-based inhibitors of different origins. These studies led to the identification of potential inhibitors which could be used in crop protection. The

very first gene of plant source which was transferred successfully to another plant species was the gene encoding the double-headed trypsin inhibitor (CpTI) isolated from cowpea, which resulted in enhanced insect resistance (Hilder 1987). Following this, other types of proteins such as alpha amylase inhibitors and lectins form different sources were utilized to test their efficacy and they proved useful for their application in crop protection of important crops (**Table 6**). Inhibitors of digestive enzymes ie the proteases and the amylases as well as lectins which are toxic to a wide range of insect orders have found application as insecticidal genes. In some cases chitinase has also been found to be effective (Peter 2010).

An effective strategy to attain durable resistance involves the simultaneous introduction into the same plant of different resistance genes, each with a different mode of action such that the biochemical and physiological processes in the insect are adversely affected by mechanisms other than the mechanism by which the *Bt* toxins act (de Maagd *et al.*1999). Development of resistance to such transgenic plants in which different resistance genes have been pyramided, would then require several mutations to occur simultaneously in the same insect or atleast the accumulation of several independent mutations through sexual crossing in the susceptible insect population (Wunn *et al*.1996).

Several examples are available, where a plant has been transformed with more than one insecticidal gene. The strategy of pyramiding genes has been adopted by biotechnologists to increase the protection efficiency, range of activity and sustainablity of resistance. In this technique packages of different genes are introduced into crops and each components of such packages acts on a different target within the insect and in doing so they mimick the multi-mechanistic resistance that occurs in nature. In this respect protease inhibitors are particularly valuable, because along with the insecticidal effects, they can protect other introduced gene products from premature digestion in the insect gut. The very first example of this approach was the introduction of both CpTI (cowpea trypsin inhibitor) and pea lectin into tobacco (Boulter 1990). Other than protease inhibitors, other insect-resistance genes have been used such as transgenic potato plants have been produced which express both snowdrop lectin and bean chitinase. The results obtained showed that pyramiding the genes encoding these particular proteins had a synergistic effect (Down 1996, 2000).

Table 6. Transgenic Plants Expressing Insecticidal Plant Genes*

Genes	Plant	Insect	Reference
Cowpea trypsin inhibitor	Tobacco, Potato, Rice, Strawberry, Lettuce, Apple, Sunflower, Sweet potato	*Heliothis virescens, Lacanobia oleracea, S. inferens* a, *C. suppressalis, Otiorhynchus sulcatus, Cydia pomenella,* Coleoptera and Lepidoptera	Hilder 1987; James *et al.* 1992; Newell *et al.* 1995; Xu *et al.* 1996; Graham *et al.* 1996; Gatehouse *et al.* 1997; Ussuf *et al.* 2001
Potato proteinase inhibitor I	Tomato	*Teleogryllus commodus*	Burgess *et al*. 1994
Potato proteinase inhibitor II	Tobacco, Rice	Lepidoptera, *Sesamia inferens* and *Chilo suppressalis*	Johnson 1989; McManus 1994; Jongsma 1995; Duan *et al.* 1996
Potato proteinase inhibitor I and II	Lettuce	*T. commodus*	Gatehoue and Gatehouse 1998
Snowdrop lectin	Tobacco, Potato	*Myzus persicae, Aulacorthum solani*	Hilder 1995; Down *et al*. 1996
α-amylase inhibitor	Tobacco, Pea, Azuki bean	*Agrotis ipsilon, Zabrotes subfaciatus, Callosobruchus chinensis*	Carbonero 1993; Shade *et al.* 1994; Ishimoto *et al.* 1996
Beam chitinase	Potato	*A. solani*	Down *et al.* 1996
Snowdrop lectin+beam chitinase	Potato	*M. persicae* and *A. solani*	Gatehouse *et al.* 1996; Down *et al*. 1996
Snowdrop lectin and cowpea trypsin inhibitor	Potato	*L. oleracea*	Birch *et al*. 1999
Rice cysteine inhibitor	Oilseed rape, Poplar	Coleoptera, *Chrysomela tremulae*	Leple *et al.* 1995; Burgess *et al.* 1997
Double headed serine protease inhibitor from soyabean	Oilseed rape, Poplar	Lepidoptera and Diptera	Leple *et al.* 1995; Burgess *et al.* 1997
Sweet potato TI	Cauliflower	*Pieris conidia*	Ding *et al.* 1998
Barley TI	Wheat	*Sitotroga cerealella*	Altpeter *et al.* 1999

*Sourec: Gatehouse & Gatehouse 1998

Inhibitors of α-amylases

Amylases catalyse the breakdown of starch into sugars. α-amylases (a-1,4-glucan-4-glucanohydrolase; EC 3.2.1.1) are endo-amylases that constitutes a family of hydrolases and cleaves a-D-(1,4)-glucan linkages in starch components, glycogen and various other related carbohydrates in starch in a random manner. Thermostability and the pH optima for insect α-amylases vary from acidic to alkaline pH. The sequences of several insect α-amylases are known, but only the three-dimensional structure of the α-amylase of *Tenebrio molitor* (TMA) (Periera *et al.* 1999) the α-amylase of the yellow meal worm (*Tenebrio molitor*), is the most extensively studied enzymes from insect origin. Its catalytic properties and inhibition by several inhibitors *in vitro* and *in vivo* have been investigated (Applebaum *et al*. 1964).

Plant seeds are rich sources of a large number of different proteinaceous inhibitors which act on key insect gut digestive enzymes. Franco *et al.* (2002) have reviewed plant inhibitors active towards insect α-amylases and their use in transgenic protection of plants against insect predators. Some α-amylase inhibitors show strict target enzyme specificity and recognise only one out of several closely related isozymes or enzymes from different species. Other inhibitors have high affinity for both mammalian and insect α-amylases and share around 35% sequence identity. Several examples are found of inhibitors acting on insect, but not on mammalian enzymes and vice versa (Juge & Svensson 2006). Proteinaceous inhibitors that act on α-amylases and proteinases have been identified and characterized from the seeds belonging to several plant families. α-amylase inhibitors are of two types –proteinaceous and non proteinaceous.

Non proteinaceous inhibitors

This class of inhibitors contains diverse types of organic compounds such as acarbose, isocarbose, acarviosine glucose, hibiscus acid and the cyclodextrins. The two hibiscus acid forms, purified from Rosselle tea (*Hibiscus sabdariffa*), the acarviosine glucose, the isocarbose and α-, β- and γ-cyclodextrins are highly active against porcine

and human pancreatic α-amylase (PPA and HPA) (Payan 2004). Inhibitory activity of such compound is due to cyclic structure, which resembles the α-amylase substrate. Nevertheless, the use of nonproteinaceous inhibitors for the production of insect resistant transgenic plants is difficult as these are the producys of pathways involving more than one gene.

Proteinaceous inhibitors

Proteinaceous α-amylase inhibitors are classified into seven types (Svensson *et al*. 2004) based on similarities in sequence and three-dimensional structures (**Table 7**) (Garcia-Olmedo *et al.* 1987; Ho *et al.* 1993; Lu *et al.* 1999; Franco *et al.* 2002). Six of them are from higher plants (Marshall and Lauda 1975; Blanco-Labra and Iturbe-Chinas 1980; Campos and Richardson 1983; Mundy and Rogers 1986; Bloch Jr & Richardson 1991; Garcia-Maroto *et al.* 1991; Ohtsubo and Richardson 1992; Ho and Whitaker 1993; Chagolla-Lopez *et al.* 1994; Barber *et al.* 1986; Fakhoury and Woloshuk 2001; Franco *et al.* 2002) whereas the seventh type occurs in *Streptomyces* species (Murao *et al.* 1983)

Table 7. Classes of α-amylase inhibitors based on tertiary structure

Structural class	Source
Legume lectin type	Common bean (*Proteus vulgaris*)
Knottin type	Amaranth (*Amaranthus hypocondriacus*)
Cereal type	Wheat (*Triticum aestivum*), Barely (*Hordeum vulgareum*)
Kunitz	Barely (*hordeum vulgarium*)
Thaumotin type	Maize (*zye mays*)
Gamma-Purothionin type	Sorghum (*sorghum bicolor*)
Microbial	*Streptomyces species*

Streptomyces inhibitors: Tendamistat (Vertesy *et al.* 1984), Haim (Murao *et al.* 1980) and Paim (Murao *et al.* 1983) and related proteins constitute a family of small proteins of approximately 75 amino acid residues which have been purified from different Streptomyces species. These inhibitors possess about 30% sequence identity and show

conserved disulfide topology (Hoffman *et al.* 1985).

AAI from the Mexican crop plant amaranth (*Amaranthus hypochondriacus*): AAI is the smallest known natural proteinaceous α-amylase inhibitor (Chagolla-Lopez *et al.* 1994). Protein knots are extremely rich in disulfides and occur both as individual miniproteins of around 32 amino acid residues and as domains in larger molecules. In a recent dendrogram, knottins group in eight clusters according to the disulfide bond topology (Carugo *et al.* 2001). The binding of AAI to TMA presents a new inhibition mode for a-amylases. Due to its unique specificity towards insect a-amylases, AAI might represent a valuable tool for protecting crop plants from predatory insects (Pereira *et al.* 1999).

c-Thionins*:* SIa1, SIa2, and SIa3 from sorghum (*Sorghum bicolor*) are small proteins of 47–48 amino acid residues that strongly inhibit insect a-amylases (Bloch & Richardson 1991). The three isoforms contain four disulfide bridges and have 42–87% sequence identity and helps in plant defence (Franco *et al.* 2002).

CM-proteins*:* A large protein family from cereal seeds containing 120-160 amino acid residues and five disulfide bonds includes inhibitors known for their action on a-amylases from mammals and insects (Barber *et al*. 1986). The name CM-protein refers to the appearance in chloroform/methanol extracts of flour. These inhibitors are also referred to as ‘‘cereal-type’’ (Franco *et al.* 2002). Many different CM proteins are found to be abundant in aqueous extracts of developing and mature barley seeds (Finnie *et al.* 2002) and several are double-headed α-amylase/trypsin inhibitors whereas others inhibit only a-amylase or trypsin (Alagiri & Singh 1993).

Kunitz-type*:* The capacity to inhibit subtilisin was earlier discovered for a barley (Yoshikawa *et al*. 1976) protein of sequence similarity to Kunitz soybean trypsin inhibitor. This barley protein was later identified in complex with an endogenous a-amylase andhence named a-amylase/subtilisin inhibitor (BASI) (Weselake *et al*. 1983). Homologous proteins with 92% and 58% sequence identity are present in wheat (WASI) and rice (RASI) respectively (Mundy *et al.* 1984).

Thaumatin-like inhibitors*:* Zeamatin from maize inhibits insect, but not mammalian α-amylases (Schmioler-O'Rourke *et al*. 2001). It is a homologue of the sweet protein thaumatin. However thaumatin and other related proteins do not inhibit α-amylases. Proteins of this family are grouped in pathogen-related group 5 (PR-5) and alter the properties of fungal cell walls. Zeamatin thus binds to β-1, 3-glucan and permeabilizes fungal cells resulting in cell death and has been applied in antifungal drugs (Roberts & Selitrennikoff 1990).

Lectin-like inhibitors: αAI1 and αAI2 are α-amylase inhibitors from common white, red and black kidney beans (Ho *et al*. 1994). Three inhibitors are encoded by two different alleles, and only αAI1 inhibits both mammalian and insect α-amylases whereas αAI2 inhibits different insect α-amylases (Svensson *et al.* 2004).

Table 8. Different structural classes of α-amylase inhibitors, based on a classification by Franco *et al.* (2002)*

Inhibitor	**Source**	**Inhibitor activity**		**References**
		Mammalian	**Insect**	
αA11	*Phaseolus vulgaris*	Porcine pancreatic α-amylase	*Callosobruchus maculates* (coleoptetra) *Diabrotica virgifera virgifera* (coleoptetra) *Hypothenemua hampei* (coleoptetra) *Tenebrio molitor* (coleoptetra)	Ishimoto & Kitamura (1989); Titarenko & Chrispeels (2000)
α- A12	*Phaseolus vulgaris*	No activity	*Zabrotes subfasciatus* (coleoptetra)	Ishimoto & Chrispeels (1996)
Wheat Extract	*Triticum aestivum*	Porcine pancreatic α-amylase and human salivery α-amylase	*Diabrotica virgifera virgifera* (coleoptetra) *Lygushesperus* (coleoptetra)*Lyguslieoralis* (coleoptetra)	Feng *et al*. (1996)
0.19	*Triticum aestivum*	Porcine pancreatic α-amylase and human salivery α-	*Diabrotica virgifera virgifera* (coleoptetra) *Callosobruchus maculates* (coleoptetra) *Zabrotes subfasciatus* (coleoptetra)	Titarenko & Chrispeels (2000); Franco *et al*. (2000)

		amylase	*Acanthoscelides obstectus* (coleoptetra) *Sitophilus oryzae* (coleoptetra) *Tribolium castaneum* (coleoptetra)	
0.53	*Triticum aestivum*	human salivery α-amylase and pancreatic α-amylase	*Tenebrio molitor* (coleoptetra) *Callosobruchus maculates* (coleoptetra) *Zabrotes subfasciatus* (coleoptetra) *Acanthoscelides obstectus* (coleoptetra)	Franco *et al*. (2000)
0.28	*Triticum aestivum*	Porcine pancreatic α-amylase and human salivery α-amylase	*Tenebrio molitor* (coleoptetra)	Sanchez-Monge *et al.* (1989)
WRP25	*Triticum aestivum*	None	*Sitophilus oryzae* (coleoptetra) *Tribolium castaneum* (coleoptetra) *Tenebrio molitor* (coleoptetra) *Callosobruchus maculates* (coleoptetra)	Feng *et al.* (1996); Franco *et al.* (2000)
WRP26	*Triticum aestivum*	None	*Tenebrio molitor* (coleoptetra) *Sitophilus oryzae* (coleoptetra) *Tribolium castaneum* (coleoptetra) *Callosobruchus maculates* (coleoptetra)	Feng *et al*. (1996); Franco *et al*. (2000)
WRP27	*Triticum aestivum*	None	*Tenebrio molitor* (coleoptetra) (low) *Sitophilus oryzae* (coleoptetra)	Feng *et al*. (1996)
1, 2 and 3	*Amaranthus hypochondriacus*	None	*Tenebrio molitor* (coleoptetra) *Hypothenemus hampei* (coleoptetra) *Prosephanus truncates* (coleoptetra)	Chagolla-Lopez *et al.* (1994)
AAI	*Vigna unguiculata*	None	*Callosobruchus maculates* (low) (coleoptetra)	Melo *et al.* (1996)
CAI	*Cjanus cajan*	human salivery α-amylase and Porcine pancreatic α-amylase	*Helicoverpa armigera* (Lepidopteran) (low)	Giri & Kachole (1998)
PAI	*Zea mays*	No activity	*Tribolium castaneum* (coleoptetra) *Sitophilus zeamais* (coleoptetra) *Rhyzoperta dominica* (coleoptetra)	Ishimoto & Kitamura (1989)

***** Source: Franco *et al.* 2002

Wheat α-amylase inhibitors **(Cereal type inhibitors)**

Wheat kernels contain a number of albumin components that actively inhibit many α-amylases from sources other than wheat but that are inactive with α-amylases in wheat (Buonocore *et al.* 1976). They have extensively characterized several albumin α-amylase inhibitors and thoroughly investigated their interactions with the inhibited amylases. In wheat a substantial fraction of the total endosperm protein content is represented by inhibitors. More than 20 different members of a single multiple family of α-amylase inhibitors, specifically expressed in endosperm, have been characterized. α-amylase inhibitors from wheat are classified according to their degree of aggregation into monomeric, dimeric and tetrameric forms. These inhibitors have exhibit monomeric molecular masses of 5 kDa, 9kDa and 13 kDa, homodimeric and heterodimeric masses of approximately 26 kDa and tetrameric masses of 50 kDa (Wang *et al.* 2006).

In some cases, the α-amylase inhibitors act only against mammalian α-amylases or, on the contrary, just against insect α-amylase. In the latter case, this provides a highly specific potential weapon in plant defence. The known α-amylase inhibitors selective for insect enzymes and inactive against mammalian enzymes include WRP25, WRP26 and WRP27 from the cereal-type class (Feng *et al.* 1996; Franco *et al.* 2000)

The biological role in the seed of such a large number of closely related protein molecules and the possible significance of their presence in foods in an active form has been extensively discussed. The most thoroughly studied albumin inhibitors from wheat kernel have been coded according to their gel electrophoretic mobility with reference to bromphenol blue, as 0.19 and 0.28. The 0.28 protein is a monomer with molecular weight 12 000 that lacks phenylalanine and histidine. The molecular weight of 0.19 is about 24 000 in non dissociating solvents, but drops to about 12 500 in 6 M guanidine hydrochloride or in 1% sodium dodecyl sulphate. Gel electrophoresis in dissociating solvents has revealed a slight difference between the molecular weights of the two 0.19 subunits 3. When the dissociating solvent is removed, the inhibitor reassociates and regains its amylase inhibitory activity (Buonocore *et al.* 1980).

The relative inefficacy of alpha-amylase inhibitors in affecting human digestion of starch has been highlighted by recent scientific and public controversy over the

commercial sales of so-called starch-blockers or slimming pills (Bloch & Richardson 1991). Alpha-amylase and its inhibitors is drug-design targets for the development of compounds for treatment of diabetes Obesity and Hyperlipaemia (Octivio & Rigden 2000). Studies of the structures of the numerous enzyme inhibitors found in cereal grains have led to the recognition of a super family of homologous proteins which includes inhibitors of alpha-amylase, proteinase and bifunctional inhibitors active against two or more classes of enzymes (Richardson 1991; Alam *et al.* 2001; Octavio & Rigden 2002). The first alpha-amylase inhibitor determined was that of the monomeric 13 kD known as 0.31 form, from wheat (Kashlan & Richardson 1981). Other dimeric 0.19, 0.23, 0.28, 0.53 forms of wheat inhibitors of exogenous alpha-amylase was later shown (Richardson 1991; Kondo & Ida 1995; Roy & Gupta 2000; Octavio & Rigden 2002; Oneda *et al.* 2004).

The favoured hypotheses about physiological roles of the enzyme inhibitors in seeds is that they act as storage or reserve proteins, as regulators of endogenous enzyme or as defensive agents against the attacks of animal predators and insect or microbial pests. It seems likely that in certain species these proteins may fulfill a combination of these functions (Richardson 1991; Octivio & Rigden 2000; Octavio & Rigden 2002) Also plant alpha-amylase inhibitors show great potential as tools to engineer resistance of crop plant against pests (Octavio & Rigden 2002). Nutritional and metabolic effects of enzyme inhibitors certainly some of inhibitors found in cereal and legume seeds can inactivate the salivary and pancreatic enzymes of humans (*pick* and *wober*) and their susceptibility to inactivation in the stomach appears to be rather variable (Singh & Blundel 2001). Many are destroyed by cooking but some retain inhibitory activity even after baking (Richardson 1991). The amylase inhibitors present in seeds currently used as food present few nutritional problems for healthy people but may have some toxicological significance in the diets of infants who have a lower production of pancreatic alpha-amylase than adults and for patients with impaired peptic or gastric function (Richardson 1991; Shewry *et al.* 2001; Brieteneder & Radauer 2004). Also one inhibitor of insect alpha-amylase isolated from barley flour is the major allergen associated with baker's asthma disease (Barber *et al.* 1989).

The inhibition is strictly competitive and in the 1:1 complexes all of the activities of the enzyme are completely abolished. Crystallographic, nuclear magnetic resonance (NMR) and mechanistic studies all indicate that the inhibitors act as highly specific substrates for the enzyme. They inhibit at a unique peptide bond called the reactive site peptide bond. The reaction mechanisms involved in the inhibition of alpha-amylase by plant protein inhibitors are not clearly understood (Silano 1987). But there are suggestion that reducing sugars which are covalently bound to the inhibitor polypeptide chain may play a major role in the mechanism or that the inhibitor may induce conformational changes in the enzyme molecule. Richardson (1991) demonstrated that when the barley bifunctional inhibitor binds to the endogenous α-amylase it affects a specific tryptophan residue of the enzyme which is essential for productive enzyme-substrate binding In a study albumin like proteins were extracted from wheat (*Triticum aestivum* var. *zarrin*) seeds and amylase inhibitor precipitated with ammonium sulfate before use in *AKTA FPLC* system followed by purification by exchange chromatography applied on AKTa FPLC (Heidari *et al.* 2005).

The interaction between the target digestive enzyme in the pest and its inhibitor

Plant derived inhibitors to the insect digestive enzymes are of particular interest because these inhibitors have always been a part of the plants' natural defense against insect predation (Ryan 1973, 1990; Franco *et al*. 2002). Transgenic plants that express such inhibition to digestive enzymes of insects such as proteinases and alpha-amylases, have been shown to be effective in controlling Lepidopteran, Coleopteran and other insect pests that feed by chewing or biting (Johnson *et al*. 1989; James *et al*. 1992; McManus *et al.* 1994, 1995; Thomas *et al.* 1994, 1995; Newell *et al.* 1995; Urwin *et al.* 1995, 1997, 2001; Duan *et al*. 1996; Irie *et al.* 1996; Xu *et al*. 1996; Gatehouse *et al.* 1997; Grossi de Sa & Chrispeels 1997; Heath *et al.* 1997; Sane *et al.* 1997; Yeh *et al.* 1997; Chrispeels *et al.* 1998; Vain *et al.* 1998; Alpteter *et al.* 1999; Charity *et al.* 1999; Lee *et al.* 1999; Mochizuki *et al*. 1999; Marchetti *et al.* 2000; Morton *et al.* 2000; Bell *et al.* 2001; De Leo *et al.* 2001; De Leo & Gallerani 2002; Fan & Guo 2002; Xie *et al.* 2003).

Alpha-amylase inhibitors from bean, peas and wheat have been shown to inhibit the amylases of certain weevils (Buonocore *et al*. 1976; Ishimoto & Kitamura 1989; Chen *et al.* 1992; Shade *et al.* 1994; Schroeder *et al.* 1995; Ishimoto *et al.* 1996; Grossi de Sa *et al*. 1997; Grossi de Sa & Chrispeels 1997; Le Berre-Anton *et al.* 1997; Chrispeels *et al.* 1998; Pereira *et al.* 1999; Franco *et al.* 2000; Morton *et al.* 2000; Titarenko & Chrispeels 2000; Amirhusin *et al.* 2004; Kluh *et al.* 2005). Insect pests adapt to host plant inhibitors by synthesizing digestive enzymes that are insensitive to these inhibitors (Bolter & Jongsma 1995; Broadway 1995, 1996, 1997; Jongsma *et al*. 1995, 1996; Michaud 1996; Giri *et al.* 1998; Girard *et al.* 1998). Hence it is appropriate to identify potential inhibitors from non-host plants.

Alpha-amylase inhibitors from different sources show considerable specificity towards their target enzymes such that a protein may inhibit the activity of one amylase and may not have any effect on a different amylase (da Silva *et al.* 2000; Franco *et al.* 2002). For example, the well characterised inhibitor AI-1 inhibits porcine pancreatic amylase (PPA), human salivary amylase (HAS) and the alpha-amylases of several bruchids, but not the alpha-amylase of the Mexican bean weevil *Zabrotes subfasciatus* (ZSA) (Powers & Whitaker 1977; Powers & Culbertson 1983; Ishimoto & Kitamura 1989). Similarily, the alpha-amylase inhibitor isolated from amaranthus seeds, AAI, strongly inhibits the larvae of red flour beetle (*Tribolium castaneum*), the yellow meal worm (*Tenebrio molitor*) and the grain borer (*Prostephanus truncatus*), but it does not inhibit mammalian alpha-amylases or proteases (Chagolla-Lopez *et al.* 1994).

An improved understanding of the specificity of the interaction between the target alpha-amylase and the potential inhibitor will help to design inhibitors with more desirable characteristics. An ideal inhibitor is one which specifically inhibits the insect amylase and has no action on mammalian amylases. Such an inhibitor would provide a highly specific potential weapon in plant defense. An attractive alternative is to develop a rational redesign of known inhibitors in order to confer the required specificity towards the target amylase. This approach requires a full understanding of the structural basis of amylase-inhibitor interaction. The specificity of interaction between the target enzyme and inhibitor can be addressed by different methods such as biochemical and biophysical methods and also by sequence analysis and modeling using available crystal structures.

REFERENCES

Abbot P (2001) *Journal of Insect Science*, **1(8)**, 1-3.

Abdallah Z, Mezghani-Kkhemakhem M, Bouktila D, Makni H and Makni M (2012) *African Journal of Agricultural Research*, **7(7)**, 1170-1176.

Albalat R and Gonzalez-Duarte R (1993) *Gene*, **126(2)**, 171-178.

Al-Barrak MS (2001) PhD Thesis, Cardiff University, UK.

Allender CJ, Clancy K M, Degomez TE., Mcmillin JD, Woolbright SA, Keim P and Wagner DM (2008) *Environtal Entomology,* **37(3)**, 817-824.

Alpteter F, Diaz I, McAuslane H, Gaddour K, Carbonero P and Vasil IK (1999) *Molecular Breeding*, **5**, 53-63.

Aikhionbare FO, Pruess KP and Mayo ZB (1998) *Genetic Analysis: Biomolecular Engineering*, **14**, 105–108.

Al-Otaibi SA (2008) *Arabian Journal of Biotechnology*, **11(2)**, 241-252.

Altpeter F, Diaz I, McAuslane H, Gaddour K, Carbonero P and Vasil IK (1999) *Molecular Breeding*, **5(1)**, 53-63.

Amirhusin B, Shade RE, Koiwa H, Hasegawa PM, Bressan RA, Murdock LL and Zhu-Salzman K (2004) *Journal of Economic Entomology*, **97**, 2095–2100.

Andreev D, Breilid H, Kirkendall L, Brun LO and ffrench-Constant RH (1998) *Insect Molecular Biology*, **7(2)**, 197–200.

Armbruster GFJ (2001) *Cannadian Journal of Zoology*, **79**, 334–345.

Asokana R , Nageshaa SN, Manamohana M, Krishnakumarb NK, Mahadevaswamya HM, Rebijitha KB, Prakasha MN and Chandraa GS (2011) *Oriental Insects*, **46(2)**, 130–143.

Austin JW, Szalanski AL and Cabrera BJ (2004) *Annuals of Entomological Society of America*, **97**, 548–555.

Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, Reeb CA and Saunders NC (1987) *Annual Review of Ecological Systematics*, **18**, 489-522.

Avise JC (1994) *Chapman & Hall. New York, New York.*

Avise JC, Walker D and Johns GC (1998) *Proceedings of the Royal Society of London. Series B: Biological Sciences,* **265(1407)**, 1707-1712.

Avise JC and Johns GC (1999) *Proceedings of the National Academy of Sciences*, **96(13)**, 7358-7363.

Avise JC (2000) *Phylogeography: The History and Formation of Species*. Harvard University Press, Cambridge, MA.

Avise JC (2004) *Molecular markers, natural history, and evolution*, 2nd edition. *Sinauer, Sunderland, USA.*

Ayres CFJ, Melo–Santos MAV, Solé–Cava AM and Furtado AF (2003) *Journal of Medical Entomology*, **40**, 430–5.

Bachtrog D, Agis M, Imhof M and Schlotterer C (2000) *Molecular Biology and Evolution*, **17**, 1277–1285.

Bailly X, Migeon A and Navajas M (2004) *Biological Journal of the Linnean Society*, **82**, 69–78.

Bajpai N and Tewari RR (2010) *Journal of Genetics*, **89**, 51–54.

Baker RH and DeSalle R (1997) *Systematic Biology*, **46(4)**, 654-673.

Baker RH, Wilkinson GS and DeSalle R (2001) *Systematic Biology*, **50**, 87–105.

Ballard JWO, Olsen GJ, Faith DP, Odgers WA, Rowell DM and Atkinson PW (1992) *Science* **258(5086)**, 1345–48.

Ballard JWO (1994) *Journal of Australian Entomological Society*, **33(2)**, 131–35.

Barber D, Sanchez-Monge R, Mendez E, Lazaro A, Garcia-Olmedo F and Salcedo G (1986) *Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology*, **869(1)**, 115-118.

Bartlett AC, Wuson N and Mattix EB (1968) *Journal of Economic Entomology*, 61(3), 808-812.

Bartlett AC and Butler GD (1975) *Journal of Economic Entomology*, **68(3)**, 331-335.

Bartlett AC (1981) *Annals of the Entomological Society of America*, **74(1)**, 9-13.

Baruffi L, Damiani G, Guglielmino CR, Bandi C, Malacrida AR and Gasperi G (1995) *Heredity*, **74 (1995)**, 425—437.

Bas B, Dalkilic Z, Peever TL, Nigg HN, Simpson SE, Gmitter FG and Adair RC (2000) *Annals of the Entomological Society of America*, **93**, 459–467.

Baskaran M (2011a) *Biotechnology Bioinformatics and Bioengetics*, **1(2)**, 221-227.

Baskaran M (2011b) *Advanced Bio Technology*, **11(5)**, 12-15.

Baumann R, Schubert R, Heitland W, Auger-Rozenberg M, Faivre-Rimpant P and Muller-Starck G (2003) *Journal of Applied Entomology*, **127**, 258–264.

Baur A, Sanetra M, Chalwatzis N, Buschinger A and Zimmermann FK (1996) *Insectes Society*, **43**, 53–67.

Beard CB, Hamm DM and Collins FH (1993) *Insect Molecular Biology*, **2**, 103–24.

Beckenbach AT, Wei YW and Liu H (1993) *Molecular Biology and Evolution,* **10(3)**, 619–34.

Beebe NW, Ellis JT, Cooper RD and Saul A (1999) *Insect Molecular Biology*, **8**, 381–390.

Bell HA, Fitches EC, Down RE, Ford L, Marris GC, Edwards JP, Gatehouse JA and Gatehouse AMR (2001) *Pest Management Science,* **57**, 57–65.

Belshaw R and Quicke DL (1997) *Molecular Phylogenetics and Evolution*, **7(3)**, 281-293.

Belshaw R, Dowton M, Quicke DLJ and Austin AD (2000) *Proceedings of the Royal Society of London Series Biological Sciences,* **267**, 491 – 496.

Beltran M, Jiggins CD, Bull V, Linares M, Mallet J, McMillan WO and Bermingham E (2002) *Molecular Biology and Evolution*, **19**, 2176–2190.

Bensch S and Akesson M (2005) *Molecular Ecology*, **14**, 2899–2914.

Bernasconi MV, Valsangiacomo C, Piffaretti JC and Ward PI (2000) *Insect Molecular Biology*, **9**, 67–74.

Bhau BS and Koul V (1998) *Current Science*, **75(8)**, 771-777.

Birch ANE, Geoghegan IE, Majerus MEN, McNicol JW, Hackett CA, Gatehouse AMR and Gatehouse JA (1999) *Molecular Breeding*, **5**, 75–83.

Birkle LM and Douglas AE (1999) Heredity, 82, 605–612.

Black WC IV, DuTeau NM, Puterka GJ, Nechols JR and Pettorini JM (1992) *Bulletien of Entomological Research*, **82**, 151–159.

Blanco-Labra A and Iturbe-Chinas FA (1981) *Journal of Food Biochemistry,* **5(1)**, 1-17.

Blears MJ, De Grandis SA, Lee H and Trevors JT (1998) *Journal of Indian Microbiology and Biotechnology*, **21**, 99–114.

Bloch Jr C and Richardson M (1991) *FEBS letters*, **279(1)**, 101-104.

Bologna MA, Oliverio M, Pitzalis M and Mariottini P (2008) *Molecular Phylogenetics and Evolution*, **48**, 679–693.

Bolter CJ and Jongsma MA (1995) *Journal of Insect Physiology*, **41**, 1071–1078.

Bond HA and Fay RW (1970) *Mosquito News*, **30(3)**, 394-402.

Bonebrake T C, Watt WB, Perez A and Boggs CL (2011) *European Journal of Entomology*, **108**, 529–535.

Boulter D, Edwards GA, Gatehouse AMR, Gatehouse JA and Hilder VA (1990) *Crop Protection*, **9**, 351-354.

Boulter D (1993) *Phytochemistry*, 34, 1453-1466.

Boyce TM, Zwick ME and Aquadro CF (1989) *Genetics*, **123**, 825-836.

Broadway RM and Duffey SS (1986a) *Journal of Insect Physiology*, **32**, 673–680.

Broadway RM and Duffey SS (1986b) *Journal of Insect Physiology*, **32**, 827-833.

Broadway RM (1995) *Journal of. Insect Physiology*, **41**, 107–116.

Broadway RM. (1996) *Insect Biochemistry and Physiology*, **32**, 39–53.

Broadway RM (1997) *Journal of Insect Physiology*, **43**, 855–874.

Brookes G and Barfoot P (2010) PG Economics Ltd UK.http://fastfacts.intraspin.com/?fact=over-the-thirteen-years-1996-2008-the cumulative-farm-income-gain-derived-by-developingcountry-frrom-gm-crops-was-26-2-billion.

Brower AVZ (1994) *Procedings of National Acadamy of Science* USA, **91**, 6491-6495.

Brower AVZ and DeSalle R (1994) *Annual Entomological Society of America*, **87**, 702-716.

Brower AVZ and Egan MG (1997) *Proceedings of Royal Society of London Series B* **264**, 969–77.

Brower AVZ and DeSalle R. (1998) *Insect Molecular Biology* **7**, 73–82.

Brown JM, Pellmyr O, Thompson JN and Harrison RG. (1994) *Molecular Biology and Evolution*, **11**, 128–41.

Brown WM (1985) In: *Molecular Evolutionary Genetics* (ed. MacIntyre R.), 95-130. Plenum, New York.

Buntjer JB (1997) Academic Thesis, Utrecht University, Utrecht, The Netherlands.

Buonocore V, Caporale C, DeRosa M and Gambacorta A (1976) *Journal of Bacteriology*, **128**, 515–521.

Buonocore V, Gramenzi F, Pace W, Petrucci T, Poerio E and Silano V (1980) *Biocheistry Journal*, **187**, 637-645.

Burke T (1998) *Molecular Ecolology*, **7**, 367-545.

Burgess EPJ, Steven PS, Keen GK, Laing WA and Christeller JT (1991) *Entomologia Experimentalis et Applicata*, **61**, 123-130.

Burgess EPJ, Main CA, Stevens PS, Christeller JT, Gatehouse AMR and Laing WA (1994) *Journal of Insect Physiology*, **40(9)**, 803–811.

Burgess EPJ, Gatehouse AMR, McKersie BD and Brown DCW (1997) *Biotechnology and the improvement of forage legumes*, 229-258.

Callejas CA, Velasco A, Gobbi F, Beitia and Ochando MD (2005) *Journal of Applied Entomology*, **129**, 382–385.

Cameron SA (1993) *Proceedings of. National Academy of Science USA*, **90**, 8687–91.

Campbell BC, Steffen- Campbell JD, Sorensen JT and Gill RJ (1995) *Systematic Entomology*, **20(3)**, 175-194.

Campbell DL, Brower AVZ and Pierce NE (2000) *Molecular Biology and Evoution*, **17**, 684–696.

Cameron SL, Lambkin CL, Barker SC and Whiting MF (2007) *Systematic Entomology*, **32**, 40–59.

Cameron SL and Whiting MF (2008) *Gene*, **408**, 112–123.

Campos FAP and Richardson M (1983) *FEBS Letters*, **152(2)**, 300-304.

Carbonero P, Royo J, Diaz I, Garcíamaroto F, Gonzalez-Hidalgo E, Gutierrez C and Casanera P (1993) In: *Workshop on engineering plants against pests and pathogens* (1st–13rd January, 1993, Madrid, Spain). Instituto Juan March de Estudios Investigaciones, Bruening GJ, Garcíaolmedo F and Ponz FJ eds, 1993.

Carisio L, Cervella P, Palestrini C, Delpero M and Rolando A (2004) *Journal of Biogeography*, **31**, 1149-1162.

Carletto J, Lombaert E, Chavigny P, Brevault T, Lapchin L andVanlerberghe-Masutti F (2009) *Molecular Ecology*, 18, 2198–2212.

Carmean D, Kimsey LS and Berbeet ML (1992) *Molecular Phylogenetics and Evolution*, **1**, 270–278.

Carpenter JM, Strassmann JE, Turillazzi S, Hughes CR, Solis CR and Cervo R (1993) *Cladistics*, **9(2)**, 129-146.

Castro LR, Ruberu K and Dowton M (2006) *Genome,* **49**, 752–766.

Caterino MS, Cho S and Sperling FAH (2000) *Annual Review of Entomology*, **45**, 1–54.

Cervera MT, Cabezas JA, Simon B, Martínez-Zapater JM, Beitia F and Cenis JL (2000) *Bulletin of Entomological Research*, **90**, 391–396.

Chagolla-Lopez A, Blanco-Labra A, Patthy A, Sanchez R and Pongor S (1994) *Journal of Biological Chemistry*, **269(38)**, 23675–23680.

Chalwaltzis N, Hauf J, van der Peer Y, Kinzelbach R and Zimmermann FK (1996) *Annals of Entomological Society of America*, **89**, 788-803.

Chandra A, Reinert JA, LaMantia J, Pond JB and Huff DR (2011) *Journal of Insect Science,* **11**, 173.

Charity JAA, Merilyn A, Bittisnich DJ, Whitecross M and Higgins TJV (1999) *Molecular Breeding*, **5(4)**, 357-365.

Chatterjee SN and Pradeep AR (2003) *Russian Journal of Genetics*, **39**, 1365-1377.

Chatterjee SN, Mohandas TP and Taraphdar T (2003) *European Journal of Entomology*, **100**, 195-200.

Chatterjee SN, Taraphdar T and Mohandas TP (2005) *Genetica*, **125**, 1–15.

Chen W, Hoy JW and Schneider RW (1992) *Experimental Mycology*, **16**, 22–34.

Chen H, Chen XM, Feng Y and YE SD (2006) *Forest Research*, **19(4)**, 423-430.

Chen N, Liu TX, Setamou M, French JV and Louzada ES. (2009) *Insect* Science, **16**, *115-124.*

Cho S, Mitchell A, Regier JC, Mitter C, Poole RW, Friedlander TP and Zhao S (1995) *Molecular Biology and Evolution*, **12(4)**, 650-656.

Choudhary M, Strassmann JE, Solís CR and Queller DC (1993) *Biochemical Genetics*, **31(1-2)**, 87-96.

Choudhary B and Gaur K (2010) ISAAA Series of Biotech Crop Profiles. ISAAA, Ithaca, NY.

Chrispeels MJ, Grossi-de-Sa MF and Higgins TJV (1998) *Seed Science Research*, **8**, 257–263.

Chu D, Wan F-H, Tao Y-L, Liu G-X, Fan Z-X and Bi Y-P (2008) *Insect Sciences*, **117**, 117–125.

Ciosi M, Miller NJ, Kim KS, Giordano R, Estoup A and Guillemaud T (2008) *Molecular Ecology*, **17**, 3614–3627.

Clark TL, Meinke LJ and Foster JE (2001) *Insect Molecular Biology,* **10**, 303–314.

Clark PL, Molina-Ochoa J, Martinelli S, Skoda SR, Isenhour DJ, Lee DJ, Krumm JT and Foster JE (2007) *Journal of Insect Science,* **7**, 05.

Clary DO and Wolstenholme DR (1985) *Journal of Molecular Evolution,* **22**, 252–271.

Clements KM, Wiegmann BM, Sorenson CE, Smith CF, Neese PA, and Michael RR (2000) *Annuals of Entomological Society of America*, **93(1)**, 31-46.

Coleman AW (2007) *Nucleic Acids Research,* **35(10)**, 3322-3329.

Collins FH and Paskewitz SM (1996) *Insect Molecular Biology,* **5**, 1–9.

Costa R, Peixoto AA, Thackeray JR, Dalgleish R and Kyriacou CP (1991) *Journal of Molcular Evolution,* **32(3)**, 238–46**.**

Crozier RH, Crozier YC and Mackinlay AG (1989) *Molecular Biology and Evolution,* **6**, 399-4 11.

Crozier RH and YC Crozier (1993) *Genetics,* **133**, 97-117.

Crozier RH, Kaufmann B, Carew ME and Crozier YC (1999) *Molecular Ecology*, ***8*(2)**, 271-276.

Cruickshank RH (2002) *Systematic and Applied Acarology,* **7**, 3–14.

da Costa-Ribeiro MCV, Lourenço-de-Oliveira R and Failloux AB (2007) *The American Journal of Ttropical Medicine and Hygiene*, ***77*(2)**, 303-309.

Dalirsefat S, Meyer A and Mirhoseini S (2009) *Journal of Insect Science,* **9**, 71.

Danforth BN and Ji S (1998) *Molecular Biology and Evolution,* **15**, 225–35.

Danforth BN, Sauquet H and Packer L (1999) *Molecular Phylogeny and Evolution,* **13**, 605–618.

Darvill B, Ellis JS, Lye GC and Goulson D (2006) *Molecular Ecology*, **15**, 601–611.

Da Silva MCM, Grossi-de-Sa MF, Chrispeels MJ, Togawa RC and Neshich G (2000) *Protein Engineering,* **13**, 167–177.

Davies N and Bermingham E (2002) *Evolution*, **56**, 573–589.

De Graaf J (2006) Doctoral dissertation, University of Pretoria.

De Leo F, Bonade-Bottino M, Ceci LR, Gallerani R and Jouanin L. (2001) *Insect Biochemistry and Molecular Biology*, **31(627)**, 593-602.

De Leo F and Gallerani R (2002) *Insect Biochemistry and Molecular Biology*, **32(5)**, 489-496.

DE Leon JH, Jones WA and Morgan DJW (2004) *Annual Entomology Society of America*, **97(4)**, 809-818.

de Maagd RA, Bosch D and Stiekema W (1999) *Trends in Plant Science*, **4(1)**, 9-13.

Demontis MA, Ortu S, Cocco A, Lentini A and Migheli Q (2007) *Journal of Applied Entomology*, **131(1)**, 59–64.

Dempster JP and McLean IFG (Eds.) (1998) In: *19th Symposium of the Royal Entomological Society, 10-11 September 1997 at the University of Newcastle*. Kluwer Academic.

Depaquit J, Ferte H, Leger N, Killick-Kendrick R, Rioux J-A, Killick-Kendrick M, Hanafi HA and Gobert S (2000) *Insect Molecular Biology*, **9**, 293–300.

Depaquit J, Ferte H, Leger N, Lefranc F, Alves-Pires C, Hanafi H, Maroli M, Morillas-Marquez F, Rioux J, Svoboda M and Volf P (2002) *International Journal of Parasitology*, **32**, 1123-1131.

DeSalle R (1992a) *Molecular Biology and Evolution*, **9(5)**, 905–16.

DeSalle R (1992b) *Molecular Phylogenetics and Evolution*, **1**, 31-40.

DeSalle R and Brower AVZ (1997) *Systematic Biology*, **46(4)**, 751–64.

Dezfouli SR N, Oshaghi MA, Vatandoost H, Djavadian, Telmadarei and Assmar M (2002) *Iranian Journal of Public Health*, **31**, 133-137.

Dhuyvetter H, Gaublomme E, Verdyck P and Desender K (2005) *Journal of Heredity*, **96(4)**, 381-387.

Di Muccio T, Marinucci M, Frusteri L, Maroli M, Pesson B and Gramiccia M (2000) *Insect Biochemistry and Molecular Biology*, **30**, 387-393.

Ding X, Gopalakrishnan B, Johnson LB, White FF, Wang X, Morgan TD, Kramer KJ and Muthukrishnan S (1998) *Transgenic Research*, **7**, 77-84.

Dong-Soo Lee P, Sang-Kyu L, Jong-Hee S, Min-Young S, Song-Yi K, Do-Yeon Y, Un-Sang J, Nam-Soo P, Soo-Kwon Y, Gihwan S, You-Chun N, Min-Hee K, Yeon-

Chung J and Jong-Seong (2007) *Theoretical and Applied Genetics,* **115**, 4, 537-547.

Down RE, Gatehouse AMR, Hamilton WDO and Gatehouse JA (1996) *Journal of Insect Physiology*, **42**, 1035–1045.

Down RE, Ford L, Woodhouse SD, Raemaekers RJM, Leitch B, Gatehouse JA and Gatehouse AMR (2000) *Journal of Insect Physiology*, **46**, 379–391.

Dowton M and Austin AD (1998) *Molecular Phylogenetics and Evolution*, **10**, 354 – 366.

Dowton M and Austin AD (1999) *Molecular Biology and Evolution*, **16(2)**, 298–309.

Duan X, Li X, Xue Q, Abo-El-Saad M, Xu D and Wu R (1996) *Nature Biotechnology*, **14(4)**, 494-498.

Dutta SR, Kar PK, Srivastava AK, Sinha MK, Shankar J and Ghosh AK (2012) *Genetics and Molecular Biology*, **35(4)**, 743-751.

Eddy SR (2001) *Nature Review of Genetics*, **2**, 919–929.

Fan SG and Guo-Jiang WU (2005) *Botanical Bulletin of Academia Sinica*, ***46***.

Fakhoury AM and Woloshuk CP (2001) *Molecular Plant-Microbe Interactions*, **14(8)**, 955-961.

Fay RW and Craig GB (1969) *Mosquito News*, ***29*(1)**, 121.

Feng GH, Richardson M, Chen M S, Kramer KJ, Morgan TD and Reeck GR (1996) *Insect Biochemistry and Molecular Biology*, **26(5)**, 419-426.

Fenton B, Birch ANE, Malloch G, Lanham PG and Brennan RM (2000) *Experimental and Applied Acarology*, **24**, 831-861.

Figueroa CC, Simon J-C, Le Gallic J-F and Niemeyer HM (1999) *Entomology Experimental and Applied*, **92**, 217–225.

Finn DS, Blouin MS and Lytle DA (2007) *Freshwater Biology*, **52**, 1881–1897.

Forneck A, Walker MA and Blaich R (2000) *Genome*, **43**, 669–678.

Franco OL, Rigden DJ, Melo FR, Bloch JC, Silva CP and Grossi de Sa MF (2000) *European Journal of Biochemistry*, **267**, 1466–1473.

Franco OL, Rigden DJ, Melo FR and Grossi-de-Sa MF (2002) *European Journal of Biochemistry*, **269**, 397–412.

Frati F and Carapelli A (1999) *Antarctic Science*, ***11*(02)**, 160-174.

Friedlander TP, Regier JC and Mitter C (1992) *Systematic Biology*, **41**, 483–490.

Friedlander TP, Regier JC and Mitter C (1994) *Systematic Biology*, **43**, 511–525.

Friedrich M and Muqim N (2003) *Molecular Phylogenetics and Evolution*, **26**, 502-512.

Fritz GN, Conn J, Cokburn A and Seawright J (1994) *Molecular Biology and Evolution*, **11**, 406–416.

Froi-Ilich DR, Robinson AS and Wells MA (1993) *Insect Molecular Biology*, **1**, 165—169.

Funk DJ, Futuyma DJ, Orti G and Meyer A (1995) *Molecular Biology and Evolution*, **12**, 627–40.

Funk DJ (1999) *Molecular Biology and Evolution*, **16**, 67–82.

Furlong JC and Maden BEH (1983) *European Molecular Biology Organization Journal*, **2**, 443–448.

Garcia BA, Caccone A, Mathiopoulos KD and Powell JR (1996) *Genetics*, **143(3)**, 1313–20.

Garcia-Olmedo F, Salcedo Duran G, Sanchez-Monge Laguna de Rins R, Gomez L, Royo J and Carbonero Zalduegui P (1987), *Oxford Surveys of Plant Molecular and Cell Biology*, 275-334.

Garcia-Maroto F, Carbonero P and GArcia-Olmedo F (1991) *Plant Molecular Biology*, **17(5)**, 1005-1011.

Garnery L, Cornuet J-M and Solignac M (1992) *Molecular Ecology*, **1**, 145–154.

Garner KJ and Slavicek JM (1996) *Insect Molecular Biology*, **5**, 81-91.

Gatehouse AMR, Gatehouse JA, Dobie P, Kilminster AM and Boulter D (1979) *Journal of the Science of Food and Agriculture*, **30**, 948–958.

Gatehouse AMR, Dewey FM, Dove J, Fenton KA and Pusztai A (1984) *Journal of Science and Food Agriculture*, **35**, 373–380.

Gatehouse AMR, Barbieri L, Stirpe F and Croy RRD (1990) *Entomology Experimental and Applied*, **54**, 43–51.

Gatehouse AMR, Howe DS, Flemming JE, Hilder VA and Gatehouse JA (1991) *Journal of the Science of Food and Agriculture*, **55**, 63–74.

Gatehouse AM, Hilder VA, Powell K, Boulter D and Gatehouse JA (1992, January) In *Proceedings of the 8th International Symposium on Insect-Plant Relationships* (pp. 221-234). Springer Netherlands.

Gatehouse AMR, Down RE, Powell KS, Sauvion N, Rahbe Y, Newell CA, Merryweather A, Hamilton WDO and Gatehouse JA (1996) *Entomology Experimental and Applied*, **34**, 295-307.

Gatehouse AMR, Davidson GM, Newell CA, Merryweather A, Hamilton WDO, Burgess EPJ, Gilbert RJC and Gatehouse JA (1997) *Molecular Breeding*, **3**, 49-63.

Gatehouse AMR and Gatehouse JA (1998) *Pest Science*, **52**, 165–175.

Gauthier N and Rasplus JY (2004) *Molecular Ecology Notes,* **4**, 294-296.

Gaviria DA, Aguilar E, Serrano HJ and Alegria AH (2006) *Journal of Insect Science,* **6**, 15.

Ghikas DV, Kouvelis VN and Typas MA (2010) *BioMedCentral Microbiology*, **10(174)**, 1471-2180.

Giri AP, Harsulkar AM, Deshpande VV, Sainani MN, Gupta VS and Ranjekar PK (1998) *Plant Physiology*, **116**, 393–401.

Girard C, Le-Metayer M, Zaccomer B, Bartlet E, Williams I, Bonade-Bottino M, Pham-Delegue MH and Ouanin L (1998) *Journal of Insect Physiology*, **44**, 263- 270.

Giribet G, Rambla M, Carranza S, Riutort M, Baguna J and Ribera C (1999) *Molecular Phylogenetics and Evolution*, **11**, 296-307.

Gleason JM and Powell JR (1997) *Molecular Biology and Evolution,* **14(7)**, 741–53.

Gleason JM, Griffith EC and Powell JR (1998) *Evolution*, **52(4)**, 1093–103

Goldstein DB and Schlotterer C (Eds) (1999) *Microsatellites: Evolution and Applications.* Oxford University Press: Oxford.

Gomez-Zurita J, Juan C and Petitpierre E (2000) *Molecular Phylogenetics and Evolution*, **14**, 304–317.

Gonzalez IL, Chambers C, Gorski JL, Stambolian D, Schmickel RD and Sylvester JE (1990) *Journal of Molecular Biology*, **212**, 27–35.

Gonzalez-Rodriguez A, Benrey B, Castaneda A and Oyama K (2000) *Annals of the Entomological Society of America*, **93(5)**, 1100-1107.

Gottschling M and Plotner J (2004) *Nucleic Acids Research,* **32(1)**, 307-315.

Gowri-Shankar V and Rattray M (2006) *Molecular Biology and Evolution*, **23**, 352–364.

Gozlan S, Millot P, Rousset A and Fournier D (1997) *Entomophaga,* **42(4)**, 593-604.

Gupta PK, Varshney RK, Sharma PC and Ramesh B (1999) *Plant Breeding*, **118(5)**, 369-390.

Gutell RR, Lee JC and Cannone JJ (2002) *Current Opinions in Structutral Biology*, **12**, 301–310.

Graham J, Gordon SC and Williamson B (1996) Brighton Crop Protection Conference: Pests and Diseases. BCPC, Farnham, UK, 777–782.

Graybeal A (1994) *Systematic Biology*, **43**, 174–93.

Grossi-de-Sa MF and Chrispeels MJ (1997) *Insect Biochemistry and Molecular Biology*, **27**, 271–281.

Grossi-de-Sa MF, Mirkov TE, Ishimoto M, Colucci G, Bateman KS and Chrispeels MJ (1997) *Planta*, **203**, 295–303.

Hackett BJ, Gimnig J, Guelbeogo W, Constantini C, Koekemoer LL, Coetzee M, Collins FH and Besansky NJ (2000) *Insect Molecular Biology*, **9**, 369–374.

Hadrys H, Balick M and Schierwater B (1992) *Molecular Ecology*, **1(1)**, 55-63.

Hallahan DL, Pickett JA, Wadham LJ, Wallsgrove RM and Woodcock CM (1992) In: Plant genetic manipulation for crop protection, A.M.R. Gatehouse, V.A. Hilder and D. Boulter (eds.), pp. 215-248. CAB International, Wallingford, UK.

Hamarsheh O, Presber W, Abdeen Z, Swalha S, Al-Lahem A and Schonian G (2007) *Medical and Veterinary Entomol*, **21**, 270-277.

Hamby RK and Zimmer EA (1992) In: Soltis, P., Soltis, D., Doyle, J. (Eds.), Molecular Systematics of Plants. Chapman & Hall, New York, 50–91.

Hancock JM, Tautz, D and Dover AG (1988) *Molecular Biology and Evolution*, **5(4)**, 393–414.

Heath RL, McDonald G, Christeller JT, Lee M, Bateman K, West J, Van Heeswijck R and Anderson MA (1997) *Journal of Insect Physiology*, **43(9)**, 833-842.

Hepburn HR, Smith DR, Radloff SE and Otis GW (2001) *Apidologie*, **32**, 3-23.

Hernandez-Vera G, Mitrovic M, Jovic J ,Tosevski I, Caldara R, Gassmann A and Emerson BC (2010) *Molecular Ecology*, **19**, 2286–2300.

Hillis DM and Dixon M (1991) *The Quarterly Review of Biology*, **66,** 411–453.

Ho MF, Yin X, Finardi-Filho F, Lajolo F and Whitaker JR (1994) In: Yada, R.Y., Jackman, R.L., Smith, J.L. (Eds.), Protein Structure–Function Relationships in Foods, Chapman & Hall, London, pp. 89–119.

Ho M and Whitaker JR (1993) *Journal of Food Biochemistry*, **17(1)**, 15-33.

Hovmoller R, Pape T and Kallersjo M (2002) *Cladistics*, **18**, 313–323.

Hung G-C, Chilton NB, Beveridge I and Gasser RB (1999) *International Journal of Parasitology*, **29**, 1949–1964.

Harris H (1966) *Series B, Biological Sciences*, 298-310.

Hartfield EA, Marvin KH and Raul FM (2012) *Agricultural and Forest Entomology*, **14**, 119–125.

Hasan HS (2006) *Journal of Entomology*, **3(4)**, 290-297, 2006.

Heath RG, McDonald JT, Christeller M, Lee K, Bateman J, West R, Van Heeswijck and Anderson MA (1997) *Journal of Insect Physiology*, **9**, 833-842.

Heckel DG (2003) *Annual Review of Entomology*, 48(1), 235-260.

Heideman C, Munhoz REF, Pattaro Junior JR and Fernandez MA (2010) *Genetics and Molecular Research*, **9(4)**, 2343-2348.

Hellmich RL, Albajes R, Bergvinson D, Prasifka JR, Wang ZY and Weiss MJ (2008) In: Romeis J, Shelton AM, Kennedy GG, eds. *Intergation of Insect-Resistant Genetically Modified Crops With IPM systems*. Springer, Berlin, Germany, 119–158.

Hentschel CC (1982) *Nature*, **295**, 714–716.

Henry Y, Wood H, Morrissey JP, Petfalski E, Kearsey S and Tollervey D (1994) *European Molecular Biology Organization Journal*, **13**, 2452–2463

Hickey DA, Benkel BF and Magoulas C (1989) *Genome*, **31(1)**, 272-283.

Hilder VA, Gatehouse AMR, Sheerman SE, Barker RF and Boulter D (1987) *Nature*, **330**, 160-163.

Hilder VA, Powell KS, Gatehouse AMR, Gatehouse JA, Gatehouse LN, Shi Y, Hamilton WD, Merryweather A, Newell CA, Timans JC, Peumans WJ, Van Damme E and Boulter D (1995) *Transgenic Research*, **4**, 18–25.

Hillis DM, Moritz C and Mable BK (1996) *Molecular Systematics.*, (Ed. 2).

Hillis DM and Dixon MT (1991) *The Quantitative Review of Biology*, **66**, 411– 453.

Hill M, Witsenboer H, Zabeau M, Vos P, Kesseli R and Michelmore R (1996) *Theoretcal and Applied Genetics*, **93**, 1202-1210.

Hindenach BR and Stafford DW (1983) *Nucleic Acids Research*, **12**, 1737- 1747.

Horn A, Roux-Morabito G, Lieutier F and Kerdelhue C (2006) *Molecular Ecology*, **15**, 1603-1615.

Hoy MA, Jeyaprakash A, Morakote R, Lo PC and Nguyen R (2000) *Biological Control*, **17**, 1–10.

Hoy MA (2003) *Insect Molecular Genetics – An introduction to principles and applications, second edition*. Academic Press, San Diego, CA.

HsuChen C-C, Kotin RM and Dubin DT (1984) *Nucleic Acids Research*, **12**, 7771–7785.

Hsuchen C-C and Dubin DT (1984) *Biochemistry International*, **8**, 385-39 1.

Huang JK, Mi JW, LIN H, Wang ZJ, Chen RJ, Hu RF, Rozelle S and Pray C (2010) *Science China Life Sciences*, **53**, 981–991.

Huber K, Ba Y, Dia I, Mathiot C, Sall AA and Diallo M (2008) *The American journal of tropical medicine and hygiene*, **79(2),** 218-229.

Hughes WOH and Boomsma JJ (2004) *Evolution*, **58(6),** 1251–1260.

Hundsdoerfer AK and Wink M (2006) *Journal of Zoological Systematics and Evolutionary Research*, **44**, 316–322.

Hurme P and Savolainen O (1999) *Molecular Ecology*, **8(1)**, 15-22.

Hutchison WD, Burkness EC, Mitchell PD, Moon RD, Leslie TW, Fleischer SJ, Abrahamson M, Hamilton KL, Steffey KL, Gray ME, Hellmich RL, Kaster LV and Hunt TE, Wright RJ, Pecinovsky K, Rabaey TL, Flood BR and Raun ES (2010) *Science*, **330**, 222–225.

Ibrahim SA, El-Mergawy RG and Moghaieb RE (2009) *Australian Journal of Basic and Applied Sciences*, **3(2)**, 1287-1295.

Inomata N, Shibata H, Okuyama E and Yamazaki T (1995) *Genetics*, **141(1)**, 237-244.

Irie K, Hosoyama H, Takeuchi T, Iwabuchi K, Watanabe H, Abe M and Arai S (1996) *Plant Molecular Biology*, **30**, 149–157.

Ishimoto M and Kitamura K (1989) *Applied Entomology and Zoology*, **24**, 281–286.

Ishimoto M, Sato T, Chrispeels MJ and Kitamura K (1996) *Entomology Experimantal and Applied*, **79**, 309–315.

Jalalizand AR, Karimi A, Modaresi M and Mahmoodi E (2012) In: International Conference on Applied Life Sciences (ICALS2012) Turkey, September 10-12, 2012.

James DJ, Passey AJ, Easterbrook MA, Solomon MG and Barbara DJ (1992) *Phytoparasitica*, **20(1)**, S83-S87.

James C (2009) *Global Status of Commercialized Biotech/GM Crops: 2009*. ISAAA, Ithaca, NY.

Janssen P, Maquelin K, Coopman R, Tjernberg I, Bouvet P, Kersters K, and Dijkshoorn L (1997) *International Journal of Systematic Bacteriologoy*, 1179-1187

Jarne P and Lagoda PJ (1996) *Trends in Ecology & Evolution*, **11(10)**, 424-429.

Jermiin LS and Crozier RH (1994) Journal of Molecular Evolution, **38**, 282–294.

Jia W and Higgs PG. (2007) *Molecular Biology and Evolution*, **25**, 339–351.

Jiang X-F, Cao W-J, Zhang L and Luo L-Z (2010) *Environmental Entomology*, **39**, 232–242.

Johnson FM, Kanapi CG, Richardson RH, Wheeler MR and Stone WS (1966) *Proceedings of the National Academy of Sciences of the United States of America*, **56(1)**, 119.

Johns GC and Avise JC (1998) *Molecular Biology and Evolution*, **15(11)**, 1481-1490.

Johnson R, Narraez J, An G and Ryan CA (1989) *Proceedings of the National Academy of Sciences United States of America*, **86**, 9871- 9875.

Johnston KA, Lee MJ, Gatehouse JA and Anstee JH (1991) *Insect Biochem*istry, **21**, 389-397.

Johnston KA, Gatehouse JA and Anstee JH (1993) *Journal of Insect Physiology*, **39**, 657-664.

Johnston KA, Lee MJ, Brough C, Hilder VA, Gatehouse AMR and Gatehouse JA (1995) *Insect Biochemistry and Molecular Biology*, **25**, 375-383.

Jones CJ, Edwards KJ, Castaglione S, Winfield MO, Sala F, Van de Wiel C and Karp A (1997) *Molecular Breeding*, **3(5)**, 381-390.

Jongsma MA, Bakker PL, Stiekema WJ and Bosch D (1995) *Molecular Breeding*, **1**, 181-191.

Jongsma MA, Stiekema WJ and Bosch D (1996) *Trends in Biotechnology*, **14**, 331-333.

Joshi SP, Ranjekar PK and Gupta VS (1999) *Current Sciience*, **77(2)**, 230-240.

Juan C, Oromi P and Hewitt GM (1996) *Heredity*, **76**, 392–403.

Julio N B, Chiapper M B, Rossi HJ, Rondan Dueñas JC and Gardenal C N (2009) *Memorias do Instituto Oswaldo Cruz,* **104(4)**, 626-631.

Ishimoto M, Sato T, Chrispeels MJ and Kitamura K (1996) *Entomologia Experimentalis et Applicata*, **79**, 309-315.

Kambhampati S (1996) *Systematic Entomology*, **21**, 89-98.

Kambhampati S, Voelkl W and Mackauer M (2000) *Systematic Entomology,* **25**, 437–445.

Kar PK, Vijayan K, Nair CV, Mohandas TP, Saratchandra B and Thangavelu K (2005) *Genetica*, **125**, 173-183.

Katiyar SK, Chandel G, Tan Y, Zhang Y, Huang B, Nugaliyaddar L, Fernando K, Bentur JS, Inthavong S, Constantino S and Bennett J (2000) *Genome,* **43**, 322-332.

Kawamura K, Sugimoto T, Matsuda Y and Toyoda H (2007) *Applied Entomology Zoology*, **42(2)**, 297–303 (2007).

Kazachkova N, Meijer J and Ekbom B (2007) *Agricultural Forest Entomology,* **9**, 259-269.

Kazachkova N, Meijer J and Ekbom B (2008) *European Journal of Entomology*, **105**, 807–814

Keyghobadi N, Unger KP, Weintraub JD and Fonseca DM (2006) *Conservation Genetics*, **7**, 309–313.

Khemakhem MM, Marrakchi M and Makni H (2005) *African Journal of Biotechnology,* **4**, 601-606.

Kim FJ, Battini JL, Manel N and Sitbon M (2004) *Virology*, **318**, 183–191.

Kim KS and Sappington TW (2004) *Insect Molecular Biology,* **13**, 293–303.

Kjer K (2004) *Systematic Biology,* **53**, 506–14.

Kluh I, Horn M, Hyblova J, Hubert J, Doleckova-Maresova L, Voburka Z, Kudlíkova I, Kocourek F and Mares M (2005) *Phytochemistry,* **66**, 31- 39.

Kostia S, Ruohonen-Lehto M, Vainola R and Varvio SL (2000) *Heredity,* **84**, 37–45.

Kumar S, Tamura K, Jakobsen IB and Nei M (2001) *Bioinformatics*, **17**, 1244–1245.

Kuwata R, Tokuda M, Yamaguchi D and Yukawa J (2005) *Journal of Applied Entomology*, **129(4)**, 191-197.

Krumm JT, Hunt TE, Skoda SR, Hein GL, Lee DJ, Clark PL and Foster JE (2008) *Journal of Insect Science,* **8**, 72.

Lall GK, Darby AC, Nystedt B, MacLeod ET, Bishop RP and Welburn SC (2010) *Parasites & Vectors*, **3**, 47.

Latif MA, Soon Guan T, Mohd OY and Siraj SS (2008) *Biochemical Genetics*, **46**, 520–537.

Lavrov DV, Brown WM and Boore JL (2004) *Proceedings of Royal Society of London Series B*, **271**, 537–544.

Le Berre-Anton V, Bompard-Gilles C, Payan F and Rouge P (1997) *Biochemistry and Biophysics Acta*, **1343**, 31–40.

Lee MK, Aguda RM, Cohen MB, Gould FL and Dean DH (1997) *Applied Environmental Microbiology*, **63(4)**, 1453-1459.

Lee SI, Lee SH, Koo JC, Chun HJ, Lim CO, Mun JH, Song YH and Cho MJ (1999) *Molecular Breeding*, **5**, 1-9.

Leple JC, Bonade-Botinno M, Augustin S, Pilate G, Ce Tan VD, Delplanque A, Cornu D and Jouanin L (1995) *Molecular Breeding*, **1**, 319–328.

Lessinger AC, Martins JAC, Lemos TA, Kemper EL, da Silva FR, Vettore AL, Arruda P and Azeredo-Espin AM (2000) *Insect Molecular Biology*, **9**, 521–529.

Lewis WJ, Van Lenteren JC, Phatak S C and Tumlinson JH (1997) *Proceedings of the National Academy of Sciences*, **94(23)**, 12243-12248.

Lewis KG, El-Kassaby YA, Alfaro RI and Barnes S (2000) *Annals of the Entomological Society of America*, **93(4)**, 807-818.

Lewontin RC and Hubby JL (1966) *Genetics*, **54(2)**, 595.

Li C and Wilkerson RC (2005) *Memorias do Instituto Oswaldo Cruz*, **100**, 495–500.

Liu H and Beckenbach AT (1992) *Molecular Phylogenetics and Evolution*, **41**, 41–52.

Liu JN, Gui FR and Li ZY (2010) *Journal of Insect Science*, **10**, 52.

Loukas M, Krimbas CB and Mavragani-Tsipidou P (1979) *Journal of Heredity*, **70(1)**, 17-26.

Loxdale HD, Tarr IJ, Weber CP, Brookes CP, Digby PGN and Castanera P (1985) *Bulletin of Entomological Research*, **75(01)**, 121-142.

Loxdale HD and den Hollander J (Eds) (1989) *Systematics Association Special Volume* No. 39, Oxford, Clarendon Press.

Loxdale HD and Lushai G (1998) *Bulletin of Entomological Research*, **88(06)**, 577-600.

Lu R and Rank GH (1996) *Genome,* **39**, 655–663.

Lu S, Deng P, Liu X, Luo J, Han R, Gu X and Pongor S (1999) *Journal of Biological Chemistry*, **274(29),** 20473-20478.

Luckinbill LS and Golenberg EM (2002) *Genetica*, **114(2)**, 147-156.

Luikart G and England PR (1999) *Trends in Ecology & Evolution*, **14(7)**, 253-256.

Lunt DH, Ibrahim KM and Hewitt GM (1998) *Heredity,* **80**, 633–641.

Luque C, Legal L, Staudter H, Gers C and Wink M (2002) *Hereditas,* **136**, 251-253.

Maa WC and Terriere LC (1983) *Comparative Biochemistry and Physiology Part C: Comparative Pharmacology,* **74(2)**, 451-460.

MacDonald C and Loxdale H (2004) *International Journal of Pest Management*, **50(3)**, 215-224.

de Maagd RA, Bosch D and Stiekema W (1999) *Trends in Plant Science*, **4(1)**, 9-13.

Mai JC and Coleman AW (1997) *Journal of Molecular Evolution*, **44**, 258–271.

Magana C, Beroiz B, Hernandez-Crespo P, de Oca MMA, Carnero A, Ortego F and Castanera P (2007) *Bulletein of Entomology Research*, **97**, 585-590.

Malloch G, Fenton B and Goodrich MA (2001) *Insect Molecular Biology*, **10**, 281–291.

Malone LA, Gatehouse AMR and Barratt BIP (2008) In: Romeis J, Shelton AM, Kennedy GG, eds. *Intergation of Insect- Resistant Genetically Modified Crops With IPM Programs*. Springer, Berlin, Germany, 357–417.

Mange A and Prudhomme JC (1999) *Molecular Biology and Evolution*, **16(2)**, 165-172.

Manonmani AM, Townson H, Adeniran T, Jambulingam P, Sahu SS and Vijayakumar T (2001) *Acta Trop, **78**,* 3-9.

Manwell C and Baker CA (1970) *Molecular biology and the origin of species: heterosis, protein polymorphism and animal breeding,* **394**.

Marchetti S, Delledonne M, Fogher C, Chiaba C, Chiesa F, Savazzini F and Giordano A (2000) *Theoretical and Applied Genetics*, **101**, 519–526.

Marinucci M, Romi R, Mancini P, Di Luca M and Severini C (1999) *Insect Molecular Biology*, **8**, 469–480.

Marrelli MT, Floeter-Winter LM, Malafronte RS, Tadei WP, Lourenco-De-Oliveira R, Flores-Mendoza C and Marinotti O (2005) *Medical and Veterinary Entomology*, **19**, 208 – 218.

Markert GR, Paxaton RJ and Hartfelder K (2006) *Apidologie,* **37**, 687-698.

Marshall JJ and Lauda CM (1975) *Journal of Biological Chemistry*, **250(20)**, 8030-8037.

Martins WF, Ayres CF and Lucena WA (2007) *Genetics and Molecular Research,* **6(1)**, 23–32.

Mba C and Tohme J (2005) In: *Molecular Evolution: Producing the Biochemical Data, Part B* (eds Zimmer EA, Roalson EH), 177–201. Academic Press, San Diego, California.

McHugh D (1998) *Zoologist*, **38(6)**, 859–66.

McKern JA, Szalanski AL, Johnson DT and Dowling APG (2008) *Journal of Agriculture and Urban Entomology*, **25(3)**, 165–177.

Mclain DK, Wesson DM, Oliver JHJ and Collins FH (1995) *Journal of Medical Entomology*, **32**, 353–360.

McManus MT, White DWR and McGregor PG (1994) *Transgenic Researc*h, **3**, 50–58.

McManus MT and Burgess EPJ (1995) *Journal of Insect Physiology*, **41**, 731-738.

Meglecz E, Anderson SJ, Bourguet D, Butcher R, Caldas A, Cassel- Lundhagen, A and Wilcock HR (2007) *Insect Molecular Biology*, **16(2)**, 175-185.

Meudt HM and Clarke AC (2007) *Trends in Plant Science*, **12**, 106–117.

Memon N, Meier R, Manan A and Feng-Yi SUK (2006) *Systematic Entomology*, **31**, 703–710.

Mezghani-Khemakhem Maha, Bouktila D, Kharrat I, Makni M and Makni H (2012) *Entomological Science,* **15**, 171–179.

Michaud D, Cantin L and Vrain TC (1996) *Archives of Insect Biochemistry and Physiology,* **31**, 451–464.

Michot B, Bachellerie JP and Raynal F (1983) *Nucleic Acids Research,* **11**, 3375–3390.

Miller BR, Crabtree MB and Savage HM (1997) *Insect Molecular Biology*, **6(2)**, 105-114.

Miller M, Blinn M and Keim P (2002) *Freshwater Biology*, **47**, 1660–1673.

Mindell DP and Honeycutt RL (1990) *Annual Review of Ecology and Systematics*, **21**, 541-566.

Minney BHP. Gatehouse AMR, Dobie P, Dendy J, Cardona C and Gatehouse JA (1990) *Journal of Insect Physiology*, **36**, 757–767.

Mirhoseini SZ, Dalirsefat SB and Khairandish MP (2007) *Journal of Economic Entomology*, **100(3)**, 939-945.

Mirhoseini SZ, Rabiei B, Potki P and Dalirsefat SB (2010) *Journal of Insect Science*, **10**, 153.

Mitchell A, Chow S, Regier JC, Mitter C, Poole RW and Matthews M (1997) *Molecular Biology and Evolution*, **14**, 381–390.

Mock K E, Bentz BJ, O'Neill EM,Chong JP, Orwin J and Pfrender ME (2007) *Molecular Ecology*, ***1*6**, 553–568.

Mochizuki A, Nishizawa Y, Onodera H, Tabei Y, Toki S, Habu Y and Ohashi Y (1999) *Entomologia Experimentalis et Applicata*, **93(2)**, 173-178.

Moin-Vaziri V, Depaquit J, Yaghoobi-Ershadi M-R, Oshagi M-A, Derakhshandeh-Peykar P, Ferte H., Kaltenbach M, Bargues MD, Leger N and Nadim A (2007) *Acta Tropica*, **102**, 29-37.

Monteiro A and Pierce NE (2001) *Molcular Phylogenetics and Evolution*, **18**, 264–281.

Moran NA, Kaplan ME, Gelsey MJ, Murphy TG and Scholes EA. (1999) *Systematic Entomology*, **24**, 85–93.

Moritz C, Dowling TE and Brown WM (1987) *Annual Review of Ecology and Systematics*, **18**, 269-292.

Morton RL, Schroeder HE, Bateman KS, Chrispeels MJ, Armstrong E and Higgins TJV (2000) *Proceedings of National Academy of Science USA*, **97**, 3820–3825.

Moya A, Guirao P, Cifuentes D, Beitia F and Cenis JL (2001) *Molecular Ecology*, **10**, 891–897.

Mozaffarian F, Mardi M, Sarafrazi A and Ganbalani G (2007) *Journal of Insect Science*, **8(6)**.

Mueller UG and Wolfenbarger LL (1999) *Trends in Ecology and Evolution*, **14**, 389–394.

Mueller RL and Boore JL (2005) *Molecular Biology and Evolution*, **22**, 2104–2112.

Mukabayire O, Boccolini D, Louchouarn L, Fontenille D and Besansky NJ (1999) *Molecular Ecology*, **8**, 289-297.

Mullis K, Faloona FA, Scharf SJ, Saiki RK, Horn GT and Erlich H (1986) *Biotechnology Series*, 17-17.

Mundy J and Rogers JC (1986) *Planta*, **169(1)**, 51-63.

Murao S, Oouchi N, Goto A and Arai M (1983) *Agricultural and Biological Chemistry*, **47(2)**, 453-454.

Murrell A, Campbell NJ and Barker SC (2001) *Insect Molecular Biology,* **10**, 587–596.

Murthy BCK, Prakash BM and Puttaraju HP (2006) *Cytologia,* **71(4)**, 331–335.

Nagaraja GM and Nagaraju J (1995) *Electrophoresis,* **16**, 1633-1638.

Naranjo SE, Ruberson JR, Sharma HC, Wilson L and Wu KM (2008) In: Romeis J, Shelton AM, Kennedy GG, eds. *Integration of Insect-Resistant Genetically Modified Crops With IPM Programs.* Springer, Berlin, Germany, 159–194.

Nardi F, Spinsanti G, Boore JL, Carapelli A, Dallai R and Frati F (2003) *Science,* **299**, 1887–89.

Navajas M, Cotton D, Kreiter S and Gutierrez J (1992) *Experimental and Applied Acarology,* **15**, 211-218.

Navajas M, Gutierrez J, Bonato O, Bolland HR and Mapangou-Divasse S (1994) *Experimental and Applied Acarology,* **18**, 351-360.

Navajas M, Lagnel J, Gutierrez J and Boursot P (1998) *Heredity,* **80**, 742–752.

Navajas M, Lagnel J, Fauvel G and de Moraes G (1999) *Experimental and Applied Acarology,* **23**, 851-859.

Navajas M, Tsagkarakov A, Lagnel J and Perrot-Minnot MJ (2000) *Experimental and Applied Acarology,* **24**, 365-376.

Nayar JK and Knight JW (2002) *Medical and Veterinary Entomology*, **16(4)**, 424-429.

Neog K, Singh HR, Unni B and Sahu AK (2010) *African Journal of Biotechnology,* **9(12)**, 1746-1752.

Newell CA, Lowe JM, Merryweather A, Rooke LM and Hamilton WDO (1995) *Plant Science,* **107**, 215-227.

Nice CC and Shapiro AM (2001) *Annals of the Entomological Society of America*, **94(2)**, 257-267.

Nielsen J, Peixoto AA, Piccin A, Costa R, Kyriacou CP and Chalmers D (1994) *Molecular Biology and Evolution,* **11(6)**, 839–53.

Nigro L and Grapputo A (1993) *Systematic Ecology,* **21(1)**, 79–83.

Nigro L, Solignac M and Sharp PM (1991) *Journal of Molecular Evolution*, **33**, 156- 162.

Ocampo CB and Wesson DM (2004) *The American jJournal of Tropical Medicine and Hygiene*, **71(4)**, 506-513.

Ohtsubo KI and Richardson M (1992) *FEBS letters*, **309(1)**, 68-72.

Okuyama E, Shibata H, Tachida H and Yamazaki T (1996) *Molecular Biology and Evolution,* **13(4)**, 574-583.

Paplauskien V, Ceksteryt V, Pasakinskien I, Tamasauskien D and Racys J (2006) *Biologia,* **3**, 16-20.

Paris M and Despres L (2012) *Molecular Ecology,* **21**, 1672–1686.

Parsons YM and Shaw L (2001) *Molecular Ecology*, **10**, 1765–1772.

Pashley DP and Ke LD (1992) *Molecular Biology and Evolution*, **9**, 1061-1075.

Pashley DP, McPheron BA and Zimmer EA (1993) *Molecular Phylogenetics and Evolution,* **2**, 132–142.

Paupy C, Vazeille-Falcoz M, Mousson L, Rodhain F and Failloux AB (2000) *The American Journal of Tropical Medicine and Hygiene,* **62(2)**, 217-224.

Payan F (2004) *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, **1696(2)**, 171-180.

Peferoen M (1997) *Trends in Biotechnology*, **15**, 173-177.

Pereira PJB, Lozanov V, Patthy A, Huber R, Bode W, Pongor S and Strobl S (1999) *Structure*, **7**, 1079-1088.

Perez T, Albornoz J and Dominguez A (1998) *Molecular Ecology*, **7(10)**, 1347-1357.

Peter M, Lindfeld A and Nentwig W (2010) *Pedobiologia*, **53(4)**, 271-279.

Philips CB, Cane RP, Mee J, Chapman HM, Hoelmer KA and Coutinot D (2002) *New Zealand Journal of Agricultural Research*, **45**, 295–303.

Phuc HK, Ball AJ, Son L, Hanh NV, Tu ND, Lien NG, Verardi A and Townson H (2003) *Medical and Veterinary Entomology*, **17**, 423–428.

Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey SS and Rafalski JA (1996) *Molecular Breeding*, **2**, 225-238.

Powers JR and Whitaker JR (1977) *Journal of Food Biochemistry*, **1**, 217–238.

Powers JR and Culberston JD (1983) *Cereal Chemistry*, **60**, 427-429.

Puterka GJ, Black WC IV, Steiner WM and Burton RL (1993) *Heredity,* 70, 604-618.

Quicke DLJ and Belshaw R (1999) *Systematic Biology*, **48**, 436 – 454.

Rabouam C, Comes A, Bretagnolle V, Humbert JF, Periquet G and Bigot Y (1999) *Molecular Ecology*, **8(3)**, 493-503.

Rademaker JLW, Louws FJ, Schultz MH, Rossbach U, Vauterin L, Swings J and de Bruijn FJ (1997) *Phytopathology*, **87**, S81.

Reddy KD, Nagaraju J and Abraham EG (1999) *Heredity,* **83**, 681-687.

Reed RD and Sperling FA (1999) *Molecular Biology and Evolution*, **16**, 286–297.

Reed DH and Frankham R (2001) *Evolution*, **55(6)**, 1095-1103.

Regier JC and Shultz JW (1997) *Molecular Biology and Evolution,* **14(9)**, 902-913.

Regier JC, Fang QQ, Mitter C, Peigler RS, Friedlander TP and Solis AM (1998) *Molecular Biology and. Evolution,* **15**, 1172–82

Richardson BJ, PR Baverstock and Adams M (1986) Allozyme Electrophoresis: A handbook for animal systematics and population studies, Academic Press, Sydney

Roderick GK (1996) *Annual Review of Entomology*, **41**, 325-352.

Roehrdanz RL, Szalanski AL and Levine E (2003) A*nnuals of the Entomological Society of America,* **96(6)**, 901-913.

Rokas A, Nylander JAA, Ronquist F and Stone GN (2002) *Molecular Phylogenetics and Evolution*, **22**, 206–219.

Ross KG, Shoemaker DD, Krieger MJ, DeHeer CJ and Keller L (1999) *Molecular Biology and Evolution*, **16**, 525–543.

Roux O, Gevrey M, Arvanitakis L, Gers C, Bordat D and Legal L (2007) *Molecular Phylogenetics and Evolution*, **43**, 240–250.

Roux-Morabito G, Gillette NE, Roques A, Dormont L, Stein J and Sperling FAH (2008) *Annals of the Entomological Sciety of America*, **101(5)**, 845-859.

Ryan CA (1973) *Annual Review in Plant Physiology*, **24**, 173-196.

Ryan CA (1990) *Annual Review in Phytopatholgy,* **28**, 425–449.

Saha D, Ranjan SK, Mallick CB, Vidyarthi AS and Ramani R (2011) *Biochemical Systematics and Ecology,* **39**, 112–120.

Samara R, Monje JC, Reineke A and Zebitz CPW (2008) *Journal of Applied Entomology*, **132**, 230–238.

Sanchis A, Latorre A, Gonzalez-Candelas F and Michelena JM (2000) *Molecular Phylogenetics and Evolution*, **2**, 180–194.

Sanchez-Gea J-F, Galian J and Serrano J (2004) *European Journal of Entomology*, **101**, 503–511.

Sane VA, Nath P, Aminuddin and Sane PV (1997) *Current Science*, **72**, 741-747.

Santos SR, Kinzie RA III, Sakai K and Coffroth MA (2003) *Journal of Eukaryotic Microbiology*, **50**, 417–421.

Saravanakumar R, Ponnuvel KM and Qadri SMH (2010) *Entomologia Experimentalis et Applicata*, **135**, 170–176.

Satta Y, Ishiwa H and Chigusa SI (1987) *Molecular Biology and Evolution*, **4**, 638-650.

Satta Y and Takahata N (1990) *Proceedings of National Academy of Science USA*, **87**, 9558-9562.

Savelkoul PH, Aarts HJ, de Haas J, Dijkshoorn L, Duim B, Otsen M, Rademaker JL, Schouls L and Lenstra JA (1999) *Journal of Clinical Microbiology*, **37(10)**, 3083–3091.

Schierwater B, Streit B, Wagner GP and DeSalle R (1994) Molecular Ecology and Evolution: Approaches and applications. Birkhauser Verlag AG. (eds B Schierwater, B Streit, GP Wagner & R DeSalle) pp. 455–477. Birkhauser Verlag, Basel.

Schiffer M, Kennington WJ, Hoffmann AA and Blacket MJ (2007) *Molecular Ecology*, **16**, 1687–1700

Schlotterer C and Tautz D (1992) *Nucleic Acids Research*, **25**, 211–215.

Schlotterer C, Hauser MT, von Haeseler A and Tautz D (1994) *Molecular Biology and Evolution*, **11(3)**, 513-522.

Schlotterer C (2000) *Chromosoma*, **109**, 365-371.

Schmidt ER, Foley DH, Hartel GF, Williams GM and Bryan JH (2001) *Bulletin of Entomological Research*, 91(05), 389-410.

Schulenburg JHG, Englisch U and Wagele JW (1999) *Journal of Molecular Evolution*, **48**, 2–12.

Schroeder HE, Gollash S, Moore A, Tabe LM, Craig S, Hardie D, Chrispeels MJ, Spencer D and Higgins TJV (1995) *Plant Physiology*, **107**, 1233–1239.

Scribner KT and Pearce JM (2000) In: Baker, A. J. (ed): Molecular Methods in Ecology – Blackwell Sci., Oxford, 235-273.

Severini C, Silvestrini F, Mancini P, LaRosa G and Marinucci M (1996) *Insect Molecular Biology*, **5**, 181–186.

Severson DW, Brown SE and Knudson DL (2001) *Annual Review of Entomology*, **46(1)**, 183-219.

Seyahooei MA, van Alphen JJM and Kraaijeveld K (2011) *BioMedCentral Ecology*, **11**, 4.

Shade RE, Schroeder HE, Pueyo JJ, Tabe LM, Murdoch LL, Higgins TJV and Chrispeels MJ (1994) *Bio/Technology*, **12**, 793-796.

Shao R, Campbell NJ and Barker SC (2001) *Molecular Biology and Evolution*, **18**, 858–865.

Shao R and Barker SC (2003) *Molecular Biology and Evolution*, **20**, 362–370.

Sharma HC and Norris DM (1991) *Journal of Science of Food and Agriculture*, **55**, 353-365.

Sharma KK, Jaiswal AK and Kumar KK (2006) *Current Science*, **91**, 894–898.

Sheffield NC, Song H, Cameron SL and Whiting MF (2008) *Molecular Biology and Evolution*, **25(11)**, 2499–2509.

Shibata H and Yamazaki T (1995) *Genetics*, **141(1)**, 223-236.

Sidorenko AP and Berezovska OP (2002) *Genetika*, **38(11)**, *1485–1491*.

Sielezniew M, Ponikwicka-Tyszko D, Ratkiewicz M, Dziekanska I, Kostro-Ambroziak A and Rutkowski R (2011) *European Journal of Entomology*, **108**, 537–545.

Silvestre D, Dowton M and Arias M C (2008) *Genetics and Molecular Biology*, **31**, 2, 451-460.

Simon C, Frati F, Beckenbach A, Crespi B, Liu H and Flook P (1994) *Annals of the entomological Society of America*, 87(6), 651-701.

Simon JC and Hebert PD (1995) *Heredity*, **74(4)**, 346-353.

Smith TB and Wayne RK (Eds.) (1996) *Molecular Genetic Approaches in Conservation* (p. 483). New York: Oxford University Press.

Smith JJ and Bush GL (1997) *Molecular Phylogenetics and Evolution*, **7(1)**, 33–43.

Smith PT, Kambhampati S, Volkl W and Mackauer M (1999) *Molecular Phylogenetics and Evolution*, **2**, 236–245.

Smith DR, Villafuerte L, Otis GW and Palmer MR (2000) *Apidologie,* **31**, 265-280.

Smith PT, Kambhampati S and Armstrong KA (2003) *Molecular Phylogenetics and Evolution*, **26**, 8–17.

Smithson JB and Lenne JM (1996) *Annals of Applied Biology*, **128(1)**, 127-158.

Solomon MG, Fitzerald JD, Murray RA (1996) In: *The Ecology of Agricultural Pests: Biochemical Approaches* (eds Symondson WOC, Liddell JE), pp. 457–468. Chapman & Hall, London.

Sosa-Gomez DR, da Silva JJ, Costa F, Binneck E, Marin SRR and Nepomuceno AL (2005) *Journal of Insect Science*, **5**, 23.

Sperling FAH, Anderson GS and Hickey DA (1994) *Journal of Forensic Science*, **39(2)**, 418–27.

Sperling FAH and Hickey DA (1994) *Molecular Biology and Evolution,* **11**, 656–65.

Steffens R, Fox FR and Kassel B (1978) *Journal of Agricultural and Food Chemistry,* **26**, 170-175.

Steiner WWM and Joslyn DJ (1979) *Mosquito News*, **39**, 35–54.

Streisinger G and Owen J (1985) *Genetics*, **109**, 633-659.

Subrahmanyamc S, Cassidy B, Busch H and Rothblum LI (1982) *Nucleic Acids Research*, **10**, 3667-3680.

Subramanian S and Mohankumar S (2006) *Journal of Insect Science*, **6(26)**.

Su-fang Z, Xiao-wen Y and Ji-sheng M (2008) *Insect Science*, **7(3)**, 235-242.

Suinaga FA, Casali VWD, Picanço M and Foster J (2004) *Pesquisa Agropecuária Brasileira*, **39 (7)**.

Sula J and Spitzer K (2000) *European Journal of Entomology*, **97(1)**, 7-12.

Susmani-Brown S, Cohnstaedt L and Munsterma NN (2009) *Annals of Entomological Society of America*, **102(1)**, 144-150.

Stewart JB and Beckenbach AT (2003) *Molecular Phylogenetics and Evolution*, 26, 513–526.

Strong DR (1979) *Annual Review of Entomology,* **24**, 89 119.

Symondson WOC and Liddell JE (*Eds*) (1996) *The Ecology of Agricultural Pests: Biochemical Approaches*. London, Chapman & Hall.

Szalanski AL, Roehrdanz RL, Taylor DB and Chandler L (1999) *Insect Molecular Biology*, **8**, 519-526.

Szalanski AL, Sikes DS, Bischof R and Fritz M (2000) *Annals of the Entomological Society of America*, **93**, 589–594.

Szalanski AL and Owens CB (2003) *Florida Entomologist*, **86**, 329-333.

Szalanski AL and McKern JA (2007) *Sociobiology,* **50(3)**, 939-945.

Tabashnik BE (2010) *Science*, **330**, 189–190.

Tan SH, Aris EM, Surin J, Omar B, Kurahashi H and Mohamed Z (2009) *Tropical Biomedicine*, **26(2)**, 173–181.

Templeton AR (1998) *Molecular Ecology*, ***7*(4)**, 381-397.

Thaler R, Brandstatter A, Meraner A, Chabicovski M, Parson W, Zelger R, Dalla Via J and Dallinger R (2008) *Molecular Phylogenetics and Evolution*, **48**, 838–849.

Thao ML, Baumann L and Baumann P (2004) *BioMedCentral Evolutionary Biology,* **4**, 1471-2148.

Thomas MB, Michael EZ and Charles FA (1989) *Genetics*, **123**, 825-836.

Thomas JC, Wasmann CC, Echt C, Dunn RL, Bohnert HJ and Mccoy TJ (1994) *Plant Cell Reports*, **14**, 31-36.

Thomas JC, Adams DG, Keppenne VD, Wasmann CC, Brown JK, Kanosh MR and Bohnert HJ (1995) *Plant Cell Reports*, **14**, 758-762.

Timm AE, Pringle KL and Warnich L (2005) *Bulletin of Entomological Research*, **95**, 187–191.

Timm AE, Geertsema H and Warnich L (2008) *Annals of Entomological Society of America*, **101(1)**, 197-203.

Tishkoff SA, Dietzsch E, Speed W, Pakstis AJ, Kidd JR, Cheung K and Kidd KK (1996) *Science*, **271(5254)**, 1380-1387.

Titarenko E and Chrispeels MJ (2000) *Insect Biochemistry and Molecular Biology,* **30**, 979–990.

Torres RA, Ganal M and Hemleben V (1990) *Journal of Molecular Evolution*, **30**, 170-181.

Trewick SA , Goldberg J and Morgan-Richards M (2005) *Zoologica Scripta*, **34(5)**, 483–491.

Trizzino M, Audisio P, Antonini G, De Biase A and Mancini E (2009) *Molecular Phylogenetics and Evolution*, **51**, 215–226.

Uhlenbusch I, McCracken A and Gellissen G (1987) *Current Genetics*, **11(8)**, 631-638.

Ullrich B, Reinhold K, Niehuis O and Misof B (2009) *Journal of Zoology Systematics Evolution and Research*, **48(3)**, 219–228.

Urbanelli S (2002) *Heredity*, ***88*(5)**, 333-341.

Urwin PE, Atkinson HJ and McPherson MJ (1995) *Protein Engineering*, **8**, 1303–1307.

Urwin PE, Lilley CJ, Mcpherson MJ and Atkinson HJ (1997) *Plant Journal*, **12(2)**, 455-461.

Urwin PE, Troth KM, Zubko EI and Atkinson HJ (2001) *Molecular Breeding*, **8**, 95-110.

Ussuf KK, Laxmi NH and Mitra R (2001) *Current Science-Bangalore*, **80(7)**, 847-853.

Vain P, Worland B, Clarke MC, Richard G, Beavis M, Liu H, Kohli A, Leech M, Snake J and Christou P (1998) *Theoretical and Applied Genetics*, **96**, 266-271.

Valenzuela I, Hoffmann A A, Malipatil MB, Ridland PM and Weeks AR (2007) *Australian Journal of Entomology*, **46**, 305–312.

Van der Sande CA, Kwa M, Van Nues RW, Van Heerikhuizen H, Raue HA and Planta RJ (1992) *Journal of Molecular Biology*, **223**, 899–910.

Van Nues RW, Rientjes JMJ, Van der Sande CAFM, Zerp SF, Sluiter C, Venema J, Planta RJ and Raue HA (1994) *Nucleic Acids Research*, **22**, 912–919.

Vanlerberghe-Masutti and Chavigny (1998) *Molecular Ecology*, **7**, 905-914.

Vaeck M, Reynaerts A, Hofte H, Jansens S, De Beuckeleer M, Dean C, Zabeau M, Van Montagu M and Leemans J (1987) *Nature*, **328**, 33–37.

Vasconcelos T, Horn A, Lieutier F, Branco M and Kerdelhue C (2006) *Agricultural and Forest Entomology*, *Agricultural and Forest Entomology*, **8**, 103–111.

Vazeille M, Mousson L, Rakatoarivony I, Villeret R, Rodhain F, Duchemin JB and Failloux AB (2001) *The American Journal of Tropical Medicine and Hygiene*, **65(5)**, 491-497.

Veldman GM, Klootwijk J, vanHeerikhuizen H and Planta RJ (1981) *Nucleic Acids Research*, **9**, 4847–4862.

Velu D, Ponnuvel KM, Muthulakshmi M, Sinha RK and Qadri SMH (2008) *Journal of Genetics and Genomics*, **35**, 1-7.

Verovnik R and Glogovcan P (2007) *European Journal of Entomology*, ***104*(4)**.

Vijayan K, Tikader A, Kar PK, Srivastava PP, Awasthi AK, Thangavelu K and Saratchandra B (2006) *Genetic Resources and Crop Evolution*, **53**, 873–882.

Vogler AP and DeSalle R (1993) *Evolution*, **47**, 1192-1202.

Vogler AP, DeSalle R, Assmann T, Knisley CB and Schultz TD (1993) *Annals of Entomological Society of America*, **86**, 142-152.

Vogler AP and DeSalle R (1993) In: Desender, K. (ed.) Carabid Beetles: Ecology and Evolution, pp.79—85. Kiuwer, Dordrecht, The Netherlands.

von der Schulenberg JHG, Hancock JM, Pagnamenta A, Sloggett JJ, Marjerus ME and Hurst GD (2001) *Molecular Biology Evolution*, **18**, 648-660.

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M and Zabeau M (1995) *Nucleic Acids Research*, **23**, 4407–4414.

Walldorf U, Hovemann BT and Bautz EKF (1985) *Proceedings of National Academy of Science USA*, **82(17)**, 5795–99.

Walton C, Handley JM, Kuvangkadilok C, Collins FH, Harbach RE, Baimai V and Butlin RK (1999) *Medical and Veterinary Entomology*, **13(1)**, 24-32.

Walton C, Somboon P, O'Loughlin SM, Zhang S, Harbach RE, Linton Y-M, Chen B, Nolan K, Duong S, Fong M-Y, Vythilingum I, Mohammed ZD, Dinh Trung Ho and Butlin RK (2007) *Infection, Genetics and Evolution*, **7**, 93–102.

Wang ZY, Wu Y, He KL and Bai SX (2007) *Bulletin of Insectology*, **60**, 49–55.

Watts PC, Saccheri IJ, Kemp SJ and Thompson DJ (2006) *Freshwater Biology*, **51**, 193–205.

Waegele JW and Stanjek G (1995) *Journal of Zoology Systematics and Evolutionar Research*, **33(2)**, 75–80.

Wei Shu-jun S, Sharkey M, van Achterberg MJ, Chen C and Xue-xin (2010) *BioMedCentral Genomics*, **11**, 371.

Weng Y, Azhaguvel P, Michels GJ and Rudd JC (2007) *Insect Molecular Biology*, **16(5)**, 613-622.

Wheeler W and Honeycut R (1988) *Molecular Biology Evolution*, **5**, 90-96.

Wheeler WC, Whiting M, Wheeler QD and Carpenter JM (2001) *Cladistics*, **17**, 113–169.

Whiting MF, Carpenter JC, Wheeler QD and Wheeler WC (1997) *Systematic Biology*, **46(1)**, 1-68.

Whiting MF (2002) *Zoologica Scripta*, **31(1)**, 3-15.

Wilding N, Mardell SK, Brookes CP and Loxdale HD (1993) *Journal of Invertebrate Pathology*, **62(3)**, 268-272.

Williams J, Kubelik A, Livac K, Rafalski J and Tingey S (1990) *Nucleic acids Research*, **18**, 6531-6535.

Willis LG, Winston ML and Honda BM (1992) *Molecular Phylogenetics and Evolution*, **1(3)**, 169–78.

Winter P and Kahl G (1995) *World Journal of Microbiology and Biotechnology*, **11(4)**, 438-448.

Wolfe M (1985) *Annual Review of Phytopathology*, **23(1)**, 251-273.

Wolfe AD and Liston A (1998) *Molecular Systematics of Plants II: DNA Sequencing*, 43–86.

Wolfe AD, Xiang QY and Kephart SR (1998) *Molecular Ecology*, **7**, 1107–1125.

Wolf M, Achtziger M, Schultz J, Dandeker T and Muller T (2005) *Rna*, **11**, 1616–1623.

Wolstenholme DR (1992) *Animal mitochondrial DNA: structure and evolution*. In: *International Review of Cytology* (eds Wolstenholme DR, Jeon KW), pp. 173-216.

Wu KM, Lu YH, Feng HQ, Jiang YY and Zhao JZ (2008) *Science*, **5896**, 1676–1678.

Wunn J, Kloti A, Burkhardt PK, Biswas GCG, Launis K, Iglesias VA and Potrykus I (1996) *Nature Biotechnology*, **14(2)**, 171-176.

Xie L-X, Xu P-L, Nie Y-F and Tian Y-C (2003) *Chinese Journal of Biotechnology*, **19(5)**, 545-550.

Xiong B and Kocher TD (1991) *Genome*, **34**, 306–311.

Xiong B and TD Kocher (1993) *Molecular Phylogeny and Evolution*, **2**, 293-303.

Xiao-feng C, Sheng-jiang TAN, Ren-yi LIU, Ying W and Dian-mo L (2000) *Entomologia Sinica*, **7(3)**, 243-249.

Xie Q, Tian Y, Zheng L and Bu W (2008) *Molecular Phylogenetics and Evolution*, **47**, 463–471.

Xu D, Xue Q, McElroy D, Mawal Y, Hilder VA and Wu R (1996) *Molecular Breeding*, **2**, 167–173.

Yan XH, De Bondt HL, Powell CC, Bullock RC and Borovsky D (1999) *European Journal of Biochemistry*, **262**, 627–636.

Yang Z, Nielsen R, Goldman N and Pedersen AM (2000) *Genetics*, **155**, 431–449.

Yahia H, Ready PD, Hamdani A, Testa JM and Guessous-Idrissi N (2004) *Parasite*, **11**, 189-199.

Yebakima A, Charles C, Mousson L, Vazeille M, Yp- Tcha MM and Failloux AB (2004) *Tropical Medicine & International Health*, **9(5)**, 582-587.

Yeh L-CC and Lee JC (1990) *Journal of Molecular Biology*, **211**, 699- 712.

Yeh KW, Lin ML, Tuan SJ, Chen YM, Lin CY and Kao SS (1997) *Plant Cell Reporter*, **16**, 696-699.

Yli-Mattila T, Paavanen-Huhtala S, Fenton B and Tuovinen T (2000) *Experimental & applied acarology*, **24**, 10-11, 863-880.

Yu HL, Li YH and Wu KM (2011) *Journal of Integrative Plant Biology*, **53(7)**, 520-538.

Zhang D-X and Hewitt GM (1997) *Insect Molecular Biology*, **6**, 143–150.

Zhang D-X and Hewitt GM (2003) *Molecular Ecology*, 12, 563 /584.

Zhang DX (2004) *Trends in Ecology & Evolution*, **19(10)**, 507-509.

Zietkiewicz E, Rafalski A and Labuda D (1994) *Genomics*, **20**, 176–183.

CHAPTER 3

Genetic diversity analysis of *O. longicollis* (Oliver) populations using RAPDs, ISSRs and AFLPs

SUMMARY

The banana pseudostem weevil (O*doiporus longicollis* Oliver) is a serious pest of bananas in India and south-east Asia. In the present work, RAPDs, ISSRs and AFLPs have been used to assess the genetic diversity within and among six Indian populations of this pest. The data generated has been analysed phylogenetically and statistically to determine the utility of each of these marker systems in genetic diversity analysis. The RAPDs and ISSRs based UPGMA dendrograms did not reveal any phylogeographic clustering of the populations. The AFLP based dendrogram showed a strong correlation between geographic and genetic distance, suggesting that AFLPs are more useful than RAPDs and ISSRs in the present genetic diversity study.

INTRODUCTION

To date, there are no reports on any molecular genetic studies on banana pseudostem weevil. In the present work, the potential of RAPDs, ISSRs and AFLPs in estimating genetic relationships between six populations of stem weevils has been compared. These PCR based markers require no prior sequence information and are especially useful in systems where little or no molecular genetics information is available which is true in case of the banana stem weevil. These markers offer the advantage of being able to evaluate a large number of loci across the genome and have been used to study genetic diversity in several insect populations (Krumm *et al*. 2008; Sharma 2009; Jain *et al*. 2010; Kar *et al*. 2010; Ranjan *et al*. 2011). Examples of the use of RAPD markers in genetic diversity studies of Coleoptera include the work by Abdallah *et al*. (2012) (*Orycetes agamemmon* Burmeister), Bas *et al*. (2000) (*Diaprepes abbreviates*), Kim & Sappington (2004) and Martins *et al*. (2007) (*Anthonomus grandis* Boheman), Sidorenko *et al*. (2000) (Colorado potato beetle, *Leptinotarsa decemlineata*), Sidorenko & Berezovska (2002) (*Leptinotarsa decemlineata*) and Ibrahim *et al*. (2009) (three scarab beetles). Magana *et al*. (2007) used RAPD markers for the study of population structure of banana weevil (*Cosmopolites sordidus*), another pest of bananas, and determined the source of introduction of this pest in the Canary Islands.

Although ISSR markers have been used for the study of genetic diversity of various insect orders which has been described in Chapter 2 of this thesis, there are limited references available for the order Curculionidae. The genetic diversity of populations of *Zabrotes subfasciatus* was evaluated using SSR molecular markers. Twelve populations from eight Brazilian states were sampled and five ISSR primers were used and a total of 51 polymorphic bands were obtained. The genetic differentiation in Brazilian populations of *Z. subfasciatus* was found to be low and there was a weak geographic structuring (de Souza *et al*. 2008). There are few references available on the use of AFLP markers for Coleoptera. Johan de Graaf (2006) used AFLP for the study of population genetics of banana root corm weevil (*Cosmopolites sordidus*) for intended use in integrated pest management. Population genetics of bark beetle (Allender *et al*. 2008), flea beetle (Breuker *et al*. 2005), dung beetle (Carisio *et al*. 2004), colorado potato beetle

(Hawthrone 2001), European pollen beetle (Kazachkova *et al*. 2007, 2008), red flour beetle (Zhong *et al*. 2004) have also been studied using AFLPs.

In the absence of any sequence information on the genome of this pest, RAPDs, ISSRs and AFLPs are ideal markers and have been used in the present work to study the genetic diversity of six Indian populations of banana pseudostem weevil. This study showed that AFLPs are more useful in deciphering meaningful estimates of genetic diversity among these six populations as compared with RAPDs and ISSRs.

MATERIALS AND METHODS

Sample collection

A total of thirty adult beetles of *Odoiporus longicollis* (Oliver) *i.e.* five adults from each locations in India were collected from infested banana stems from banana fields in Assam (Kamrup), Bihar (Vaishali), Kerala (Wayanad), Maharashtra (Jalgaon, Narayangaon) and Tamil Nadu (Trichy) (**Table 1** and **Fig. 1**). In each location, the five individuals were collected within an area of 500 m.

Table 1. Geographical coordinates of sampled locations

Location	**Latitude**	**Longitude**
Kamrup (Assam)	26.080287° N	91.5596° E
Jalgaon (Maharashtra)	21.013321° N	75.564° E
Narayangaon (Maharashtra)	19.117429° N	73.9723° E
Trichy (Tamilnadu)	16.8222° N	78.7047° E
Vaishali (Bihar)	26.003818° N	85.0818° E
Wayand (Kerala)	11.709446° N	76.0955° E

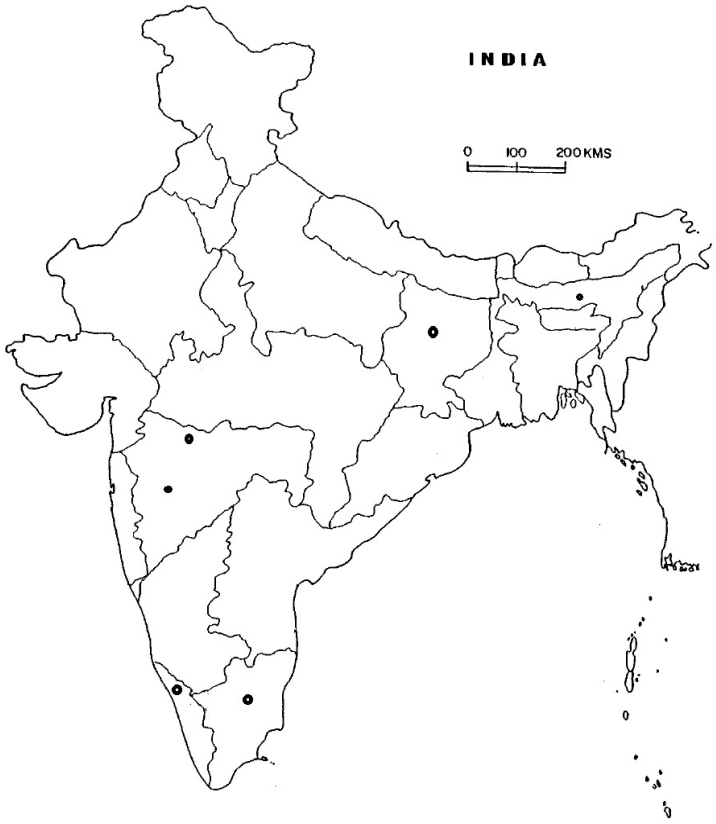


Fig. 1. Map of India showing locations from where the weevil samples were collected. Labels represent: K- Kamrup (Assam); J- Jalgaon and N- Narayangaon (Maharashtra); T- Trichy (Tamilnadu); V- Vaishali (Bihar); W- Wayanad (Kerala). Name of the state is indicated in parenthesis.

DNA extraction

Total genomic DNA was extracted from individual beetles using a modified Dellaporta *et al*. (1983) protocol as follows: The entire beetle was crushed to a fine powder in liquid nitrogen. The powder was suspended for 15 min at 65^0C in pre-warmed extraction buffer (100 mM of Tris (pH8), 50 mM of EDTA 500 mM NaCl, 1.25% SDS, 10 mM β-mercaptoethanoll). This was followed by the addition of 400 µl of chilled 5M potassium acetate per ml of the suspension. The suspension was then kept on ice for 1 h followed by centrifugation at 10,000 rpm for 10 min at 10^0C. Equal volume of chloroform: isoamyl alcohol mix was added to the supernatant followed by centrifugation at 10,000 rpm for 10 min at 10^0C. The DNA was precipitated from the aqueous layer by the addition of 750 µl of

isopropyl alcohol per ml of the aqueous layer. The tubes were kept overnight at room temperature to allow the DNA to precipitate. The precipitated DNA was collected by centrifugation at 10,000 rpm for 10 min at 10^0C and given a wash with 70% chilled alcohol. The DNA was air dried and dissolved in $T_{10}E_1$ buffer (pH 8.0). RNase treatment was given to the DNA solution (50µg/ml) and the solution was deproteinised with equal volume of phenol: (chloroform: isoamyl alcohol) and precipitated as described above. The DNA was dissolved in $T_{10}E_1$ (pH 8.0). The concentration of the extracted DNA was determined by agarose gel electrophoresis using λ DNA of known concentration. Five individuals from each location were used in the present work.

PCR amplification using RAPDs and ISSRs

A total of sixty-eight decanucleotide Operon RAPD primers of the series A, B, G (20 primers per series) and E (8 primers) and twenty two ISSR primers ie UBC Nos. 808, 811, 815, 836, 840, 845, 848, 852, 856, 859, 881, 887, 809, 810, 812, 814, 817 and HB12, HB15, 17901, 17899A, 17899B were initially screened for identifying primers that would give clear amplification products for the present work. Of these primers, nine each of RAPD and UBC primers were selected for inter and intra-population analysis (**Table 2 and Figs. 2 a, b and c**).

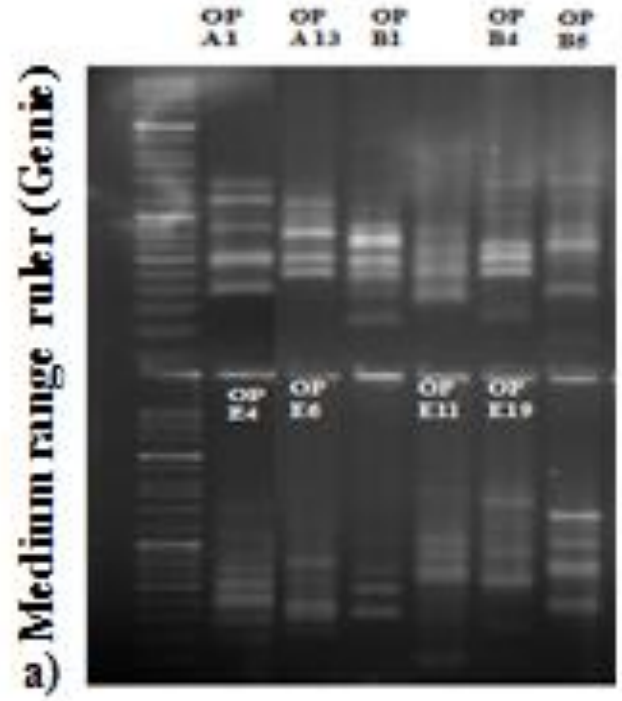
Table 2. The list of RAPD and ISSR primers used in the present study, the number and size range of bands amplified and Resolving power (Rp), PIC (Polymorphism Information Content), **EMR (effective multiplex ratio) and MI (Marker Index) of the primers**

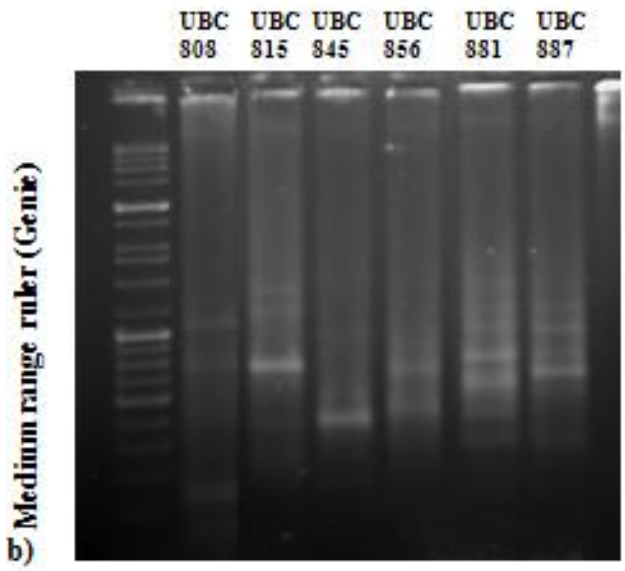
RAPD /ISSR primers	No. of bands studied	Range of fragment sizes (bp)	Resolving power (Rp)	Average of informativeness bands (AvIb)	PIC	EMR	MI
OPA-01	7	250-1200	3.712	0.53	0.051	3.76	0.191
OPA-13	10	450-1050	2.861	0.286	0.016	3.59	0.057
OPB-01	11	250-900	4.53	0.411	0.030	5	0.15
OPB-04	19	250-1500	5.258	0.276	0.010	5.23	0.052
OPB-05	10	250-1300	3.268	0.326	0.025	4.16	0.104
OPE-04	11	250-900	7.668	0.697	0.039	4.83	1.88
OPE-06	16	250-1000	6.996	0.437	0.016	5.3	0.0848

OPE-11	11	150-1185	3.86	0.350	0.033	2.33	0.076
OPE-19	17	150-1500	5.25	0.308	0.013	2.6	0.033
Average					**0.026**	**4.09**	**0.292**
UBC-815	4	450-700	1.87	0.467	0.0825	3.06	0.252
UBC-856	9	100-1050	3.252	0.361	0.027	5.92	0.159
UBC-808	8	100-1000	4.87	0.608	0.0375	4.43	0.166
UBC 881	6	250-1100	3.132	0.522	0.0421	2.770	0.116
UBC-845	9	100-850	5.866	0.651	0.026	3.611	0.093
UBC 887	6	350-1000	4.068	0.678	0.044	2.915	0.128
UBC-812	7	200-700	2.116	0.302	0.034	4.236	0.144
HB12	13	200-1200	8.666	0.666	0.025	5.907	0.147
UBC-809	9	200-1050	4.866	0.540	0.031	3.910	0.024
Average					**0.039**	**4.084**	**0.136**

The RAPD PCR reactions were carried out in a total volume of 25 µl containing 25 ng of genomic DNA, 0.6 U of Taq polymerase (Genie, India), 1X PCR buffer (200 mM tris-HCL pH 8.4 with 15 mM $MgCl_2$), 50 pmoles of primers, 10 mM spermidine and 0.25 mM dNTP mix. The PCR cycle conditions for RAPD-PCR were as follows: initial denaturation at 94^0C for 5 min; 35 cycles each of 94^0C for 30 s; 28^0C/40^0C for 1 min; 72^0C for 2 min; final extension at 72^0C for 10 min. The ISSR PCR reactions were carried out in a total volume of 25 µl containing 25 ng of genomic DNA, 2 U of Taq polymerase (Genie), 1X PCR buffer (200 mM tris-HCL pH 8.4 with 15 mM $MgCl_2$), 25 pmoles of primers and 0.25 mM dNTP mix. The PCR cycle conditions for ISSR-PCR were: initial denaturation at 94^0C for 5 min; 45 cycles each of 94^0C for 1min; 47^0C/48^0C for 45 sec; 72^0C for 2 min; final extension at 72^0C for 5 min.

The PCR products were resolved on a 1.5 % agarose gel, stained with ethidium bromide (0.5 µg/ml) and viewed on a UV transilluminater and photographed using the Gel Documentation System (Syngene, G Box EF).


OP A1
OP A13
OP B1
OP B4
OP B6
OP E4
OP E6
OP E11
OP E19
a) Medium range ruler (Genie)


UBC 808
UBC 815
UBC 845
UBC 856
UBC 881
UBC 887
b) Medium range ruler (Genie)

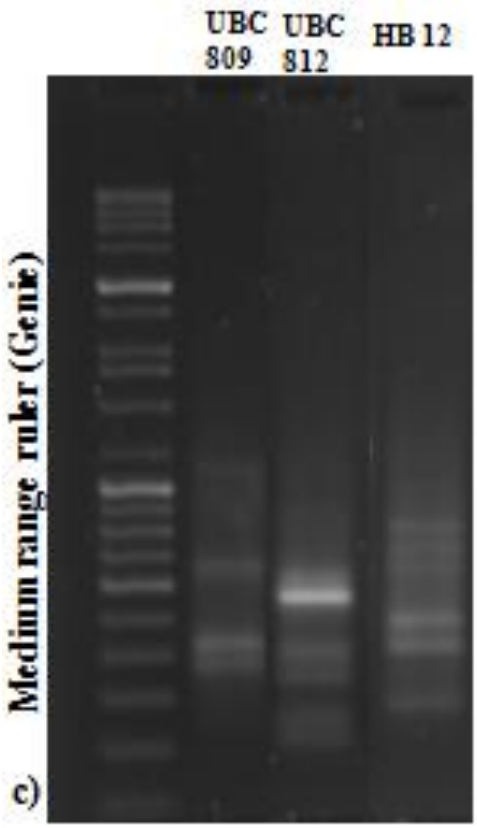


Fig. 2 a, b and c. Amplification patterns of gDNA of *O. longicollis* (Oliver) from Narayangaon with nine RAPD and nine ISSR Primers used in the present analysis

RAPD and ISSR data scoring and analysis

Binary scoring of the profiles was done on the basis of presence or absence of a band at a particular locus and matrix was generated with 1 (present) or 0 (absent) for each band. Data generated using both the markers were analysed in the following manner Diversity parameters such as the percentage of polymorphism, heterozygosity (h), number of alleles (Na), number of effective alleles (Ne),Shannon's information index (I), gene flow estimation (N*m*), coefficient of gene differentiation (G_{st}) and Nei's genetic distance were calculated according to Nei (Nei 1973, 1978). Nei's gene identity (I), Nei's genetic distance (D) and Nei's average gene diversity (H) were computed according to Nevo

(1978) and Nei (1972, 1973). The above analysis was performed using POPGENE Version 1.3 (Yeh 1998) and GenAlEx 6.4 (Peakall & Smouse 2006) software. G_{ST} values generated by POPGENE were estimated for the whole population and were expressed as the gene diversity of the single population subtracted from the gene diversity of the total population divided by the gene diversity of the total population (GST = HT – HS / HT). Gene flow was estimated from G_{ST} values and was expressed as $(Nm) = 0.5(1 – G_{ST})/G_{ST}$ (McDermott & McDonald 1993). The confidence of clustering in the dendrogram was evaluated through bootstrap analysis using the FreeTree software (available at http://www.natur.cuni.cz/~flegr/freetree.htm) with 500 replications. The significance of correlation between the matrices was tested using the normalized Mantel *Z*-statistics (Mantel 1967). Geographical distances of pairs of populations were calculated using the latitude, longitude, and elevation of each population. The Mantel test was used to find out the correlation between geographical distances and genetic distances (Mantel 1967). The Principal Coordinate Analysis (PCoA) and Analysis of Molecular Variance (AMOVA) was carried out using GenAlEx software v 6.1 (Peakall & Smouse 2006). For AMOVA analysis, total variance of the ISSR and RAPD data set was partitioned at three hierarchical levels: (1) an among-population component; (2) a regional or six population component; (3) and a within-population component. Expected heterozygosity, average heterozygosity (H_{av}) and Marker index (MI) were calculated according to Nei (1978) and Powell *et al*. (1996). Expected heterozygosity Hn for a genetic marker was calculated from the sum of the squares of allele frequencies as Hn = 1– ($p2$+ $q2$), where p is the frequency of individuals having the allele and q is the frequency of individuals lacking the allele. The arithmetic mean heterozygosity H_{av} for each marker class was calculated as the arithmetic mean heterozygosity. H_{av} was calculated for each marker class as, $H_{av} = \sum Hn/n$, where n is the number of markers or loci analyzed (Powell *et al*. 1996). The average heterozygosity for polymorphic markers (Hav)p was calculated as $(H_{av})p = \sum Hn/np$, where np is the number of polymorphic markers or loci (Powell *et al*. 1996). Banding characteristics for primers such as number of total bands (TB), number of polymorphic bands (PB) and percentage of polymorphic bands (PPB) were obtained. The suitability of each of the markers systems for suitability in evaluation of genetic diversity of *O. longicollis*, three parameters were caculated. These parameters were: polymorphic

information content (PIC), marker index (MI) and resolving power (RP). The PIC value for each locus obtained from the primer used for amplification was calculated using the formula of Roldan-Ruiz *et al.* (2000); PICi = 2fi (1 - fi), where PICi is the polymorphic information content of the locus i, fi is the frequency of the amplified fragments and 1 - fi is the frequency of nonamplified fragments. The frequency was calculated as the ratio between the number of amplified fragments at each locus and the total number of individuals. The PIC of each primer was calculated using the average PIC value from all loci of each primer. Effective multiplex ratio was calculated using formula; EMR (effective multiplex ratio) = n X β, where n is the average number of fragments amplified by individuals to a specific system marker (multiplex ratio) and β is estimated from the number of polymorphic loci (PB) and (MB) the number of nonpolymorphic loci; β = PB/(PB + MB). The capacity of each primer to detect polymorphic loci among the genotypes was characterized using Marker index for each of the markers. Marker index for each primers is calculated as a product of polymorphic information content and effective multiplex ratio (Varshney *et al.* 2007); MI = EMR X PIC. The resolving power (RP) of each primer was calculated as (Prevost & Wilkinson 1999): RP = RIb, Where Ib represents the informative fragments. The Ib can be represented on a scale of 0/1 by the following formula; Ib = 1 - (2 X |0.5 - pi|), where pi is the proportion of individuals containing the ith band.

Amplified Fragment Length Polymorphism (AFLP) analysis

The AFLP protocol consists of the restriction of the DNA, ligation of the adapters and two-step amplification (pre-amplification and selective amplification) (Vos *et al*. 1995). From each location, DNA from the five individuals was pooled for AFLP analysis. An initial screening using 60 pre-selective primer combinations were performed on the pooled DNA from six locations. Five selective primer combinations (**Table 9a**) giving clear and reproducible electrophoretic patterns and showing variation within and between populations were chosen for further analyses. The AFLP adapter sequences, pre-amplification primer sequences and selective amplification primer sequences are reported in **Table 9a**. Six PCR primer combinations (*Eco*RI–*Mse*I) with three selective nucleotides

each were used: EcoRI-C + MSeI-CA, EcoRI-C + MSeI-CC, EcoRI-C + MSeI-CG, EcoRI-A + MSeI-CC, EcoRI G JOE- MseI CA, EcoRI G JOE- MseI CC. Selective amplification products were separated with an Applied Biosystems automated sequencer. Fragments were scored automatically by the program Genemapper v3.7. For each combination, all loci showing a clear and unambiguous banding pattern were scored, whereas uncertain peaks were considered as missing data.

RESULTS

RAPD analysis

Genomic DNA was initially screened for amplification with 22 RAPD primers belonging to the Operon A, B, G and E sets to identify primers that would be useful for genetic diversity studies. Of the screened primers, nine primers exhibiting high polymorphism were selected for the population diversity analysis (**Table 2 and Fig. 2a**). The number of bands produced by individual primers ranged from 7 (OPA1) to 19 (OPB4) with an average of 12.5. The size of the polymorphic bands ranged from 100 bp to 1600 bp. The most polymorphic primer was OPB4 which generated 19 bands. The band informativeness (Ib) of each band amplified by each of the primers was determined in order to determine the resolving power of each primer. The resolving power of a primer (Rp) is the sum of band informativeness ie the sum of Ib of all the bands amplified by that primer *i.e*. Rp=ΣIb (Guasmi *et al*. 2012). The resolving power (Rp) of the primers screened ranged from 2.861 (OPA 13) to 7.668 (OPE 4) with a mean of 4.82 (**Table 2**). These nine primers amplified a total of 113 clear and discernible bands among the six populations of which 112 (99%) were polymorphic (**Table 3**).

Table 3. Estimates of number and proportion of polymorphic loci, number of observed genetic types (Na), effective number of alleles (Ne), Nei's gene diversity (Hs) and Shannon's Information index (I) for six populations of *O. longicollis* (Oliver) estimated by nine RAPD markers (Values in brackets are standard deviation values.)

Population	Na	Ne	Number of polymorphic loci	Proportion of polymorphic loci (%)	Shannon's Information index (I)	Hs
Assam	1.3982 (0.4917)	1.2715 (0.3645)	45	39.082	0.2314 (0.2908)	0.1572 (0.0402)
Jalgaon	1.4336 (0.4978)	1.3002 (0.3757)	49	43.36	0.2536 (0.2967)	0.1727 (0.0421)
Narayangaon	1.2124 (0.4108)	1.1400 (0.2898)	24	21.24	0.1216 (0.2384)	0.0821 (0.0266)
Trichy	1.3717 (0.4854)	1.2270 (0.3229)	42	37.17	0.2058 (0.2732)	0.1373 (0.342)
Vaishali	1.3805 (0.8477)	1.2271 (0.3171)	43	38.05	0.2087 (0.2718)	0.1388 (0.0336)
Wayanad	1.1062 (0.3095)	1.0780 (0.2388)	12	10.62	0.0638 (0.1881)	0.0439 (0.0170)

The percentage polymorphism per population varied from 10.62% (Wayanad) to 43.36% (Jalgaon), with an overall average polymorphism of 25.34% among the six populations. The population diversity analysis using POPGENE revealed that the average number of observed alleles varied from 1.1 in Wayanad to 1.43 in Jalgaon while the effective number of alleles varied from 1.07 (Wayanad) to 1.30 (Jalgaon). The intra-population genetic diversity was highest in Jalgaon (0.1727) and lowest in Wayanad (0.0439) with an average gene diversity within subpopulations (Hs) of 0.1220. The total genetic diversity (Ht) was 0.2682. The Shannon's information index varied from 0.0638 (Wayanad) to 0.2536 (Jalgaon) (Table 2). The pair-wise comparison of the populations showed that the genetic distance varied from 0.0948 between Assam and Narayangaon to 0.3441 between Jalgaon and Wayanad (**Table 4**).

Table 4. Pair-wise Nei's (1972) genetic identity (above diagonal) and genetic distance (below diagonal) between the six populations of *O. longicollis* (Oliver) using RAPD primers

Population	Assam	Jalgaon	Narayangaon	Trichy	Vaishali	Wayanad
Assam	****	0.7516	**0.9096**	0.8261	0.7981	0.7930
Jalgaon	0.2855	****	0.7174	0.7191	0.8626	**0.7089**
Narayangaon	**0.0948**	0.3321	****	0.8406	0.7987	0.8033
Trichy	0.1910	0.3298	0.1737	****	00.7959	0.8980
Vaishali	0.2255	0.1477	0.2248	0.2283	****	0.7845
Wayanad	0.2319	**0.3441**	0.2190	0.1075	0.2427	****

*The highest and lowest genetic distances and identities are indicated in bold.

The total genetic differentiation coefficient (G_{ST}) among the populations was 0.5450. Gene flow among the populations was 0.4174. There was a poor correlation between genetic distance and geographic distance (**Fig. 3**).

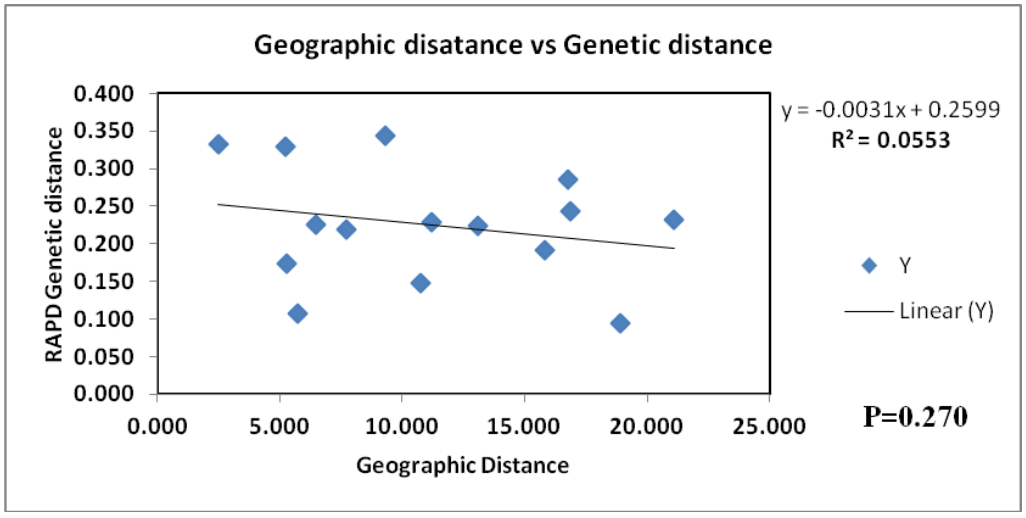


Fig. 3. Mantel test for correlation between genetic and geographical distance for RAPD

AMOVA analysis and a PhiPT value of 0.487 which was significant at a probability of 0.05 revealed that almost half (51%) of the variation occurred within the populations ie 49% of the variation accounted for the inter-population genetic variability (**Table 5**).

Table 5. AMOVA for RAPDs

Source	**df**	**SS**	**MS**	**Estimated Variance**	**% variation**
Among Pops	5	247.733	49.547	8.186	49%
Within Pops	24	206.800	8.617	8.617	51%
Total	29	454.533		16.803	100%

df : degree of freedom; SS : sum of squares; MS : mean sum of squares.

Based on the distance matrix expressed as similarity coefficients, an UPGMA dendrogram was generated (**Fig. 4**). The dendrogram clustered all individuals into two major groups and within each group individuals from each location formed a sub-cluster. Group I had individuals from Jalgaon and Vaishali locations while group II clustered individuals from the rest of the four locations *i.e* Narayangaon, Assam, Trichy and Wayanad.

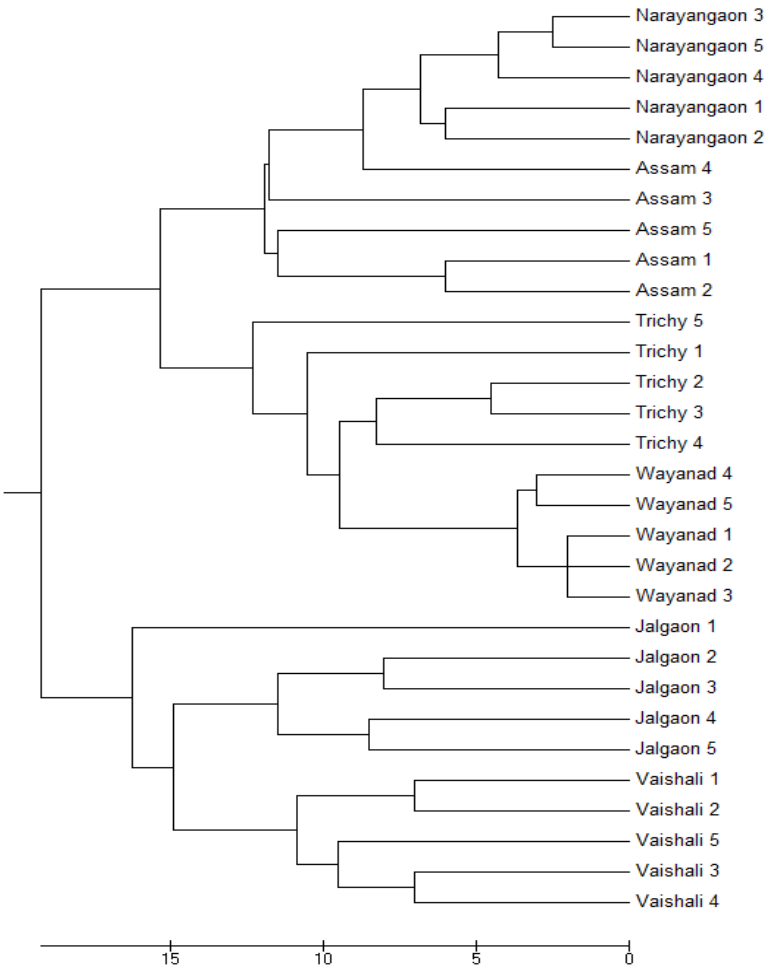


Fig. 4. UPGMA dendrogram derived from RAPDs data

The PCA ordination based on the first two components (factors) confirmed cluster analysis with the Eigen values for first and second components being 23.27% and 15.86% respectively. The first two components of PCOA explained 39.13% of the total variation, and the first three components explained 46.82% of the total variation (**Fig. 5**).

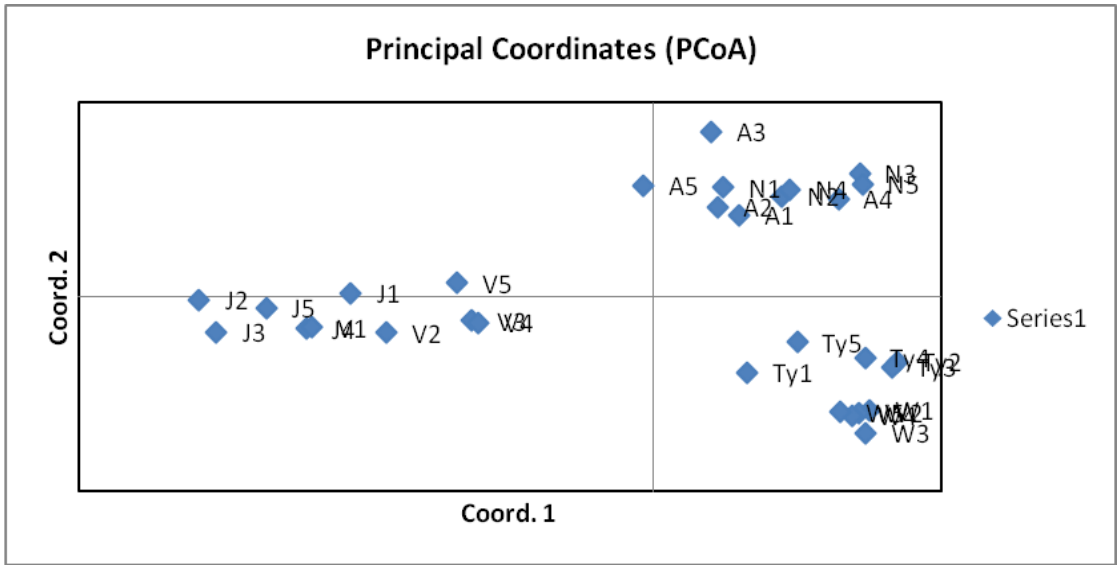


Fig. 5. PCoA of the RAPDs generated data

ISSR analysis

Seventeen UBC primers were initially screened to identify primers that would give clear amplification products. Of these, nine primers gave clear and discernible bands and hence were chosen for the genetic diversity studies (**Table 2, Figs 2b and c**). The size of the amplified bands varied from ~100 to 1200 bp. The number of bands produced by individual primers was in the range of 4 (UBC815) to 13 (HB12) with an average of 11.8. Among all populations, primer HB12 was the most polymorphic which generated 13 polymorphic bands. The resolving power (Rp) of the primers screened ranged from 1.87 (UBC 815) to 8.666 (HB 12) with a mean of 4.29 (**Table 2**). A total of seventy-one bands were amplified by the nine ISSR primers in the six populations of which 61 (85.91%) were polymorphic (**Table 6**).

Table 6. Estimates of number and proportion of polymorphic loci, number of observed genetic types (Na), effective number of alleles (Ne), Nei's gene diversity (Hs) and Shannon's Information index (I) for six populations of *O. longicollis* (Oliver) estimated by nine ISSR markers

(Values in brackets are standard deviation values.)

Population	Na	Ne	Number of polymorphic loci	Proportion of polymorphic loci (%)	Shannon's Information index (I)	Hs
Assam	1.6338 (0.4852)	1.5086 (0.4211)	45	63.38	0.3974 (0.3103)	0.2772 (0.0482)
Jalgaon	1.3944 (0.4992)	1.2748 (0.3716)	28	39.44	0.2314 (0.2939)	0.1577 (0.0413)
Narayangaon	1.5211 (0.5281)	1.3727 (0.3954)	37	52.11	0.3094 (0.3052)	0.2118 (0.0452)
Trichy	1.3521 (0.4810)	1.2167 (0.3218)	25	35.22	0.1956 (0.2718)	0.1307 (0.0339)
Vaishali	1.4789 (0.5031)	1.3464 (0.3962)	37	47.89	0.2858 (0.3062)	0.1961 (0.0455)
Wayanad	1.2535 (0.4381)	1.1512 (0.2807)	18	25.35	0.1390 (0.2434)	0.0924 (0.0268)

The percentage polymorphism per population varied from 25.35% (Wanayad) to 63.38% (Assam), with an average polymorphism of 43.88 among the six populations (**Table 6**). Population diversity analysis using POPGENE revealed that the average number of observed alleles varied from 1.25 in Wayanad to 1.63 in Assam while the effective number of alleles varied from 1.15 (Wayanad) to 1.50 (Assam). The heterozygosity present within a population was highest in Assam (0.272) and lowest in Wayanad (0.092) (**Tables 6**). The total genetic diversity (Ht) was 0.3208 while the average Hs was 0.177. The Shannon's information index varied from 0.139 (Wayanad) to 0.397 (Assam) (**Table 6**). The total genetic differentiation coefficient (G_{ST}) among the populations was 0.4461 and gene flow (Nm) among the populations was 0.6207. Pair-wise comparison of the populations showed that the genetic distances varied from 0.1377 between Assam and Trichy to 0.3718 between

Jalgaon and Wayanad (**Table 7**). There was a weak correlation between genetic distance and geographic distance (**Fig. 6**).

Table 7. Pair-wise Nei's (1972) genetic identity (above diagonal) and genetic distance (below diagonal) between the six populations of *O. longicollis* (Oliver) using ISSR primers

Population	Assam	Jalgaon	Narayangaon	Trichy	Vaishali	Wayanad
Assam	****	0.8101	0.8196	0.8714	**0.8292**	0.7964
Jalgaon	0.2106	****	0.7973	0.7987	0.7826	**0.6895**
Narayangaon	0.1990	0.2266	****	0.7834	0.8231	0.7500
Trichy	**0.1377**	0.2248	0.2441	****	0.8269	0.7611
Vaishali	0.1873	0.2452	0.1947	0.1901	****	0.7677
Wayanad	0.2276	**0.3718**	0.2876	0.2730	0.2644	****

*****The highest and lowest genetic distances and identities are indicated in bold.

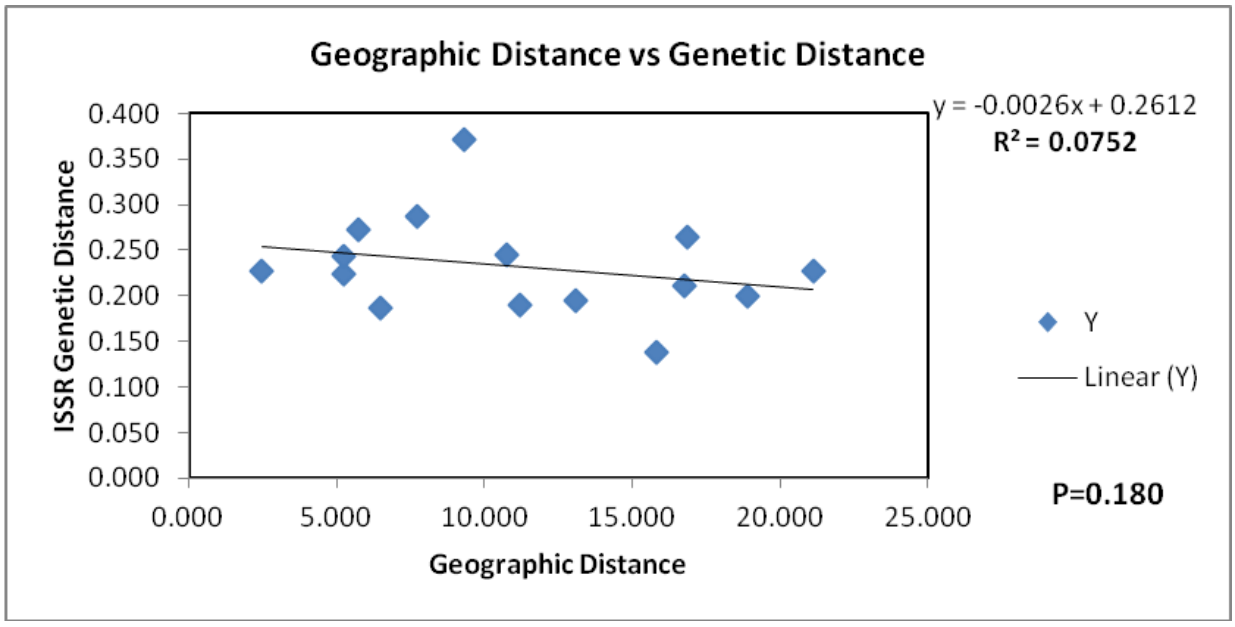


Fig. 6. Mantel test for correlation between genetic and geographical distance for ISSR

The UPGMA dendrogram generated based on the distance matrix clustered the thirty individuals into two groups – a minor group I having individuals 4 and 5 from Assam, and a major group II comprising the remaining twenty eight individuals. Within this major group, with the exception of individuals 1, 2 and 3 from Assam, all the individuals grouped as per the location from where they were collected (**Fig. 7**).

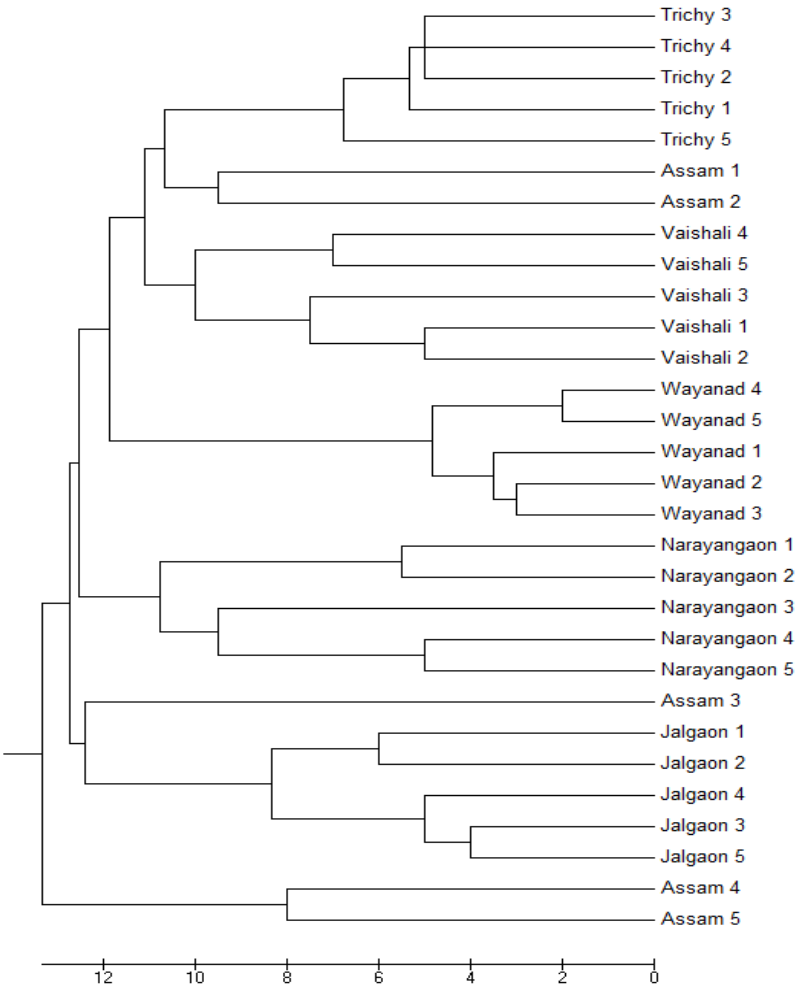


Fig. 7. UPGMA dendrogram derived from ISSRs data

The PCOA ordination based on the first two components (factors) confirmed this cluster analysis and the Eigen values for the first and second components were 15.86% and 10.98% respectively (**Fig. 8**).

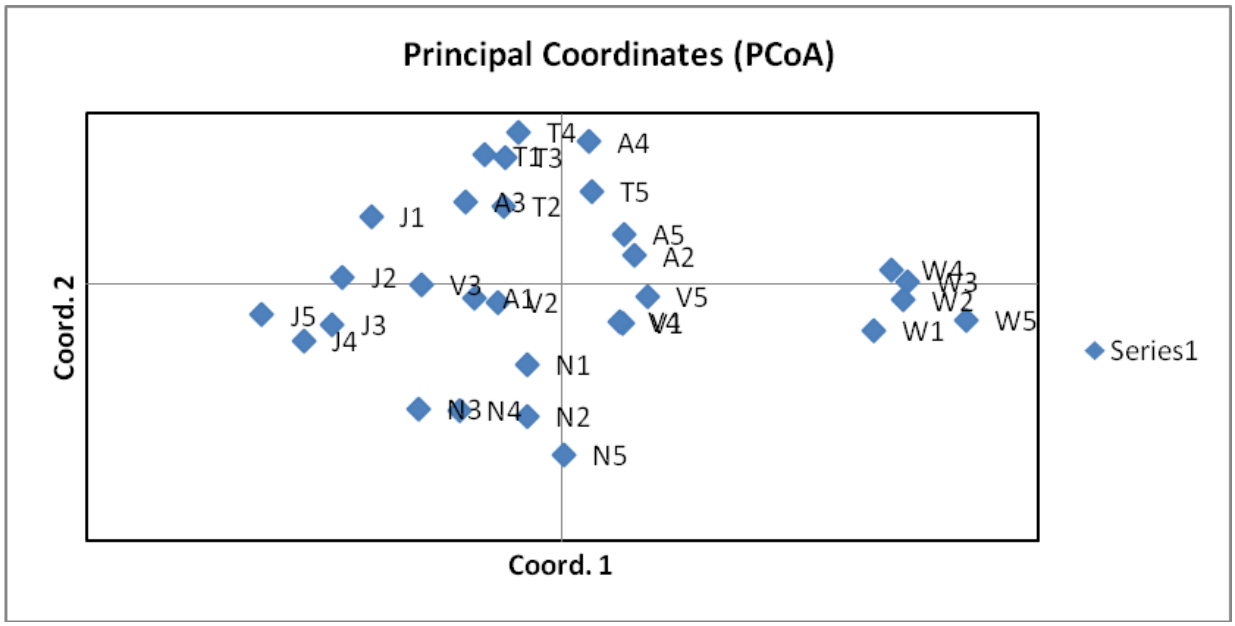


Fig. 8. PCoA of the ISSRs generated data

The first two components of PCOA explained 26.84% of the total variation, and the first three components explained 36.31% of the total variation. AMOVA revealed an inter-population genetic variability of 36% and a significant PhiPT value of 0.364 (p= 0.001), indicating that almost 64% of the variation could be attributed to variation within the populations (**Table 8**).

Table 8. AMOVA for ISSRs

Source	df	SS	MS	Est. Var.	%
Among Pops	5	152.400	30.480	4.519	36%
Within Pops	24	189.200	7.883	7.883	64%
Total	29	341.600		12.403	100%

df : degree of freedom; SS : sum of squares; MS : mean sum of squares.

Analysis based on the combined data i.e. RAPDs + ISSRs

The dendrogram generated using the distance matrix of the combined data (RAPDs + ISSRs) had a topography similar to that generated using the RAPD data set wherein the individuals clustered into two major groups and individuals from each location formed separate sub-groups (**Fig. 9**).

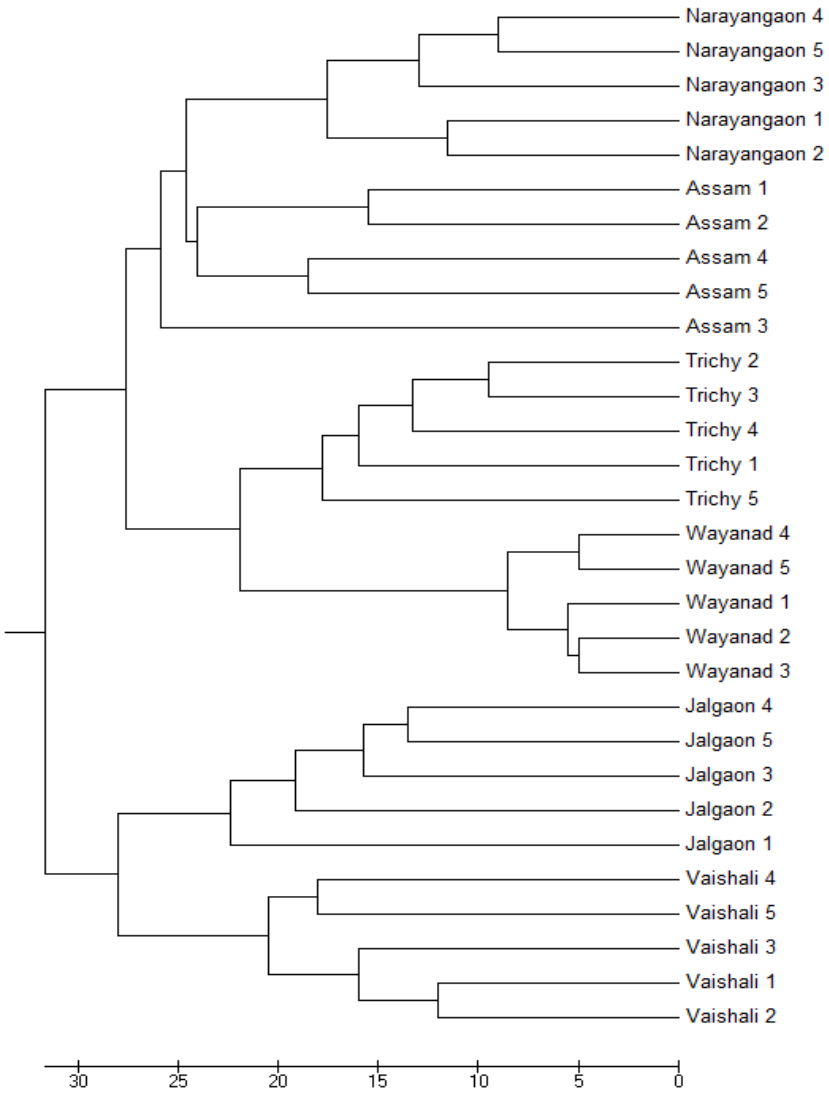


Fig. 9. UPGMA tree based on RAPD + ISSR data

AFLP analysis

Six AFLP primer combinations were used for studying the genetic diversity of the six *O. longicollis* populations (**Table 9a**).

Table 9a. Adapters and pre-amplification primers used in AFLP analysis of the six Indian populations of *O. longicollis* (Oliver)

Primer Pair + Selective Ext.		Number of Polymorphic Loci	Fragment Size Range (bp)	PIC	EMR	MI
EcoR-I	MSe-I					
C-NED	CA	6	74-261	0.341	1	0.341
C-NED	CC	19	69-225	0.009	3.16	0.0284
C-NED	CG	7	119-226	0.032	1.16	0.037
A-FAM	CC	7	81-89	0.05	1.015	0.005
G-JOE	CA	4	308	0.0285	0.66	0.188
G-JOE	CC	8	63-113	0.033	1.33	0.043
Total		50	63-308	--	-	-
Average/Primer Pair		9		0.0822	1.387	0.107

A total of 51 bands were amplified by the six primer combinations in the six populations of which 50 (98%) were polymorphic (**Table 9b**). The maximum no of bands (21) were amplified in the Assam population and the least number of bands (3) were present in the Jalgaon population. The size of the amplified bands varied from 63 bp to 308 bp (**Table 9b**). The A-FAM/MseI CC primer pair amplified a single monomorphic band in all the populations. In addition, this primer pair amplified a 89 bp band that was specific for the Wayanad population. The remaining five primer combinations did not amplify bands in a few populations. Population specific bands were amplified with four of the primer combinations. For eg. of the 21 bands amplified in the Assam population, 12 were specific for this population. Similarly a single band specific for the Narayangaon and Trichy populations were observed. No such population specific bands were observed for the Jalgaon and Vaishali

populations. The resolving power (Rp) of the primer pairs screened ranged from 0.32 to 4.92 with a mean of 2.07 (**Table 9b**).

Table 9b. Selective primers used for AFLP analysis, the number and size range of bands amplified as detected by GeneMapper and Resolving power (Rp) of the primers

Primer Combination									
EcoR-I	**MSe-I**	**Assam**	**Jalgaon**	**Narayangaon**	**Trichy**	**Vaishali**	**Wayanad**	**Resolving power (Rp)**	**Average band informativness (AvIb)**
C-NED	CA	3 (**74**, 229, 261)	0	0	1(261)	2 (229, 261)	0	1.98	0.33
C-NED	CC	7 (69, 102, **104**, 121, **125, 178,** 199)	1 (121)	2 (**70**, 121)	5(69, 102, 121, 199, **225**)	0	4 (69, 121, **183**, 199)	4.92	0.259
C-NED	CG	5 (**114**, 119, **148, 210,** 226)	0	0	1 (119)	1 (226)	0	1.94	0.277
A-FAM	CC	1 (89)	1 (89)	1 (89)	1 (89)	1 (89)	2 (**81**, 89)	0.32	0.045
G-JOE	CA	0	1 (308)	1 (308)	1 (308)	0	1 (308)	0.68	0.17
G-JOE	CC	6 (64, **66, 68, 79, 86, 113)**	0	0	1 (64)	0	1 (63)	2.6	3.25
Total No. of bands		**21**	**3**	**4**	**10**	**5**	**7**		

An UPGMA dendrogram which was generated based on the distance matrix showed high bootstrap values at the branch points (**Fig. 10).**

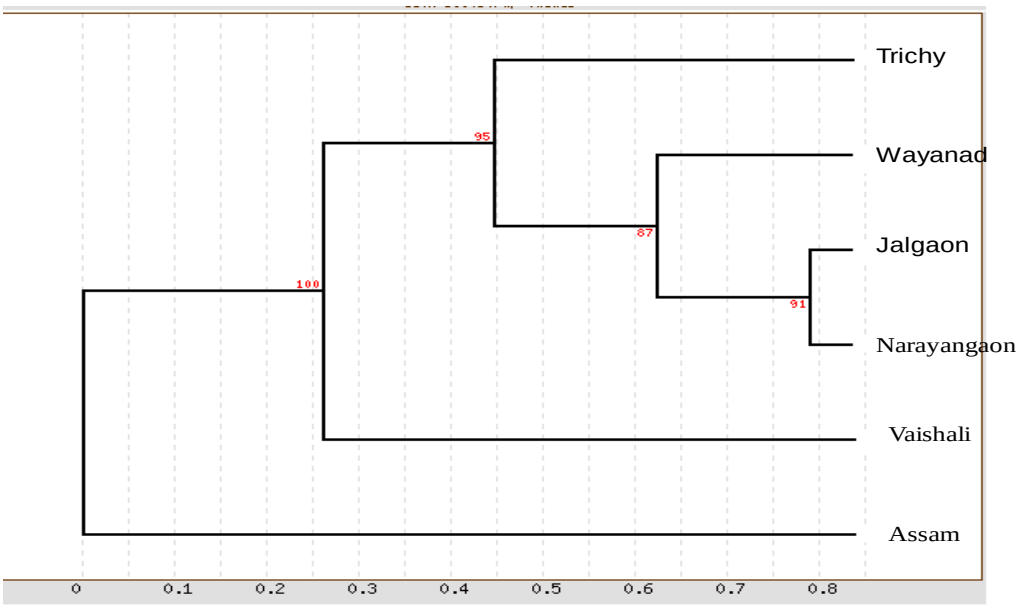
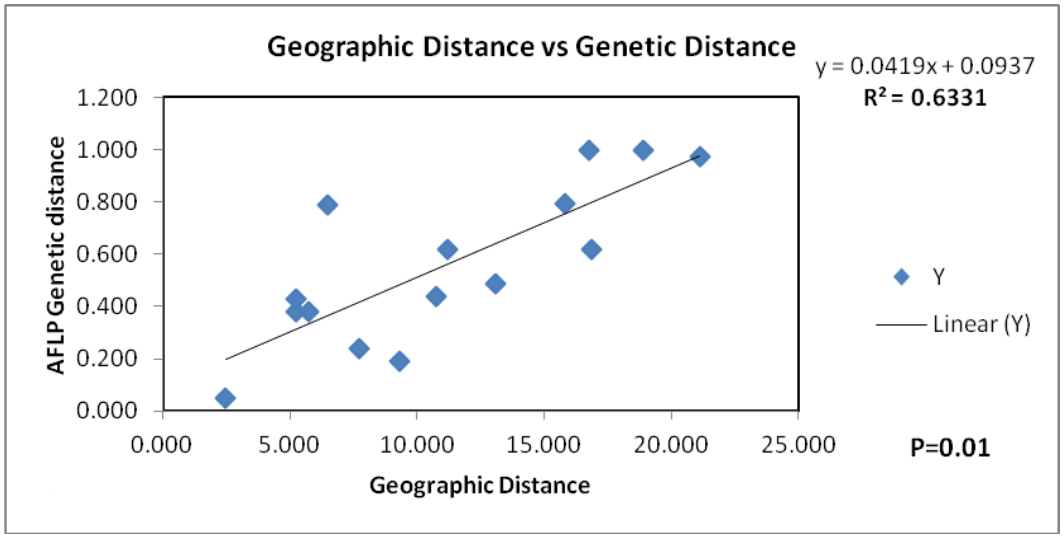


Fig. 10. UPGMA dendrogram derived from AFLPs data

The Assam population separates from the remaining populations at a genetic distance of ~ 26% while the Vaishali population separates from the remaining four populations at a genetic distance of ~ 45%. A strong correlation between genetic distance and geographic distance was observed (**Fig. 11**).



* P-value <0.05

Fig. 11. Mantel test for correlation between genetic and geographical distance for AFLP

DISCUSSION

Though *O. longicollis* (Oliver) is a serious pest in banana growing regions in the world, this pest has neither been characterised using molecular markers nor has the genetic diversity of pest populations been assessed. The present work was aimed to determine the level and distribution of genetic diversity among Indian populations of this pest collected from six hot-spot locations, using AFLP, RAPD and ISSR marker systems. The utility of each of these markers in the estimation of genetic diversity was done by comparing:

(i) different marker parameters (**Table 10).**

Table 10. Comparison of RAPDs, ISSRs and AFLPs in genetic diversity analysis of Indian populations of *O. longicollis* (Oliver)

Parameter	**RAPDs**	**ISSRs**	**AFLPs**
Total number of bands	113	71	51
Total number of polymorphic bands	112	61	50
% polymorphic bands	99%	85.91%	98%
Average number of bands /primer	12.5	11.8	9
Average number of polymorphic bands /primer	12.44	6.77	8.16
Rp	4.82	4.29	2.07
Total Heterozygosity (Ht)	0.2682	0.3208	-----
Average Heterozygosity (Hs)	0.1220	0.177	-----
Gst	0.5450	0.4461	----
Gene Flow (Nm)	0.4174	0.6207	----
H_{av}	0.2725	0.3686	0.3424
MI	0.292	0.136	0.107

(ii) the UPGMA dendrograms **(Figs 4, 7, 9 and 10)**

(iii) regression analysis (**Figs. 12 a-f**).

In the present study, the data for individual weevils was generated for RAPD and ISSR analysis. AFLP analysis was done using pooled DNA samples from the same thirty individuals used for RAPD and ISSR analysis as pooled samples have been reported to give consistent results in AFLP analysis (Katiyar *et al*. 2000). 112, 61 and 50 polymorphic bands were generated by RAPDs, ISSRs and AFLPs respectively, satisfying the general recommendation of having atleast 50 polymorphic bands that are required for estimating

precise genetic distances (Nei 1978). The H_{av} and Marker index (MI) values for these three marker systems are shown in **Table 11**.

Table 11. Heterozygosity and Marker Index analyses of polymorphisms revealed by RAPD and ISSR Markers in Indian Populations of *O. longicollis* (Oliver)

Marker	Number of primers	Total number of amplification products (*n*)	Number of polymorphic amplification products (*n*p)	Mean number of polymorphic products per primer (*n*pp)	$\sum H_p$	$H_{av(p)}$	β	H_{av}	MI
RAPD	9	113	112	12.44	30.524	0.2725	0.991	0.2700	0.292
ISSR	9	71	61	6.777	22.49	0.3686	0.859	0.3166	0.136
AFLP	6	51	59	3.11	9.247	0.3424	0.964	0.3300	0.107

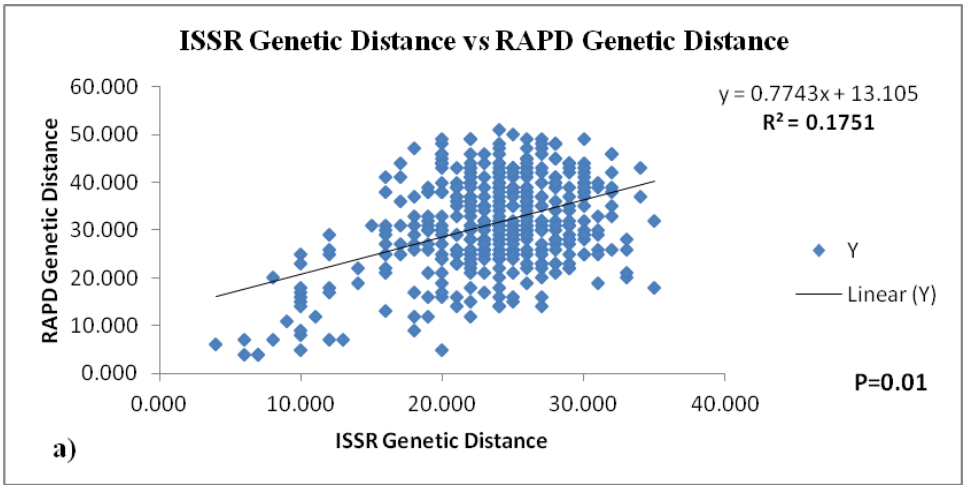
Values of different parameters which give quantitative estimates of the observed polymorphism and utility of the marker systems are shown in **Table 10**. Values of parameters such as average number of bands per primer, average number of polymorphic bands per primer, Rp and MI were higher for RAPDs than for ISSRs and AFLPs, but estimates for Shannon's Informative Index, Ht, Hs, Nm and H_{av} were higher for ISSRs than RAPDs. However, `t' test reveals no significant differences in Hav, PIC and MI values generated by each of these marker systems **(Table 12)**.

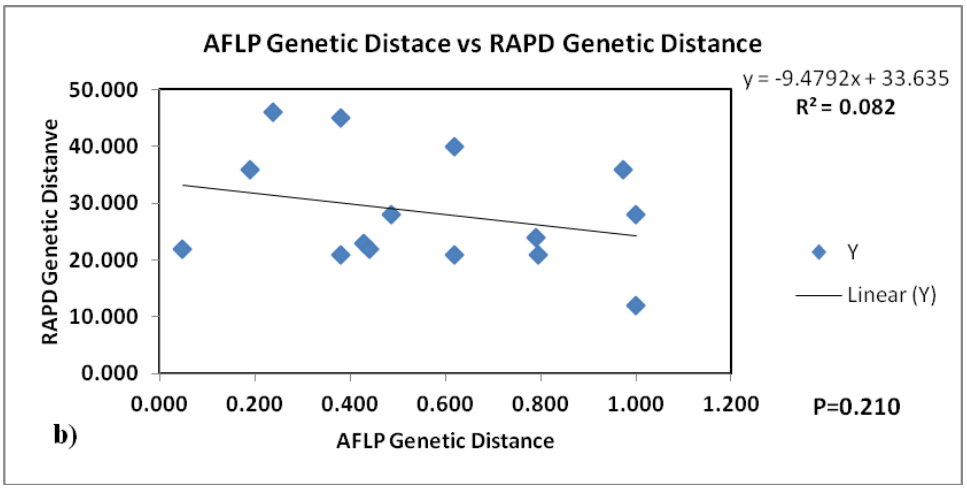
Table 12. Comparison of H_{av}, PIC, EMR, MI values derived from RAPDs, ISSRs and AFLPs using the standard two sample t-test

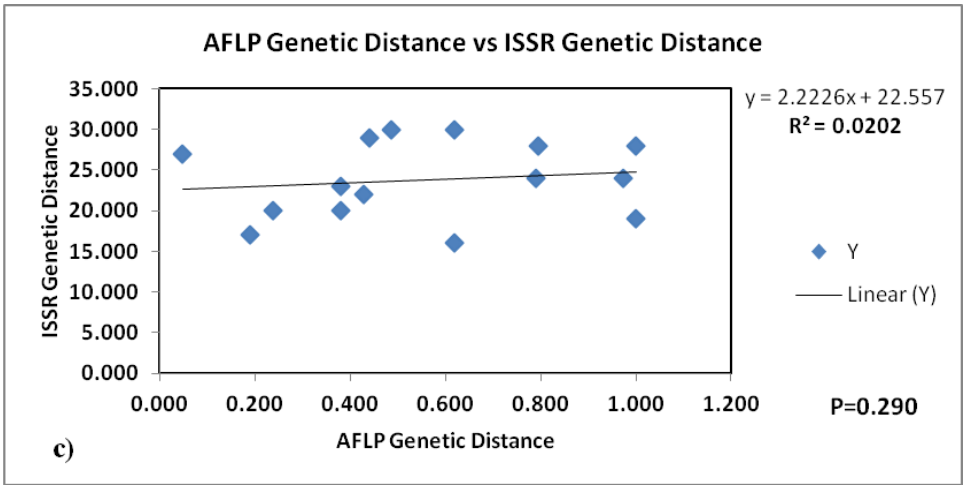
Marker	**t-test Values** H_{av} (P value)	**t-test Values PIC** (P value)	**t-test Values MI** (P value)
RAPD and ISSR	1.1611 (0.1087)	1.5760 (0.1359)	0.7762 (0.4490)
AFLP and RAPD	1.6287 (0.1056)	0.0091 (0.9929)	0.7527 (0.4661)
AFLP and ISSR	0.3396 (0.7349)	1.0164 (0.2668)	0.3096 (0.7622)

The UPGMA dendrograms generated using RAPDs and the concatenated data set have the same topology wherein individuals from each collection site group together ((**Figs. 4 and 9).** However, no phylogeographic clustering of the populations was observed. For *e.g*. populations of Assam and Narayangaon were comparatively close in genetic space (0.095), although they are distant in geographic space (**Fig. 1**). Similarly, populations of Jalgaon and Narayagaon which are close in geographic space were separated by a comparatively high genetic distance (0.332). In the UPGMA dendrogram generated from the ISSR data, all the populations with the exception of Assam group as per the collection site and no phylogeographic clustering of the populations was observed (**Fig. 7)**. The overall topology of clustering in the ISSR dendrogram is different from that observed in the dendrograms based on RAPD and RAPD+ISSR data. The dendrogram generated from the AFLP data shows a phylogeographic clustering of the populations (**Fig. 10**) which is supported by a significant positive correlation between the genetic distance and geographic distance **(Fig. 11).**

To assess the significance of these differences between the dendrograms derived from the four data sets ie RAPDs, ISSRs, RAPDs+ISSRs and AFLPs, the genetic distance matrices based on RAPDs, ISSRs, RAPDs+ISSR and AFLPs were compared by regression analysis using Mantel test (**Figs 12 a-f**).


ISSR Genetic Distance vs RAPD Genetic Distance
y = 0.7743x + 13.105
R² = 0.1751
RAPD Genetic Distance
ISSR Genetic Distance
Y
Linear (Y)
P=0.01
a)


AFLP Genetic Distace vs RAPD Genetic Distance
y = -9.4792x + 33.635
R² = 0.082
RAPD Genetic Distanve
AFLP Genetic Distance
Y
Linear (Y)
P=0.210
b)


AFLP Genetic Distance vs ISSR Genetic Distance
y = 2.2226x + 22.557
R² = 0.0202
ISSR Genetic Distance
AFLP Genetic Distance
Y
Linear (Y)
P=0.290
c)

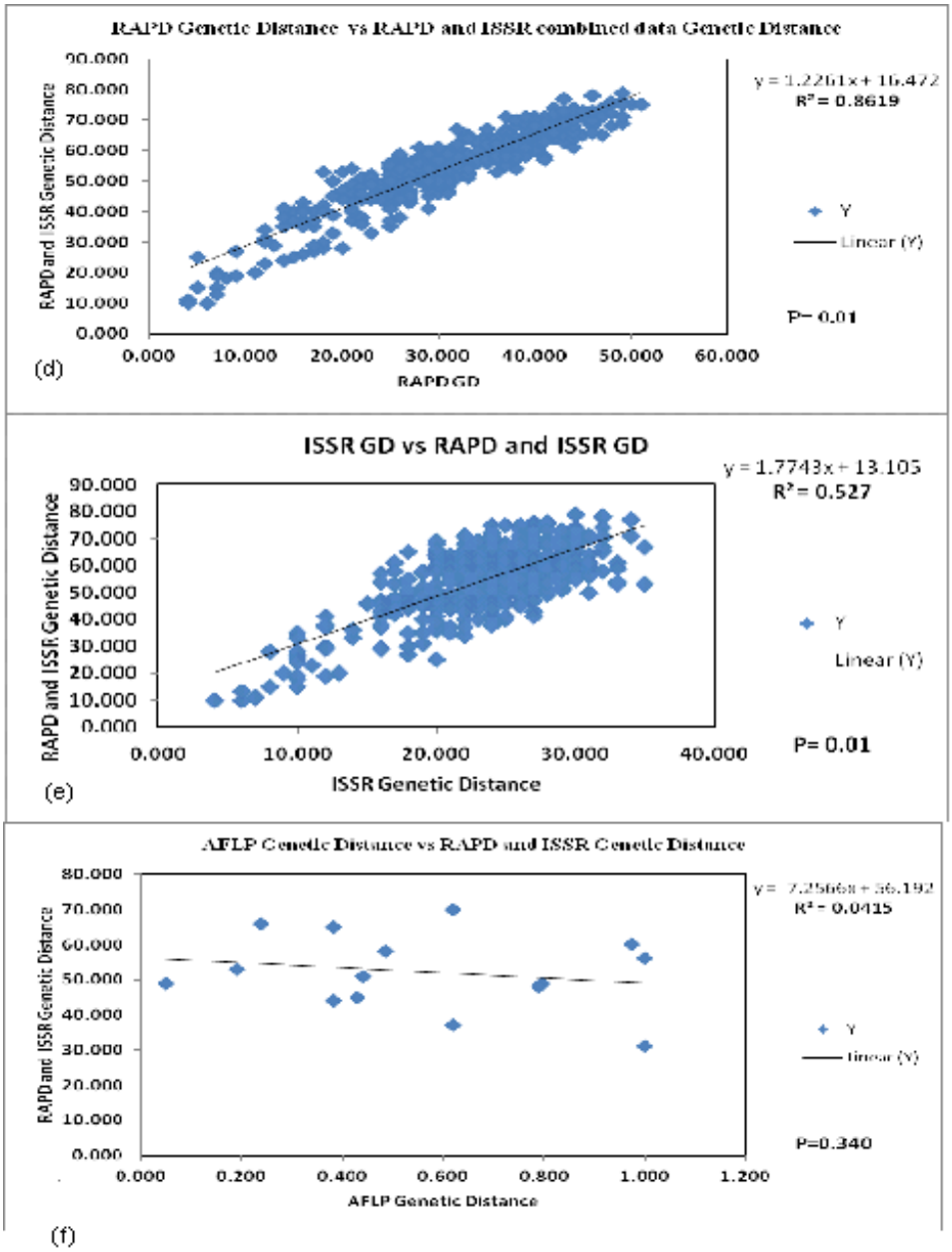


Fig. 12. Mantel test for correlation between genetic distances of (a) RAPD and ISSRs (b) RAPDs and AFLPs (c) ISSRs and AFLPs (d) RAPDs + ISSRs and RAPDs (e) RAPDs + ISSRs and ISSRs (f) RAPDs + ISSRs and AFLPs

From these figure it is observed that a strong correlation exists between the genetic distance matrices generated by RAPDs and RAPDS + ISSRs (R^2=0.8619) whereas the correlation between ISSRs and RAPDs + ISSRs (R^2=0.527) was moderate. A

poor correlation is revealed between the genetic distance matrices derived from RAPDs and ISSRs (R^2= 0.1751).

Weak correlation is also observed in the remaining cases (**Figs 12 a, b, c and f**). These results along with the clustering observed in the UPGMA dendrograms suggest that RAPDs alone or in conjunction with ISSRs provide reliable estimates of genetic variation within populations. The genetic distance matrix derived from AFLPs shows a poor correlation with the matrices generated from RAPDs, ISSRs and RAPDs + ISSRs. The data generated by AFLPs reveals a meaningful phylogeographic distribution of these populations which is supported by Mantel test (**Fig. 11**). In this study, RAPDs alone or in combination with ISSRs revealed diversity at the within population' level. In the AFLP based dendrogram, the Assam population is clearly distinct from the other populations as revealed by the longer branch length. A similar observation has been made when the sequences of the ITS1 and ITS2 regions of the same thirty individuals used in the present work were analysed (Chapter 4 of this thesis). The ITS1 and ITS2 regions are known to provide reliable estimates of diversity at the inter- and intra-population level (Hamasheh *et al*. 2007; Moin-Vaziri *et al*. 2007). Hence in the present study, AFLPs are more informative in revealing genetic relationships between the six *O. longicollis* populations. A possible explanation for the above observations is that these markers target different regions of the genome. ISSRs consist of tandemly repeated short stretches of oligonucleotides which find complimentarity with microsatellite sequences within the genome and amplify the region in between the microsatellite regions. Lu *et al*. (1996) have suggested that microsatellites are mainly clustered together and are associated with other repetitive sequences within the genome. Evolution of microsatellites occurs by processes such as amplification and transposition and is faster than that of single copy sequences (Delseny *et al*. 1988). This could explain the low level of correlation between the genetic distance matrices generated by RAPDs, ISSRs and AFLPs. RAPDs and AFLPs reveal polymorphisms in the coding as well as the non-coding regions and can potentially cover the entire genome. However, in case of AFLPs the selective nucleotides at the 3' end of each primer could potentially cover additional regions in the genome as compared to RAPDs due to which fewer AFLP primer combinations and relatively small

number of polymorphic bands as compared with RAPDs and ISSRs were able to reveal phylogeographic relation between the populations.

The present analysis based on nine RAPD (113 loci) and nine ISSR markers (71 loci) suggests that there is restricted gene flow between the populations (For RAPDs : G_{ST} = 0.545, Nm = 0.4174; For ISSRs : G_{ST} = 0.4461, Nm = 0.6207). However, these results are in contrast to those obtained using the rDNA markers *i.e.* ITS1 and ITS2 (Chapter 4 of this thesis). Possible explanations to this observed discrepancy is discussed in Chapter 7 of this thesis.

Bananas originated in south-east Asia which is considered to be the primary centre of diversification and earliest domestication of this crop (Simmonds 1962). *O. longicollis* (Oliver) being a monophagous pest of bananas would also have originated in this region, coincident with the centre of origin of bananas, before migrating to other regions. The only land route by which this pest could have entered the Indian sub-continent is via north-east India. The UPGMA dendrogram based on AFLPs reveals a phylogeographic distribution of the populations wherein the populations from Assam and Vaishali separate from the remaining populations with longer branch lengths, suggesting that among these populations, the population from Assam is the oldest followed by the population from Vaishali and that the populations from the other four locations have been colonised later than the Assam and Vaishali populations.

Knowledge about genetic variation patterns within and between populations of the banana stem weevil populations is very important for understanding possible local adaptation and migration patterns of the pest. Study of the population dynamics of the pest population ie rates of mutation and migration might provide useful data for an Integrated Pest Management program for the control of this pest. This work is the first report of the study of the genetic diversity of this important banana pest.

REFERENCES

Abdallah Z, Mezghani-Kkhemakhem M, Bouktila D, Makni H and Makni M (2012) *African Journal of Agricultural Research,* **7(7)**, 1170-1176.

Allender CJ, Clancy K M, Degomez TE., Mcmillin JD, Woolbright SA, Keim P and Wagner DM (2008) *Environtal Entomology,* **37(3)**, 817-824.

Bas B, Dalkilic Z, Peever TL, Nigg HN, Simpson SE, Gmitter FG and Adair RC (2000) *Annals of the Entomological Society of America,* **93**, 459–467.

Breuker CJ, Victoir K, De Jong PW, van der Meijden E, Brakefield PM and Vrieling K (2005) *Journal of Insect Science*, **5.**

Carisio L, Cervella P, Palestrini C, Delpero M and Rolando A (2004) *Journal of Biogeography,* **31**, 1149-1162.

Cerritos R, Wegier A and Alavez V (2012) ISBM: 978-953-51-0050-8. InTech.

Dellaporta SL, Wood J and Hicks JB (1983) *Plant Molecular Biology Reporter* **1(4)**, 19-21.

Delseny M, Grellet F, Tremousaygue D, Raynal M and Panabieres F (1988) *Bulletin de la Société Botanique de France. Actualités Botaniques*, **135(2)**, 23-38.

De Graaf J (2006) Doctoral dissertation, University of Pretoria.

De Souza GAD, Carvalho MRDO, Martins ER, Guedes RNC and Oliveira LOD (2008) *Pesquisa Agropecuária Brasileira*, **43(7)**, 843-849.

Gailce Leo Justin C, Leelamathi M and Nirmaljohson SB (2008) *Agriculture Review*, **29(3)**, 185-192.

Gold CS, Rukazambuga NDTR, Karamura EB, Nemeye P and Night G (1999) In: Proceedings of a workshop on Banana IPM, Nelspruit, South Africa. November 1998, 33-50.

Guasmi F, Elfalleh W, Hannachi H, Feres K, Touil L, Marzougui N, Triki T and Ferchichi A (2012) *ISRN Agronomy Article* ID 952196, 10 pages.

Hamarsheh O, Presber W, Abdeen Z, Sawalha S, Al-Lahem A and Schonian G (2007) *Medical and Veterinary Entomology*, **21(3)**, 270-277.

Hawthorne DJ (2001) *Genetics*, **158**, 695–700.

Ibrahim SA, El-Mergawy RG and Moghaieb RE (2009) *Australian Journal of Basic and Applied Sciences*, **3(2)**, 1287-1295.

Jain SK, Neekhra B, Pandey D and Jain K (2010) *Indian Journal of Biotechnology*, **9**, 7-12.

Kar PK, Srivastava AK, Sinha MK, Sinha AK and Prasad BC (2010) *The Bioscan*, **3**, 627-634.

Katiyar SK, Chandel G, Tan Y, Zhang Y, Huang B, Nugaliyadde L,Fernando K, Bentur JS, Inthavong S, Constantino S and Bennett J (2000) *Genome*, **43**, 322–332.

Kazachkova N, Meijer J and Ekbom B (2007) *Agricultural Forest Entomology*, **9**, 259-269.

Kazachkova N, Meijer J and Ekbom B (2008) *European Journal of Entomology*, **105**, 807–814.

Kim KS and Sappington TW (2004) *Insect Molecular Biology*, **13**, 293–303.

Krumm JT, Hunt TE, Skoda SR, Hein GL, Lee DJ, Clark PL and Foster JE (2008) *Journal of Insect Science*, **8.**

Kumar LS, Sawant AS, Gupta VS and Ranjekar PK (2001) *Biochemical Genetics*, **39(9/10)**, 297-309

Lu J, Knox MR, Ambrose MJ, Brown JKM and Ellis THN (1996) *Theoretical and Applied Genetics*, **93**, 1103.

Magana C, Beroiz B, Hernandez-Crespo P, de Oca MMA, Carnero A, Ortego F and Castanera P (2007) *Bulletin of Entomology Research*, **97**, 585-590.

Martins WF, Ayres CF and Lucena WA (2007) *Genetics and Molecular Research*, **6(1)**, 23–32.

Mantel N (1967) *Cancer Research*, **27**(2 Part 1):209-220.

McDermott JM and McDonald BA (1993) *Annual Review of Phytopathology*, **31(1)**, 353-373.

Moin-Vaziri V, Depaquit J, Yaghoobi-Ershadi MR, Oshaghi MA, Derakhsh eh-Peykar P, Ferte H and Nadim A (2007) *Acta Tropica*, **102(1)**, 29-37.

Mustaffa MM and Sathiamoorthy S (2004) Advancing banana and plantain R&D in Asia and the Pacific - Vol.12. Proceedings of the 2st BAPNET Steering Committee

meeting. Jakarta (IDN), 2003/10/06-07. Los Baños (PHL):INIBAP-AP 2004. ISSN: 1729-0805.

Nei M (1972) *American Naturalist*, 283-292.

Nei M and Roychoudhury A K (1972) *Science* **177**, 434-436.

Nei M (1973) *Proceedings of Natural Academy of Science USA*, **70**, 3321-3323.

Nei M (1978) *Genetics*, **89**, 583-590.

Nevo E (1978) *Theoretical Population Biology*, **13(1)**, 121-177.

Ostmark HE (1974) *Annual Review of Entomology*, **19**, 161-176.

Padmanaban B and Sathiamoorthy S (2001) *Musa Pest Fact Sheet* No. 5.

Padmanaban B, Sundararaju P, Velayudhan KC and Sathiamoorthy S (2001a) *InfoMusa*, **10(1)**, 26-28.

Padmanaban B, Sundararaju P and Sathiamoorthy S (2001b) *Indian Journal of Entomology*, **63**, 204-206.

Padmanaban B and Kandasamy M (2003) *Indian Journal of Entomology*, **65(3)**, 424-425.

Peakall ROD and Smouse PE (2006) *Molecular Ecology Notes*, **6(1)**, 288-295.

Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S and Rafalski A (1996) *Molecular Breeding*, **2,** 225.

Prevost A and Wilkinson MJ (1999) *Theoretical and Applied Genetic*, **98**, 107–112.

Ranjan SK, Mallick CB, Saha D, Vidyarthi AS and Ramani R (2011) *Genetics and Molecular Biology*, **34(3)**, 511-519.

Roldan-Ruiz I, Dendauw J, Vanbockstaele E, Depicker A and De Loose M (2000) *Molecular Breeding*, **6**, 125–134.

Sharma HC (2009) *Biotechnological Approaches for Pest Management and Ecological Sustainability*. Google eBook. CRC Press. Taylor and Francis Group.

Sidorenko AP, Berezovskaia OP and Sozinov AA (2000) *Genetika*, **36(5)**, 651-656

Sidorenko AP and Berezovska OP (2002) *Genetika*, **38(11)**, 1485–1491.

Sikora RA, Bafokuzaka ND, Mbwana ASS, Oloo GW, Uronu B and Seshu-Reddy KV (1989) FAO *Plant Protocol Bulletin*, **37**, 151–157.

Simmonds NW (1962) The Evolution of the Bananas. Longmans, London, Great Britain.

Sripriya C, Padmanaban B and Uma S (2000) *Indian Journal of Entomology*, **62(4)**, 382-390.

Valmayor RV, Davide RG, Stanton JM, Treverrow NL and Rao VN (eds) In: Proceedings of a conference-workshop on nematodes and weevil borers affecting bananas in Asia and the Pacific, 18-22 April 1994, Serdang, Selangor, Malaysia. INIBAP/ASPNET, Los Banos, Philippines, 258 pp.

Varshney RK, Chabane K, Hendre PS, Aggarwal RK and Graner A (2007) *Plant Science*, **173**, 638–649.

Vos P, Hogers R, Bleeker M, Reijans M, Van De Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M and Zabeau M (1995) *Nucleic Acids Res*, **23**, 4407-4414.

Yeh FC, Rong-Cai Y and Boyle T (1998) POPGENE version 1.31. University of Alberta. Center for International Forestry Research, Edmonton, Alberta, Canada

Zhong D, Pai A and Yan G (2004) *Journal of Heredity*, **95(1)**, 53-61.

CHAPTER 4

Genetic diversity analysis of *O. longicollis* (Oliver) populations using rDNA markers *i.e.* ITS1 and ITS2

SUMMARY

An important step towards developing a successful Integrated Pest Management (IPM) program for the control of *O. longicollis* (Oliver), a serious pest of bananas, is the study of the population structure of the pest. In the present study, the primary nucleotide sequences of the ITS1 and ITS2 regions of thirty individual weevils representing six Indian populations of *O. longicollis* were used to assess the genetic variation amongst these populations. Results of the phylogenetic analyses of these two regions were congruent and indicated that the phylogenetic trees were influenced by haplotypes. These results indicate gene flow and no phylogeographical distribution of these six populations.

A consensus secondary structure of the ITS1 and ITS2 regions of *O. longicollis* (Oliver) is also presented. The consensus secondary structures of the ITS2 region conforms to the pan-eukaryotic model. The present study of the secondary structure of the ITS1 and ITS2 regions of *O. longicollis* and its use in assessing the phylogenetic relationships is the first detailed report for insects, in particular family Curculionidae. The topology of the phylogenetic trees based on the secondary structure of the ITS1 and ITS2 regions bear striking similarity to each other and also to the trees generated based on primary sequence data suggesting that conserved secondary structures are important `markers' in phylogenetic analysis. The phylogenetic analyses of the ITS 1 and 2 regions suggest that this pest has migrated into India via the north-east region.

INTRODUCTION

The ribosomal gene cluster has become a useful tool for classifying/diagnosing organisms at various taxonomic levels. The low level of evolutionary constraint in the spacers, combined with the rapid tendency of these arrays to be homogenized leads to the rapid spread of a new sequence variant which results in genetically isolated populations to be homogenized for different variants of rDNA. This process leads to high levels of intra-specific and intra-individual variation, which makes these regions highly useful for studies at the inter- and intra-population level (Bebee *et al*. 2001). The utility of the ITS1 and ITS2 regions in phylogeny varies in different organisms. In several studies these regions have proved useful for assessing phylogenetic relationships, but in a few cases ie *Anopheles punctulatus* (Beebe *et al*. 1999), *Metastrongylus,* Strongylida, Metastrongyloidea (Conole *et al*. 2001), *Chrysotoxum* species, Diptera, Syrphidae (Masetti *et al*. 2006), mites and ticks (Cruickshank 2002) and *Argosarchus* and *Clitarchus* stick insects, Phasmida, Phasmatinae (Trewick *et al*. 2005), these markers were not useful in decipheric phylogenetic relationships.

The secondary structure of the ITS1 and ITS2 regions have been implicated to play an important role in the final processing of the rRNA and hence their secondary structure has also been deciphered for several insects. To date, very few reports on the secondary structure of the ITS1 region are available. The very first insect ITS1 secondary structure was proposed for *D. simulans* (Armbruster 2001), and later for the bush cricket *Poecilimon chopardi* (Ullrich *et al*. 2010). However due to significant variation in the length of the spacer region of the above species, no common structure for the ITS1 molecule could be proposed. Examples of other organisms whose secondary structure of ITS1 has been reported are hookworms (Zhu 1999), Digenea, Trematoda (Schulenburg *et al*. 1999) and dinoflagellates (Gottschling and Plötner 2004). As compared to ITS1, ITS2 has received more attention which has resulted in more than 2,10,000 ITS2 sequences in GenBank, and numerous studies have used this region for phylogenetic analyses. A conserved `pan-eukaryotic model for the secondary structure of ITS2 was proposed by Coleman (2007). Schlotterer *et al*. (1994) studied eight *Drosophila* species and proposed an ITS2 secondary structure. In case of *Anopheles*, Beebe (1999) and Marrelli *et al*.

(2005) have studied ITS2 structure for species identification. Severini *et al*. (1996) studied sequence and secondary structure of the rDNA ITS2 to differentiate sibling species of two mosquito species. Several beetles species were also studied using ITS2 secondary structure, e.g. meloid beetles (Bologna *et al*. 2008), *Timarcha* leaf beetles (Gomez-Zurita *et al*. 2000), raspberry beetles (Malloch *et al*. 2001) and pollen beetles (Trizzino *et al*. 2009). Wolf *et al*. (2005) studied the extent of the conservation in ITS2 structure based on homology modeled secondary structures of more than 20,000 ITS2 covering about 14,000 species. Coleman (2007) studied different aspects of evolution and higher relationships across eukaryotes (including algae, plants, fungi, animals, molluscs and arthropods) using the secondary structure of the ITS2 region.

The primary sequence of the ITS1 ad ITS2 regions have been used extensively in deciphering phylogenetic relationships in a wide range of organisms. The secondary structure is highly conserved despite mutations occurring freely in these regions because the secondary structure provides docking signals during the maturation of the precursor of ribosomal RNA. The secondary structure elements such as helices, bulges, loops, etc are evolutionary conserved phylogenetic characters and can be treated as conserved `markers' in phylogenetic analysis (Billoud *et al*. 2000). Since multiple sequence alignment techniques sometimes neglect such vital structural information, Schultz & Wolf (2009) and Keller *et al*. (2010) emphasized the inclusion of RNA secondary structure, as predicted by bioinformatics tools or by homology modeling, in phylogenetic analysis. Phylogenetic analysis based on secondary structures have yielded more robust alignments and trees as compared to methods that do not use structural information. Improved reconstruction of phylogenetic trees by including secondary structure information has been documented in several studies including *Syzygium* and Myrtacea (Biffin *et al.* 2007), Carribean gorgonian corals (Grajales *et al*. 2007), Sphaeropleales (Keller *et al*. 2008), *Coelastrum* (Hegewald *et al*. 2010) and *Agrodiaetus* (Wiemers *et al*. 2009).

In the present work, the primary sequences of both these ITS regions including indel polymorphisms, haplotype data, concatenated data as well as the secondary structure data have been used to study the genetic diversity of the different Indian populations of *O. longicollis* (Oliver). The congruency in the results obtained using both these regions indicate their utility in deciphering phylogenetic relationships and the

importance of secondary structure of these two regions in phylogenetic analysis. The consensus secondary structures of the ITS1 and ITS2 regions of *O. longicollis* (Oliver) has also been determined. The present study of the use of secondary structure of these two regions in assessing the phylogenetic relationships is the first detailed report for insects, in particular family Curculionidae.

MATERIALS AND METHODS

Sample collection

A total of thirty adult beetles of *O. longicollis (Oliver)* i.e. five adults from each locations in India were collected from infested banana stems from banana fields in Assam (Kamrup-K), Bihar (Vaishali-V), Kerala (Wayanad-W), Maharashtra (Jalgaon-J, Narayangaon-N) and Tamilnadu (Trichy-T) (As described in Chapter 3). In each location, the five individuals were collected within an area of 500 m.

DNA extraction

Total genomic DNA was extracted from individual beetles using a modified Dellaporta *et.al*. (1983) protocol as described in 1Materials and Methods' in Chapter 3.

Amplification of ITS1 and ITS2 fragments

ITS1 and ITS2 fragments were amplified using the primer pairs R5/R6 and R8/R9 respectively (**Table 1**).

Table 1. Sequence of primers used for amplifying the ITS1 and ITS2 regions

Primer	Primer Sequence (5'-3')	Reference
R5-F (ITS1)	TTGATTACGTCCCTGCCCTTT	Szalanski & Owens 2003
R6-R (ITS1)	ACGAGCCGAGTGATCCACCG	Szalanski & Owens 2003
R8-F (ITS2)	GTTGAAAGGGAGTGCCATGAAC	Malloch *et al.* 2001
R9-R (ITS2)	CAACTTTCCCTCACGGTACTTG	Malloch *et al.* 2001

PCR was carried out in a total reaction volume of 25 µl consisting of 25 ng DNA, 0.1 mM dNTPs, 100 pmol of each primer and 0.5 U Taq DNA Polymerase. PCR conditions were as follows: pre-denaturation at 94^0C for 5 min followed by thirty cycles of denaturation at 94^0C for 45 sec; annealing at 54^0C (ITS1)/(ITS2) for 45 sec; extension at 72^0C for 2 min; a final extension at 72^0C for 20 min. The PCR product of appropriate size was cloned into PCR4 TOPO (Invitrogen) vector and transformed into *E. coli* Top10 host cells (Invitrogen), as per the manufacturer's instructions. The sequences have been deposited in GenBank (**Table 2**).

Table 2. List of *O. longicollis* (Oliver) individuals examined showing their ITS1 and ITS2 spacer lengths, GC content and accession numbers

Location (Individual No.)	Genbank Accession ITS1	ITS1 Length (bp)	ITS1 GC content (%)	Genbank Accession ITS2	ITS2 Length (bp)	ITS2 GC content (%)
Assam 1	KF225405	976	38.4	KF225435	478	50.0
Assam 2	KF225406	980	38.4	KF225436	491	50.2
Assam 3	KF225407	968	38.2	KF225437	486	50.3
Assam 4	KF225408	970	38.0	KF225438	487	49.9
Assam 5	KF225409	961	37.8	KF225439	491	49.9
Jalgaon 1	KF225410	970	37.6	KF225440	489	49.9
Jalgaon 2	KF225411	976	37.8	KF225441	489	49.9
Jalgaon 3	KF225412	986	37.6	KF225442	489	50.2
Jalgaon 4	KF225413	973	37.4	KF225443	489	50.1
Jalgaon 5	KF225414	976	37.6	KF225444	489	50.1
Narayangaon 1	KF225415	974	37.6	KF225445	493	50.1
Narayangaon 2	KF225416	976	37.8	KF225446	493	49.9
Narayangaon 3	KF225417	972	37.6	KF225447	493	50.1
Narayangaon 4	KF225418	974	37.6	KF225448	493	49.8
Narayangaon 5	KF225419	974	37.6	KF225449	489	49.9
Trichy 1	KF225420	974	37.6	KF225450	491	49.9
Trichy 2	KF225421	975	37.8	KF225451	489	49.9
Trichy 3	KF225422	975	37.6	KF225452	491	49.9
Trichy 4	KF225423	980	37.6	KF225453	489	49.9
Trichy 5	KF225424	979	37.8	KF225454	486	50.3
Vaishali 1	KF225425	974	37.6	KF225455	493	49.9
Vaishali 2	KF225426	986	37.6	KF225456	493	49.8
Vaishali 3	KF225427	989	37.6	KF225457	488	49.9

Vaishali 4	KF225428	976	37.6	KF225458	488	50.3
Vaishali 5	KF225429	989	37.8	KF225459	491	49.8
Wayanad 1	KF225430	965	37.6	KF225460	488	50.1
Wayanad 2	KF225431	965	37.8	KF225461	486	49.6
Wayanad 3	KF225432	965	37.8	KF225462	471	49.8
Wayanad 4	KF225433	965	37.8	KF225463	486	49.9
Wayanad 5	KF225434	966	37.8	KF225464	488	50.0

Sequence analysis

DNA sequence chromatograms were read and discrepancies between forward and reverse sequences were resolved using the Chromas software version 2.01 (http:// www.technelysium.com.au/chromas.html). EBI CLUSTALW was used to generate the alignments of the sequences (http,//www.ebi.ac.uk/Tools/msa/clustalw2/) (Thompson *et al*. 1994). The sequences were imported into MEGA v 5.2.1 for phylogenetic analysis (Kumar *et al*. 2001). MEGA v 5.2.1 was used to calculate genetic distances between the different populations. The phylogenetic trees were generated using UPGMA algorithm. Gaps in the alignments were excluded from the phylogenetic analysis. Statistical support for the inferred nodes was obtained by bootstrapping (1000 replicates) in MEGA version 5.2.1. DnaSP v 5.10 was used to determine haplotype data *i.e*. the haplotype diversity, number of haplotypes and insertion-deletion polymorphism (indels). DAMBE v 5.3.27 was used to calculate Ts/Tv and divergences ratio. ITS2 structure database was used to compare the secondary structure and other parameters of all the sequences. RNalifold (http,//rna.tbi.univie.ac.at/cgi-bin/RNAalifold.cgi) (Hofacker 2003) was used to derive the consensus secondary structure. RNAdist (http//mobyle.pasteur.fr/cgi-bin/portal.py#forms,rnadistance) was used to perform phylogenetic analysis based on pairwise distance matrix built based on the secondary structure. Correlation between genetic distance and geographical distance was calculated by Mantel test using the GenAlEx v 6.501 software (Peakall & Smouse 2006). RNA v 5.5 was used to generate the 2D structures of 5.8S and 18S (flanking region of ITS2) (Reuter *et al*. 2010).

RESULTS

Analysis of the ITS1 region

Sequence analysis

A total of thirty adult beetles of *O. longicollis (Oliver)* i.e. five adults from each locations were collected from infested banana stems from banana fields in Assam (Kamrup), Bihar (Vaishali), Kerala (Wayanad), Maharashtra (Jalgaon, Narayangaon) and Tamilnadu (Trichy) (As described in Chapter 3). The presumptive boundaries between 18S and ITS1 and ITS1 and 5.8S were determined by comparison with the ITS1 sequences of southern corn rootworm. (Genus Diabrotica; Coleoptera, Chrysomelidae) (Szalanski & Owens 2003). The length of the ITS1 sequences ranged between 885 bp and 909 bp and the base composition was biased with a mean GC content of 37.64% (**Table 2**).

The ClustalW alignment of the ITS1 sequences (including gaps) from the thirty individuals generated 513 characters of which 476 (92.78%) were conserved, 33 (6.43%) were parsimony informative and 13 (2.5%) were singleton sites (**Fig. 1**).

```
ASSAM1            GTACACACCGCCCGTCGCTACTACCGATTGAATGATTTAGTGAGGTCTTCGGATCGACAC 60
TRICHY3           GTACACACCGCCCGTCGCTACTACCGATTGAATGATTTAGTGAGGTCTTCGGATCGACAC 60
TRICHY5           GTACACACCGCCCGTCGCTACTACCGATTGAATGATTTAGTGAGGTCTTCGGATCGACAC 60
JALGAON5          GTACACACCGCCCGTCGCTACTACCGATTGAATGATTTAGTGAGGTCTTCGGATCGACAC 60
TRICHY2           GTACACACCGCCCGTCGCTACTACCGATTGAATGATTTAGTGAGGTCTTCGGATCGACAC 60
JALGAON2          GTACACACCGCCCGTCGCTACTACCGATTGAATGATTTAGTGAGGTCTTCGGATCGACAC 60
VAISHALI3         GTACACACCGCCCGTCGCTACTACCGATTGAATGATTTAGTGAGGTCTTCGGATCGACAC 60
VAISHALI5         GTACACACCGCCCGTCGCTACTACCGATTGAATGATTTAGTGAGGTCTTCGGATCGACAC 60
TRICHY4           GTACACACCGCCCGTCGCTACTACCGATTGAATGATTTAGTGAGGTCTTCGGATCGACAC 60
WAYANAD2          GTACACACCGCCCGTCGCTACTACCGATTGAATGATTTAGTGAGGTCTTCGGATCGACAC 60
WAYANAD4          GTACACACCGCCCGTCGCTACTACCGATTGAATGATTTAGTGAGGTCTTCGGATCGACAC 60
WAYANAD5          GTACACGCCGCCCGTCGCTACTACCGATTGAATGATTTAGTGAGGTCTTCAGATCGACAC 60
WAYANAD1          GTACACACCGCCCGTCGCTACTACCGATTGAATGATTTAGTGAGGTCTTCGGATCGACAC 60
WAYANAD3          GTACACACCGCCCGTCGCTACTACCGATTGAATGATTTAGTGAGGTCTTCGGATCGACAC 60
NARAYNAGAON2      GTACACACCGCCCGTCGCTACTACCGATTGAATGATTTAGTGAGGTCTTCGGATCGACAC 60
VAISHALI4         GTACACACCGCCCGTCGCTACTACCGATTGAATGATTTAGTGAGGTCTTCGGATCGACAC 60
VAISHALI1         GTACACACCGCCCGTCGCTACTACCGATTGAATGATTTAGTGAGGTCTTCGGATCGACAC 60
JALGAON4          GTACACACCGCCCGTCGCTACTACCGATTGAATGATTTAGTGAGGTCTTCGGATCGACAC 60
NRAYANGAON1       GTACACACCGCCCGTCGCTACTACCGATTGAATGATTTAGTGAGGTCTTCGGATCGACAC 60
NARAYANGAON4      GTACACACCGCCCGTCGCTACTACCGATTGAATGATTTAGTGAGGTCTTCGGATCGACAC 60
NARAYANGAON5      GTACACACCGCCCGTCGCTACTACCGATTGAATGATTTAGTGAGGTCTTCGGATCGACAC 60
TRICHY1           GTACACACCGCCCGTCGCTACTACCGATTGAATGATTTAGTGAGGTCTTCGGATCGACAC 60
NARAYAGAON3       GTACACACCGCCCGTCGCTACTACCGATTGAATGATTTAGTGAGGTCTTCGGATCGACAC 60
JALGAON1          GTACACACCGCCCGTCGCTACTACCGATTGAATGATTTAGTGAGGTCTTCGGATCGACAC 60
VAISHLAI2         GTACACACCGCCCGTCGCTACTACCGATTGAATGATTTAGTGAGGTCTTCGGATCGACAC 60
JALGAON3          GTACACACCGCCCGTCGCTACTACCGATTGAATGATTTAGTGAGGTCTTCGGATCGACAC 60
ASSAM2            GTACACACCGCCCGTCGCTACTACCGATTGAATGATTTAGTGAGGTCTTCGGATCGACAC 60
ASSAM4            GTACACACCGCCCGTCGCTACTACCGATTGAATGATTTAGTGAGGTCTTCGGATCGACAC 60
```

```
ASSAM3          GTACACACCGCCCGTCGCTACTACCGATTGAATGATTTAGTGAGGTCTTCGGATCGACAC 60
ASSAM5          GTACACACCGCCCGTCGCTACTACCGATTGAATGATTTAGTGAGGTCTTCGGATCGACAC 60
                ****** ****************************************** *********

ASSAM1          GCGATGGCGCTTCGGCGTCGTCGATGTGTAAGAAAAGATGACCAAACTTGATCATTTAGA 120
TRICHY3         GCGATGGCGCTTCGGCGTCGTCGATGTGTAAGAAAAGATGACCAAACTTGATCATTTAGA 120
TRICHY5         GCGATGGCGCTTCGGCGTCGTCGATGTGTAAGAAAAGATGACCAAACTTGATCATTTAGA 120
JALGAON5        GCGATGGCGCTTCGGCGTCGTCGATGTGTAAGAAAAGATGACCAAACTTGATCATTTAGA 120
TRICHY2         GCGATGGCGCTTCGGCGTCGTCGATGTGTAAGAAAAGATGACCAAACTTGATCATTTAGA 120
JALGAON2        GCGATGGCGCTTCGGCGTCGTCGATGTGTAAGAAAAGATGACCAAACTTGATCATTTAGA 120
VAISHALI3       GCGATGGCGCCTCGGCGTCGTCGATGTGTAAGAAAAGATGACCAAACTTGATCATTTAGA 120
VAISHALI5       GCGATGGCGCCTCGGCGTCGTCGATGTGTAAGAAAAGATGACCAAACTTGATCATTTAGA 120
TRICHY4         GCGATGGCGCTTCGGCGTCGTCGATGTGTAAGAAAAGATGACCAAACTTGATCATTTAGA 120
WAYANAD2        GCGATGGCGCTTCGGCGTCGTCGATGTGTAAGAAAAGATGACCAAACTTGATCATTTAGA 120
WAYANAD4        GCGATGGCGCTTCGGCGTCGTCGATGTGTAAGAAAAGATGACCAAACTTGATCATTTAGA 120
WAYANAD5        GCGATGGCGCTTCGGCGTCGTCGATGTGTAAGAAAAGATGACCAAACTTGATCATTTAGA 120
WAYANAD1        GCGATGGCGCTTCGGCGTCGTCGATGTGTAAGAAAAGATGACCAAACTTGATCATTTAGA 120
WAYANAD3        GCGATGGCGCTTCGGCGTCGTCGATGTGTAAGAAAAGATGACCAAACTTGATCATTTAGA 120
NARAYNAGAON2    GCGATGGCGCTTCGGCGTCGTCGATGTGTAAGAAAAGATGACCAAACTTGATCATTTAGA 120
VAISHALI4       GCGATGGCGCTTCGGCGTCGTCGATGTGTAAGAAAAGATGACCAAACTTGATCATTTAGA 120
VAISHALI1       GCGATGGCGCTTCGGCGTCGTCGATGTGTAAGAAAAGATGACCAAACTTGATCATTTAGA 120
JALGAON4        GCGATGGCGCTTCGGCGTCGTCGATGTGTAAGAAAAGATGACCAAACTTGATCATTTAGA 120
NRAYANGAON1     GCGATGGCGCTTCGGCGTCGTCGATGTGTAAGAAAAGATGACCAAACTTGATCATTTAGA 120
NARAYANGAON4    GCGATGGCGCTTCGGCGTCGTCGATGTGTAAGAAAAGATGACCAAACTTGATCATTTAGA 120
NARAYANGAON5    GCGATGGCGCTTCGGCGTCGTCGATGTGTAAGAAAAGATGACCAAACTTGATCATTTAGA 120
TRICHY1         GCGATGGCGCTTCGGCGTCGTCGATGTGTAAGAAAAGATGACCAAACTTGATCATTTAGA 120
NARAYAGAON3     GCGATGGCGCTTCGGCGTCGTCGATGTGTAAGAAAAGATGACCAAACTTGATCATTTAGA 120
JALGAON1        GCGATGGCGCTTCGGCGTCGTCGATGTGTAAGAAAAGATGACCAAACTTGATCATTTAGA 120
VAISHLAI2       GCGATGGCGCTTCGGCGTCGTCGATGTGTAAGAAAAGATGACCAAACTTGATCATTTAGA 120
JALGAON3        GCGATGGCGCTTCGGCGTCGTCGATGTGTAAGAAAAGATGACCAAACTTGATCATTTAGA 120
ASSAM2          GCGATGGCGCTTCGGCGTCGTCGATGTGTAAGAAAAGATGACCAAACTTGATCATTTAGA 120
ASSAM4          GCGATGGCGCTTCGGCGTCGTCGATGTGTAAGAAAAGATGACCAAACTTGATCATTTAGA 120
ASSAM3          GCGATGGCGCTTCGGCGTCGTCGATGTGTAAGAAAAGATGACCAAACTTGATCATTTAGA 120
ASSAM5          GCGATGGCGCTTCGGCGTCGTCGATGTGTAAGAAAAGATGACCAAACTTGATCATTTAGA 120
                ********** ************************************************

ASSAM1          GGAAGTAAAAGTCGTAACAAGGTTTCCGTAGGTGAACCTGCGGAAGGATCATTACAGTGT 180
TRICHY3         GGAAGTAAAAGTCGTAACAAGGTTTCCGTAGGTGAACCTGCGGAAGGATCATTACAGTGT 180
TRICHY5         GGAAGTAAAAGTCGTAACAAGGTTTCCGTAGGTGAACCTGCGGAAGGATCATTACAGTGT 180
JALGAON5        GGAAGTAAAAGTCGTAACAAGGTTTCCGTAGGTGAACCTGCGGAAGGATCATTACAGTGT 180
TRICHY2         GGAAGTAAAAGTCGTAACAAGGTTTCCGTAGGTGAACCTGCGGAAGGATCATTACAGTGT 180
JALGAON2        GGAAGTAAAAGTCGTAACAAGGTTTCCGTAGGTGAACCTGCGGAAGGATCATTACAGTGT 180
VAISHALI3       GGAAGTAAAAGTCGTAACAAGGTTTCCGTAGGTGAACCTGCGGAAGGATCATTACAGTGT 180
VAISHALI5       GGAAGTAAAAGTCGTAACAAGGTTTCCGTAGGTGAACCTGCGGAAGGATCATTACAGTGT 180
TRICHY4         GGAAGTAAAAGTCGTAACAAGGTTTCCGTAGGTGAACCTGCGGAAGGATCATTACAGTGT 180
WAYANAD2        GGAAGTAAAAGTCGTAACAAGGTTTCCGTAGGTGAACCTGCGGAAGGATCATTACAGTGT 180
WAYANAD4        GGAAGTAAAAGTCGTAACAAGGTTTCCGTAGGTGAACCTGCGGAAGGATCATTACAGTGT 180
WAYANAD5        GGAAGTAAAAGTCGTAACAAGGTTTCCGTAGGTGAACCTGCGGAAGGATCATTACAGTGT 180
WAYANAD1        GGAAGTAAAAGTCGTAACAAGGTTTCCGTAGGTGAACCTGCGGAAGGATCATTACAGTGT 180
WAYANAD3        GGAAGTAAAAGTCGTAACAAGGTTTCCGTAGGTGAACCTGCGGAAGGATCATTACAGTGT 180
NARAYNAGAON2    GGAAGTAAAAGTCGTAACAAGGTTTCCGTAGGTGAACCTGCGGAAGGATCATTACAGTGT 180
VAISHALI4       GGAAGTAAAAGTCGTAACAAGGTTTCCGTAGGTGAACCTGCGGAAGGATCATTACAGTGT 180
VAISHALI1       GGAAGTAAAAGTCGTAACAAGGTTTCCGTAGGTGAACCTGCGGAAGGATCATTACAGTGT 180
JALGAON4        GGAAGTAAAAGTCGTAACAAGGTTTCCGTAGGTGAACCTGCGGAAGGATCATTACAGTGT 180
NRAYANGAON1     GGAAGTAAAAGTCGTAACAAGGTTTCCGTAGGTGAACCTGCGGAAGGATCATTACAGTGT 180
NARAYANGAON4    GGAAGTAAAAGTCGTAACAAGGTTTCCGTAGGTGAACCTGCGGAAGGATCATTACAGTGT 180
NARAYANGAON5    GGAAGTAAAAGTCGTAACAAGGTTTCCGTAGGTGAACCTGCGGAAGGATCATTACAGTGT 180
TRICHY1         GGAAGTAAAAGTCGTAACAAGGTTTCCGTAGGTGAACCTGCGGAAGGATCATTACAGTGT 180
NARAYAGAON3     GGAAGTAAAAGTCGTAACAAGGTTTCCGTAGGTGAACCTGCGGAAGGATCATTACAGTGT 180
JALGAON1        GGAAGTAAAAGTCGTAACAAGGTTTCCGTAGGTGAACCTGCGGAAGGATCATTACAGTGT 180
VAISHLAI2       GGAAGTAAAAGTCGTAACAAGGTTTCCGTAGGTGAACCTGCGGAAGGATCATTACAGTGT 180
JALGAON3        GGAAGTAAAAGTCGTAACAAGGTTTCCGTAGGTGAACCTGCGGAAGGATCATTACAGTGT 180
ASSAM2          GGAAGTAAAAGTCGTAACAAGGTTTCCGTAGGTGAACCTGCGGAAGGATCATTACAGTGT 180
ASSAM4          GGAAGTAAAAGTCGTAACAAGGTTTCCGTAGGTGAACCTGCGGAAGGATCATTACAGTGT 180
ASSAM3          GGAAGTAAAAGTCGTAACAAGGTTTCCGTAGGTGAACCTGCGGAAGGATCATTACAGTGT 180
ASSAM5          GGAAGTAAAAGTCGTAACAAGGTTTCCGTAGGTGAACCTGCGGAAGGATCATTACAGTGT 180
                ************************************************************

ASSAM1          GAAATTATCGTACGATAATAACGATCACGTCTGTTGTGCAGAGATCATACCTTGCACGAA 240
TRICHY3         GAAATTATCGTACGATAATAACGATCACGTCTGTTGTGCAGAGATCATACCTTGCACGAA 240
```

```
TRICHY5          GAAATTATCGTACGATAATAACGATCACGTCTGTTGTGCAGAGATCATACCTTGCACGAA 240
JALGAON5         GAAATTATCGTACGATAATAACGATCACGTCTGTTGTGCAGAGATCATACCTTGCACGAA 240
TRICHY2          GAAATTATCGTACGATAATAACGATCACGTCTGTTGTGCAGAGATCATACCTTGCACGAA 240
JALGAON2         GAAATTATCGTACGATAATAACGATCACGTCTGTTGTGCAGAGATCATACCTTGCACGAA 240
VAISHALI3        GAAATTATCGTACGATAATAACGATCACGTCTGTTGTGCAGAGATCATACCTTGCACGAA 240
VAISHALI5        GAAATTATCGTACGATAATAACGATCACGTCTGTTGTGCAGAGATCATACCTTGCACGAA 240
TRICHY4          GAAATTATCGTACGATAATAACGATCACGTCTGTTGTGCAGAGATCATACCTTGCACGAA 240
WAYANAD2         GAAATTATCGTACGATAATAACGATCACGTCTGTTGTGCAGAGATCATACCTTGCACGAA 240
WAYANAD4         GAAATTATCGTACGATAATAACGATCACGTCTGTTGTGCAGAGATCATACCTTGCACGAA 240
WAYANAD5         GAAATTATCGTACGATAATAACGATCACGTCTGTTGTGCAGAGATCATACCTTGCACGAA 240
WAYANAD1         GAAATTATCGTACGATAATAACGATCACGTCTGTTGTGCAGAGATCATACCTTGCACGAA 240
WAYANAD3         GAAATTATCGTACGATAATAACGATCACGTCTGTTGTGCAGAGATCATACCTTGCACGAA 240
NARAYNAGAON2     GAAATTATCGTACGATAATAACGATCACGTCTGTTGTGCAGAGATCATACCTTGCACGAA 240
VAISHALI4        GAAATTATCGTACGATAATAACGATCACGTCTGTTGTGCAGAGATCATACCTTGCACGAA 240
VAISHALI1        GAAATTATCGTACGATAATAACGATCACGTCTGTTGTGCAGAGATCATACCTTGCACGAA 240
JALGAON4         GAAATTATCGTACGATAATAACGATCACGTCTGTTGTGCAGAGATCATACCTTGCACGAA 240
NRAYANGAON1      GAAATTATCGTACGATAATAACGATCACGTCTGTTGTGCAGAGATCATACCTTGCACGAA 240
NARAYANGAON4     GAAATTATCGTACGATAATAACGATCACGTCTGTTGTGCAGAGATCATACCTTGCACGAA 240
NARAYANGAON5     GAAATTATCGTACGATAATAACGATCACGTCTGTTGTGCAGAGATCATACCTTGCACGAA 240
TRICHY1          GAAATTATCGTACGATAATAACGATCACGTCTGTTGTGCAGAGATCATACCTTGCACGAA 240
NARAYAGAON3      GAAATTATCGTACGATAATAACGATCACGTCTGTTGTGCAGAGATCATACCTTGCACGAA 240
JALGAON1         GAAATTATCGTACGATAATAACGATCACGTCTGTTGTGCAGAGATCATACCTTGCACGAA 240
VAISHLAI2        GAAATTATCGTACGATAATAACGATCACGTCTGTTGTGCAGAGATCATACCTTGCACGAA 240
JALGAON3         GAAATTATCGTACGATAATAACGATCACGTCTGTTGTGCAGAGATCATACCTTGCACGAA 240
ASSAM2           GAAATTATCGTACGATAATAACGATCACGTCTGTTGTGCAAAGATCATACCTTGCACGAA 240
ASSAM4           GAAATTATCGTACGATAATAACGATCACGTCTGTTGTGCAAAGATCATACCTTGCACGAA 240
ASSAM3           GAAATTATCGTACGATAATAACGATCACGTCTGTTGTGCAAAGATCATACCTTGCACGAA 240
ASSAM5           GAAATTATCGTACGATAATAACGATCACGTCTGTTGTGCAAAGATCATACCTTGCACGAA 240
                 ************************************* ******************
                                                   I
ASSAM1           AGCGTTTTACGCTTTTCTCGAACTAACTTTTGATGATGATGATG---------TTGTATA 291
TRICHY3          AGCGTTTTACGCTTTTCTCGAACTAACTTTTGATGATGATGATG---------TTGTATA 291
TRICHY5          AGCGTTTTACGCTTTTCTCGAACTAACTTTTGATGATGATGATG---------TTGTATA 291
JALGAON5         AGCGTTTTACGCTTTTCTCGAACTAACTTTTGATGATGATGATGATGATG---TTGTATA 297
TRICHY2          AGCGTTTTACGCTTTTCTCGAACTAACTTTTGATGATGATGATGATGATG---TTGTATA 297
JALGAON2         AGCGTTTTACGCTTTTCTCGAACTAACTTTTGATGATGATGATGATGATG---TTGTATA 297
VAISHALI3        AGCGTTTTACGCTTTTCTCGAACTAACTTTTGATGATGATGATGATGATG---TTGTATA 297
VAISHALI5        AGCGTTTTACGCTTTTCTCGAACTAACTTTTGATGATGATGATGATGATG---TTGTATA 297
TRICHY4          AGCGTTTTACGCTTTTCTCGAACTAACTTTTGATGATGATGATGATGATG---TTGTATA 297
WAYANAD2         AGCGTTTTACGCTTTTCTCGAACTAACTTTTGATGATGATGATG---------TTGTATA 291
WAYANAD4         AGCGTTTTACGCTTTTCTCGAACTAACTTTTGATGATGATGATG---------TTGTATA 291
WAYANAD5         AGCGTTTTACGCTTTTCTCGAACTAACTTTTGATGATGATGATG---------TTGTATA 291
WAYANAD1         AGCGTTTTACGCTTTTCTCGAACTAACTTTTGATGATGATGATG---------TTGTATA 291
WAYANAD3         AGCGTTTTACGCTTTTCTCGAACTAACTTTTGATGATGATGATG---------TTGTATA 291
NARAYNAGAON2     AGCGTTTTACGCTTTTCTCGAACTAACTTTTGATGATGATGATGATGATG---TTGTATA 297
VAISHALI4        AGCGTTTTACGCTTTTCTCGAACTAACTTTTGATGATGATGATGATGATG---TTGTATA 297
VAISHALI1        AGCGTTTTACGCTTTTCTCGAACTAACTTTTGATGATGATGATGATGATG---TTGTATA 297
JALGAON4         AGCGTTTTACGCTTTTCTCGAACTAACTTTTGATGATGATGATGATGATG---TTGTATA 297
NRAYANGAON1      AGCGTTTTACGCTTTTCTCGAACTAACTTTTGATGATGATGATGATGATG---TTGTATA 297
NARAYANGAON4     AGCGTTTTACGCTTTTCTCGAACTAACTTTTGATGATGATGATGATGATG---TTGTATA 297
NARAYANGAON5     AGCGTTTTACGCTTTTCTCGAACTAACTTTTGATGATGATGATGATGATG---TTGTATA 297
TRICHY1          AGCGTTTTACGCTTTTCTCGAACTAACTTTTGATGATGATGATGATGATG---TTGTATA 297
NARAYAGAON3      AGCGTTTTACGCTTTTCTCGAACTAACTTTTGATGATGATGATGATGATG---TTGTATA 297
JALGAON1         AGCGTTTTACGCTTTTCTCGAACTAACTTTTGATGATGATGATGATGATG---TTGTATA 297
VAISHLAI2        AGCGTTTTACGCTTTTCTCGAACTAACTTTTGATGATGATGATGATGATG---TTGTATA 297
JALGAON3         AGCGTTTTACGCTTTTCTCGAACTAACTTTTGATGATGATGATGATGATGATGTTGTATA 300
ASSAM2           AGCGTTTTACGCTTTTCTCGAACTAACTTTTGATGATGATGTTGT---------TGTATA 291
ASSAM4           AGCGTTTTACGCTTTTCTCGAACTAACTTTTGATGATGATGTTG------------TATA 288
ASSAM3           AGCGTTTTACGCTTTTCTCGAACTAACTTTTGATGATGATGTTG------------TATA 288
ASSAM5           AGCGTTTTACGCTTTTCTCGAACTAACTTTTGATGATGATGTTG------------TATA 288
                 ************************************* **           ****
                                                              (II)
ASSAM1           GTAGTAATAGTTTCTTACTAGTACT--ACTACT-CAACAACATCAACA------ACAACA 342
TRICHY3          GTAGTAATAGTTTCTTACTAGTACT--ACTACT-CAACAACATCAACA------ACAACA 342
TRICHY5          GTAGTAATAGTTTCTTACTAGTACT--ACTACT-CAACAACATCAACA------ACAACA 342
JALGAON5         GTAGTAATAGTTTCTTACTAGTACT--ACTACT-CAACAACATCAACA------ACAAC- 347
TRICHY2          GTAGTAATAGTTTCTTACTAGTACT--ACTACT-CAACAACATCAACA------ACAAC- 347
JALGAON2         GTAGTAATAGTTTCTTACTAGTACT--ACTACT-CAACAACATCAACA------ACAAC- 347
VAISHALI3        GTAGTAATAGTTTCTTACTAGTACT--ACTACT-CAACAACATCAACA------ACAACA 348
VAISHALI5        GTAGTAATAGTTTCTTACTAGTACT--ACTACT-CAACAACATCAACA------ACAACA 348
TRICHY4          GTAGTAATAGTTTCTTACTAGTACT--ACTACT-CAACAACATCAACA------ACAAC- 347
```

```
WAYANAD2          GTAGTAATAGTTTCTTACTAGTACT--ACTACT-CAACAACA---ACA------ACAAC- 338
WAYANAD4          GTAGTAATAGTTTCTTACTAGTACT--ACTACT-CAACAACA---ACA------ACAAC- 338
WAYANAD5          GTAGTAATAGTTTCTTACTAGTACT--ACTACT-CAACAACA---ACA------ACAAC- 338
WAYANAD1          GTAGTAATAGTTTCTTACTAGTACT--ACTACT-CAACAACA---ACA------ACAAC- 338
WAYANAD3          GTAGTAATAGTTTCTTACTAGTACT--ACTACT-CAACAACA---ACA------ACAAC- 338
NARAYNAGAON2      GTAGTAATAGTTTCTTACTAGTACT--ACTACT-CAACAACATCAACA------ACAAC- 347
VAISHALI4         GTAGTAATAGTTTCTTACTAGTACT--ACTACT-CAACAACATCAACA------ACAAC- 347
VAISHALI1         GTAGTAATAGTTTCTTACTAGTACT--ACTACT-CAACAACATCAACA------ACAAC- 347
JALGAON4          GTAGTAATAGTTTCTTACTAGTACT--ACTACT-CAACAACATCAACA------ACAAC- 347
NRAYANGAON1       GTAGTAATAGTTTCTTACTAGTACT--ACTACT-CAACAACATCAACA------ACAAC- 347
NARAYANGAON4      GTAGTAATAGTTTCTTACTAGTACT--ACTACT-CAACAACATCAACA------ACAAC- 347
NARAYANGAON5      GTAGTAATAGTTTCTTACTAGTACT--ACTACT-CAACAACATCAACA------ACAAC- 347
TRICHY1           GTAGTAATAGTTTCTTACTAGTACT--ACTACT-CAACAACATCAACA------ACAAC- 347
NARAYAGAON3       GTAGTAATAGTTTCTTACTAGTACT--ACTACT-CAACAACATCAACA------ACAAC- 347
JALGAON1          GTAGTAATAGTTTCTTACTAGTACT--ACTACT-CAACAACATCAACA------ACAAC- 347
VAISHLAI2         GTAGTAATAGTTTCTTACTAGTACT--ACTACT-CAACAACATCAACA------ACAACA 348
JALGAON3          GTAGTAATAGTTTCTTACTAGTACTCTACTACT-CAACAACATCAACA------ACAACA 353
ASSAM2            GTAGAAATAGTTTCTTACTAGTACT—-ACTACTACGACAACAACAACA ACAACAACAAC 348
ASSAM4            GTAGAAATAGTTTCTTACTAGTACT--ACTACTACGACAACAACAACAACAACAACAAC- 345
ASSAM3            GTAGAAATAGTTTCTTACTAGTACT--ACTACTACGACAACAACAACA------ACAAC- 339
ASSAM5            GTAGAAATAGTTTCTTACTAGTACT--ACTACTACGACAACAACAACA------ACAAC- 339
                  **** ********************  ****** * ******   ***      *****

ASSAM1            ACAACAACCGTATATACGTAGACGGTATATTAATAATAATTCGGGTACCTAGAAACCGTT 402
TRICHY3           ACAACAACCGTATATACGTAGACGGTATATTAATAATAATTCGGGTACCTAGAAACCGTT 402
TRICHY5           ACAACAACCGTATATACGTAGACGGTATATTAATAATAATTCGGGTACCTAGAAACCGTT 402
JALGAON5          ---------CGTATATACGTAGACGGTATATTAATAATAATTCGGGTACCTAGAAACCGTT 399
TRICHY2           ---------CGTATATACGTAGACGGTATATTAATAATAATTCGGGTACCTAGAAACCGTT 399
JALGAON2          ---------CGTATATACGTAGACGGTATATTAATAATAATTCGGGTACCTAGAAACCGTT 399
VAISHALI3         ACAACAACCGTATATACGTAGACGGTATATTAATAATAATTCGGGTACCTAGAAACCGTT 408
VAISHALI5         ACAACAACCGTATATACGTAGACGGTATATTAATAATAATTCGGGTACCTAGAAACCGTT 408
TRICHY4           ---------CGTATATACGTAGACGGTATATTAATAATAATTCGGGTACCTAGAAACCGTT 399
WAYANAD2          ---------CGTATATACGTAGACGGTATATTAATAATAATTCGGGTACCTAGAAACCGTT 390
WAYANAD4          ---------CGTATATACGTAGACGGTATATTAATAATAATTCGGGTACCTAGAAACCGTT 390
WAYANAD5          ---------CGTATATACGTAGACGGTATATTAATAATAATTCGGGTACCTAGAAACCGTT 390
WAYANAD1          ---------CGTATATACGTAGACGGTATATTAATAATAATTCGGGTACCTAGAAACCGTT 390
WAYANAD3          ---------CGTATATACGTAGACGGTATATTAATAATAATTCGGGTACCTAGAAACCGTT 390
NARAYNAGAON2      ---------CGTATATACGTAGACGGTATATTAATAATAATTCGGGTACCTAGAAACCGTT 399
VAISHALI4         ---------CGTATATACGTAGACGGTATATTAATAATAATTCGGGTACCTAGAAACCGTT 399
VAISHALI1         ---------CGTATATACGTAGACGGTATATTAATAATAATTCGGGTACCTAGAAACCGTT 399
JALGAON4          ---------CGTATATACGTAGACGGTATATTAATAATAATTCGGGTACCTAGAAACCGTT 399
NRAYANGAON1       ---------CGTATATACGTAGACGGTATATTAATAATAATTCGGGTACCTAGAAACCGTT 399
NARAYANGAON4      ---------CGTATATACGTAGACGGTATATTAATAATAATTCGGGTACCTAGAAACCGTT 399
NARAYANGAON5      ---------CGTATATACGTAGACGGTATATTAATAATAATTCGGGTACCTAGAAACCGTT 399
TRICHY1           ---------CGTATATACGTAGACGGTATATTAATAATAATTCGGGTACCTAGAAACCGTT 399
NARAYAGAON3       ---------CGTATATACGTAGACGGTATATTAATAATAATTCGGGTACCTAGAAACCGTT 399
JALGAON1          ---------CGTATATACGTAGACGGTATATTAATAATAATTCGGGTACCTAGAAACCGTT 399
VAISHLAI2         ACAACAACCGTATATACGTAGACGGTATATTAATAATAATTCGGGTACCTAGAAACCGTT 408
JALGAON3          ACAACAACCGTATATACGTAGACGGTATATTAATAATAATTCGGGTACCTAGAAACCGTT 413
ASSAM2            ---------CGTATATACGTAGACGGTATATTAATAATAATTCGGGTACCTAGAAACCGTT 400
ASSAM4            ---------CGTATATACGTAGACGGTATATTAATAATAATTCGGGTACCTAGAAACCGTT 397
ASSAM3            ---------CGTATATACGTAGACGGTATATTAATAATAATTCGGGTACCTAGAAACCGTT 391
ASSAM5            ---------CGTATATACGTAGACGGTATATTAATAATAATTCGGGTACCTAGAAACCGTT 391
                          **************************************************
                            (III)                                      (IV)
ASSAM1            ATCTCCGATCGCGCGCATAGCTATGCGATCGTAGTTTGACGATCGAAGAATAATGTTGTT 462
TRICHY3           ATCTCCGATCGCGCGCATAGCTATGCGATCGTAGTTTGACGATCGAAGAATAATGTTGTT 462
TRICHY5           ATCTCCGATCGCGCGCATAGCTATGCGATCGTAGTTTGACGATCGAAGAATAATGTTGTT 462
JALGAON5          ATCTCCGATCGCGC--ATAGCTATGCGATCGTAGTTTGACGATCGAAGAATAATGTTGTT 457
TRICHY2           ATCTCCGATCGCGC--ATAGCTATGCGATCGTAGTTTGACGATCGAAGAATAATGTTGTT 457
JALGAON2          ATCTCCGATCGCGC--ATAGCTATGCGATCGTAGTTTGACGATCGAAGAATAATGTTGTT 457
VAISHALI3         ATCTCCGATCGCGCGAATAGCTATGCGATCGTAGTTTGACGATCGAAGAATAATGTTGTT 468
VAISHALI5         ATCTCCGATCGCGCGAATAGCTATGCGATCGTAGTTTGACGATCGAAGAATAATGTTGTT 468
TRICHY4           ATCTCCGATCGCGCGCATAGCTATGCGATCGTAGTTTGACGATCGAAGAATAATGTTGTT 459
WAYANAD2          ATCTCCGATCGCGCGCATAGCTATGCGATCGTAGTTTGACGATCGAAGAATAATGTTGTT 450
WAYANAD4          ATCTCCGATCGCGCGCATAGCTATGCGATCGTAGTTTGACGATCGAAGAATAATGTTGTT 450
WAYANAD5          ATCTCCGATCGCGCGCATAGCTATGCGATCGTAGTTTGACGATCGAAGAATAATGTTGTT 450
WAYANAD1          ATCTCCGATCGCGCGCATAGCTATGCGATCGTAGTTTGACGATCGAAGAATAATGTTGTT 450
WAYANAD3          ATCTCCGATCGCGCGCATAGCTATGCGATCGTAGTTTGACGATCGAAGAATAATGTTGTT 450
NARAYNAGAON2      ATCTCCGATCGCGCGCATAGCTATGCGATCGTAGTTTGACGATCGAAGAATAATGTTGTT 459
VAISHALI4         ATCTCCGATCGCGCGCATAGCTATGCGATCGTAGTTTGACGATCGAAGAATAATGTTGTT 459
```

```
VAISHALI1         ATCTCCGATCGCGCGCATAGCTATGCGATCGTAGTTTGACGATCGAAGAATAATGTTGTT 459
JALGAON4          ATCTCCGATCGCGCGCATAGCTATGCGATCGTAGTTTGACGATCGAAGAATAATGTTGTT 459
NRAYANGAON1       ATCTCCGATCGCGCGCATAGCTATGCGATCGTAGTTTGACGATCGAAGAATAATGTTGTT 459
NARAYANGAON4      ATCTCCGATCGCGCGCATAGCTATGCGATCGTAGTTTGACGATCGAAGAATAATGTTGTT 459
NARAYANGAON5      ATCTCCGATCGCGCGCATAGCTATGCGATCGTAGTTTGACGATCGAAGAATAATGTTGTT 459
TRICHY1           ATCTCCGATCGCGCGCATAGCTATGCGATCGTAGTTTGACGATCGAAGAATAATGTTGTT 459
NARAYAGAON3       ATCTCCGATCGCGCGCATAGCTATGCGATCGTAGTTTGACGATCGAAGAATAATGTTGTT 459
JALGAON1          ATCTCCGATCGCGCGCATAGCTATGCGATCGTAGTTTGACGATCGAAGAATAATGTTGTT 459
VAISHLAI2         ATCTCCGATCGCGCGCATAGCTATGCGATCGTAGTTTGACGATCGAAGAATAATGTTGTT 468
JALGAON3          ATCTCCGATCGCGCGCATAGCTATGCGATCGTAGTTTGACGATCGAAGAATAATGTTGTT 473
ASSAM2            ATCTCCGAACGCGC--ATAGCTATGCGATCGTAGTTTGACGATCGAAGAATAATGT---T 455
ASSAM4            ATCTCCGATCGCGC--ATAGCTATGCGATCGTAGTTTGACGATCGAAGAATAATGT---T 452
ASSAM3            ATCTCCGATCGCGC--ATAGCTATGCGATCGTAGTTTGACGATCGAAGAATAATGT---T 446
ASSAM5            ATCTCCGATCGCGC--ATAGCTATGCGATCGTAGTTTGACGATCGAAGAATAATGT---T 446
                  ******** *****  ****************************************   *

ASSAM1            AATAATAATAACAACGTTATTCGATGGTATTTAAATGATTTTCGCCCGAATATTATACGC 522
TRICHY3           AATAATAATAACAACGTTATTCGATGGTATTTAAATGATTTTCGCCCGAATATGATACGC 522
TRICHY5           AATAATAATAACAACGTTATTCGATGGTATTTAAATGATTTTCGCCCGAATATTATACGC 522
JALGAON5          AATAATAATAACAACGTTATTCGATGGTATTTAAATGATTTTCGCCCGAATATTATACGC 517
TRICHY2           AATAATAATAACAACGTTATTCGATGGTATTTAAATGATTTTCGCCCGAATATTATACGC 517
JALGAON2          AATAATAATAACAACGTTATTCGATGGTATTTAAATGATTTTCGCCCGAATATTATACGC 517
VAISHALI3         AATAATAATAACAACGTTATTCGATGGTATTTAAATGATTTTCGCCCGAATATTATACGC 528
VAISHALI5         AATAATAATAACAACGTTATTCGATGGTATTTAAATGATTTTCGCCCGAATATTATACGC 528
TRICHY4           AATAATAATAACAACGTTATTCGATGGTATTTAAATGATTTTCGCCCGAATATTATACGC 519
WAYANAD2          AATAATAATAACAACGTTATTCGATGGTATTTAAATGATTTTCGCCCGAATATTATACGC 510
WAYANAD4          AATAATAATAACAACGTTATTCGATGGTATTTAAATGATTTTCGCCCGAATATTATACGC 510
WAYANAD5          AATAATAATAACAACGTTATTCGATGGTATTTAAATGATTTTCGCCCGAATATTATACGC 510
WAYANAD1          AATAATAATAACAACGTTATTCGATGGTATTTAAATGATTTTCGCCCGAATATTATACGC 510
WAYANAD3          AATAATAATAACAACGTTATTCGATGGTATTTAAATGATTTTCGCCCGAATATTATACGC 510
NARAYNAGAON2      AATAATAATAACAACGTTATTCGATGGTATTTAAATGATTTTCGCCCGAATATTATACGC 519
VAISHALI4         AATAATAATAACAACGTTATTCGATGGTATTTAAATGATTTTCGCCCGAATATTATACGC 519
VAISHALI1         AATAATAATAACAACGTTATTCGATGGTATTTAAATGATTTTCGCCCGAATATTATACGC 519
JALGAON4          AATAATAATAACAACGTTATTCGATGGTATTTAAATGATTTTCGCCCGAATATTATACGC 519
NRAYANGAON1       AATAATAATAACAACGTTATTCGATGGTATTTAAATGATTTTCGCCCGAATATTATACGC 519
NARAYANGAON4      AATAATAATAACAACGTTATTCGATGGTATTTAAATGATTTTCGCCCGAATATTATACGC 519
NARAYANGAON5      AATAATAATAACAACGTTATTCGATGGTATTTAAATGATTTTCGCCCGAATATTATACGC 519
TRICHY1           AATAATAATAACAACGTTATTCGATGGTATTTAAATGATTTTCGCCCGAATATTATACGC 519
NARAYAGAON3       AATAATAATAACAACGTTATTCGATGGTATTTAAATGATTTTCGCCCGAATATTATACGC 519
JALGAON1          AATAATAATAACAACGTTATTCGATGGTATTTAAATGATTTTCGCCCGAATATTATACGC 519
VAISHLAI2         AATAATAATAACAACGTTATTCGATGGTATTTAAATGATTTTCGCCCGAATATTATACGC 528
JALGAON3          AATAATAATAACAACGTTATTCGATGGTATTTAAATGATTTTCGCCCGAATATTATACGC 533
ASSAM2            AATAATAATAACAACGTTATTCGATGGTATTTAAATGATTTTCGCCCGAATATTATACG- 514
ASSAM4            AATAATAATAACAACGTTATTCGATGGTATTTAAATGATTTTCGCCCGAATATTATACG- 511
ASSAM3            AATAATAATAACAACGTTATTCGATGGTATTTAAATGATTTTCGCCCGAATATTATACG- 505
ASSAM5            AATAATAATAACAACGTTATTCGATGGTATTTAAATGATTTTCGCCCGAATATTATACG- 505
                  ************************************************** *****

ASSAM1            GTCTCGTATGGTTGAAAAATTTAACCGTTTTTCATAATTATATATACGTCTCGGATGATA 582
TRICHY3           GTCTCGTATGGTTGAAAAATTTAACCGTTTTTCATAATTATATATACGTCTCGGATGATA 582
TRICHY5           GTCTCGTATGGTTGAAAAATTTAACCGTTTTTCATAATTATATATACGTCTCGGATGATA 582
JALGAON5          GTCTCGTATGGTTGAAAAATTTAACCGTTTTTCATAATTATATATACGTCTCGGATGATA 577
TRICHY2           GTCTCGTATGGTTGAAAAATTTAACCGTTTTTCATAATTATATATACGTCTCGGATGATA 577
JALGAON2          GTCTCGTATGGTTGAAAAATTTAACCGTTTTTCATAATTATATATACGTCTCGGATGATA 577
VAISHALI3         GTCTCGTATGGTTGAAAAATTTAACCGTTTTTCATAATTATATATACGTCTCGGATGATA 588
VAISHALI5         GTCTCGTATGGTTGAAAAATTTAACCGTTTTTCATAATTATATATACGTCTCGGATGATA 588
TRICHY4           GTCTCGTATGGTTGAAAAATTTAACCGTTTTTCATAATTATATATACGTCTCGGATGATA 579
WAYANAD2          GTCTCGTATGGTTGAAAAATTTAACCGTTTTTCATAATTATATATACGTCTCGGATGATA 570
WAYANAD4          GTCTCGTATGGTTGAAAAATTTAACCGTTTTTCATAATTATATATACGTCTCGGATGATA 570
WAYANAD5          GTCTCGTATGGTTGAAAAATTTAACCGTTTTTCATAATTATATATACGTCTCGGATGATA 570
WAYANAD1          GTCTCGTATGGTTGAAAAATTTAACCGTTTTTCATAATTATATATACGTCTCGGATGATA 570
WAYANAD3          GTCTCGTATGGTTGAAAAATTTAACCGTTTTTCATAATTATATATACGTCTCGGATGATA 570
NARAYNAGAON2      GTCTCGTATGGTTGAAAAATTTAACCGTTTTTCATAATTATATATACGTCTCGGATGATA 579
VAISHALI4         GTCTCGTATGGTTGAAAAATTTAACCGTTTTTCATAATTATATATACGTCTCGGATGATA 579
VAISHALI1         GTCTCGTATGGTTGAAAAATTTAACCGTTTTTCATAATTATATATACGTCTCGGATGATA 579
JALGAON4          GTCTCGTATGGTTGAAAAATTTAACCGTTTTTCATAATTATATATACGTCTCGGATGATA 579
NRAYANGAON1       GTCTCGTATGGTTGAAAAATTTAACCGTTTTTCATAATTATATATACGTCTCGGATGATA 579
NARAYANGAON4      GTCTCGTATGGTTGAAAAATTTAACCGTTTTTCATAATTATATATACGTCTCGGATGATA 579
NARAYANGAON5      GTCTCGTATGGTTGAAAAATTTAACCGTTTTTCATAATTATATATACGTCTCGGATGATA 579
TRICHY1           GTCTCGTATGGTTGAAAAATTTAACCGTTTTTCATAATTATATATACGTCTCGGATGATA 579
NARAYAGAON3       GTCTCGTATGGTTGAAAAATTTAACCGTTTTTCATAATTATATATACGTCTCGGATGATA 579
```

```
JALGAON1          GTCTCGTATGGTTGAAAAATTTAACCGTTTTTCATAATTATATATACGTCTCGGATGATA 579
VAISHLAI2         GTCTCGTATGGTTGAAAAATTTAACCGTTTTTCATAATTATATATACGTCTCGGATGATA 588
JALGAON3          GTCTCGTATGGTTGAAAAATTTAACCGTTTTTCATAATTATATATACGTCTCGGATGATA 593
ASSAM2            -TCTCGTATGGTTGAAGAATTTAACCGTTTTTCATAATTATATATACGTCTCGGATGATA 573
ASSAM4            -TCTCGTATGGTTGAAAAATTTAACCGTTTTTCATAATTATATATACGTCTCGGATGATA 570
ASSAM3            -TCTCGTATGGTTGAAAAATTTAACCGTTTTTCATAATTATATATACGTCTCGGATGATA 564
ASSAM5            -TCTCGTATGGTTGAAAAATTTAACCGTTTTTCATAATTATATATACGTCTCGGATGATA 564
                   *************** ****************************************
                                                                  (V)
ASSAM1            ATTAATCATAATCATCTCTCATTCGACTACGACGAATATTATATATATATA--AAAACGA 640
TRICHY3           ATTAATCATAATCATCTCTCATTCGACTACGACGAATATTATATATATATA--AAAACGA 640
TRICHY5           ATTAATCATAATCATCTCTCATTCGACTACGACGAATATTATATATATATATAAAAACGA 642
JALGAON5          ATTAATCATAATCATCTCTCATTCGACTACGACGAATATTATATATATATATAAAAACGA 637
TRICHY2           ATTAATCATAATCATCTCTCATTCGACTACGACGAATATTATATATATATATAAAAACGA 637
JALGAON2          ATTAATCATAATCATCTCTCATTCGACTACGACGAATATTATATATATATATAAAAACGA 637
VAISHALI3         ATTAATCATAATCATCTCTCATTCGACTACGACGAATATTATATATATATATAAAAACGA 648
VAISHALI5         ATTAATCATAATCATCTCTCATTCGACTACGACGAATATTATATATATATATAAAAACGA 648
TRICHY4           ATTAATCATAATCATCTCTCATTCGACTACGACGAATATTATATATATATA-AAAAACGA 638
WAYANAD2          ATTAATCATAATCATCTCTCATTCGACTACGACGAATATTATATATATA----AAAACGA 626
WAYANAD4          ATTAATCATAATCATCTCTCATTCGACTACGACGAATATTATATATATA----AAAACGA 626
WAYANAD5          ATTAATCATAATCATCTCTCATTCGACTACGACGAATATTATATATATA----AAAACGA 626
WAYANAD1          ATTAATCATAATCATCTCTCATTCGACTACGACGAATATTATATATATA----AAAACGA 626
WAYANAD3          ATTAATCATAATCATCTCTCATTCGACTACGACGAATATTATATATATA----AAAACGA 626
NARAYNAGAON2      ATTAATCATAATCATCTCTCATTCGACTACGACGAATATTATATATATATA--AAAACGA 637
VAISHALI4         ATTAATCATAATCATCTCTCATTCGACTACGACGAATATTATATATATATA--AAAACGA 637
VAISHALI1         ATTAATCATAATCATCTCTCATTCGACTACGACGAATATTATATATATAT---AAAACGA 636
JALGAON4          ATTAATCATAATCATCTCTCATTCGACTACGACGAATATTATATATATATA--AAAACGA 637
NRAYANGAON1       ATTAATCATAATCATCTCTCATTCGACTACGACGAATATTATATATATATA--AAAACGA 637
NARAYANGAON4      ATTAATCATAATCATCTCTCATTCGACTACGACGAATATTATATATATATA--AAAACGA 637
NARAYANGAON5      ATTAATCATAATCATCTCTCATTCGACTACGACGAATATTATATATATATA--AAAACGA 637
TRICHY1           ATTAATCATAATCATCTCTCATTCGACTACGACGAATATTATATATATATA--AAAACGA 637
NARAYAGAON3       ATTAATCATAATCATCTCTCATTCGACTACGACGAATATTATATATATATA--AAAACGA 637
JALGAON1          ATTAATCATAATCATCTCTCATTCGACTACGACGAATATTATATATATATA--AAAACGA 637
VAISHLAI2         ATTAATCATAATCATCTCTCATTCGACTACGACGAATATTATATATATATA--AAAACGA 646
JALGAON3          ATTAATCATAATCATCTCTCATTCGACTACGACGAATATTATATATATATA--AAAACGA 651
ASSAM2            ATTAATCATAATCATCTCTCATTCGACTGCGACGAATATTATATATATA----AAAACGA 629
ASSAM4            ATTAATCATAATCATCTCTCATTCGACTGCGACGAATATTATATATATA----AAAACGA 626
ASSAM3            ATTAATCATAATCATCTCTCATTCGACTGCGACGAATATTATATATATA----AAAACGA 620
ASSAM5            ATTAATCATAATCATCTCTCATTCGACTGCGACGAATATTATATATATA----AAAACGA 620
                  *************************** ********************    *******

ASSAM1            GAAAATACATACACCGCTCGTCTGTTATTTTTACAATATGACGGTGTTGGTTTTTCGTTT 700
TRICHY3           GAAAATACATACACCGCTCGTCTGTTATTTTTACAATATGACGGTGTTGGTTTTTCGTTT 700
TRICHY5           GAAAATACATACACCGCTCGTCTGTTATTTTTACAATATGACGGTGTTGGTTTTTCGTTT 702
JALGAON5          GAAAATACATACACCGCTCGTCTGTTATTTTTACAATATGACGGTGTTGGTTTTTCGTTT 697
TRICHY2           GAAAATACATACACCGCTCGTCTGTTATTTTTACAATATGACGGTGTTGGTTTTTCGTTT 697
JALGAON2          GAAAATACATACACCGCTCGTCTGTTATTTTTACAATATGACGGTGTTGGTTTTTCGTTT 697
VAISHALI3         GAAAATACATACACCGCTCGTCTGTTATTTTTACAATATGACGGTGTTGGTTTTTCGTTT 708
VAISHALI5         GAAAATACATACACCGCTCGTCTGTTATTTTTACAATATGACGGTGTTGGTTTTTCGTTT 708
TRICHY4           GAAAATACATACACCGCTCGTCTGTTATTTTTACAATATGACGGTGTTGGTTTTTCGTTT 698
WAYANAD2          GAAAATACATACACCGCTCGTCTGTTATTTTTACAACATGACGGTGTTGGTTTTTCGTTT 686
WAYANAD4          GAAGATACATACACCGCTCGTCTGTTATTTTTACAATATGACGGTGTTGGTTTTTCGTTT 686
WAYANAD5          GAAAATACATACACCGCTCGTCTGTTATTTTTACAATATGACGGTGTTGGTTTTTCGTTT 686
WAYANAD1          GAAAATACATACACCGCTCGTCTGTTATTTTTACAATATGACGGTGTTGGTTTTTCGTTT 686
WAYANAD3          GAAAATACATACACCGCTCGTCTGTTATTTTTACAATATGACGGTGTTGGTTTTTCGTTT 686
NARAYNAGAON2      GAAAATACATACACCGCTCGTCTGTTATTTTTACAATATGACGGTGTTGGTTTTTCGTTT 697
VAISHALI4         GAAAATACATACACCGCTCGTCTGTTATTTTTACAATATGACGGTGTTGGTTTTTCGTTT 697
VAISHALI1         GAAAATACATACACCGCTCGTCTGTTATTTTTACAATATGACGGTGTTGGTTTTTCGTTT 696
JALGAON4          GAAAATACATACACCGCTCGTCTGTTATTTTTACAATATGACGGTGTTGGTTTTTCGTTT 697
NRAYANGAON1       GAAAATACATACACCGCTCGTCTGTTATTTTTACAATATGACGGTGTTGGTTTTTCGTTT 697
NARAYANGAON4      GAAAATACATACACCGCTCGTCTGTTATTTTTACAATATGACGGTGTTGGTTTTTCGTTT 697
NARAYANGAON5      GAAAATACATACACCGCTCGTCTGTTATTTTTACAATATGACGGTGTTGGTTTTTCGTTT 697
TRICHY1           GAAAATACATACACCGCTCGTCTGTTATTTTTACAATATGACGGTGTTGGTTTTTCGTTT 697
NARAYAGAON3       GAAAATACATACACCGCTCGTCTGTTATTTTTACAATATGACGGTGTTGGTTTTTCGTTT 697
JALGAON1          GAAAATACATACACCGCTCGTCTGTTATTTTTACAATATGACGGTGTTGGTTTTTCGTTT 697
VAISHLAI2         GAAAATACATACACCGCTCGTCTGTTATTTTTACAATATGACGGTGTTGGTTTTTCGTTT 706
JALGAON3          GAAAATACATACACCGCTCGTCTGTTATTTTTACAATATGACGGTGTTGGTTTTTCGTTT 711
ASSAM2            GAAAAGACATACACCGCTCGTCTGTTATTTTTACAATATGACGGTGTTGGTTTTTCGTTT 689
ASSAM4            GAAAAGACATACACCGCTCGTCTGTTATTTTTACAATATGACGGTGTTGGTTTTTCGTTT 686
ASSAM3            GAAAAGACATACACCGCTCGTCTGTTATTTTTACAATATGACGGTGTTGGTTTTTCGTTT 680
ASSAM5            GAAAAGACATACACCGCTCGTCTGTTATTTTTACAATATGACGGTGTTGGTTTTTCGTTT 680
```

```
                 *** * ***************************** ***********************
                       (VI)
ASSAM1           TTCTTCGTCGTCG---------AATATGATGATGATAATGTCCTGTAGACGTTTTCGAAA 751
TRICHY3          TTCTTCGTCGTCG---------AATATGATGATGATAATGTCCTGTAGACGTTTTCGAAA 751
TRICHY5          TTCTTCGTCGTCG---------AATATGATGATGATAATGTCCTGTAGACGTTTTCGAAA 753
JALGAON5         TTCTTCGTCGTCA---------AATATGATGATGATAATGTCCTGTAGACGTTTTCGAAA 748
TRICHY2          TTCTTCGTCGTCG---------AATATGATGATGATAATGTCCTGTAGACGTTTTCGAAA 748
JALGAON2         TTCTTCGTCGTCG---------AATATGATGATGATAATGTCCTGTAGACGTTTTCGAAA 748
VAISHALI3        TTCTTAGTCGTCG---------AATATGATGATGATAATGTCCTGTAGACGTTTTCGAAA 759
VAISHALI5        TTCTTAGTCGTCG---------AATATGATGATGATAATGTCCTGTAGACGTTTTCGAAA 759
TRICHY4          TTCTTCGTCGTCG---------AATATGATGATGATAATGTCCTGTAGACGTTTTCGAAA 749
WAYANAD2         TTCTTCGTCGTCG---------AATATGATGATTATAATGTCCTGTAGACGTTTTCGAAA 737
WAYANAD4         TTCTTCGTCGTCG---------AATATGATGATTATAATGTCCTGTAGACGTTTTCGAAA 737
WAYANAD5         TTCTTCGTCGTCG---------AATATGATGATTATAATGTCCTGTAGACGTTTTCGAAA 737
WAYANAD1         TTCTTCGTCGTCG---------AATATGATGATTATAATGTCCTGTAGACGTTTTCGAAA 737
WAYANAD3         TTCTTCGTCGTCG---------AATATGATGATTATAATGTCCTGTAGACGTTTTCGAAA 737
NARAYNAGAON2     TTCTTCGTCGTCG---------AATATGATGATGATAATGTCCTGTAGACGTTTTCGAAA 748
VAISHALI4        TTCTTCGTCGTCG---------AATATGATGATGATAATGTCCTGTAGACGTTTTCGAAA 748
VAISHALI1        TTCTTCGTCGTCG---------AATATGATGATGATAATGTCCTGTAGACGTTTTCGAAA 747
JALGAON4         TTCTTCGTCGTCG---------AATATGATGATGATAATGTCCTGTAGACGTTTTCGAAA 748
NRAYANGAON1      TTCTTCGTCGTCG---------AATATGATGATGATAATGTCCTGTAGACGTTTTCGAAA 748
NARAYANGAON4     TTCTTCGTCGTCG---------AATATGATGATGATAATGTCCTGTAGACGTTTTCGAAA 748
NARAYANGAON5     TTCTTCGTCGTCG---------AATATGATGATGATAATGTCCTGTAGACGTTTTCGAAA 748
TRICHY1          TTCTTCGTCGTCG---------AATATGATGATGATAATGTCCTGTAGACGTTTTCGAAA 748
NARAYAGAON3      TTCTTCGTCGTCG---------AATATGATGATGATAATGTCCTGTAGACGTTTTCGAAA 748
JALGAON1         TTCTTCGTCGTCG---------AATATGATGATGATAATGTCCTGTAGACGTTTTCGAAA 748
VAISHLAI2        TTCTTCGTCGTCG---------AATATGATGATGATAATGTCCTGTAGACGTTTTCGAAA 757
JALGAON3         TTCTTCGTCGTCG---------AATATGATGATGATAATGTCCTGTAGACGTTTTCGAAA 762
ASSAM2           TTCTTCGTCGTCGTCGTCGTCGAATATGATGATGATAGTGTCCTGTAGACGTTTTCGAAA 749
ASSAM4           TTCTTCGTCGTCGTCGTCG---AATATGATGATGATAGTGTCCTGTAGACGTTTTCGAAA 743
ASSAM3           TTCTTCGTCGTCGTCGTCGTCGAATATGATGATGATAGTGTCCTGTAGACGTTTTCGAAA 740
ASSAM5           TTCTTCGTCGTCGTCGTCG---AATATGATGATGATAGTGTCCTGTAAACGTTTTCGAAA 737
                 ***** ******          ************ *** ********* ************
                              (VII)                                  (VIII)
ASSAM1           AACGGGAGATATATATATATATAT----CGAATGCCAATCACCCTTAAAGAAATTATG-- 805
TRICHY3          AACGGGAGATATATATATATATATA--TCGAATGCCAATCACCCTTAAAGAAATTATG-- 807
TRICHY5          AACGGGAGATATATATATATATATA--TCGAATGCCAATCACCCTTAAAGAAATTATG-- 809
JALGAON5         AACGGGAGATATATATATATATATA--TCGAATGCCAATCACCCTTAAAGAAATTATG-- 804
TRICHY2          AACGGGAGATATATATATATATA----TCGAATGCCAATCACCCTTAAAGAAATTATG-- 802
JALGAON2         AACGGGAGATATATATATATATATA--TCGAATGCCAATCACCCTTAAAGAAATTATG-- 804
VAISHALI3        AACGGGAGATATATATATATATATA--TCGAATGCCAATCACCCTTAAAGAAATTATG-- 815
VAISHALI5        AACGGGAGATATATATATATATATA--TCGAATGCCAATCACCCTTAAAGAAATTATG-- 815
TRICHY4          AACGGGAGATATATATATATATATA--TCGAATGCCAATCACCCTTAAAGAAATTATG-- 805
WAYANAD2         AACGGGAGATATATATATATATATATCGAATGCCAATCACCCTTAAAGAAATTATG-- 795
WAYANAD4         AACGGGAGATATATATATATATATATCGAATGCCAATCACCCTTAAAGAAATTATG-- 795
WAYANAD5         AACGGGAGATATATATATATATATATCGAATGCCAATCACCCTTAAAGAAATTATG-- 795
WAYANAD1         AACGGGAGATATATATATATATATATCGAATGCCAATCACCCTTAAAGAAATTATG-- 795
WAYANAD3         AACGGGAGATATATATATATATATATCGAATGCCAATCACCCTTAAAGAAATTATG-- 795
NARAYNAGAON2     AACGGGAGATATATATATATATATATCGAATGCCAATCACCCTTAAAGAAATTATG-- 806
VAISHALI4        AACGGGAGATATATATATATATATATCGAATGCCAATCACCCTTAAAGAAATTATG-- 806
VAISHALI1        AACGGGAGATATATATATATATATA--TCGAATGCCAATCACCCTTAAAGAAATTATG-- 803
JALGAON4         AACGGGAGATATATATATATATATA--TCGAATGCCAATCACCCTTAAAGAAATTATG-- 804
NRAYANGAON1      AACGGGAGATATATATATATATATA--TCGAATGCCAATCACCCTTAAAGAAATTATG-- 804
NARAYANGAON4     AACGGGAGATATATATATATATATA--TCGAATGCCAATCACCCTTAAAGAAATTATG-- 804
NARAYANGAON5     AACGGGAGATATATATATATATATA--TCGAATGCCAATCACCCTTAAAGAAATTATG-- 804
TRICHY1          AACGGGAGATATATATATATATATA--TCGAATGCCAATCACCCTTAAAGAAATTATG-- 804
NARAYAGAON3      AACGGGAGATATATATATATATA----TCGAATGCCAATCACCCTTAAAGAAATTATG-- 802
JALGAON1         AACGGGAGATATATATATATATA----TCGAATGCCAATCACCCTTAAAGAAATTATG-- 802
VAISHLAI2        AACGGGAGATATATATATATATA----TCGAATGCCAATTACCCTTAAAGAAATTATG-- 811
JALGAON3         AACGGGAGATATATATATATATA----TCGAATGCCAATCACCCTTAAAGAAATTATG-- 816
ASSAM2           AACGGGAGATATATATGTATATA----TCGAATGCCAATCACCCTTAAAGAAATTATGAT 805
ASSAM4           AACGGGAGATATATATGTATATA----TCGAATGCCAATCACCCTTAAAGAAATTATGAT 799
ASSAM3           AACGGGAGATATATATGTATATA----TCGAATGCCAATCACCCTTAAAGAAATTATGAT 796
ASSAM5           AACGGGAGATATATATGTATATA----TCAAATGCCAATCACCCTTAAAGAAATTATGAT 793
                 *************** ******     * ********* ******************
                                                                       (IX)
ASSAM1           -TAATTTCTTAGGAAAACGACGCATTGGCGAATCGTATTGACGACGAC-ACACACACACA 863
TRICHY3          -TAATTTCTTAGGAAAACGACGCATTGGCGAATCGTATTGACGACGAC-ACACACACACA 865
TRICHY5          -TAATTTCTTAGGAAAACGACGCATTGGCGAATCGTATTGACGACGAC-ACACACACACA 867
```

```
JALGAON5          -TAATTTCTTAGGAAAACGACGCATTGGCGAATCGTATTGACGACGAC-ACACACACACA 862
TRICHY2           -TAATTTCTTAGGAAAACGACGCATTGGCGAATCGTATTGACGACGAC-ACACACACACA 860
JALGAON2          -TAATTTCTTAGGAAAACGACGCATTGGCGAATCGTATTGACGACGAC-ACACACACACA 862
VAISHALI3         -TAATTTCTTAGGAAAACGACGCATTGGCGAATCGTATTGACGACGAC-ACACACACACA 873
VAISHALI5         -TAATTTCTTAGGAAAACGACGCATTGGCGAATCGTATTGACGACGAC-ACACACACACA 873
TRICHY4           -TAATTTCTTAGGAAAACGACGCATTGGCGAATCGTATTGACAACGAC-ACACACACACA 863
WAYANAD2          -TAATTTCTTAGGAAAACGACGCATTGGCGAATCGTATTGACGACGAC-ACACACACACA 853
WAYANAD4          -TAATTTCTTAGGAAAACGACGCATTGGCGAATCGTATTGACGACGAC-ACACACACACA 853
WAYANAD5          -TAATTTCTTAGGAAAACGACGCATTGGCGAATCGTATTGACGACGAC-ACACACACACA 853
WAYANAD1          -TAATTTCTTAGGAAAACGACGCATTGGCGAATCGTATTGACGACGAC-ACACACACACA 853
WAYANAD3          -TAATTTCTTAGGAAAACGACGCATTGGCGAATCGTATTGACGACGAC-ACACACACACA 853
NARAYNAGAON2      -TAATTTCTTAGGAAAACGACGCATTGGCGAATCGTATTGACGACGAC-ACACACACACA 864
VAISHALI4         -TAATTTCTTAGGAAAACGACGCATTGGCGAATCGTATTGACGACGAC-ACACACACACA 864
VAISHALI1         -TAATTTCTTAGGAAAACGACGCATTGGCGAATCGTATTGACGACGAC-ACACACACACA 861
JALGAON4          -TAATTTCTTAGGAAAACGACGCATTGGCGAATCGTATTGACGACGAC-ACACACACACA 862
NRAYANGAON1       -TAATTTCTTAGGAAAACGACGCATTGGCGAATCGTATTGACGACGAC-ACACACACACA 862
NARAYANGAON4      -TAATTTCTTAGGAAAACGACGCATTGGCGAATCGTATTGACGACGAC-ACACACACACA 862
NARAYANGAON5      -TAATTTCTTAGGAAAACGACGCATTGGCGAATCGTATTGACGACGAC-ACACACACACA 862
TRICHY1           -TAATTTCTTAGGAAAACGACGCATTGGCGAATCGTATTGACGACGAC-ACACACACACA 862
NARAYAGAON3       -TAATTTCTTAGGAAAACGACGCATTGGCGAATCGTATTGACGACGAC-ACACACACACA 860
JALGAON1          -TAATTTCTTAGGAAAACGACGCATTGGCGAATCGTATTGACGACGAC-ACACACACACA 860
VAISHLAI2         -TAATTTCTTAGGAAAACGACGCATTGGCGAATCGTATTGACGACGACGACACACACACA 870
JALGAON3          -TAATTTCTTAGGAAAACGACGCATTGGCGAATCGTATTGACGGCGAC-ACACACACACA 874
ASSAM2            GTAATTTCTTAGGAAAACGACGCATTGGCGAATCGTATTGACGACGAC-ACACACACACA 864
ASSAM4            GTAATTTCTTAGGAAAACGGCGCATTGGCGAATCGTATTGACGACGAC-ACACACACAC- 857
ASSAM3            GTAATTTCTTAGGAAAACGACGCATTGGCGAATCGTATTGACGACGAC-ACACACACACA 855
ASSAM5            GTAATTTCTTAGGAAAACGACGCATTGGCGAATCGTATTGACGACGAC-ACACACACAC- 851
                   ****************** **********************  **** **********



ASSAM1            CAC----GTGTTCGACGTTACGAACGATCAGCTGAG 895
TRICHY3           CAC----GTGTTCGACGTTACGAACGATCAGCTGAG 897
TRICHY5           CAC----GTGTTCGACGTTACGAACGATCAGCTGAG 899
JALGAON5          CACAC--GTGTTCGACGTTACGAACGATCAGCTGAG 896
TRICHY2           CACAC--GTGTTCGACGTTACGAACGATCAGCTGAG 894
JALGAON2          CACAC--GTGTTCGACGTTACGAACGATCAGCTGAG 896
VAISHALI3         CACACACGTGTTCGACGTTACGAACGATCAGCTGAG 909
VAISHALI5         CACACACGTGTTCGACGTTACGAACGATCAGCTGAG 909
TRICHY4           CACACACGTGTTCGACGTTACGAACGATCAGCTGAG 899
WAYANAD2          CAC----GTGTTCGACGTTACGGACGATCAGCTGAG 885
WAYANAD4          CAC----GTGTTCGACGTTACGAACGATCAGCTGAG 885
WAYANAD5          CAC----GTGTTCGACGTTACGAACGATCAGCTGAG 885
WAYANAD1          CAC----GTGTTCGACGTTACGAACGATCAGCTGAG 885
WAYANAD3          CAC----GTGTTCGACGTTACGAACGATCAGCTGAG 885
NARAYNAGAON2      CAC----GTGTTCGACGTTACGAACGATCAGCTGAG 896
VAISHALI4         CAC----GTGTTCGACGTTACGAACGATCAGCTGAG 896
VAISHALI1         CAC----GTGTTCGACGTTACGAACGATCAGCTGAG 893
JALGAON4          CAC----GTGTTCGACGTTACGAACGATCAGCTGAG 894
NRAYANGAON1       CAC----GTGTTCGACGTTACGAACGATCAGCTGAG 894
NARAYANGAON4      CAC----GTGTTCGACGTTACGAACGATCAGCTGAG 894
NARAYANGAON5      CAC----GTGTTCGACGTTACGAACGATCAGCTGAG 894
TRICHY1           CAC----GTGTTCGACGTTACGAACGATCAGCTGAG 894
NARAYAGAON3       CAC----GTGTTCGACGTTACGAACGATCAGCTGAG 892
JALGAON1          C------GTGTTCGACGTTACGAACGATCAGCTGAG 890
VAISHLAI2         CACACACGTGTTCGACGTTACGAACGATCAGCTGAG 906
JALGAON3          CAC----GTGTTCGACGTTACGAACGATCAGCTGAG 906
ASSAM2            CACAC--GTGTTCGACGTTACGAACGATCAGCTGAG 898
ASSAM4            -------GTGTTCGACGTTACGAACGATCAGCTGAG 886
ASSAM3            C------GTGTTCGACGTTACGAACGATCAGCTGAG 885
ASSAM5            -------GTGTTCGACGTTACGAACGATCAGCTGAG 880
                         *************** *************
```

Fig. 1. ClustalW alignment of the ITS1 sequences from thirty individuals

Fourteen regions having simple repeats (n ≥ 2) were identified through the length of the ITS1 sequences (**Table 3**). Of these, nine regions (I to IX) exhibited variation in the number of the repeats which contribute to the variation in length between the 30 sequences.

Table 3. Simple repeats present in the ITS1 sequences

Type of repeat	Region # in the ClustalW alignment (Fig. 2)	`n' *i.e.* number of repeats
ATG	I	3-7
CAA	II	8
CG	III	2-3
GTT	IV	2
TA	V	5-7
CGT	VI	3-5
AT	VII	8-10
ATG	VIII	2
AC	IX	8-10

Fifteen haplotypes (gaps excluded) were identified amongst these 30 individuals (**Table 4**) using DnaSp v 5.10 and the mean haplotype diversity (Hd) was 0.7862. Haplotype 1 was observed in all the locations suggesting that this common haplotype could possibly represent an older haplotype.

Table 4. Haplotypes based on ITS1 sequences

Haplotype	Individual	Total No.
Hap 1	Assam 1; Jalgaon 1, 2 & 4; Narayangaon 1, 2, 3. 4 & 5; Trichy 1, 2 & 5; Vaishali 1, 4.	14
Hap 2	Assam 2	1
Hap 3	Assam 3	1
Hap 4	Assam 4	1
Hap 5	Assam 5	1
Hap6	Jalgaon 3	1
Hap 7	Jalgaon 5	1
Hap 8	Trichy 3	1
Hap 9	Trichy 4	1
Hap 10	Vaishali 2	1
Hap11	Vaishali 3, 5	2
Hap12	Wayanad 1. 3	2
Hap_13	Wayanad 2	1

Hap_14	Wayanad 4	1
Hap_15	Wayanad 5	1

InDel calculations of ITS1 were done using Model 1 with the option of Diallelic (non-overlapping). Total number of InDel sites analysed were eighteen, total number of InDel sites were sixty-seven, total number of excluded overlapping InDel sites were forty-nine and total number of InDel and non-InDel sites analysed were (862+18) 880. Total number of InDel events analysed were eight. Average InDel length event was 2.250 and average InDel length was 2.056.

Ts/Tv ratio of ITS1

An important parameter to be considered for phylogenetic analysis is sequence divergence and extent of saturation of substitution in the sequences being examined. Sequence divergence should be such that the sequences under consideration should not be too conserved to contain few substitutions nor should they be too diverged to exhibit saturation that information required for building phylogenetic trees is lost. The extent of saturation in the ITS1 data set was determined by the Ts/Tv ratio and the Iss values (Xia *et al*. 2003). The Ts/Tv ratio was found to be 2.72. The I_{ss} value of 0.0056 suggested that substitutions in this data set have not reached saturation. The transitions and transversions increase with genetic distance confirming that there is less saturation and hence this data can be used for phylogenetic analysis (**Fig. 2**).

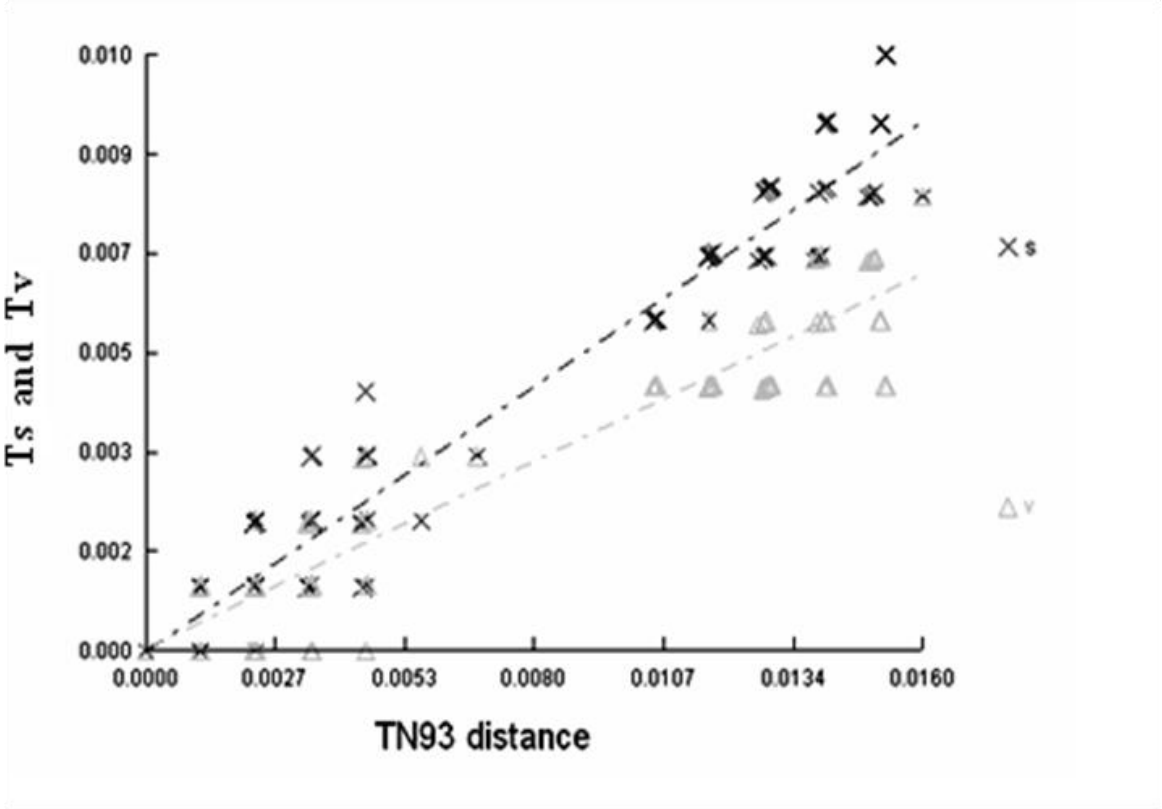


Fig. 2. Graph showing relationship between Ts and Tv and divergence in the ITS1 sequences

The T92 model was selected as the best fit model as it gave the lowest BIC score of 3418.9 with lnL = -1409.3. Substitution pattern and rates estimated under the Tamura-Nei (1993) model using MEGA v5.2.1 revealed that the most frequent substitutions were transitions of the G-A and C-T (U) types (**Table 5**).

Table 5. Maximum Likelihood Estimate of Substitution Matrix based on T92 model

	A	T/U	C	G
A	-	*4.66*	*3.03*	**13.47**
T/U	*3.35*	-	**17.29**	*3.15*
C	*3.35*	**26.57**	-	*3.15*
G	**14.29**	*4.66*	*3.03*	-

Genetic Distance based on ITS1

The intra-population genetic distance ranged from 0.0009 (Jalgaon and Trichy) to 0.0061 (Assam) with an overall mean genetic divergence of 0.0040678 (S.E. 0.0009608). The mean diversity within populations was 0.0020293 (S.E. 0.000411). The inter-population divergence ranged from 0.0004650 (between Jalgaon and Narayangaon and also between Narayangaon and Trichy) to 0.0111 (between Assam and Wayanad) with a mean diversity of 0.0040678 (S. E. 0.0009412) and mean inter-population diversity of 0.0020385 (S.E. 0.0006680).

Plylogenetic analyses based on the ITS1 region

The T92 model was selected as the best fit model as it gave the lowest BIC score of 3418.9 with lnL = -1409.3. The UPGMA tree of the 30 ITS1 sequences derived using MEGA v. 5 is shown in **Fig**. **3**.

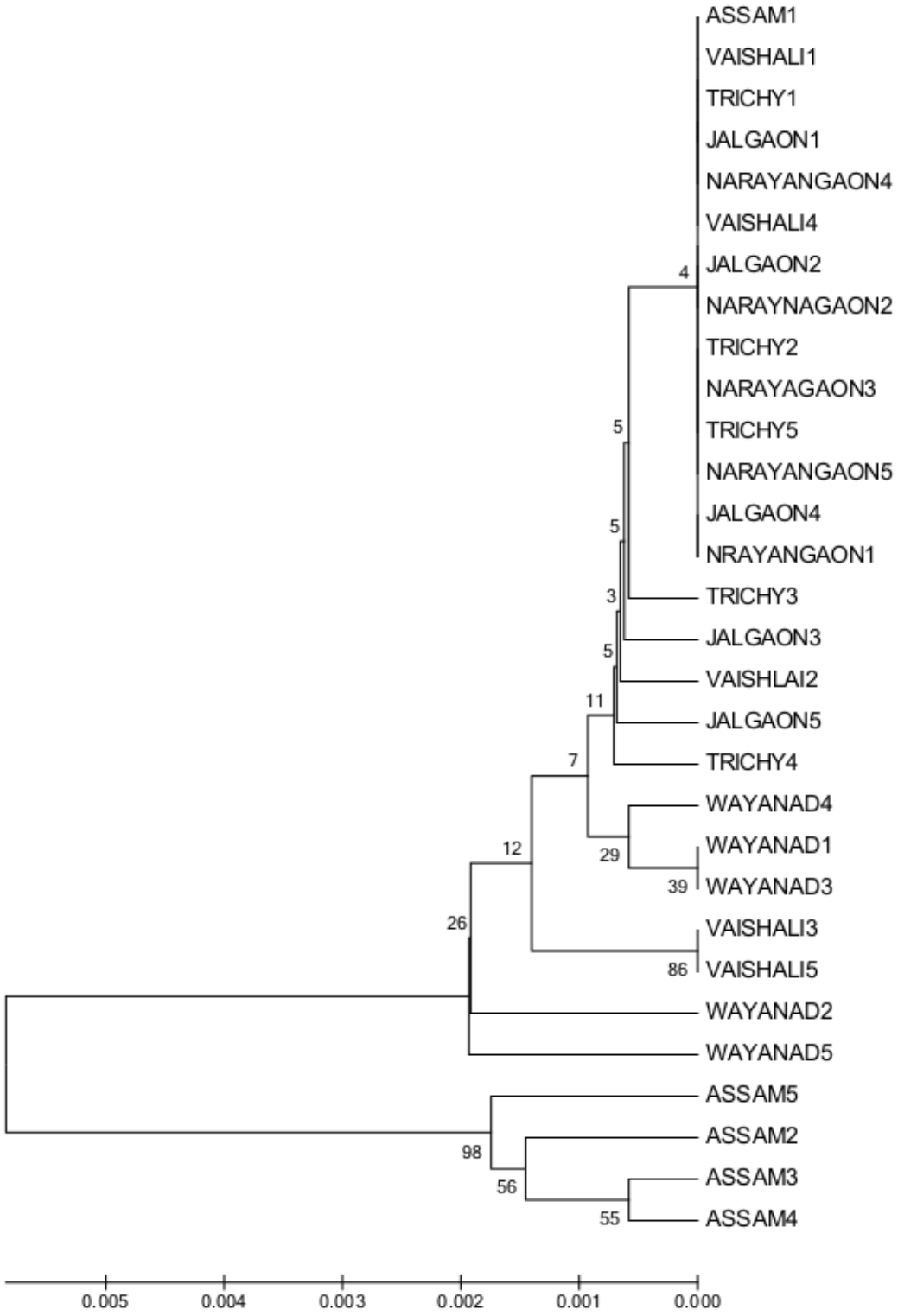


Fig. 3. UPGMA tree based on the ITS1 sequences

Since the microsatellite regions could affect the choice of model, UPGMA trees were also derived based on the indel polymorphism and haplotype data (**Figs. 4 and 5**).

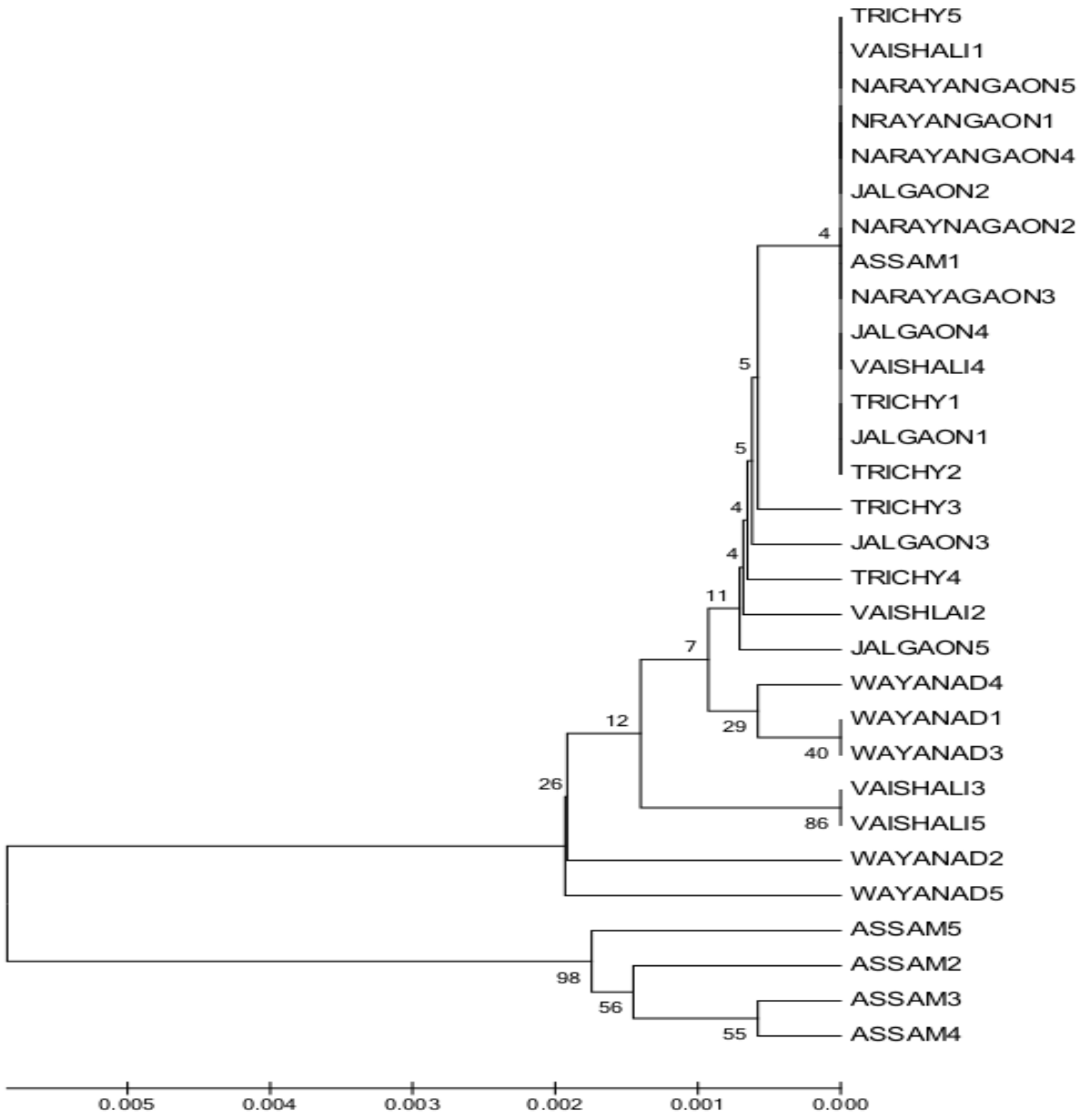


Fig. 4. UPGMA tree based on indel polymorphisms in the ITS1 sequences

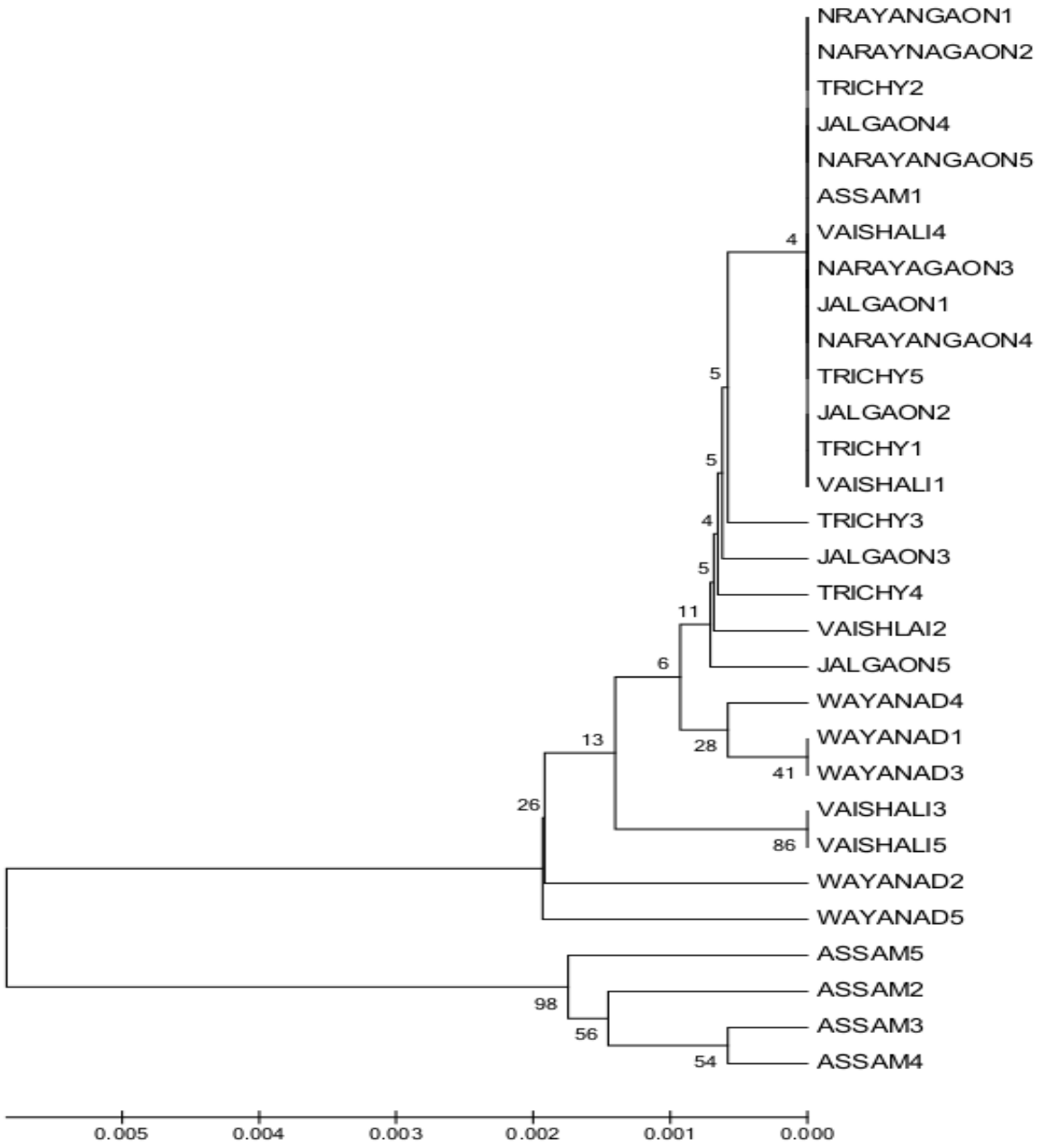


Fig. 5. UPGMA tree based on haplotype data of the ITS1 sequences

The topology of the UPGMA trees generated using each of the above data sets were strikingly similar. In all three trees, two clusters were observed. The smaller cluster comprised of individuals 2, 3 4 and 5 from Assam while the remaining twenty-six individuals formed the major cluster. Haplotype based grouping was observed within this

major cluster with all haplotype 1 individuals forming one sub-cluster. Similarly, haplotype 11 and 12 individuals also grouped together.

The ITS1 sequences of individuals 2, 3, 4 and 5 from Assam when compared with the remaining 26 sequences revealed 7 indels and nine base substitutions at positions 221, 270, 300, 316, 322, 593, 626, 714, 754 (5 transitions and 4 transversions).

The eleven individuals of haplotype 1 differed only in the number of microsatellite repeats and did not reveal any base substitution. A similar observation was also made in the ITS1 sequences from the five Narayangaon individuals. However, they differed from each other in a single region in the number of (TA) repeats. The alignment of the ITS1 sequences of the five individuals from Wayanad did not show any indels but five transitions were identified of which four were of the A-G type and one of C-T type. The ITS1 sequences from the five individuals from each of the other four locations i.e. Trichy, Jalgaon, Vaishali and Assam differed from each other not only in the number of microsatellite repeats but also in base substitutions *i.e.* transversions (A-G and C-T) and transversions (A-T, G-T, A-C). The similarity in the topology of the trees generated based on different data sets i.e. complete ITS1 sequences including gaps, indel polymorphisms and haplotype suggests that the phylogeny based on ITS1 data is strongly influenced by haplotype distribution.

The topology of all the trees did not show any geographical pattern. There was no strong correlation between genetic distance and geographical distance (R^2 = 0.4377; p= 0.020) (**Fig. 6**). Genetic differentiation estimates using DnaSp v.5.10. (Gst = 0.12704; Nm =1.72; χ^2 a= 86.571; p = 0.0872; df = 70) confirmed that the populations are not isolated thereby suggesting gene flow between the populations.

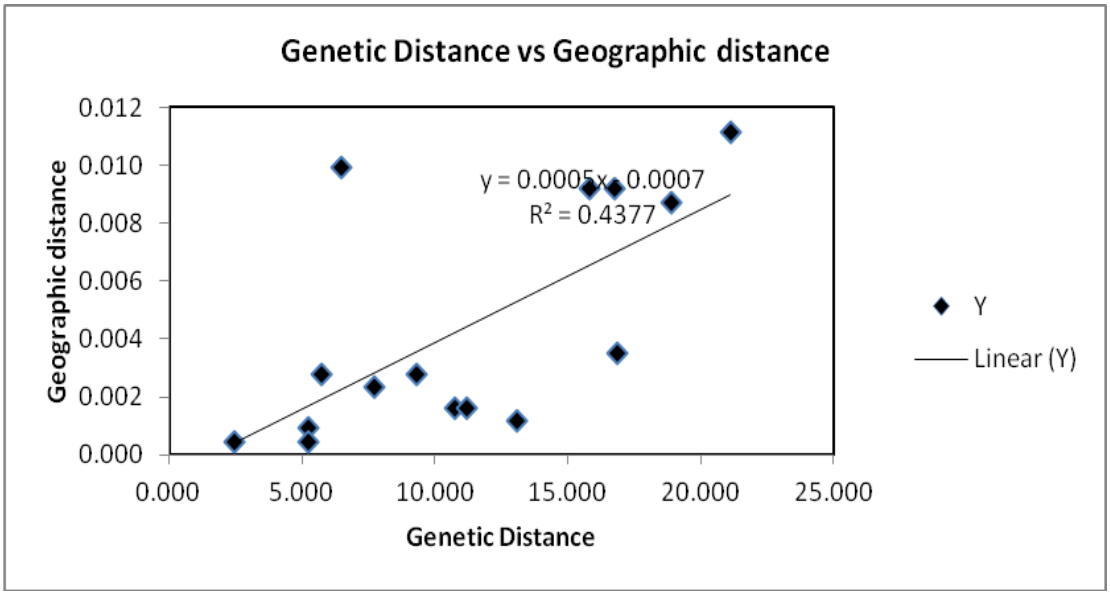


Fig. 6. Geographic vs genetic distance (Mantel test) for the ITS1 data set

Analysis of the ITS2 region

Sequence analysis

ITS2 sequences were determined for the same thirty adults of *O. longicollis* (Oliver) that were used for ITS1 analysis. The presumptive boundaries between 5.8S and ITS2 and ITS2 and 28S were determined by comparison with the ITS2 sequence of *Anthonomus grandis* (Genus Anthonomini; Coleoptera, Curculionidae) (Roehrdanz *et al*. 2010). The length of the ITS2 sequences ranged between 471 bp and 493 bp (**Table 2**). The overall base composition was T (31.8%), C (21.1%), A (22.9%) and G (24.1%).

The ClustalW alignment (**Fig. 7**) of the ITS2 sequences (including gaps) from the thirty individuals generated 513 characters of which 476 (92.78%) were conserved, 33 (6.43%) were parsimony informative and 13 (2.5%) were singleton sites**.**

```
                                                          I
ASSAM2            TATTTAAAAGACTGCTCGGTTTTCGATCGTCATCATCGTCGTCGTCGTCGTCTTTTCGAC 60
ASSAM5            TATTTAAAAGACTGCTCGGTTTTCGATCGTCATCATCGTCGTCGTCGTCGTCTTTTCGAC 60
VAISHALI2         TATTTAAAAGACTGCTCGGTTTTCGATCGTCATCATCGTCGTCGTCGTCGTCTTTTCGAC 60
NARAYANGAON2      TATTTAAAAGACTGCTCGGTTTTCGATCGTCATCATCGTCGTCGTCGTCGTCTTTTCGAC 60
NARAYANGAON4      TATTTAAAAGACTGCTCGGTTTTCGATCGTCATCATCGTCGTCGTCGTCGTCTTTTCGAC 60
NARAYANGAON3      TATTTAAAAGACTGCTCGGTTTTCGATCGTCATCATCGTCGTCGTCGTCGTCTTTTCGAC 60
TRICHY1           TATTTAAAAGACTGCTCGGTTTTCGATCGTCATCATCGTCGTCGTCGTCGTCTTTTCGAC 60
TRICHY3           TATTTAAAAGACTGCTCGGTTTTCGATCGTCATCATCGTCGTCGTCGTCGTCTTTTCGAC 60
TRICHY5           TATTTAAAAGACTGCTCGGTTTTCGATCGTCATCATCGTCGTCGTCGTCGTCTTTTCGAC 60
VAISHALI1         TATTTAAAAGACTGCTCGGTTTTCGATCGTCATCATCGTCGTCGTCGTCGTCTTTTCGAC 60
VAISHALI5         TATTTAAAAGACTGCTCGGTTTTCGATCGTCATCATCGTCGTCGTCGTCGTCTTTTCGAC 60
WAYANAD3          TATTTAAAAGACTGCTCGGTTTTCGATCGTCATCATCGTCGTCGTCGTC---TTTTCGAC 57
VAISHALI3         TATTTAAAAGACTGCTCGGTTTTCGATCGTCATCGTCGTCGTCGTCGTC---TTTTCGAC 57
WAYANAD1          TATTTAAAAGACTGCTCGGTTTTCGATCGTCATCGTCGTCGTCGTCGTC---TTTTCGAC 57
WAYANAD5          TATTTAAAAGACTGCTCGGTTTTCGATCGTCATCGTCGTCGTCGTCGTC---TTTTCGAC 57
VAISHALI4         TATTTAAAAGACTGCTCGGTTTTCGATCGTCATCGTCGTCGTCGTCGTC---TTCTCGAC 57
WAYANAD2          TATTTAAAAGACTGCTCGGTTTTCGATCGTCATCATCGTCGTCGTCGTC---TTTTCGAC 57
WAYANAD4          TATTTAAAAGACTGCTCGGTTTTCGATCGTCATCATCGTCGTCGTCGTC---TTTTCGAC 57
NARAYANGAON1      TATTTAAAAGACTGCTCGGTTTTCGATCGTCATCATCGTCGTCGTCGTCGTCTTTTCGAC 60
JALGAON1          TATTTAAAAGACTGCTCGGTTTTCGATCGTCATCATCGTCGTCGTCGTC---TTTTCGAC 57
JALGAON2          TATTTAAAAGACTGCTCGGTTTTCGATCGTCATCATCGTCGTCGTCGTC---TTTTCGAC 57
JALGAON3          TATTTAAAAGACTGCTCGGTTTTCGATCGTCATCATCGTCGTCGTCGTC---TTTTCGAC 57
JALGAON5          TATTTAAAAGACTGCTCGGTTTTCGATCGTCATCATCGTCGTCGTCGTC---TTTTCGAC 57
TRICHY2           TATTTAAAAGACTGCTCGGTTTTCGATCGTCATCATCGTCGTCGTCGTC---TTTTCGAC 57
TRICHY4           TATTTAAAAGACTGCTCGGTTTTCGATCGTCATCATCGTCGTCGTCGTC---TTTTCGAC 57
NARAYANGAON5      TATTTAAAAGACTGCTCGGTTTTCGATCGTCATCATCGTCGTCGTCGTC---TTTTCGAC 57
JALGAON4          TATTTAAAAGACTGCTCGGTTTTTGATCGTCATCATCGTCGTCGTCGTC---TTTTCGAC 57
ASSAM3            TATTTAAAAGACTGCTCGGTTTTCGATCGTCATCGTCGTC------------TTTTCGAC 48
ASSAM4            TATTTAAAAGACTGCTCGGTTTTCGATCGTCATCGTCGTC------------TTTTCGAC 48
ASSAM1            TATTTAAAAGACTGCTCGGTTTTCGATCGTCATCGTCGTC------------TTTTCGAC 48
                  ********************** ********** *****           ** *****

ASSAM2            GAGGAGACTGGTGGTGATTGGTCGATTACGTCGGAGCGTGTTGGATGTACGACGCGAAAG 120
ASSAM5            GAGGAGACTGGTGGTGATTGGTCGATTACGTCGGAGCGTGTTGGATGTACGACGCGAAAG 120
VAISHALI2         GAGGAGACTGGTGGTGATTGGTCGATTACGTCGGAGCGTGTTGGATGTACGACGCGAAAG 120
NARAYANGAON2      GAGGAGACTGGTGGTGATTGGTCGATTACGTCGGAGCGTGTTGGATGTACGACGCGAAAG 120
NARAYANGAON4      GAGGAGACTGGTGGTGATTGGTCGATTACGTCGGAGCGTGTTGGATGTACGACGCGAAAG 120
NARAYANGAON3      GAGGAGACTGGTGGTGATTGGTCGATTACGTCGGAGCGTGTTGGATGTACGACGCGAAAG 120
TRICHY1           GAGGAGACTGGTGGTGATTGGTCGATTACGTCGGAGCGTGTTGGATGTACGACGCGAAAG 120
TRICHY3           GAGGAGACTGGTGGTGATTGGTCGATTACGTCGGAGCGTGTTGGATGTACGACGCGAAAG 120
TRICHY5           GAGGAGACTGGTGGTGATTGGTCGATTACGTCGGAGCGTGTTGGATGTACGACGCGAAAG 120
VAISHALI1         GAGGAGACTGGTGGTGATTGGTCGATTACGTCGGAGCGTGTTGGATGTACGACGCGAAAG 120
VAISHALI5         GAGGAGACTGGTGGTGATTGGTCGATTACGTCGGAGCGTGTTGGATGTACGACGCGAAAG 120
WAYANAD3          GAGGAGACTGGTGGTGATTGGTCGATTACGTCGGAGCGTGTTGGATGTACGACGCGAAAG 117
VAISHALI3         GAGGAGACTGGTGGTGATTGGTCGATTACGTCGGAGCGTGTTGGATGTACGACGCGAAAG 117
WAYANAD1          GAGGAGACTGGTGGTGATTGGTCGATTACGTCGGAGCGTGTTGGATGTACGACGCGAAAG 117
WAYANAD5          GAGGAGACTGGTGGTGATTGGTCGATTACGTCGGAGCGTGTTGGATGTACGACGCGAAAG 117
VAISHALI4         GAGGAGACTGGTGGTGATTGGTCGATTACGTCGGAGCGTGTTGGATGTACGACGCGAAAG 117
WAYANAD2          GAGGAGACTGGTGGTGATTGGTCGATTACGTCGGAGCGTGTTGGATGTACGACGCGAAAG 117
WAYANAD4          GAGGAGACTGGTGGTGATTGGTCGATTACGTCGGAGCGTGTTGGATGTACGACGCGAAAG 117
NARAYANGAON1      GAGGAGACTGGTGGTGATTGGTCGATTACGTCGGAGCGTGTTGGATGTACGACGCGAAAG 120
JALGAON1          GAGGAGACTGGTGGTGATTGGTCGATTACGTCGGAGCGTGTTGGATGTACGACGCGAAAG 117
JALGAON2          GAGGAGACTGGTGGTGATTGGTCGATTACGTCGGAGCGTGTTGGATGTACGACGCGAAAG 117
JALGAON3          GAGGAGACTGGTGGTGATTGGTCGATTACGTCGGAGCGTGTTGGATGTACGACGCGAAAG 117
JALGAON5          GAGGAGACTGGTGGTGATTGGTCGATTACGTCGGAGCGTGTTGGATGTACGACGCGAAAG 117
TRICHY2           GAGGAGACTGGTGGTGATTGGTCGATTACGTCGGAGCGTGTTGGATGTACGACGCGAAAG 117
TRICHY4           GAGGAGACTGGTGGTGATTGGTCGATTACGTCGGAGCGTGTTGGATGTACGACGCGAAAG 117
NARAYANGAON5      GAGGAGACTGGTGGTGATTGGTCGATTACGTCGGAGCGTGTTGGATGTACGACGCGAAAG 117
JALGAON4          GAGGAGACTGGTGGTGATTGGTCGATTACGTCGGAGCGTGTTGGATGTACGACGCGAAAG 117
ASSAM3            GAGGAGACTGGTGGTGATTGGTCGATTACGTCGGAGCGTGTTGGATGTACGACGCGAAAG 108
ASSAM4            GAGGAGACTGGTGGTGATTCGTCGATTACGTCGGAGCGTGTTGGATGTACGACGCGAAAG 108
ASSAM1            GAGGAGACTGGTGGTGATTGGTCGATTACGTCGGAGCGTGTTGGATGTACGACGCGAAAG 108
                  ****************** ****************************************
                                                        II
ASSAM2            AAAGAGGAGAGGAGGGTGGGTGGTGGCGAATTTACTTCTCCTCCTCC-TTCTCTTTCTCT 179
ASSAM5            AAAGAGGAGAGGAGGGTGGGTGGTGGCGAATTTACTTCTCCTCCTCC-TTCTCTTTCTCT 179
VAISHALI2         AAAGAGGAGAGGAGGGTGGGTGGTGGCGAATTTACTTCTCCTCCTCC-TTCTCTTTCTCT 179
NARAYANGAON2      AAAGAGGAGAGGAGGGTGGGTGGTGGCGAATTTACTTCTCCTCCTCC-TTCTCTTTCTCT 179
NARAYANGAON4      AAAGAGGAGAGGAGGGTGGGTGGTGGCGAATTTACTTCTCCTCCTCC-TTCTCTTTCTCT 179
NARAYANGAON3      AAAGAGGAGAGGAGGGTGGGTGGTGGCGAATTTACTTCTCCTCCTCC-TTCTCTTTCTCT 179
```

```
TRICHY1          AAAGAGGAGAGGAGGGTGGGTGGTGGCGAATTTACTTCTCCTCCTCC-TTCTCTTTCTCT 179
TRICHY3          AAAGAGGAGAGGAGGGTGGGTGGTGGCGAATTTACTTCTCCTCCTCC-TTCTCTTTCTCT 179
TRICHY5          AAAGAGGAGAGGAGGGTGGGTGGTGGCGAATTTACTTCTCCTCCTCC-TTCTCTTTCTCT 179
VAISHALI1        AAAGAGGAGAGGAGG-----TGGTGGCGAATTTACTTCTCCTCCTCCGTTCTCTTTCTCC 175
VAISHALI5        AAAGAGGAGAGGAG-------GGTGGCGAATTTACTTCTCCTCCTCCGTTCTCTTTCTCC 173
WAYANAD3         AAGGG---------------TGGTGGCGAATTTACTTCTCCTCCTCC-TTCTCTTTCTCT 161
VAISHALI3        AAAGAGGAGAGGAGGGTGGGTGGTGGCGAATTTACTTCTCCTCCTCC-TTCTCTTTCTCT 176
WAYANAD1         AAAGAGGAGAGGAGGGTGGGTGGTGGCGAATTTACTTCTCCTCCTCC-TTCTCTTTCTCT 176
WAYANAD5         AAAGAGGAGAGGAGGGTGGGTGGTGGCGAATTTACTTCTCCTCCTCC-TTCTCTTTCTCT 176
VAISHALI4        AAAGAGGAGAGGAGGGTGGGTGGTGGCGAATTTACTTCTCCTCCTCC-TTCTCTTTCTCT 176
WAYANAD2         AGAGAGGAGAGGAGGGTGGGTGGTGGCGAATTTACTTCTCCTCCTCC-TTCTCTTTCTCT 176
WAYANAD4         AAAGAGGAGAGGAGGGTGGGTGGTGGCGAATTTACTTCTCCTCCTCC-TTCTCTTTCTCT 176
NARAYANGAON1     AAAGAGGAGAGGAGGGTGGGTGGTGGCGAATTTACTTCTCCTCCTCC-TTCTCTTTCTCT 179
JALGAON1         AAAGAGGAGAGGAGGGTAGGTGGTGGCGAATTTACTTCTCCTCCTCCTCCTTCTCTTTCT 177
JALGAON2         AAAGAGGAGAGGAGGGTAGGTGGTGGCGAATTTACTTCTCCTCCTCCTCCTTCTCTTTCT 177
JALGAON3         AAAGAGGAGAGGAGGGTAGGTGGTGGCGAATTTACTTCTCCTCCTCCTCCTTCTCTTTCT 177
JALGAON5         AAAGAGGAGAGGAGGGTAGGTGGTGGCGAATTTACTTCTCCTCCTCCTCCTTCTCTTTCT 177
TRICHY2          AAAGAGGAGAGGAGGGTAGGTGGTGGCGAATTTACTTCTCCTCCTCCTCCTTCTCTTTCT 177
TRICHY4          AAAGAGGAGAGGAGGGTAGGTGGTGGCGAATTTACTTCTCCTCCTCCTCCTTCTCTTTCT 177
NARAYANGAON5     AAAGAGGAGAGGAGGGTAGGTGGTGGCGAATTTACTTCTCCTCCTCCTCCTTCTCTTTCT 177
JALGAON4         AAAGAGGAGAGGAGGGTAGGTGGTGGCGAATTTACTTCTCCTCCTCCTCCTTCTCTTTCT 177
ASSAM3           AAAGAGGAGAGGAGGGTGGGTGGTGGCGAATTTACTTCTCCTCCTCC-TCCTCCTACTCT 167
ASSAM4           AAAGAGGAGAGGAGGGTGGGTGGTGGCGAATTTACTTCTCCTCCTCC-TCCTCCTACTCT 167
ASSAM1           AAAGAGGAGAGGAGGGTGGGTGGTGGCGAATTTACTTCTCCTCCTCC-TCCTCCTACTCT 167
                 *  *                ***************************    **    **
                                                 III
ASSAM2           CTCTCTCTCATTCGCGACATCTCAAAACGTGAATAC--GTTCACGTTCATATTACGTTTA 237
ASSAM5           CTCTCTCTCATTCGCGACATCTCAAAACGTGAATAC--GTTCACGTTCATATTACGTTTA 237
VAISHALI2        CTCTCTCTCATTCGCGACATCTCAAAACGTGAATACACGTTCACGTTCATATTACGTTTA 239
NARAYANGAON2     CTCTCTCTCATTCGCGACATCTCAAAACGTGAATACACGTTCACGTTCATATTACGTTTA 239
NARAYANGAON4     CTCTCTCTCATTCGCGACATCTCAAAACGTGAATACACGTTCACGTTCATATTACGTTTA 239
NARAYANGAON3     CTCTCTCTCATTCGCGACATCTCAAAACGTGAATACACGTTCACGTTCATATTACGTTTA 239
TRICHY1          CTCTCTCTCATTCGCGACATCTCAAAACGTGAATACACGTTCACGTTCATATTACGTTTA 239
TRICHY3          CTCTCTCTCATTCGCGACATCTCAAAACGTGAATACACGTTCACGTTCATATTACGTTTA 239
TRICHY5          CTCTCTCTCATTCGCGACATCTCAAAACGTGAATACACGTTCACGTTCATATTACGTTTA 239
VAISHALI1        CTCTCTCTCATTCGCGACATCTCAAAACGTGAATACACGTTCACGTTCATATTACGTTTA 235
VAISHALI5        CTCTCTCTCATTCGCGACATCTCAAAACGTGAATACACGTTCACGTTCATATTACGTTTA 233
WAYANAD3         CTCTCTCTCATTCGCGACATCTCAAAACGTGAATAC--GTTCACGTTCATATTACGTTTA 219
VAISHALI3        CTCTCTCTCATTCGCGACATCTCAAAACGTGAATAC--GTTCACGTTCATATTACGTTTA 234
WAYANAD1         CTCTCTCTCATTCGCGACATCTCAAAACGTGAATAC--GTTCACGTTCATATTACGTTTA 234
WAYANAD5         CTCTCTCTCATTCGCGACATCTCAAAACGTGAATAC--GTTCACGTTCATATTACGTTTA 234
VAISHALI4        CTCTCTCTCATTCGCGACATCTCAAAACGTGAATAC--GTTCACGTTCATATTACGTTTA 234
WAYANAD2         CTCTCTCTCATTCGCGACATCTCAAAACGTGAATAC--GTTCACGTTCATATTACGTTTA 234
WAYANAD4         CTCTCTCTCATTCGCGACATCTCAAAACGTGAATAC--GTTCACGTTCATATTACGTTTA 234
NARAYANGAON1     CTCTCTCTCATTCGCGACATCTCAAAACGTGAATACACGTTCACGTTCATATTACGTTTA 239
JALGAON1         CTCTCTCTCATTCGCGACATCTCAAAACGTGAATAC--GTTCACGTTCATATTACGTTTA 235
JALGAON2         CTCTCTCTCATTCGCGACATCTCAAAACGTGAATAC--GTTCACGTTCATATTACGTTTA 235
JALGAON3         CTCTCTCTCATTCGCGACATCTCAAAACGTGAATAC--GTTCACGTTCATATTACGTTTA 235
JALGAON5         CTCTCTCTCATTCGCGACATCTCAAAACGTGAATAC--GTTCACGTTCATATTACGTTTA 235
TRICHY2          CTCTCTCTCATTCGCGACATCTCAAAACGTGAATAC--GTTCACGTTCATATTACGTTTA 235
TRICHY4          CTCTCTCTCATTCGCGACATCTCAAAACGTGAATAC--GTTCACGTTCATATTACGTTTA 235
NARAYANGAON5     CTCTCTCTCATTCGCGACATCTCAAAACGTGAATAC--GTTCACGTTCATATTACGTTTA 235
JALGAON4         CTCTCTCTCATTCGCGACATCTCAAAACGTGAATAC--GTTCACGTTCATATTACGTTTA 235
ASSAM3           TTCTCTCTCATTCGCGACATCTCAAAACGTGAATAC--GTTCACGTTCATATTACGTTTA 225
ASSAM4           TTCTCTCTCATTCGCGACATCTCAAAACGTGAATAC--GTTCACGTTCATATTACGTTTA 225
ASSAM1           TTCTCTCTCATTCGCGACATCTCAAAACGTGAATAC--GTTCACGTTCATATTACGTTTA 225
                  *********************************  **********************

                                                   IV
ASSAM2           TCGAAATGTGAACGTGTAGAGTGTAACCGCGATTCGAGCGGGGCTTTATTAAATACGTCG 297
ASSAM5           TCGAAATGTGAACGTGTAGAGTGTAACCGCGATTCGAGCGGGGCTTTATTAAATACGTCG 297
VAISHALI2        TCGAAATGTGAACGTGTAGAGTGTAACCGCGATTCGAGCGGGGCTTTATTAAATACGTCG 299
NARAYANGAON2     TCGAAATGTGAACGTGTAGAGCGTAACCGCGATTCGAGCGGGGCTTTATTAAATACGTCG 299
NARAYANGAON4     TCGAAATGTGAACGTGTAGAGTGTAACCGCGATTCGAGCGGGGCTTTATTAAATACGTCG 299
NARAYANGAON3     TCGAAATGTGAACGTGTAGAGTGTAACCGCGATTCGAGCGGGGCTTTATTAAATACGTCG 299
TRICHY1          TCGAAATGTGAACGTGTAGAGTGTAACCGCGATTCGAGCGGGGCTTTATTAAATACGTCG 299
TRICHY3          TCGAAATGTGAACGTGTAGAGTGTAACCGCGATTCGAGCGGGGCTTTATTAAATACGTCG 299
TRICHY5          TCGAAATGTGAACGTGTAGAGTGTAACCGCGATTCGAGCGGGGCTTTATTAAATACGTCG 299
VAISHALI1        TCGAAATGTGAACGTGTAGAGTGTAACCGCGATTCGAGCGGGGCTTTATTAAATACGTCG 295
VAISHALI5        TCGAAATGTGAACGTGTAGAGTGTAACCGCGATTCGAGCGGGGCTTTATTAAATACGTCG 293
WAYANAD3         TCGAAATGTGAACGTGTAGAATGTAACCGCGATTCGAGCGGGGCTTTATTAAATACGTCG 279
```

```
VAISHALI3         TCGAAATGTGAACGTGTAGAATGTAACCGCGATTCGAGCGGGGCTTTATTAAATACGTCG 294
WAYANAD1          TCGAAATGTGAACGTGTAGAATGTAACCGCGATTCGAGCGGGGCTTTATTAAATACGTCG 294
WAYANAD5          TCGAAATGTGAACGTGTAGAATGTAACCGCGATTCGAGCGGGGCTTTATTAAATACGTCG 294
VAISHALI4         TCGAAATGTGAACGTGTAGAATGTAACCGCGATTCGAGCGGGGCTTTATTAAATACGTCG 294
WAYANAD2          TCGAAATGTGAACGTGTAGAATGTAACCGCGATTCGAGCGGGGCTTTATTAAATACGTCG 294
WAYANAD4          TCGAAATGTGAACGTGTAGAATGTAACCGCGATTCGAGCGGGGCTTTATTAAATACGTCG 294
NARAYANGAON1      TCGAAATGTGAACGTGTAGAATGTAACCGCGATTCGAGCGGGGCTTTATTAAATACGTCG 299
JALGAON1          TCGAAATGTGAACGTGTAGAATGTAACCGCGATTCGAGCGGGGCTTTATTAAATACGTCG 295
JALGAON2          TCGAAATGTGAACGTGTAGAATGTAACCGCGATTCGAGCGGGGCTTTATTAAATACGTCG 295
JALGAON3          TCGAAATGTGAACGTGTAGAATGTAACCGCGATTCGAGCGGGGCTTTATTAAATACGTCG 295
JALGAON5          TCGAAATGTGAACGTGTAGAATGTAACCGCGATTCGAGCGGGGCTTTATTAAATACGTCG 295
TRICHY2           TCGAAATGTGAACGTGTAGAATGTAACCGCGATTCGAGCGGGGCTTTATTAAATACGTCG 295
TRICHY4           TCGAAATGTGAACGTGTAGAATGTAACCGCGATTCGAGCGGGGCTTTATTAAATACGTCG 295
NARAYANGAON5      TCGAAATGTGAACGTGTAGAATGTAACCGCGATTCGAGCGGGGCTTTATTAAATACGTCG 295
JALGAON4          TCGAAATGTGAACGTGTAGAATGTAACCGCGATTCGAGCGGGGCTTTATTAAATACGTCG 295
ASSAM3            TCGAAATGTGAACGTGTAGAGTGTAACCGCGATTCGAGCGG--CTTTATTAAATACGTCG 283
ASSAM4            TCGAAATGTGAACGTGTAGAGTGTAACCGCGATTCGAGCGG--CTTTATTAAATACGTCG 283
ASSAM1            TCGAAATGTGAACGTGTAGAGTGTAACCGCGATTCGAGCGG--CTTTATTAAATACGTCG 283
                  ********************  ********************  *****************

                                                                          V
ASSAM2            TCGCCTATGTATTATACTATAGTCGTCAATACGTGTTTATCGTATAGTCTTGGAGCGCGC 357
ASSAM5            TCGCCTATGTATTATACTATAGTCGTCAATACGTGTTTATCGTATAGTCTTGGAGCGCGC 357
VAISHALI2         TCGCCTATGTATTATACTATAGTCGTCAATACGTGTTTATCGTATAGTCTTGGAGCGCGC 359
NARAYANGAON2      TCGCCTATGTATTATACTATAGTCGTCAATACGTGTTTATCGTATAGTCTTGGAGCGCGC 359
NARAYANGAON4      TCGCCTATGTATTATACTATAGTCGTCAATACGTGTTTATCGTATAGTCTTGGAGCGCGC 359
NARAYANGAON3      TCGCCTATGTATTATACTATAGTCGTCAATACGTGTTTATCGTATAGTCTTGGAGCGCGC 359
TRICHY1           TCGCCTATGTATTATACTATAGTCGTCAATACGTGTTTATCGTATAGTCTTCGAGCGCGC 359
TRICHY3           TCGCCTATGTATTATACTATAGTCGTCAATACGTGTTTATCGTATAGTCTTCGAGCGCGC 359
TRICHY5           TCGCCTATGTATTATACTATAGTCGTCAATACGTGTTTATCGTATAGTCTTCGAGCGCGC 359
VAISHALI1         TCGCCTATGTATTATACTATAGTCGTCAATACGTGTTTATCGTATAGTCTTGGAGCGCGC 355
VAISHALI5         TCGCCTATGTATTATACTATAGTCGTCAATACGTGTTTATCGTATAGTCTTGGAGCGCGC 353
WAYANAD3          TCGCCTATGTATTATACTATAGTCGTCAATACGTGTTTATCGTATAGTCTTCGAGCGCGC 339
VAISHALI3         TCGCCTATGTATTATACTATAGTCGTCAATACGTGTTTATCGTATAGTCTTGGAGCGCGC 354
WAYANAD1          TCGCCTATGTATTATACTATAGTCGTCAATACGTGTTTATCGTATAGTCTTGGAGCGCGC 354
WAYANAD5          TCGCCTATGTATTATACTATAGTCGTCAATACGTGTTTATCGTATAGTCTTGGAGCGCGC 354
VAISHALI4         TCGCCTATGTATTATACTATAGTCGTCAATACGTGTTTATCGTATAGTCTTGGAGCGCGC 354
WAYANAD2          TCGCCTATGTATTATACTATAGTCGTCAATACGTGTTTATCGTATAGTCTTCGAGCGCGC 354
WAYANAD4          TCGCCTATGTATTATACTATAGTCGTCAATACGTGTTTATCGTATAGTCTTCGAGCGCGC 354
NARAYANGAON1      TCGCCTATGTATTATACTATAGTCGTCAATACGTGTTTATCGTATAGTCTTGGGGCGCGC 359
JALGAON1          TCGCCTATGTATTATACTATAGTCGTCAATACGTGTTTATCGTATAGTCTTGGGGCGCGC 355
JALGAON2          TCGCCTATGTATTATACTATAGTCGTCAATACGTGTTTATCGTATAGTCTTGGGGCGCGC 355
JALGAON3          TCGCCTATGTATTATACTATAGTCGTCAATACGTGTTTATCGTATAGTCTTGGGGCGCGC 355
JALGAON5          TCGCCTATGTATTATACTATAGTCGTCAATACGTGTTTATCGTATAGTCTTGGGGCGCGC 355
TRICHY2           TCGCCTATGTATTATACTATAGTCGTCAATACGTGTTTATCGTATAGTCTTGGGGCGCGC 355
TRICHY4           TCGCCTATGTATTATACTATAGTCGTCAATACGTGTTTATCGTATAGTCTTGGGGCGCGC 355
NARAYANGAON5      TCGCCTATGTATTATACTATAGTCGTCAATACGTGTTTATCGTATAGTCTTGGGGCGCGC 355
JALGAON4          TCGCCTATGTATTATACTATAGTCGTCAATACGTGCTTATCGTATAGTCTCGGGGCGCGC 355
ASSAM3            TCGCCTATGTATTATACTATAGTCGTCAATACGTGTTTATCGTAAAGTCTTCGAGGGCGC 343
ASSAM4            TCGCCTATGTATTATACTATAGTCGTCAATACGTGTTTATCGTAAAGTCTTCGAGGGCGC 343
ASSAM1            TCGCCTATGTATTATACTATAGTCGTCAATACGTGTTTATCGTAAAGTCTTCGAGGGCGC 343
                  ********************************** ******** *****  * * ****


ASSAM2            GC----GCTCGTGCCGGTCGACGTGTCCGATATTATTCTTTTTGTAATATAAAGGAAATC 413
ASSAM5            GC----GCTCGTGCCGGTCGACGTGTCCGATATTATTCTTTTTGTAATATAAAGGAAATC 413
VAISHALI2         GC----GCTCGTGCCGGTCGACGTGTCCGATATTATTCTTTTTGTAATATAAAGGAAATC 415
NARAYANGAON2      GC----GCTCGTGCCGGTCGACGTGTCCGATATTATTCTTTTTGTAATATAAAGGAAATC 415
NARAYANGAON4      GC----GCTCGTGCCGGTCGACGTGTCCGATATTATTCTTTTTGTAATATAAAGGAAATC 415
NARAYANGAON3      GC----GCTCGTGCCGGTCGACGTGTCCGATATTATTCTTTTTGTAATATAAAGGAAATC 415
TRICHY1           ------GCTCGTGCCGGTCGACGTGTCCGATATTATTCTTTTTGTAATATAAAGGAAATC 413
TRICHY3           ------GCTCGTGCCGGTCGACGTGTCCGATATTATTCTTTTTGTAATATAAAGGAAATC 413
TRICHY5           ------GCTCGTGCCGGTCGACGTGTCCGATATTATTCTTTTTGTAATATAAAGGAAATC 413
VAISHALI1         GCGCGCGCTCGTGCCGGTCGACGTGTCCGATATTATTCTTTTTGTAATATAAAGGAAATC 415
VAISHALI5         GCGCGCGCTCGTGCCGGTCGACGTGTCCGATATTATTCTTTTTGTAATATAAAGGAAATC 413
WAYANAD3          ------GCTCGTGCCGGTCGACGTGTCCGATATTATTCTTTTTGTAATATAAAGGAAATC 393
VAISHALI3         ------GCTCGTGCCGGTCGACGTGTCCGATATTATTCTTTTTGTAATATAAAGGAAATC 408
WAYANAD1          ------GCTCGTGCCGGTCGACGTGTCCGATATTATTCTTTTTGTAATATAAAGGAAATC 408
WAYANAD5          ------GCTCGTGCCGGTCGACGTGTCCGATATTATTCTTTTTGTAATATAAAGGAAATC 408
VAISHALI4         ------GCTCGTGCCGGTCGACGTGTCCGATATTATTCTTTTTGTAATATAAAGGAAATC 408
WAYANAD2          ------GCTCGTGCCGGTCGACGTGTCCGATATTATTCTTTTTGTAATATAAAGGAAATC 408
```

```
WAYANAD4          ------GCTCGTGCCGGTCGACGTGTCCGATATTATTCTTTTTGTAATATAAAGGAAATC 408
NARAYANGAON1      GC----GCTCGTGCCGGTCGACGTGTCCGATATTATTCTTTTTGTAATATAAAGGAAATC 415
JALGAON1          GC----GCTCGTGCCGGTCGACGTGTCCGATATTATTCTTTTTGTAATATAAAGGAAATC 411
JALGAON2          GC----GCTCGTGCCGGTCGACGTGTCCGATATTATTCTTTTTGTAATATAAAGGAAATC 411
JALGAON3          GC----GCTCGTGCCGGTCGACGTGTCCGATATTATTCTTTTTGTAATATAAAGGAAATC 411
JALGAON5          GC----GCTCGTGCCGGTCGACGTGTCCGATATTATTCTTTTTGTAATATAAAGGAAATC 411
TRICHY2           GC----GCTCGTGCCGGTCGACGTGTCCGATATTATTCTTTTTGTAATATAAAGGAAATC 411
TRICHY4           GC----GCTCGTGCCGGTCGACGTGTCCGATATTATTCTTTTTGTAATATAAAGGAAATC 411
NARAYANGAON5      GC----GCTCGTGCCGGTCGACGTGTCCGATATTATTCTTTTTGTAATATAAAGGAAATC 411
JALGAON4          GC----GCTCGTGCCGGTCGACGTGTCCGATATTATTCTTTTTGTAATATAAAGGAAATC 411
ASSAM3            GC----GCTCGTGCCGGTCGACGTGTCCGATATTATTCTTTTTGTAATATAAAGGAAATC 399
ASSAM4            GC----GCTCGTGCCGGTCGACGTGTCCGATATTATTCTTTTTGTAATATAAAGGAAATC 399
ASSAM1            GC----GCTCGTGCCGGTCGACGTGTCCGATATTATTCTTTTTGTAATATAAAGGAAATC 399
                        ****************************************************

                                  VI                    VII   VIII       IX
ASSAM2            GACGTCTGAACGGGC---AATGATGATGCCGTCGC--CAACG--ATCGAC--------AC 458
ASSAM5            GACGTCTGAACCGGC---AATGATGATGCCGTCGC--CAACG--ATCGAC--------AC 458
VAISHALI2         GACGTCTGAACCGGC---AATGATGATGCCGTCGC--CAACG--ATCGAC--------AC 460
NARAYANGAON2      GACGTCTGAACCGGC---AATGATGATGCCGTCGC--CAACG--ATCGAC--------AC 460
NARAYANGAON4      GACGTCTGAACCGGC---AATGATGATGCCGTCGC--CAACG--ATCGAC--------AC 460
NARAYANGAON3      GACGTCTGAACCGGC---AATAATGATGCCGTCGC--CAACG--ATCGAC--------AC 460
TRICHY1           GACGTCTGAACCGGC---AATGATGATGCCGTCGC--CAACG--ATCGAC--------AC 458
TRICHY3           GACGTCTGAACCGGC---AATGATGATGCCGTCGC--CAACG--ATCGAC--------AC 458
TRICHY5           GACGTCTGAACCGGC---AATGATGATGCCGTCGC--CAACG--ATCGAC--------AC 458
VAISHALI1         GACGTCTGAACCGGC---AATGATGATGCCGTCGC--CAACG--ATCGAC--------AC 460
VAISHALI5         GACGTCTGAACCGGC---AATGATGATGCCGTCGC--CAACG--ATCGAC--------AC 458
WAYANAD3          GACGTCTGAACCGGC---AATGATGATGCCGTCGC--CAACG--ATCGAC--------AC 438
VAISHALI3         GACGTCTGAACCGGC---AATGATGATGCCGTCGC--CAACGAGATCGAC--------AC 455
WAYANAD1          GACGTCTGAACCGGC---AATGATGATGCCGTCGC--CAACGAGATCGAC--------AC 455
WAYANAD5          GACGTCTGAACCGGC---AATGATGATGCCGTCGC--CAACGAGATCGAC--------AC 455
VAISHALI4         GACGTCTGAACCGGC---AATGATGATTCCGTCGC--CAACGAGATCGAC--------AC 455
WAYANAD2          GACGTCTGAACCGGC---AATGATGATGCCGTCGC--CAACG--ATCGAC--------AC 453
WAYANAD4          GACGTCTGAACCGGC---AATGATGATGCCGTCGC--CAACG--ATCGAC--------AC 453
NARAYANGAON1      GACGTCTGAACCGGC---AATGATGATGCCGTCGC--CAACG--ATCGAC--------AC 460
JALGAON1          GACGTCTGAACCGGC---AATGATGATGCCGTCGC--CAACG--ATCGAC--------AC 456
JALGAON2          GACGTCTGAACCGGC---AATGATGATGCCGTCGC--CAACG--ATCGAC--------AC 456
JALGAON3          GACGTCTGAACCGGC---AATGATGATGCCGTCGC--CAACG--ATCGAC--------AC 456
JALGAON5          GACGTCTGAACCGGC---AATGATGATGCCGTCGC--CAACG--ATCGAC--------AC 456
TRICHY2           GACGTCTGAACCGGC---AATGATGATGCCGTCGC--CAACG--ATCGAC--------AC 456
TRICHY4           GACGTCTGAACCGGC---AATGATGATGCCGTCGC--CAACG--ATCGAC--------AC 456
NARAYANGAON5      GACGTCTGAACCGGC---AATAATGATGCCGTCGC--CAACG--ATCGAC--------AC 456
JALGAON4          GACGTCTGAACCGGC---AATGATGATGCCGTCGC--CAACG--ATCGAC--------AC 456
ASSAM3            GACGTCTGAACCGGC---AATGATGCCGC-GTCGCCCCAACG--ATCGACTCGAATCGAC 453
ASSAM4            GACGTCTGAACCGGCGGCAATGATGCCGC-GTCGC--CAACG--ATCGACTCGAATCGAC 454
ASSAM1            GACGTCTGAACCGGG--CAATGATGATGCCGTCGC--CAACG--ATCGAC--------AC 445
                  ********** **    *** ***   * *****  *****  ******        **


ASSAM2            GCGTTATAGATTTATATCTATATATTTCTATCT 491
ASSAM5            GCGTTATAGATTTATATCTATATATTTTTATCT 491
VAISHALI2         GCGTTATAGATTTATATCTATATATTTCTATCT 493
NARAYANGAON2      GCGTTATAGATTTATATCTATATATTTCTATCT 493
NARAYANGAON4      GCGTTATAGATTTATATCTATATATTTCTATCT 493
NARAYANGAON3      GCGTTATAGATTTATATCTATATATTTCTATCT 493
TRICHY1           GCGTTATAGATTTATATCTATATATTTCTATCT 491
TRICHY3           GCGTTATAGATTTATATCTATATATTTCTATCT 491
TRICHY5           GCGTTATAGATTTATATCTATATATTTCTATCT 491
VAISHALI1         GCGTTATAGATTTATATCTATATATTTCTATCT 493
VAISHALI5         GCGTTATAGATTTATATCTATATATTTCTATCT 491
WAYANAD3          GCGTTATAGATTTATATCTATATATTTCTATCT 471
VAISHALI3         GCGTTATAGATTTATATCTATATATTTCTATCT 488
WAYANAD1          GCGTTATAGATTTATATCTATATATTTCTATCT 488
WAYANAD5          GCGTTATAGATTTATATCTATATATTTCTATCT 488
VAISHALI4         GCGTTATAGATTTATATCTATATATTTCTATCT 488
WAYANAD2          GCGTTATAGATTTATATCTATATATTTCTATCT 486
WAYANAD4          GCGTTATAGATTTATATCTATATATTTCTATCT 486
NARAYANGAON1      GCGTTATAGATTTATATCTATATATTTCTATCT 493
JALGAON1          GCGTTATAGATTTATATCTATATATTTCTATCT 489
JALGAON2          GCGTTATAGATTTATATCTATATATTTCTATCT 489
JALGAON3          GCGTTATAGATTTATATCTATATATTTCTATCT 489
```

```
JALGAON5          GCGTTATAGATTTATATCTATATATTTCTATCT 489
TRICHY2           GCGTTATAGATTTATATCTATATATTTCTATCT 489
TRICHY4           GCGTTATAGATTTATATCTATATATTTCTATCT 489
NARAYANGAON5      GCGTTATAGATTTATATCTATATATTTCTATCT 489
JALGAON4          GCGTTATAGATTTATATCTATATATTTCTATCT 489
ASSAM3            GCGTTATAGATTTATATCTATATATTTCTATCT 486
ASSAM4            GCGTTATAGATTTATATCTATATATTTCTATCT 487
ASSAM1            GCGTTATAGATTTATATCTATATATTTCTATCT 478
                  ************************** *****
```

Fig. 7. ClustalW alignment of the ITS2 sequences of thirty individuals

Five regions (I, II, III, V and VI) with variable number of simple repeats were identified in the ClustalW alignment which gave rise to indels (**Table 6**).

Table 6. Simple repeats in the ITS2 region

Type of repeat	Region # in the ClustalW alignment (Fig. 7)	`n' *i.e.* number of repeats
(CGT)	I	2-5
(CTC)	II	2-3
(AC)	III	1-2
(GC)	V	5-7
(GGC)	VI	1-2

Indels in regions IV and VII were caused due to presence of short stretches of `G' and `C' respectively while indels in regions VIII and IX were due to the insertion of short stretches of nucleotides.

Nineteen haplotypes were identified (**Table 7**) and the haplotype diversity (Hd) was estimated to be 0.947 using DnaSP v.5.10. Of the nineteen haplotypes, haplotypes 6, 11 and 15 were shared by a few individuals from atleast two locations.

Table 7. Haplotypes based on the ITS2 sequences

Haplotype	Individual	Total no
Hap 1	Assam 1	1
Hap 2	Assam 2	1
Hap 3	Assam 3	1
Hap 4	Assam 4	1
Hap 5	Assam 5	1
Hap 6	Jalgaon 1, 2, 3, 5, Trichy 2, 4	6
Hap 7	Jalgaon 4	1
Hap 8	Narayangaon 1	1
Hap 9	Narayangaon 2	1
Hap 10	Narayangaon 3	1
Hap 11	Narayangaon 4, Vaishali 2	2
Hap 12	Narayangaon 5	1
Hap 13	Trichy 1, 3, 5	3
Hap 14	Vaishali 1, 5	2
Hap 15	Vaishali 3, Wayanad 1, 5	3
Hap 16	Vaishali 4	1
Hap 17	Wayanad 2	1
Hap 18	Wayanad 3	1
Hap 19	Wayanad 4	1

InDel calculations were done using Model 1 with the option of Diallelic (non-overlapping). Total number of InDel sites analysed were eighteen, total number of InDel sites were fifty-five, total number of excluded overlapping InDel sites were thirty-seven and total number of InDel and non-InDel sites analysed were (458+18) 476. Total number of InDel events analysed were 7. Average InDel length event was 2.571.

Ts/Tv ratio of ITS2

The Kimura-2 parameter model was selected as the best fit model for the ITS2 data set of the 30 individuals as it gave the lowest BIC score of 2318.6 with lnL = -882.98. Substitution pattern and rates were estimated using this model is shown in **Table 8**.

Table 8. Maximum Likelihood Estimate of Substitution Matrix for the ITS2 sequences using Kimura – parameter model

	A	T/U	C	G
A	-	*3.20*	*3.20*	**18.60**
T/U	*3.20*	-	**18.60**	*3.20*
C	*3.20*	**18.60**	-	*3.20*
G	**18.60**	*3.20*	*3.20*	-

The transitions are more frequent than transversions with the G-A and C-T transitions occurring with equal frequencies. The Ts/Tv ratio of this data set was R=2.27 and the I_{ss} value was 0.0152 which was significantly less than Iss.c (0.6978) suggesting that substitutions in this data set have not reached saturation (Xia *et al*. 2003). This was also confirmed by the observation that transitions and transversions rates increased with genetic distance (**Fig. 8**).

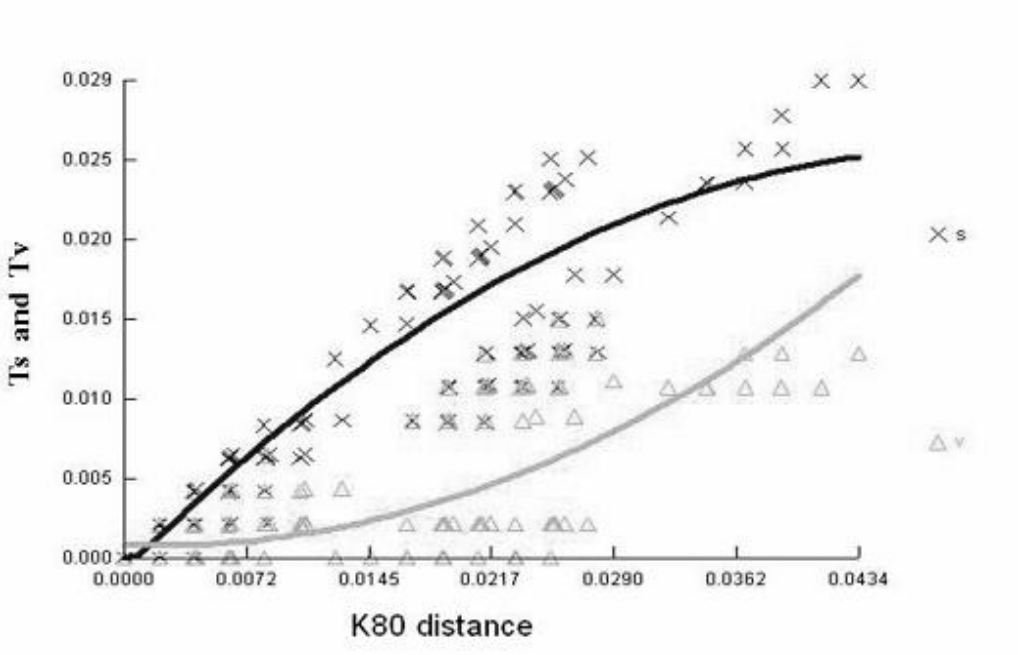


Fig. 8. Graph showing relationship between Ts and Tv and divergence in the ITS2 sequences

Genetic distance based on ITS2

The intra-population genetic distance ranged from 0.00943 (Narayangaon) to 0.01419 (Assam), with an overall mean diversity within populations of 0.008. The inter-population divergence ranged from 0.007 (between Vaishali and Wayanad) to 0.029 (between Assam and Jalgaon) with a mean inter-population diversity of 0.005 (S.E. 0.001).

Phylogenetic analysis based on the ITS2 region

The UPGMA tree derived using the Kimura-2 parameter model is shown in **Fig. 9**. The UPGMA trees based on indel polymorphism and haplotype data are shown in **Figs. 10 and 11**. In all the three trees, the overall grouping of individuals is similar with individuals 1, 3 and 4 from Assam forming a separate cluster. The grouping appears to be strongly influenced by haplotypes as individuals carrying any particular haplotype i.e. 6, 13, 14 and 15 always group together.

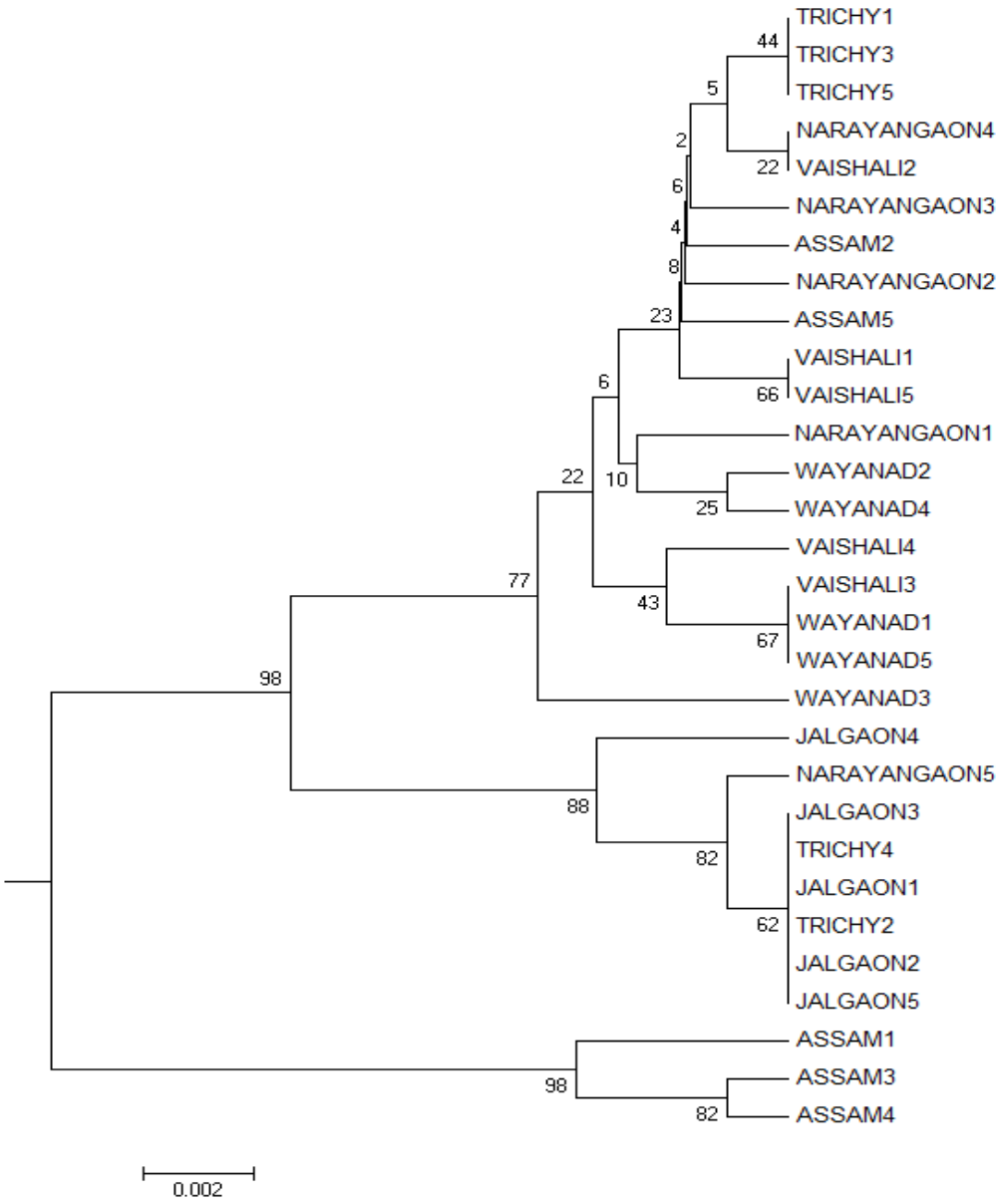


Fig. 9. UPGMA tree of the ITS2 sequences

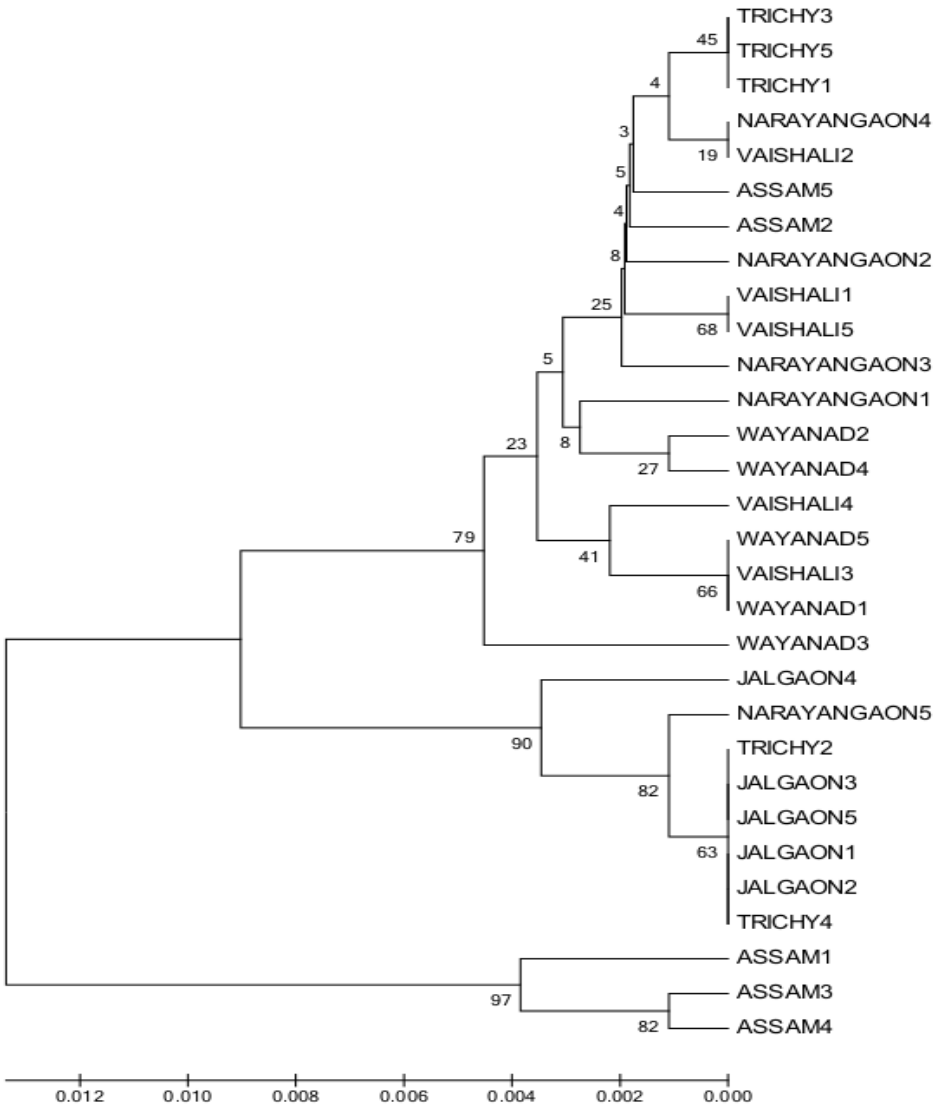


Fig. 10. UPGMA tree based on indel polymorphisms in the ITS2 sequences

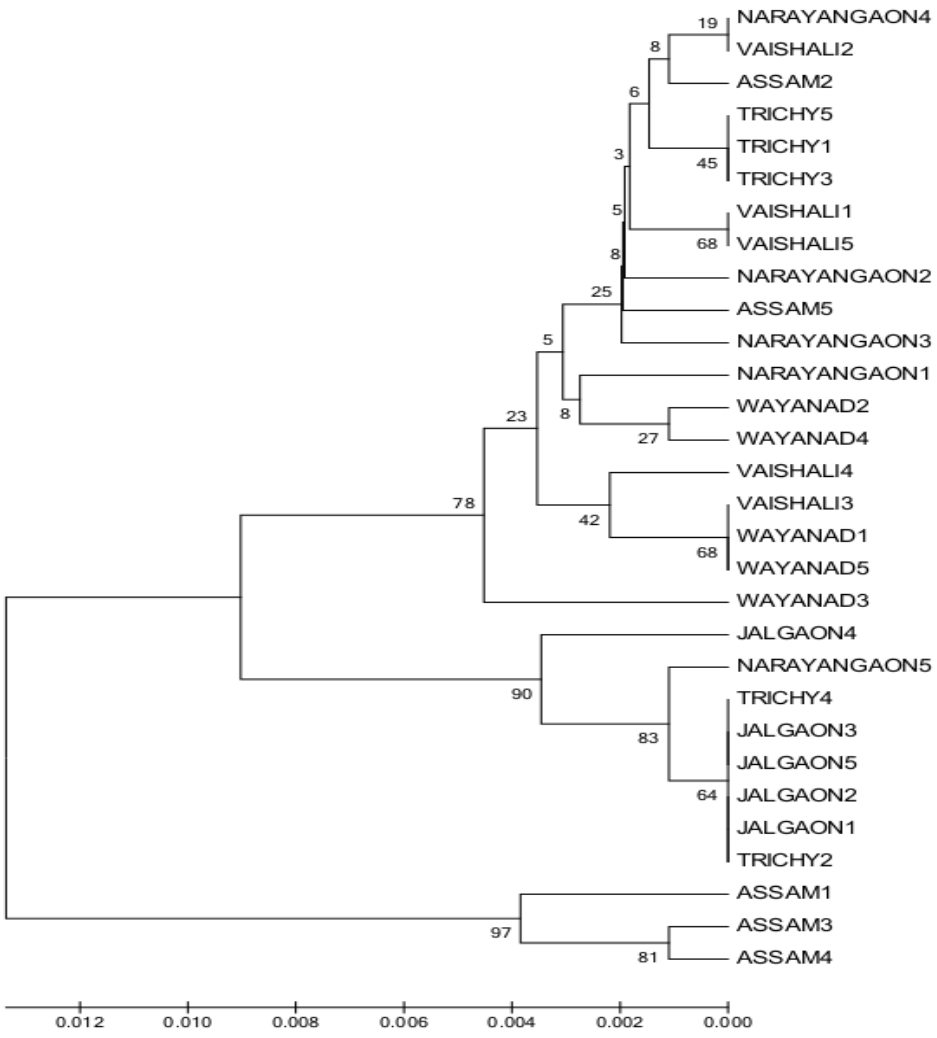


Fig. 11. UPGMA tree based on haplotye data of the ITS2 sequences

The ITS2 sequences of the thirty individuals show transitions and transversions as well as indels, some of which are caused by simple repeats. The ITS2 sequences of the Assam individuals showed maximum variation in length as compared to the rest of twenty-seven individuals due to the presence of six indels. The ITS2 sequences also revealed seven transversions (four G-C, two A-T and A-C) and three C-T transitions. Individuals 1, 2, 3 and 5 from Jalgaon were very similar with only Jalgaon 4 showing three C-T transitions and no size variation was observed amongst these five individuals. Amongst the individuals from Narayangaon, three indels and four transitions each of A-G

and C-T types were observed. Four indels, six transitions (three A-G, two C-T and one G-A) and one C-G transversion was observed in the ITS2 sequences from Trichy. Six indels, one transition each of C-T and A-G and one G-T transversion were present in the ITS2 sequences from Vaishali. Individuals 3 and 4 from Vaishali differed in a single position due to a C-T transition. The Wayanad ITS2 sequences showed two indels of which the stretch 'AGAGGAGAGGAGGGT' was absent only in individual 3. In addition, one A-G transition and one C-G transversion were observed in these sequences.

Correlation between genetic distance and geographical distance was observed (R^2=0.7319; p = 0.070) (**Fig. 12**) but no phylogeographic distribution of the populations was evident from the phylogenetic trees. Genetic differentiation estimates using DnaSP v5.10 (G_{ST}=0.15094; Nm=1.41; χ^2, 120.000; p value =0.0190 (p< 0.05); df =90) confirmed that the populations are not isolated, thereby suggesting gene flow between the populations.

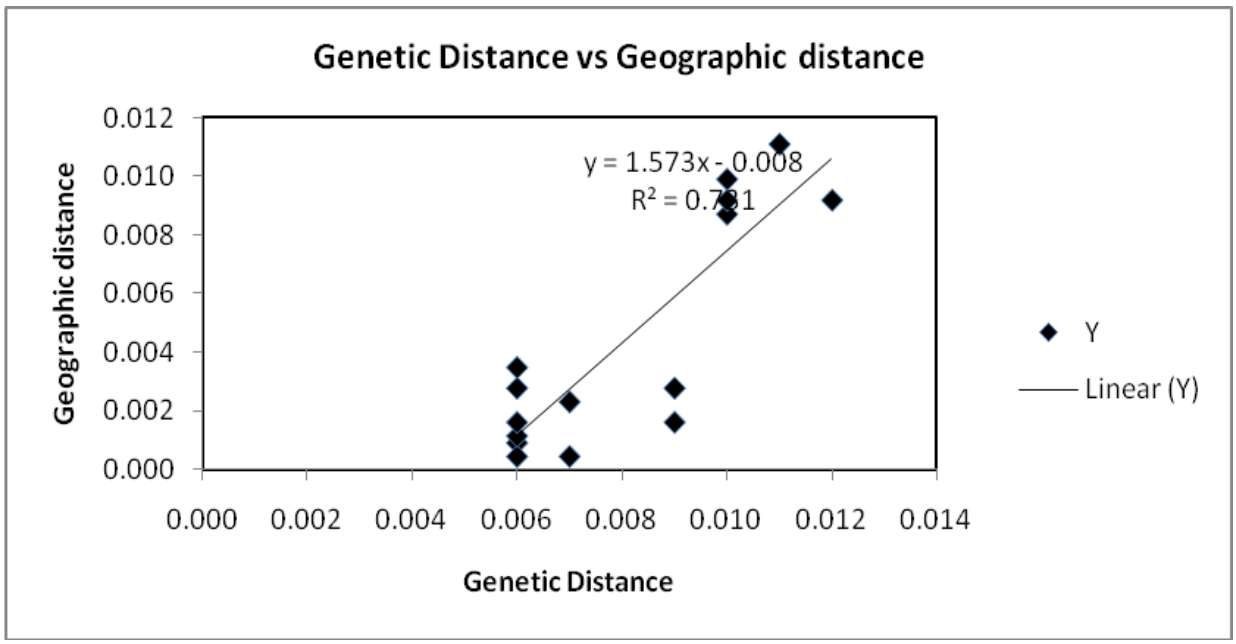


Fig. 12. Geographic vs genetic distance (Mantel test) for the ITS2 data set

Phylogenetic analysis based on concatenated data of ITS1 and ITS2

The concatenated data set which was generated by DnaSP v.5.10 from the ITS1 and ITS2 data sets had 1452 sites of which 61 were variable and 33 were parsimony informative. The best fit model for the concatenated data with the lowest BIC value of 5207.7 was the T92 model with Gamma distribution and no invariant sites. The UPGMA trees generated using this data as well as haplotype and indel polymorphism show two clusters i.e. Assam individuals 2, 3, 4 and 5 forms the smaller cluster while the remaining twenty-seven individuals form the major group (**Figs. 13, 14 and 15**). This grouping is similar to that observed in the trees derived from the ITS1 data. However, within the major cluster, grouping is according to ITS2 haplotypes.

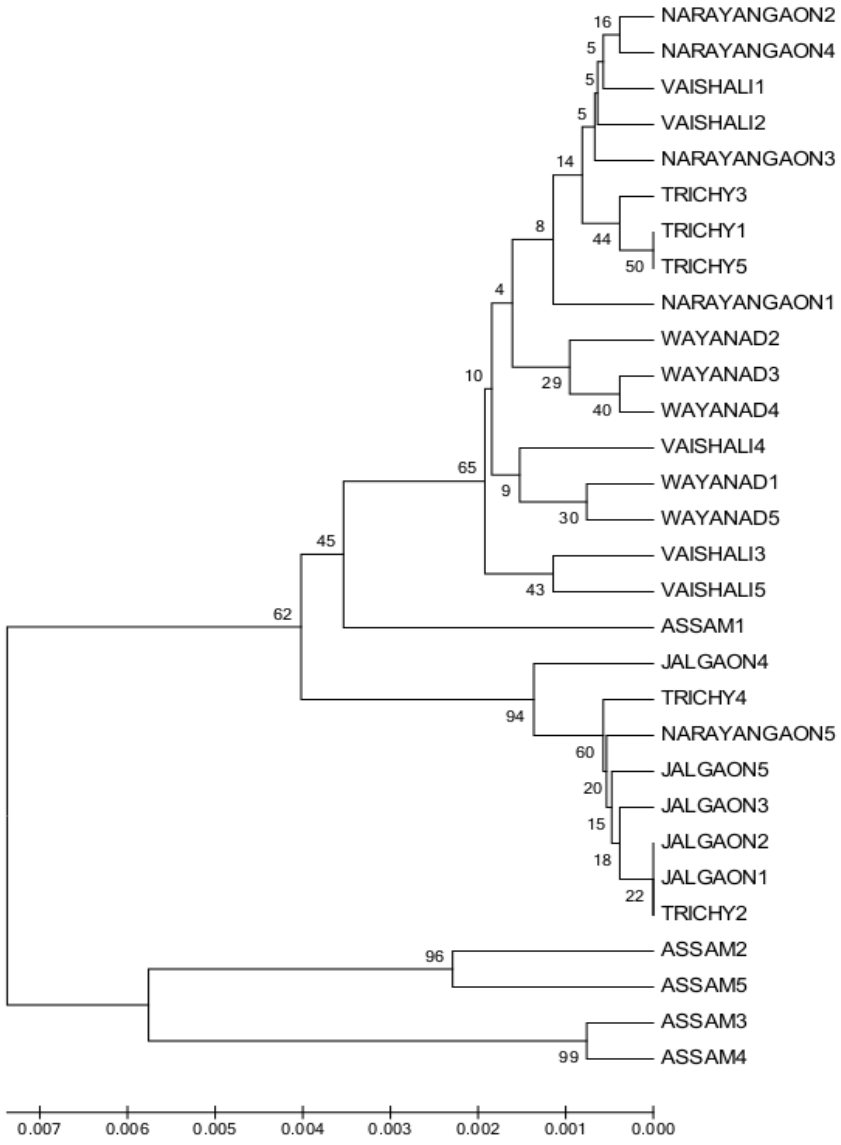


Fig. 13. UPGMA tree based on concatenated data sets of ITS1 and ITS2

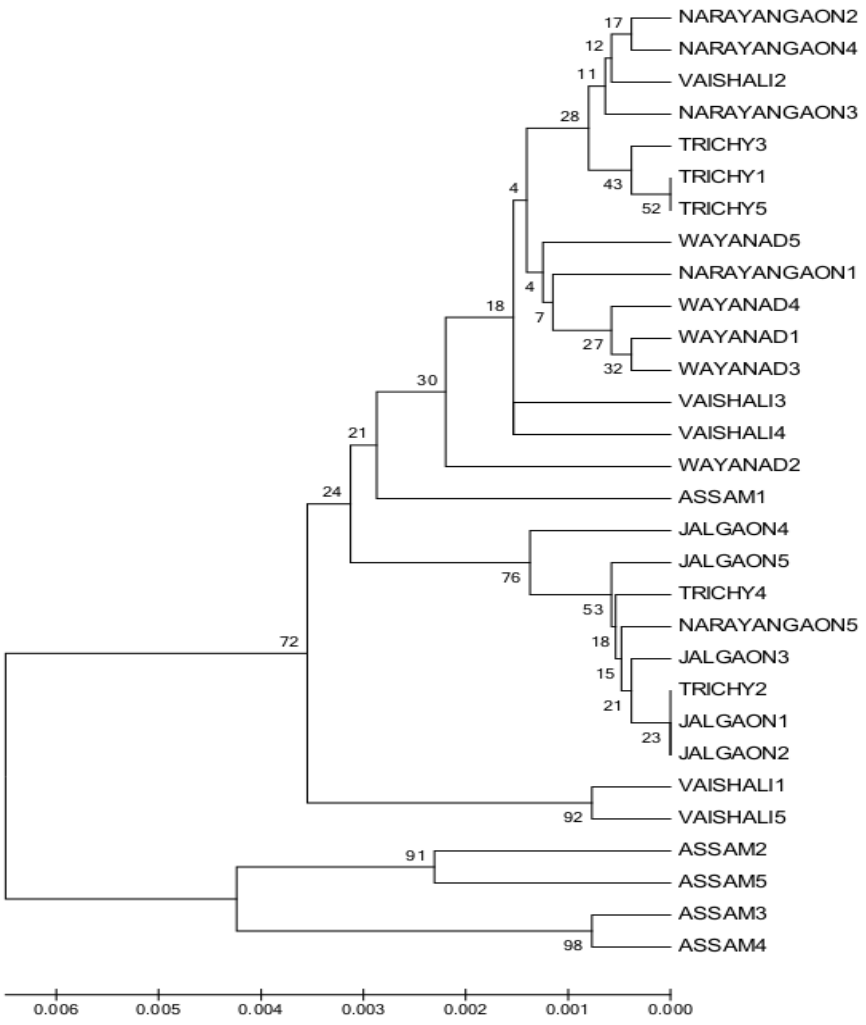


Fig. 14. UPGMA tree based on haplotype in the concatenated data set

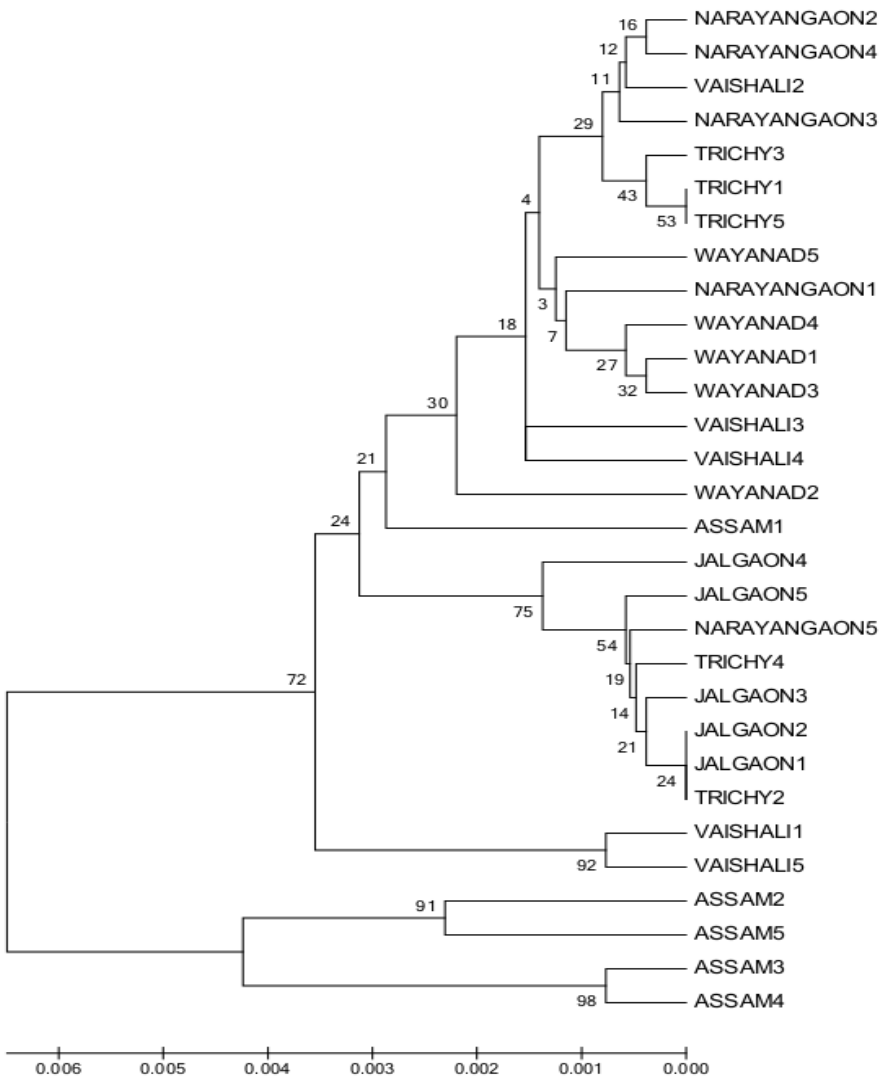


Fig. 15. UPGMA tree based on indel polymorphism in the concatenated data set

Secondary Structure of the ITS1 region

The consensus secondary structure derived by considering the ITS1 sequences from all the 30 individuals used in this study is shown in **Fig. 16**.

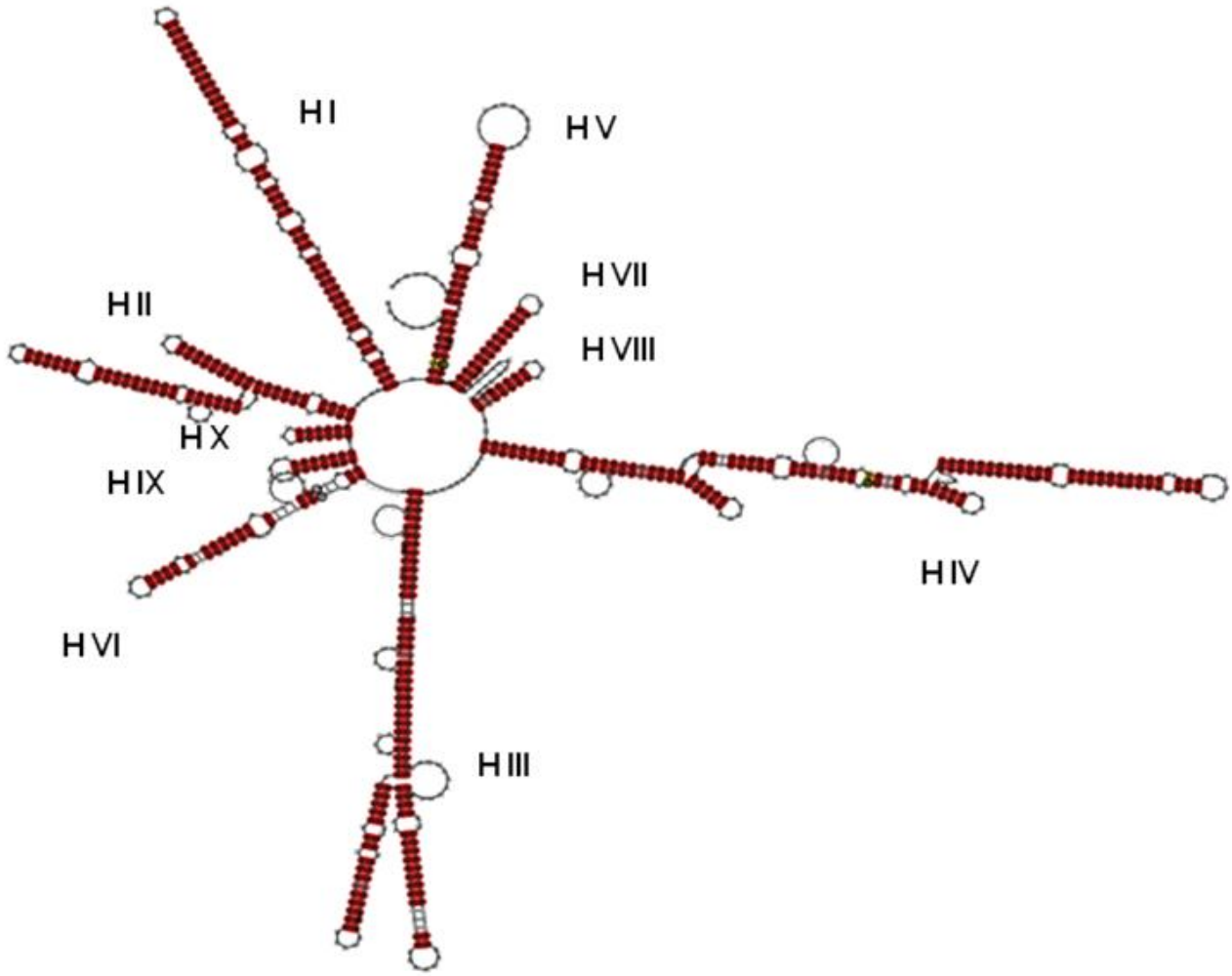


Fig. 16. Consensus RNA secondary structure of the ITS1 region

The ITS1 secondary structure consists of an open multibranch loop having a total of ten helices of paired regions. Helix IV is the longest (74 bp), Helix III (66 bp), Helix I (41 bp), Helix II (12 bp), Helix V (26 bp), Helix VI (23 bp), Helix VII (12 bp), Helix VIII (7 bp), Helix IX (7 bp) and helix X (6 bp). Of these Helix VI has a stretch which is less conserved. The proportion of AU pairs in these helices are as follows, Helix VII has a

higher proportion of AU pairs (9/12 AU) 75%), Helix VIII (6/7 AU) 85%, Helix IV (45/74 AU) 60%, Helix III (62/66 AU) 48.48%, Helix II (20/40 AU) 50%. Helices IV, VII and VIII have a high AU concentration of 60%, 75% and 85% respectively. The proportion of GU pairing is low approximately 5.8%, indicating high stability of the helices. The seven indels (1, 2, 3, 4) project as bulges in the helices with indels 5, 6 and 7 appearing as loops at the tip of helices VIII, VII and V respectively.

Adalia bipunctata (Accession no AJ272139) was used as reference organism in an attempt to identify putative homologous conserved regions in *O. longicollis* (Oliver) (Schulenberg *et al*. 1999). The sequence alignment showed few conserved and variable regions in the ITS1 of *O. longicollis* (Oliver) and *Adalia bipunctata* **(Fig. 17**).

Block D

```
Adalia bipunctata      AGGTATTAAAGATTTTCGCCCGA 1051
Assam1                 ATGTCCTGTAGACGTTTTCGAAA 773
Assam2                 GTGTCCTGTAGACGTTTTCGAAA 771
Assam3                 GTGTCCTGTAGACGTTTTCGAAA 765
Assam4                 GTGTCCTGTAGACGTTTTCGAAA 765
Assam5                 GTGTCCTGTAAACGTTTTCGAAA 762
Jalgaon1               ATGTCCTGTAGACGTTTTCGAAA 769
Jalgaon2               ATGTCCTGTAGACGTTTTCGAAA 769
Jalgaon3               ATGTCCTGTAGACGTTTTCGAAA 783
Jalgaon4               ATGTCCTGTAGACGTTTTCGAAA 769
Jalgaon5               ATGTCCTGTAGACGTTTTCGAAA 769
Narayangaon1           ATGTCCTGTAGACGTTTTCGAAA 770
Narayangaon2           ATGTCCTGTAGACGTTTTCGAAA 769
Narayangaon3           ATGTCCTGTAGACGTTTTCGAAA 769
Narayangaon4           ATGTCCTGTAGACGTTTTCGAAA 769
Narayangaon5           ATGTCCTGTAGACGTTTTCGAAA 769
Trichy1                ATGTCCTGTAGACGTTTTCGAAA 769
Trichy2                ATGTCCTGTAGACGTTTTCGAAA 770
Trichy3                ATGTCCTGTAGACGTTTTCGAAA 770
Trichy4                ATGTCCTGTAGACGTTTTCGAAA 771
Trichy5                ATGTCCTGTAGACGTTTTCGAAA 775
Vaishali1              ATGTCCTGTAGACGTTTTCGAAA 769
Vaishali2              ATGTCCTGTAGACGTTTTCGAAA 778
Vaishali3              ATGTCCTGTAGACGTTTTCGAAA 780
Vaishali4              ATGTCCTGTAGACGTTTTCGAAA 769
Vaishali5              ATGTCCTGTAGACGTTTTCGAAA 780
Wayanad1               ATGTCCTGTAGACGTTTTCGAAA 758
Wayanad2               ATGTCCTGTAGACGTTTTCGAAA 758
Wayanad3               ATGTCCTGTAGACGTTTTCGAAA 758
Wayanad4               ATGTCCTGTAGACGTTTTCGAAA 758
Wayanad5               ATGTCCTGTAGACGTTTTCGAAA 759
                         **  *  * *  **  *   *
```

Fig. 17. ClustalW alignment of the ITS1 sequences of the thirty individuals and *Adalia bipunctata* showing the Block D region

In *A. bipunctata*, the three conserved regions identified were **(i)** 8 bp of adjacent 18S rRNA gene and 4 bp at the 5'end of ITS1, **(ii)** about 20 bp in the middle of the spacer (Block D), and **(iii)** 87 bp of the adjacent 5.8S rRNA and 8 bp at the 3'end of ITS1 (Schulenberg *et al*. 1999). The forward primer used to amplify ITS1 of *O. longicollis* (Oliver) overlaps with the conserved region of 8 bp of *Adalia punctata*. The second region conserved in *A. punctata* and related taxa is the D Block. This region was highly homologous amongst the ITS1 sequences of the thirty individuals with A to G transition observed in a single position in Assam individuals 2, 3, 4 and 5. Only 10 positions are conserved in this region as compared to *A. punctata.* The third region could not be identified because of the length variation between *O. longicollis* and *A. bipunctata* ITS1 sequence. The putative Block D region which has been identified in *O.longicollis* (Oliver) individuals does not show very high similarity with *A. punctata*, possibly because these are distant taxa.

Phylogenetic tree based on ITS1 secondary structure

The phylogenetic tree based on the secondary structure of ITS1 using Wards distance is shown in **Fig. 18.**

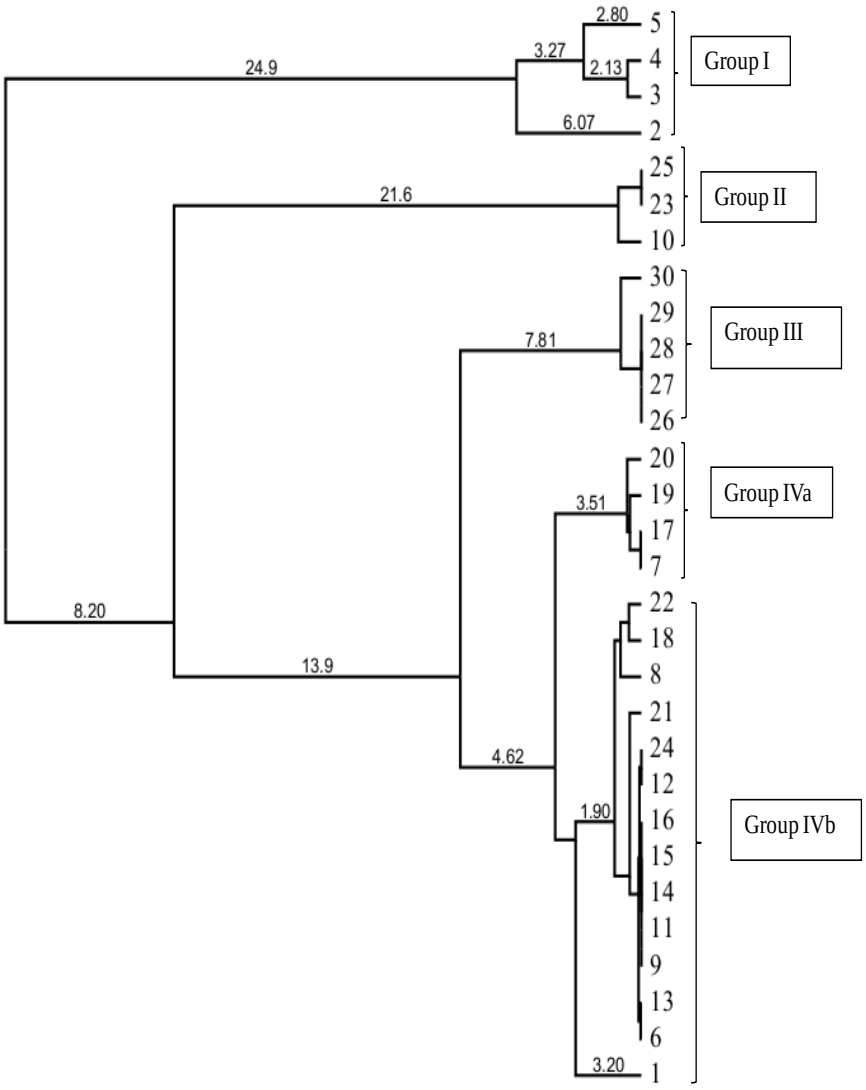
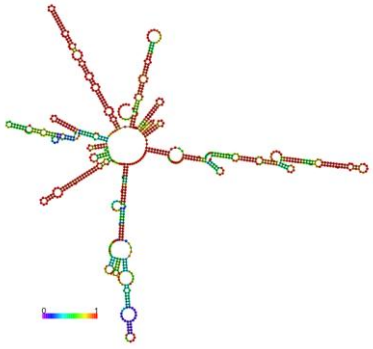


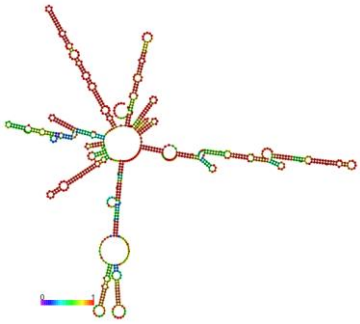
Fig. 18. Phylogenetic tree based on the secondary structure of ITS1

The thirty individuals cluster into four groups *i.e.* I, II, III, IVa and IVb. The major group (IV) comprises of all the individuals of haplotype 1, 6, 8, 9 and 10 while individuals 2, 3, 4 and 5 from Assam form a distinct cluster I. This grouping appears to be influenced by ITS1 haplotypes similar to that observed in the UPGMA trees derived from the ITS1 sequence data

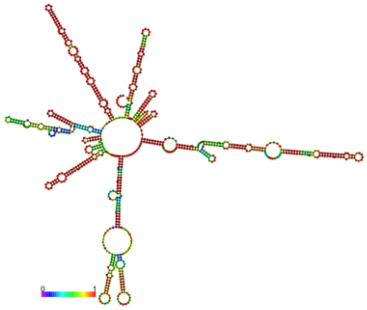
Following are the characteristics structural features specific for each of the above mentioned four groups: (**Fig. 19**).



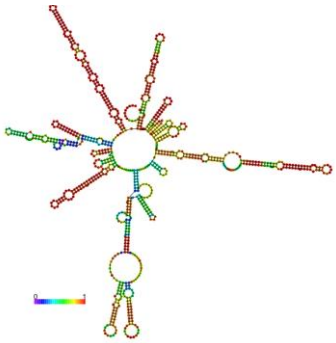
Assam 2 (Grp 1)



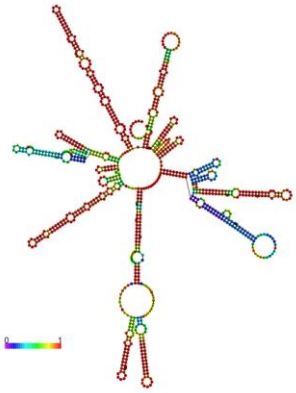
Assam 3 (Grp 1)



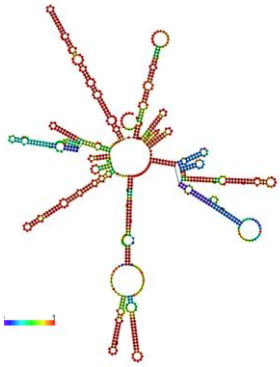
Assam 4 (Grp 1)



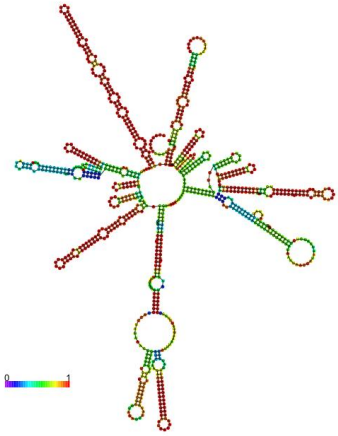
Assam 5 (Grp 1)



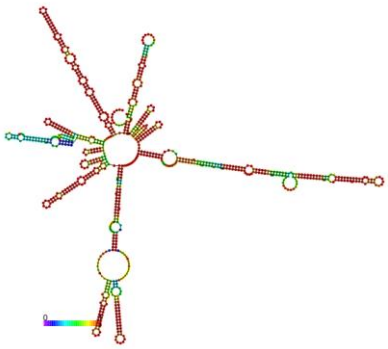
Vaishali 5 (Grp 2)



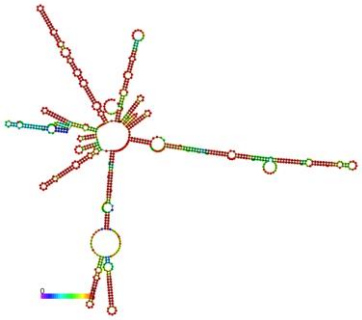
Vaishali 3 (Grp 2)



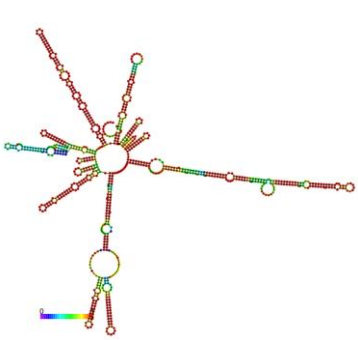
Jalgaon 5 (Grp 2)



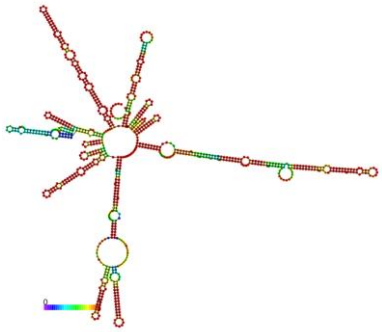
Wayanad 1 (Grp III)



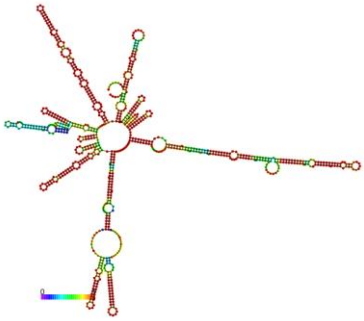
Wayanad 2 (Grp III)



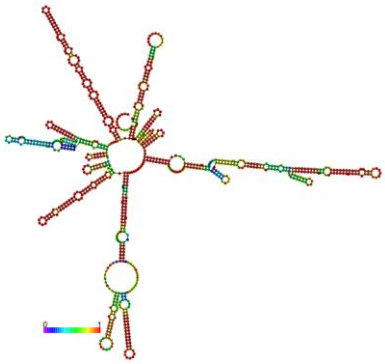
Wayanad 3 (Grp III)



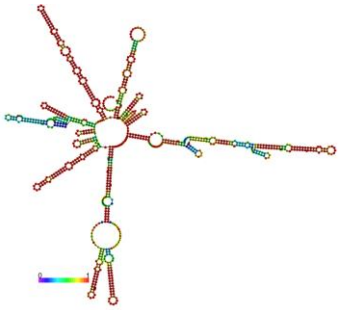
Wayanad 4 (Grp III)



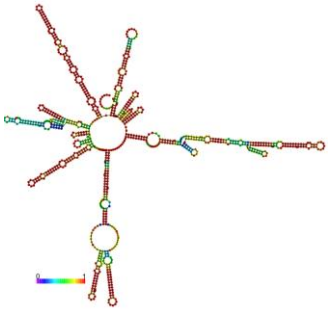
Wayanad 5 (Grp III)



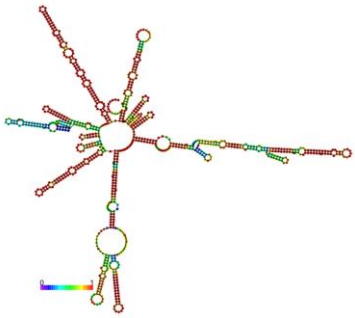
Trichy 2 (Grp IVa)



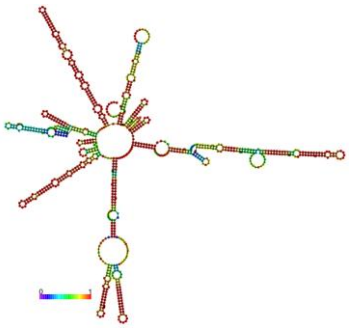
Trichy 4 (Grp IVa)



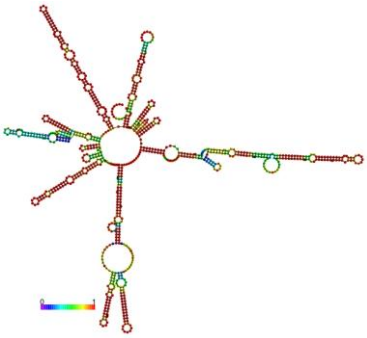
Trichy 5 (Grp IVa)



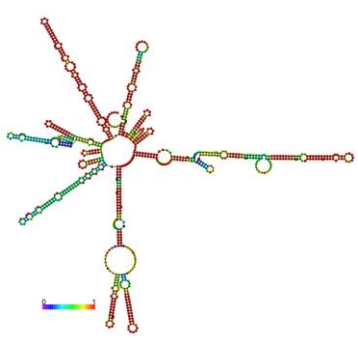
Jalgaon 2 (Grp IVa)



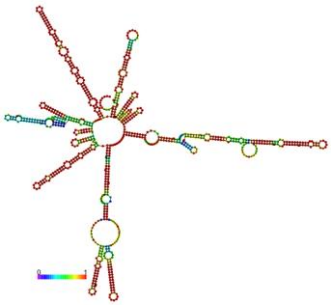
Vaishali 2 (Grp IVb)



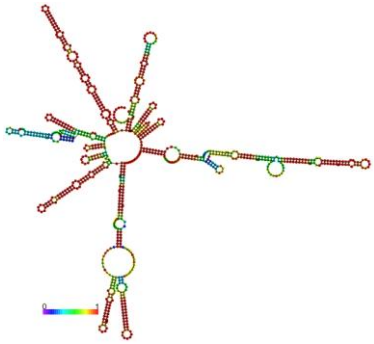
Trichy 3 (Grp IVb)



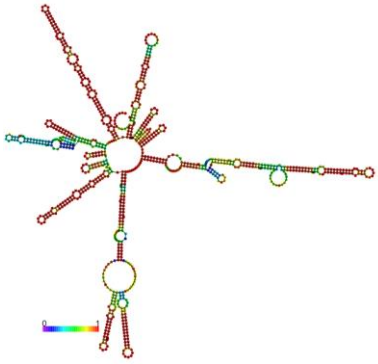
Jalgaon 3 (Grp IVb)



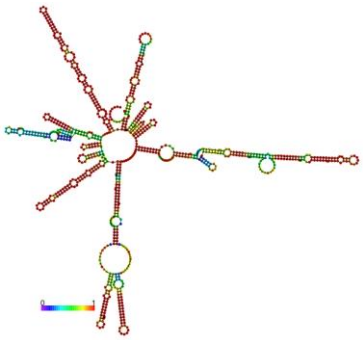
Vaishali 1 (Grp IVb)



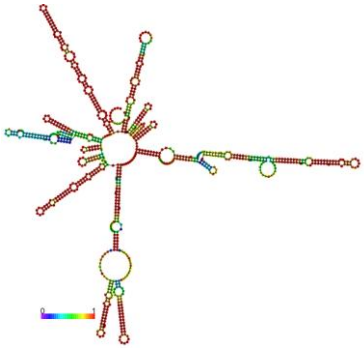
Vaishali 4 (Grp IVb)



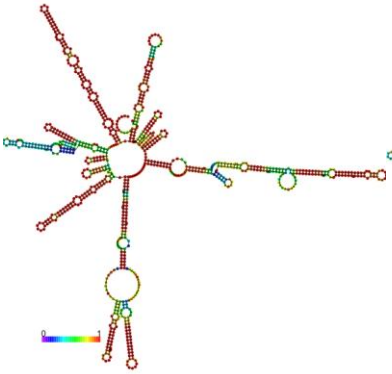
Narayangaon 2 (Grp IVb)



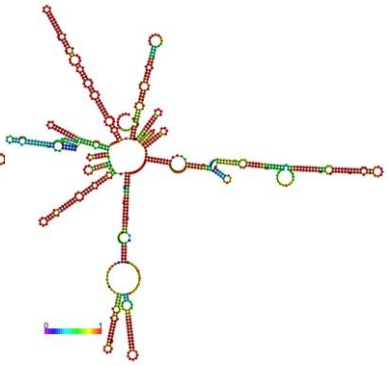
Trichy 1 (Grp IVb)



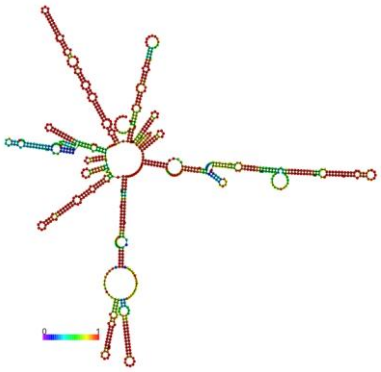
Narayangaon 5 (Grp IVb)



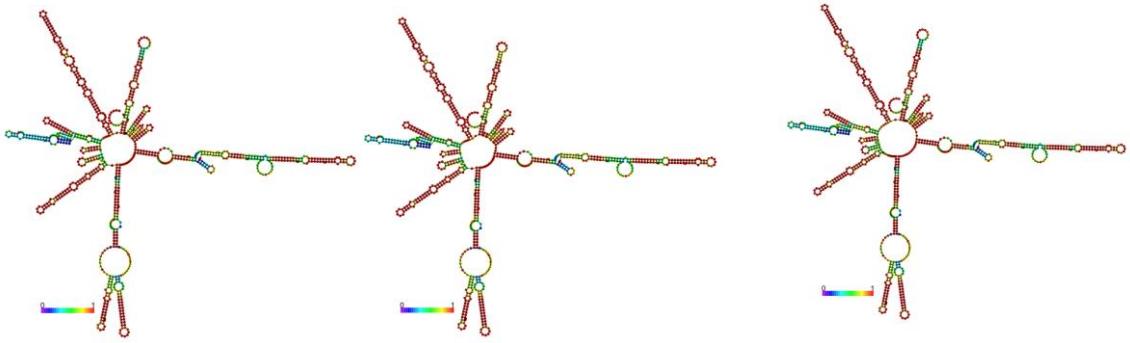
Narayangaon 4 (Grp IVb)



Narayangaon 1 (Grp IVb)



Jalgaon 4 (Grp IVb)



Narayangaon 3 (Grp IVb) **Jalgaon 1 (Grp IVb)** **Assam 1 ITS1 (Grp IVb)**

Fig. 19. Secondary structures of the ITS1 region of the thirty individuals included in the present study

(i) Individuals of group I have four bulges and one loop on Helix II while the secondary structure of the rest of the individuals exhibit only 3 bulges. The minor helix present in between Helix II and its branch is absent in this group but is present in all the other individuals. Helix VI in Group I individuals has only four bulges whereas the remaining groups show a minimum of five bulges.

(ii) Helix IV in Group II individuals has four branches emerging from a prominent loop. Such branching is absent in the other groups.

(iii) Helix IV in Group III individuals has no branches. Their secondary structure exhibits five bulges and one loop unlike individual 5 from Assam which also exhibits six bulges but no loop.

(iv) Helix IV in Group IVa individuals *i.e.* individuals 2 (Jalgaon) and 2, 4 and 5 (Trichy) has two branches similar to individuals 2 and 3 of Assam but differ from the Assam individuals in the absence of a loop at the base of the second branch. In Group IVb individuals, the second branch is replaced by a loop.

Secondary structure of ITS2 region

The regions flanking the ITS2 sequence, *i.e.* the 3' end of the 5.8S gene and the 5' end of the 28S gene have been implicated in the correct processing of these genes and hence have been suggested to play an important role in directing the folding of the ITS2 region (Keller *et al.* 2009). In the present study, the flanking 25 bp regions from each of 5.8S and 18S were observed to form a fairly stable stem structure i.e. HMM-based ITS2 annotation (Keller *et al*. 2009) (**Fig. 20**).

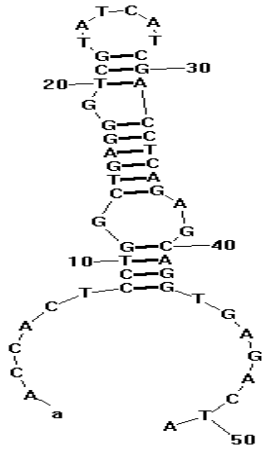


Fig. 20. 5.8S and 28S rRNA interaction and HMM-based ITS2 annotation of *O. longicollis* (Oliver). Annotation shows interaction of 25 nucleotides from the 5' end of 5.8S rRNA sequence with 25 nucleotides from the 5' end of 28S rRNA sequence

HMM is a mathematical model based on statistical approach used in computational bioinformatics for deducing the 2D and 3D structures of biological macromolecules. Most of the bioinformatics based modelling softwares use this model to build the secondary structure from the nucleotide sequences and has been proved very useful in solving the structure relates problems (Keller *et al.* 2009). This additional structural element did not change the predicted folding of the remainder of the ITS2 suggesting that formation of the secondary structure of the ITS2 region is not affected by adjacent coding regions.

The consensus secondary structure of the ITS2 region derived from thirty individuals conforms to the four helix `Pan-eukaryote ITS2 model' suggested for eukaryotes (Coleman 2007) (**Figs 21 and 22**).

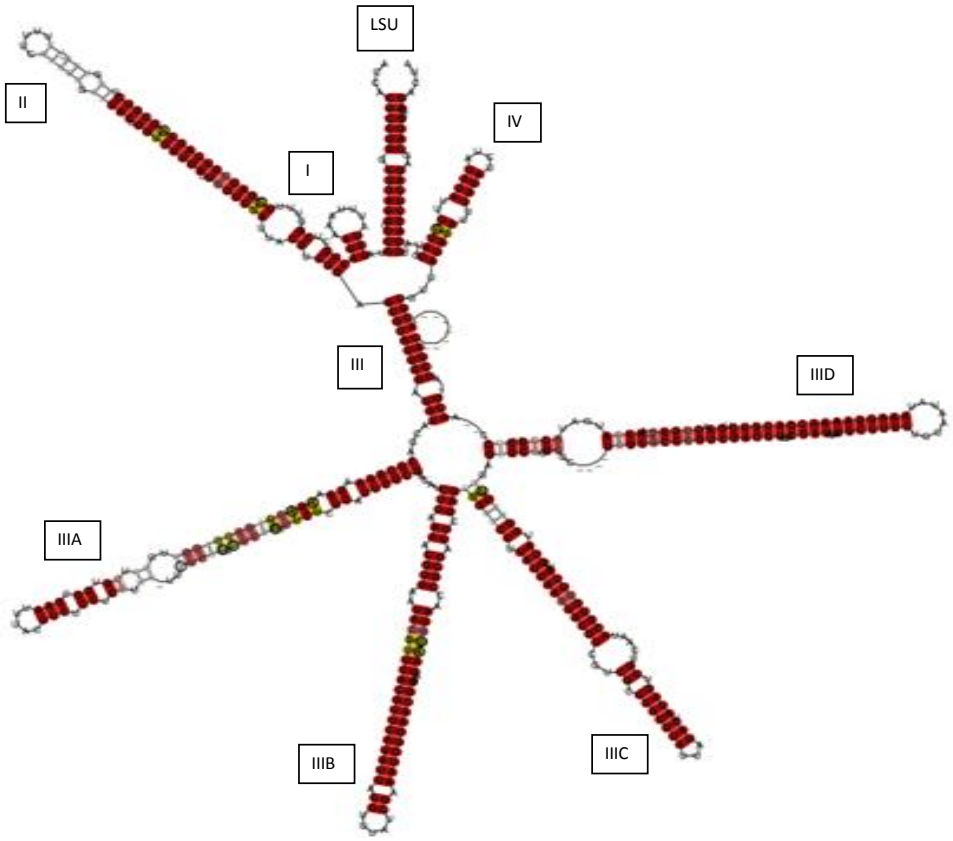
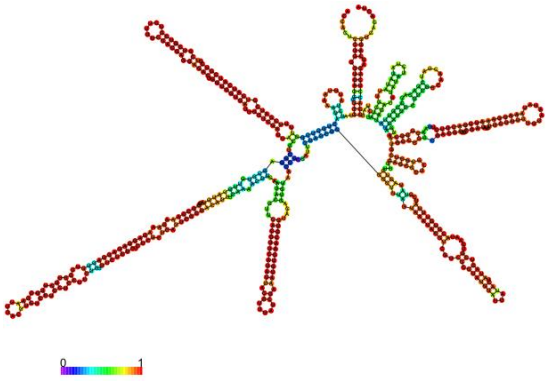
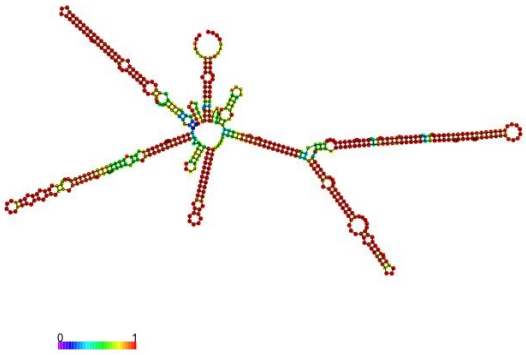


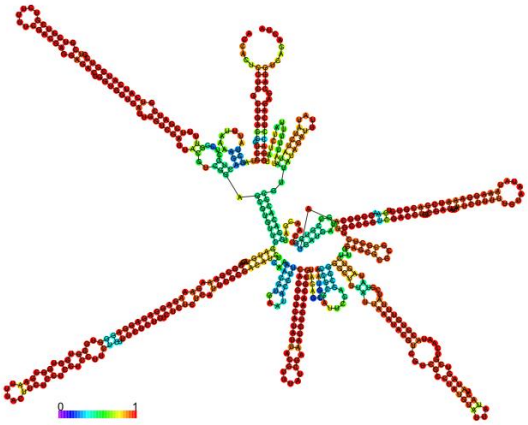
Fig. 21. Consensus RNA secondary structure of the ITS2 region



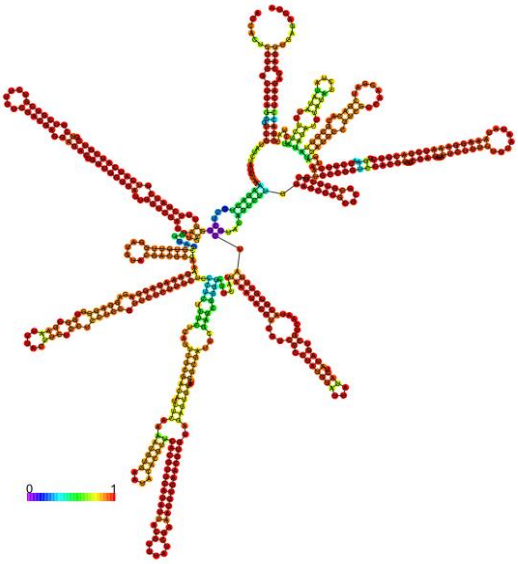
Assam 2 (Grp I)



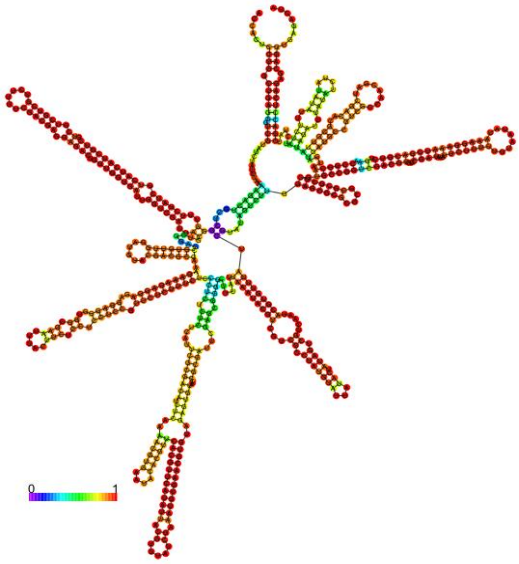
Assam 3 (Grp I)



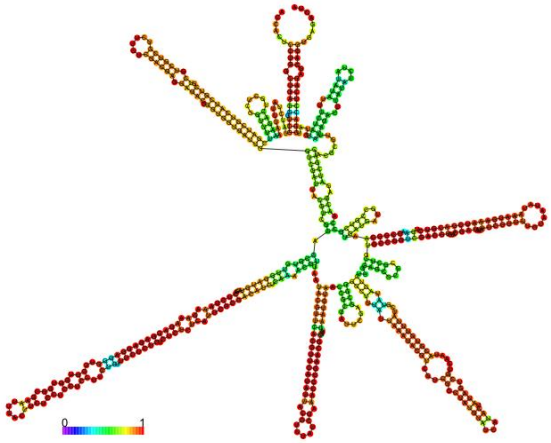
Assam 4 (Grp I)



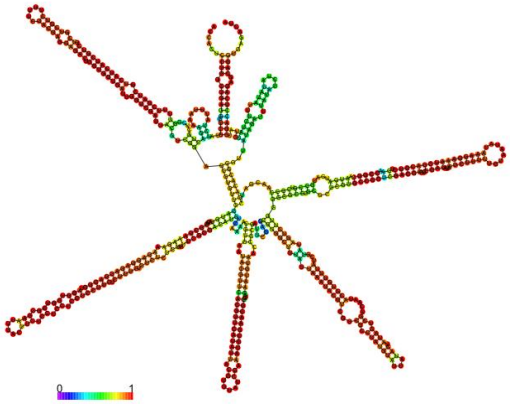
Assam 5 (Grp I)



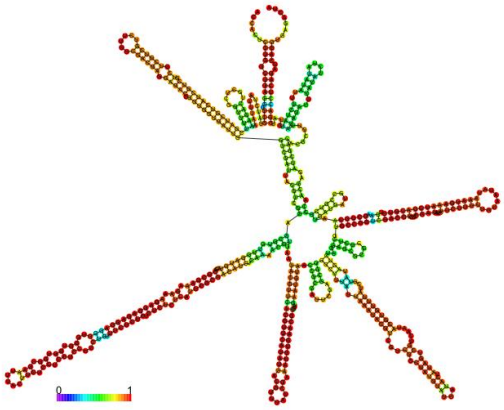
Vaishali 5 (Grp II)



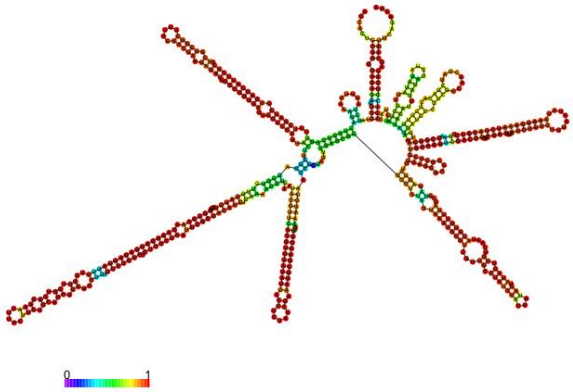
Vaishali 3 (Grp II)



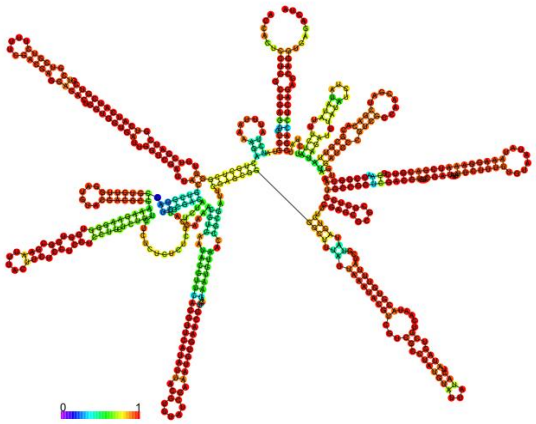
Jalgaon 5 (Grp II)



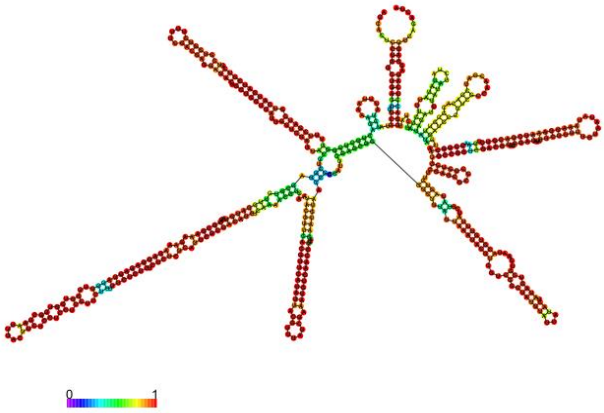
Wayanad 1 (Grp III)



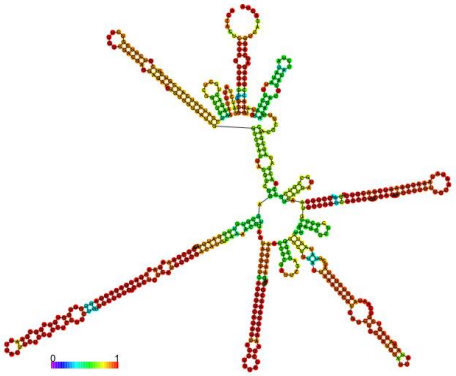
Wayanad 2 (Grp III)



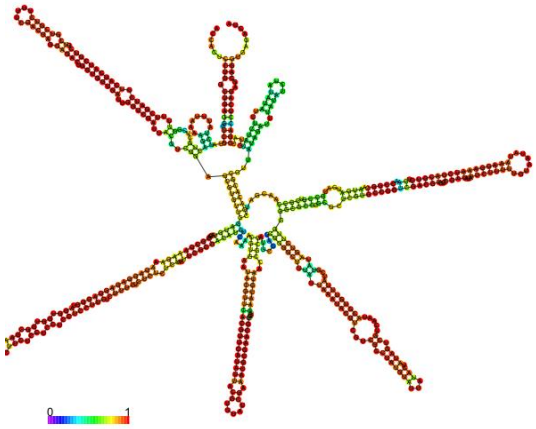
Wayanad 3 (Grp III)



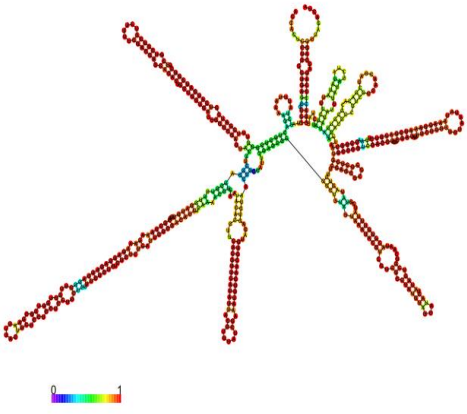
Wayanad 4 (Grp III)



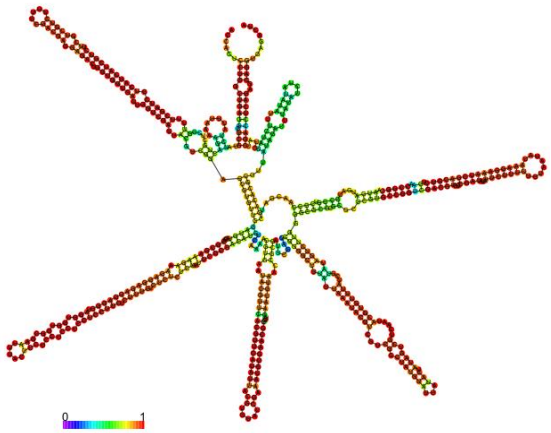
Wayanad 5 (Grp III)



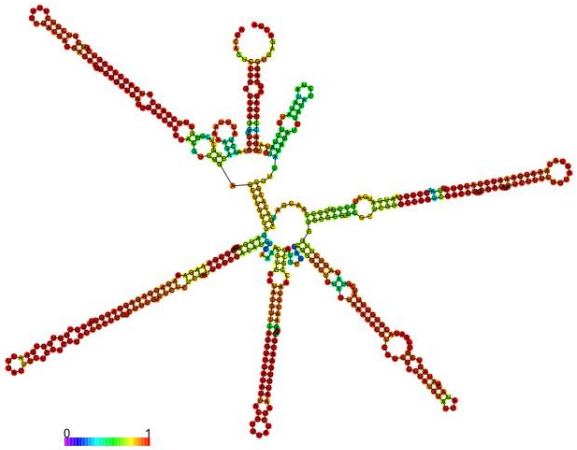
Trichy 4 (Grp IVa)



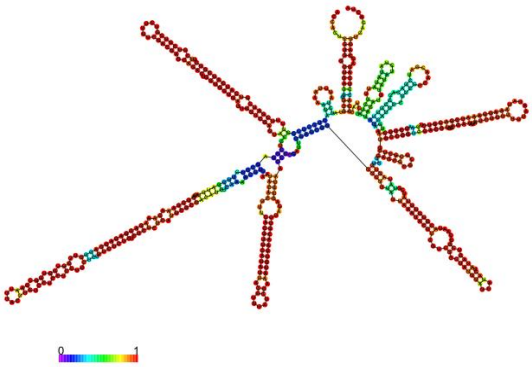
Trichy 5 (Grp IVa)



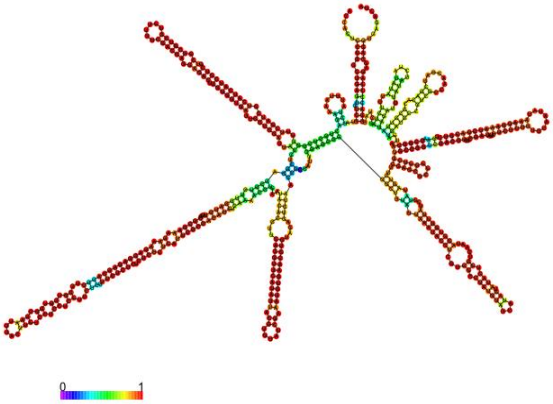
Trichy 2 (Grp IVa)



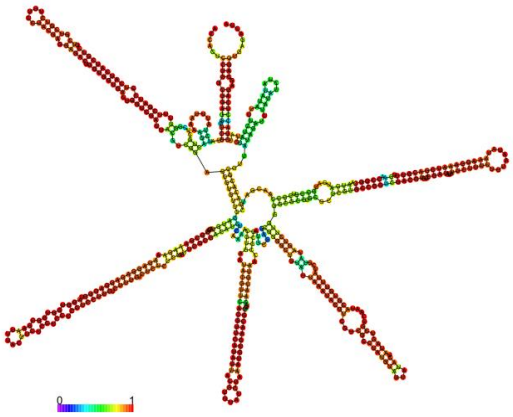
Jalgaon 2 (Grp IVa)



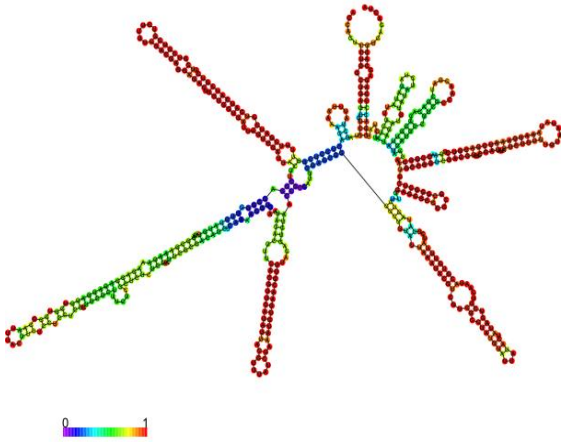
Vaishali 2 (Grp IVb)



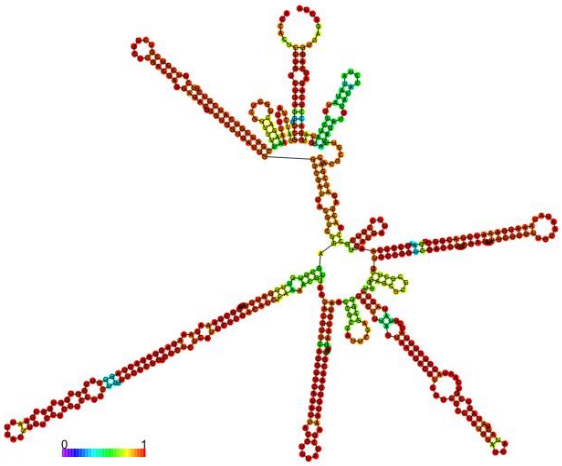
Trichy 3 (Grp IVb)



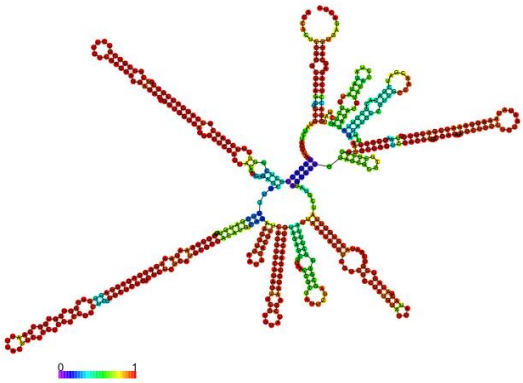
Jalgaon 3 (Grp IVb)



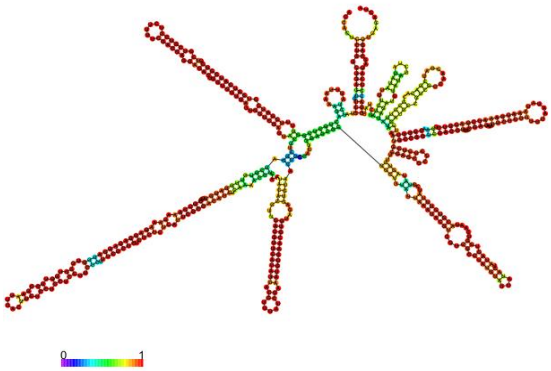
Vaishali 1 (Grp IVb)



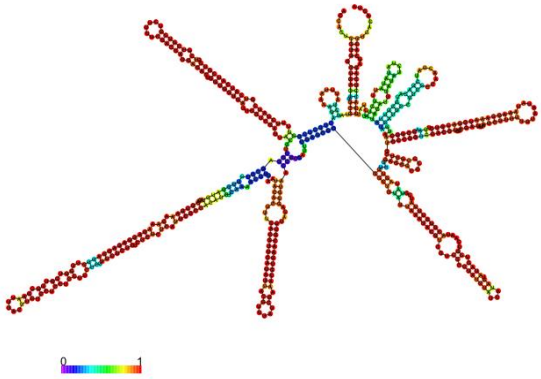
Vaishali 4 (Grp IVb)



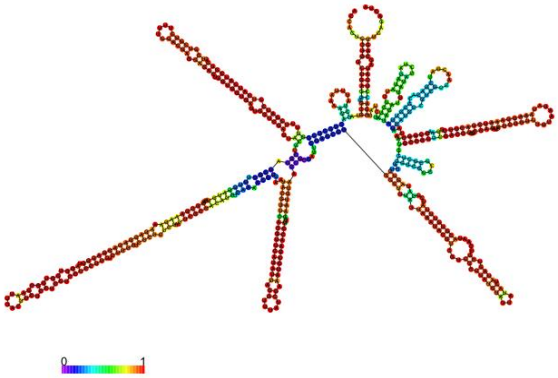
Narayangaon 2 (Grp IVb)



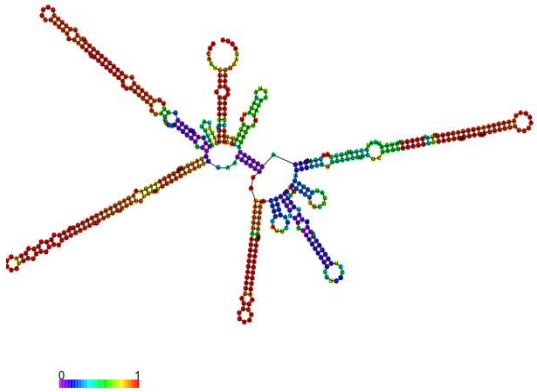
Trichy 1 (Grp IVb)



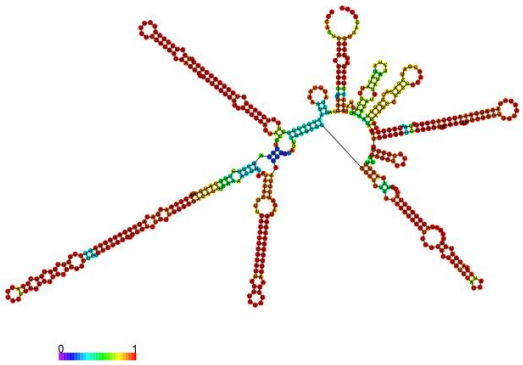
Narayangaon 4 (Grp IVb)



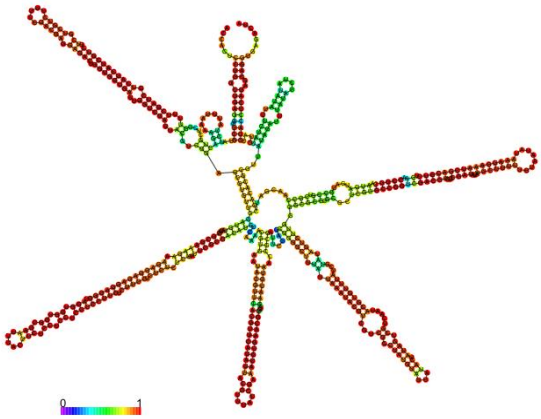
Narayangaon 5 (Grp IVb)



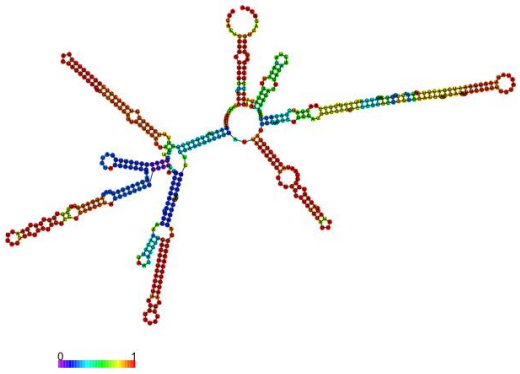
Jalgaon 4 (Grp IVb)



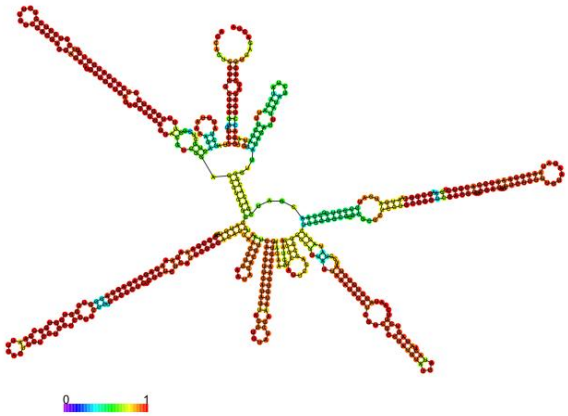
Narayangaon 3 (Grp IVb)



Jalgaon 1 (Grp IVb)



Assam 1 (Grp IVb)



Narayangaon 1 (Grp IVb)

Fig. 22. Secondary structures of the ITS2 region of the thirty individuals included in the present study

The following are the features of the consensus secondary structure,:

(i) A closed loop with four helices *i.e.* helix I, II, III (III A, III B, III C, III C and IIID) and IV

(ii) Helix I is the smallest and is 3 bp long with GC% of 33.3%.

(iii) Helix II is 30 bp long and is highly conserved excepting for the pyrimidine-pyrimidine mismatch, (U-U) which forms a bulge near the base of the helix. This bulge is a hallmark of Helix II. The % GC and % GU are 15% and 16.6% respectively. The loop region is 5 bases long and includes Indel 1.

(iv) Helix III is the longest of the four helices with four branches i.e. IIIA, IIIB, IIIC and IIID. Helix III is frequently branched as described in the pan-eukaryotic model which is also observed in the secondary ITS2 structure in the present study. The base of the helix has indel no. X. Helix branch IIIA is 31bp long with 17% GC and the loop is 5 bases long. It has indels 2 and 3. There are three small bulges in the helix and loop is 5 bases long. Helix branch IIIB branch is 29 bp long with 12% GC, 3% GU and the loop is 5 bases long. There are four small bulges caused by indel 4. Helix branch IIIC is 28 bp long with 7% GC and 3% GU. Loop is 3 bases long with one big and two small bulges. This helix has indels 5 and 6. Helix branch IIID is 34 bp long with 19% GC and 2% GU with a 8 bases long loop. This helix has indels 8 and 9 and also has a big bulge near the base of the secondary loop. Helix III, in contrast, is longer than helix II and is branched. In Coelomates and some other taxa a highly branched helix III has been reported. Helix III of ITS2 is long and multiply branched as observed in mammals and human (Denboh *et al*. 2003; Coleman 2007).

(v) Helix IV is at the 3’end of the ITS2. It is 8 bp long, has one GC pair and two U-U mismatches. The end loop is four bp long.

Phylogenetic tree based on the secondary structure of the ITS2 region

The topology of the phylogenetic tree based on the secondary structure of the ITS2 region (**Fig. 23**) is strikingly similar to the ITS1 secondary structure tree. However, no ITS2 haplotype based grouping is evident in the ITS2 secondary structure tree.

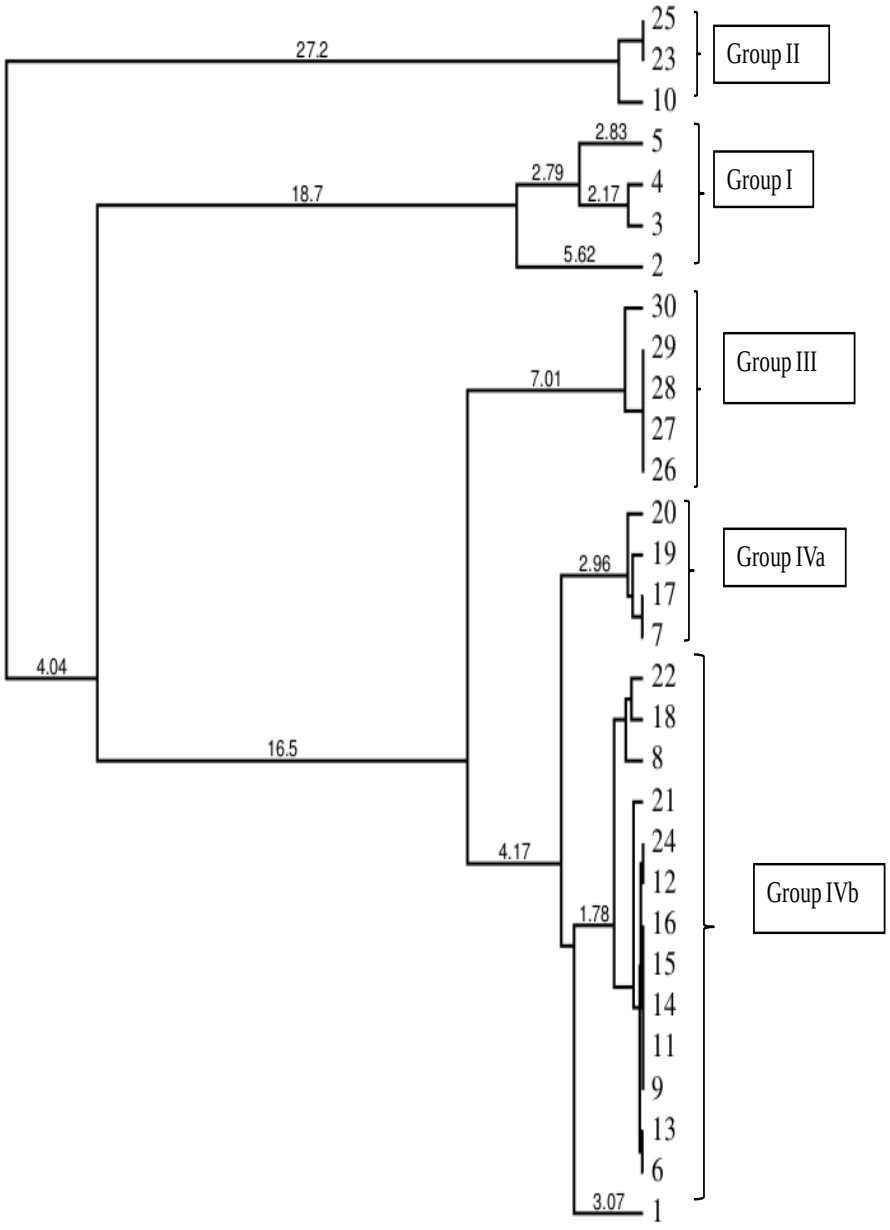


Fig. 23. Phylogenetic tree based on the secondary structure of ITS2

DISCUSSION

Molecular phylogenies of intra-species samples can only be obtained with fast-evolving DNA regions such as the ITS1 and ITS2 regions which are ideal candidates for such a study (Young & Coleman 2003). As compared to the coding regions of the rDNA, these regions show a relatively high rate of divergence. These regions are present in all eukaryotes and can be amplified with universal primers homologous to the highly conserved flanking regions.

The present work on the phylogenetic analysis of the ITS1 and ITS2 regions of thirty weevils of *O. longicollis* from six geographic locations in India has revealed that there is gene flow between these populations and there is no phylogeographical distribution of these six populations. Amongst the six populations, the Assam population shows the highest ITS1 and ITS2 diversity of 0.0061 and 0.01419 respectively. In addition, four of the five individuals from this population form a separate cluster with a longer branch length in phylogenetic analysis. These two points suggest that **(i)** this population is older than the other populations and **(ii)** the populations from the other five locations have been colonized by a subset of the Assam population. Bananas originated in south-east Asia which is considered to be the primary centre of diversification and earliest domestication of this crop (Simmonds 1962). *O. longicollis* being a monophagous pest of bananas would also have originated in this region, coincident with the centre of origin of bananas, before migrating to other regions. The only land route by which this pest could have entered the Indian sub-continent is via north-east India. Phylogenetic analysis shows that atleast 4 individuals from Assam form a separate cluster supporting the initial migration/colonization of this pest in the north-east region of India. This pest being a strong flier could have then spread to other locations within India. The strong migratory behaviour of this pest is responsible for the observed gene flow between the populations. Similar observations were made from the ITS2 data of the malaria vector *Anopheles nuneztovari* collected from different geographic locations (Fritz *et al.* 1994). In case of the bark beetle forest pests, ITS1 and ITS2 were used to differentiate the two species and groupings based on ITS2 showed a phylogenetic clustering rather than a phylogeographic structure (Gallego & Galian 2001). Szalanski *et al*. (1999) in their study of genetic

variation of the southern corn rootworm, (Coleoptera, Chrysomelidae) genus *Diabrotica* using ITS1 and mitochondrial DNA (mtDNA) cytochrome oxidase I and II genes found no genetic distinction in the studied populations suggesting either high levels of dispersal or a recent geographical expansion from a relatively small base. In the present study, the topology of trees generated using the ITS1 and ITS2 data suggests that the grouping is strongly influenced by haplotypes.

The ITS1 and ITS2 regions of thirty weevils were also examined for the folding patterns and conservation in the putative secondary structure. The main aim of this study was

(i) to determine if a conserved core structure as observed in other eukaryotes was present and

(ii) whether the secondary structure could be useful in predicting the phylogenetic relationships.

Unlike the ITS2 region, no consensus structure for the ITS1 region has been reported. The consensus secondary structure of the ITS1 region for *O. longicollis* (Oliver) reveals the presence of ten helices. The number of helices has been found to vary in different organisms (Gottschling & Plotner 2004; Ullrich *et al*. 2009). Four and seven helices have been reported in Chlorobionta and Yeast respectively, while several helices have been proposed for Digenea (Prasad *et al.* 2009). Reports from several species have shown that within the ITS1 region there is a high propensity to form secondary structure (Michot *et al*. 1983; Gonzalez *et al*. 1990; Torres *et al*. 1990; Yeh and Lee 1990). Studies on the ITS1 sequences of lady bird beetles (Coleoptera, Coccinellidae species) showed that no putatively homologous similarities could be identified in the taxa studied, but some homologous regions were found to be conserved throughout the taxa (Von der Schulenberg *et al*. 2001). ITS1 elements were found to be conserved at the 3'end for drosophilid dipterans (Insecta) (Schlotterer *et al*. 1994). Conserved secondary structure motifs primarily found at the 3'and 5'end, contain (as shown in yeasts) elements required for ribosome biogenesis (Henry *et al*. 1994; Van Nues *et al*. 1994; Weaver *et al*. 1997). Functional importance of the 3'end of the ITS1 sequences has also been suggested in the studies on metazoans, (Schlotterer *et al*. 1994; Schulenberg *et al*. 1999).

The consensus secondary structure of the ITS2 region of *O. longicollis* (Oliver) conforms to the Pan-eukaryote ITS2 model suggested for eukaryotes (Coleman 2007). It has been suggested that the conservation of the structural domains (*i.e.* Helices II and III) found in all eukaryotic organisms investigated to date could probably plays a significant functional role in the folding of the secondary structure of ITS2 during rRNA primary transcript processing (Joseph *et al*. 1999; Michot *et al*. 1999). Helices I and IV, which are close to the ITS2 ends are highly variable in length. The ITS2 region has been used to decipher phylogenetic relationships at the species level in sand fly, *Diabrotica* beetles and tomato leafminer (Belen *et al*. 2011; Clark *et al*. 2001; Cifuentes *et al.* 2011).

In the present work, the topology of the phylogenetic trees based on the secondary structure of the ITS1 and ITS2 regions bear striking similarity to each other and also to the trees derived from the primary sequence data. In particular, in case of the ITS1 region, members of each group in the phylogenetic tree show distinguishing structural features. The significant observation is that though the consensus secondary structures of each of these two regions bear no resemblance to each other, yet the topology of the secondary structure trees are congruent confirming the importance of secondary structure of these two regions in phylogenetic analysis. These observations confirm that evolutionary conserved secondary structures are important `markers' in phylogenetic analysis as suggested by Schultz & Wolf (2009) and Keller *et al*. (2010). This work on the use of the secondary structure of the ITS1 and ITS2 regions in assessing the phylogenetic relationships in *O. longicollis* (Oliver) is the first report in family Curculionidae.

REFERENCES

Armbruster GF, Moorsel CV and Gittenberger E (2000) *Journal of Molluscs Studies*, **66(4)**, 570-573.

Armbruster GF (2001) *Canadian Journal of Zoology*, **79(2)**, 334-345.

Baur A, Sanetra M, Chalwatzis N, Buschinger A and Zimmermann FK (1996) *Insectes Sociaux*, **43(1)**, 53-67.

Beebe NW, Ellis JT, Cooper RD and Saul A (1999) *Insect Molecular Biology*, **8(3)**, 381-390.

Bologna MA, Oliverio M, Pitzalis M and Mariottini P (2008) *Molecular Phylogenetics and Evolution*, **48(2)**, 679-693.

Chilton NB, Hoste H, Newton LA, Beveridge I and Gasser RB (1998) *International Journal of Parasitology*, **28(11)**, 1765-1773.

Coleman AW (2007) *Nucleic Acids Research*, **35(10)**, 3322-3329.

Conole JC, Chilton NB, Jarvis T and Gasser RB (2001) *Parasitology*, **122(2)**, 195-206.

Collins FH and Paskewitz SM (1996) *Insect Molecular Biology*, **5(1)**, 1-9.

Cruickshank RH (2002) *Systematic & Applied Acarology*, **7**, 3-14.

Dellaporta SL, Wood J and Hicks JB (1983) *Plant Molecular Biology Reporter*, **1(4)**, 19-21.

Denboh T, Ichimura T, Hendrayanti D and Coleman AW (2003) *Journal of Phycology*, **39(5)**, 960-977.

Depaquit J, Ferte H, Leger N, Killick-Kendrick R, Rioux JA, Killick-Kendrick M and Gobert S (2000) *Insect Molecular Biology*, **9(3)**, 293-300.

Di Muccio T, Marinucci M, Frusteri L, Maroli M, Pesson B and Gramiccia M (2000) *Insect Biochemistry and Molecular Biology*, **30(5)**, 387-393.

Felsenstein J (1981) *Journal of Molecular Evolution*, **17(6)**, 368-376.

Fritz GN, Conn J, Cockburn A and Seawright J (1994) *Molecular Biology and Evolution*, **11(3)**, 406-416.

Furlong JC and Maden BE (1983) *The EMBO journal*, **2(3)**, 443.

Gómez-Zurita J, Juan C and Petitpierre E (2000) *Insect Molecular Biology*, **9(6)**, 591-604.

Gonzalez IL, Chambers C, Gorski JL, Stambolian D, Schmickel RD and Sylvester JE (1990) *Journal of Molecular Biology*, **212(1)**, 27-35.

Good L, Intine RV and Nazar RN (1997) *European Journal of Biochemistry*, **247(1)**, 314-321.

Gottschling M and Plotner J (2004) *Nucleic Acids Research*, **32(1)**, 307-315.

Henry Y, Wood H, Morrissey JP, Petfalski E, Kearsey S and Tollervey D (1994) *The EMBO Journal*, **13(10)**, 2452.

Hamarsheh O, Presber W, Abdeen Z, Sawalha S, Al-Lahem A and Schonian G (2007) *Medical and Veterinary Entomology*, **21(3)**, 270-277.

Hindenach BR and Stafford DW (1984) *Nucleic Acids Research*, **12(3)**, 1737.

Hofacker I (2003) *Nucleic Acids Research,* **31**, 3429-3431.

Justin CG L, Leelamathi M and Nirmaljohson SB (2008) *Agriculture Review*, **29(3)**, 185-192.

Keller A, Schleicher T, Schultz J, Muller T, Dekar T and Wolf M (2009) *Gene,* **430(1)**, 50-57.

Kimura M (1980) *Journal of Molecular Evolution*, **16(2)**, 111-120.

Kumar S, Tamura K and Nei M (1993) Available via http://evolgen.biol.metro-u.ac.jp/MEGA/manual/default.html.

Kumar S, Tamura K, Jakobsen IB and Nei M (2001) *Bioinformatics*, **17(12)**, 1244-1245.

Mai JC and Coleman AW (1997) *Journal of Molecular Evolution*, **44(3)**, 258-271.

Malloch G, Fenton B and Goodrich MA (2001) *Insect Molecular Biology*, **10(3)**, 281-291.

Marrelli MT, Floeter-Winter LM, Malafronte RD S, Tadei WP, Lourenço-de-oliveira R, Flores-Mendoza C and Marinotti O (2005) *Medical and Veterinary Entomology*, **19(2)**, 208-218.

Michot B, Bachellerie JP and Raynal F (1983) *Nucleic Acids Research*, **11(10)**, 3375-3391.

Moin-Vaziri V, Depaquit J, Yaghoobi-Ershadi MR, Oshaghi MA, Derakhsh eh-Peykar P, Ferte H and Nadim A (2007) *Acta Tropica*, **102(1)**, 29-37.

Musters W, Boon K, Van der Se CA, Van Heerikhuizen H and Planta RJ (1990) *The EMBO journal*, **9(12)**, 3989.

Padmanaban B, Sundararaju P and Sathiamoorthy S (2001) *Indian Journal of Entomology*, **63(2)**, 204-204.

Padmanaban B, Sundararaju P, Velayudhan KC and Sathiyamoorthy S (2001) *Infomusa*, **10**, 26-28.

Padmanaban B and Kasamy M (2003) *Indian Journal of Entomology***, 65(3)**, 424-425.

Magaña C, Beroiz B, Hernández-Crespo P, de Oca MM, Carnero A, Ortego F and Castanera P (2007) *Bulletin of Entomological Research*, **97(6)**, 585-590.

Marrelli MT, Floeter-Winter LM, Malafronte RS, Tadei WP, Lourenco De-Oliveira R, Flores-Mendoza C and Marinotti O (2005) *Medical and Veterinary Entomology***, 19**, 208 – 218.

Masetti A, Luchetti A, Sommaggio D, Burgio G and Mantovani B (2006) *European Journal of Entomology*, **103(2)**, 459.

Joseph N, Krauskopf E, Vera MI and Michot B (1999) *Nucleic Acids Research*, **27(23)**, 4533-4540.

Michot B, Joseph N, Mazan S and Bachellerie JP (1999) *Nucleic Acids Research***, 27(11)**, 2271-2282.

Moin-Vaziri V, Depaquit J, Yaghoobi-Ershadi MR, Oshaghi MA, Derakhsh eh-Peykar P, Ferte H and Nadim A (2007) *Acta Tropica*, **102(1)**, 29-37.

Mukabayire O, Boccolini D, Lochouarn L, Fontenille D and Besansky NJ (1999) *Molecular Ecology*, **8(2)**, 289-297.

Ohta T and Dover GA (1983) *Proceedings of the National Academy of Sciences USA***, 80(13)**, 4079-4083.

Peakall R and Smouse PE (2006) *Molecular Ecology Notes*, **6**, 288–295.

Reuter JS and Mathews DH (2010) *BMC Bioinformatics*, **11(1)**, 129.

Roehrdanz R, Heilmann L, Senechal P, Sears S and Evenson P (2010) *Insect Molecular Biology*, **19(4)**, 463-471.

Rozas J, Sanchez-Delbarrio JC, Messeguer X and Rozas R (2003) *Bioinformatics*, **19**, 2496-2497.

Schlötterer C, Hauser MT, von Haeseler A and Tautz D (1994) *Molecular Biology and Evolution*, **11(3)**, 513-522.

Schulenburg JH G, Englisch U and Wagele JW (1999) *Insect Molecular Biology*, **48(1)**, 2-12.

Severini C, Sllvestrlnl F, Manclnl P, Rosa GL and Marlnucci M (1996) *Insect Molecular Biology***, 5(3)**, 181-186.

Shapiro BA (1988) *Cabios*, **4**, 381–393.

Shapiro BA and Zhang K (1990) *Cabios*, **6**, 309–318.

Simmonds NM (1962) The Evolution of the Bananas. Longmans, London, Great Britain.

Subrahmanyam CS, Cassidy B, Busch H and Rothblum LI (1982) *Nucleic Acids Research*, **10(12)**, 3667-3680.

Szalanski AL, Owens CB (2003) *The Florida Entomologist*, **86(3)**, 329-333.

Tamura K and Nei M (1993) *Molecular Biology and Evolution***, 10(3)**, 512-526.

Tamura T, Thibert C, Royer C, Ka T, Eappen A, Kamba M and Couble P (2000) *Nature Biotechnology*, **18(1)**, 81-84.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M and Kumar S (2011) *Molecular Biology and Evolution***, 28(10)**, 2731-2739.

Thompson JD, Desmond GH and Toby JG (1994) *Nucleic Acids Research*, **22.22**, 4673-4680.

Torres RA, Ganal M and Hemleben V (1990) *Journal of Molecular Evolution*, **30(2)**, 170-181.

Trewick SA, Goldberg J and Morgan-Richards M (2005) *Zoologica Scripta*, **34(5)**, 483-491.

Trizzino M, Audisio P, Antonini G, De Biase A and Mancini E (2009) *Molecular Phylogenetics and Evolution*, **51(2),** 215-226.

Ullrich B, Reinhold K, Niehuis O and Misof B (2010) *Journal of Zoological Systematics and Evolutionary Research*, **48(3)**, 219-228.

van Nues RW, Rientjes JM, van der Sande CA, Zerp SF, Sluiter C, Venema J and Raue HA (1994) *Nucleic Acids Research*, **22(6)**, 912-919.

Veldman GM, Klootwijk J, van Heerikhuizen H and Planta RJ (1981) *Nucleic Acids Research*, **9(19)**, 4847-4862.

Vogler AP and DeSalle R (1994) *Molecular Biology and Evolution***, 11(3)**, 393-405.

von der Schulenburg JH G, Hancock JM, Pagnamenta A, Sloggett JJ, Majerus ME and Hurst GD, (2001) *Molecular Biology and Evolution***, 18(4)**, 648-660.

Weaver PL and Sun Cm Chang TH (1997) *Molecular Cell Biology*, **17(3)**, 1354-1365.

Wolf M, Achtziger M, Schultz J, Dekar T and Muller T (2005) *Rna, **11(11)***, 1616-1623.

Xia Q, Guo Y, Zhang Z, Li D, Xuan Z, Li Z and Zhang X (2009) *Science*, **326(5951)**, 433-436.

Yahia H, Ready PD, Hamdani A, Testa JM and Guessous-Idrissi N (2004) *Parasite*, **11(2)**, 189.

Yeh LCC and Lee JC (1990) *Journal of Molecular Biology*, **211(4)**, 699-712.

CHAPTER 5

Genetic diversity analysis of *O. longicollis* (Oliver) populations using the Mitochondrial COI-tRNA^{Leu}-COII region

SUMMARY

The genetic diversity amongst thirty weevils representing six Indian populations of banana pseudostem weevil *i.e*. *Odoiporus longicollis* (Oliver) was estimated by sequence analysis of the partial COI-tRNA^{Leu}- COII region. The sequences exhibited AT bias typical of insect mitochondrial DNA which was highest in the first codon position of COI and in the third codon position of COII. Majority of the amino-acids in the COI and COII proteins reflected a bias against codons ending in C or G. There was no phylogeographic distribution of the populations. The Fu and Li's D and F tests were non-significant for this mitochondrial region. No *Wolbachia* infection was detected in any of the populations. The genetic differentiation amongst the populations was highly significant ($p<0.001$; $\chi^2 = 123.333$; df=75), suggesting restricted gene flow between the populations. This result did not correlate with that obtained with nuclear markers *i.e*., ITS1 and ITS2 suggesting a male-biased gene flow between the populations.

INTRODUCTION

Mitochondrial genes has been used in the population studies of several insects because they have several important characteristics i.e. they are haploid, maternally inherited, lack introns, evolve more rapidly than nuclear coding genes, lack recombination and are assumed to vary in a neutral manner. The rate of evolution of the mitochondrial protein coding regions is much faster than the 12S and 16S rDNA genes and hence these regions serve as useful markers for deciphering evolutionary history at the levels of family, genera and species (Wan *et al.* 2004). Mitochondrial genes have been used to decipher phylogeny and phylogeography in several insects (Orsini *et al.* 2007) including *Heliconius* butterflies (Brown 1994), *Halys fabriciusm* (Memon *et al.* 2006), *Diabrotica* (Szalanski *et al.* 2000), *Adelges cooleyi* (Ahern *et al.* 2009), *Aphidus ervi* (Hufbauer *et al.* 2004) and *Apis cerana indica* F (Baskaran 2011). In few studies, mitochondrial genes have not proved useful for the estimation of phylogeography at the intraspecies level. For *e.g.* intraspecific phylogeography of *Apis cerana* did not correlate with geographic distribution when COI/COII region was used for comparison of different geographic populations (Hepburn *et al.* 2001). Horn *et al.* (2006) obtained no clear phylogeographic pattern within geographic populations of *Tomincus destruens* Woll.

Incongruencies observed in the phylogenies based on mitochondrial and nuclear genes have made analyses based on both sets of genes important in insect molecular systematics. In this chapter, the genetic diversity of six Indian populations of *Odoiporus longicollis* (Oliver) represented by thirty individuals has been assessed using the mitochondrial COI-tRNA^{Leu}-COII fragment. The results suggest that this mitochondrial sequence is not useful in assessing the genetic diversity between the populations of *O. longicollis* (Oliver). A comparison of the present analysis based on the mitochondrial DNA fragment with the results obtained using nuclear markers *i.e.* rDNA markers ITS1 and ITS2, suggests a male biased gene flow amongst these populations.

MATERIALS AND METHODS

Sample collection

A total of thirty adult beetles of *Odoiporus longicollis* (Oliver) *i.e*. five adults from each location in India were collected from infested banana stems from banana fields in Assam (Kamrup), Bihar (Vaishali), Kerala (Wayanad), Maharashtra (Jalgaon, Narayangaon) and Tamil Nadu (Trichy) (As described in Chapter 3). In each location, the five individuals were collected within an area of 500 m.

DNA extraction

Total genomic DNA was extracted from individual beetles using a modified Dellaporta *et al.* (1983) protocol as described in `Materials and Methods' of Chapter 3.

Amplification of mt DNA COI-tRNA^{Leu}-COII fragments

The Mt DNA COI-tRNA^{Leu}-COII fragment was amplified from individual weevils using the primer pair COI-RLR 5'TTGATTTTTTGGTCATCCAGAAGTTTATA3 and COII-Croz 5'CCACAAATTTCTGAACATTGACC3' (Crozier *et al.* 1989). PCR was performed in a total reaction volume of 25 µl consisting of 25 ng DNA, 0.1 mM dNTPs, 100 pmol primer and 0.5 U Taq DNA Polymerase. PCR conditions were as follows: pre-denaturation at 94^{0}C for 5 min followed by thirty cycles of denaturation at 94^{0}C for 45 sec, annealing at 45^{0}C for 45 sec, extension at 72^{0}C for 2 min; followed by a final extension at 72^{0}C for 20 min. The PCR product was cloned into PCR4 TOPO (Invitrogen) vector and transformed into *E. coli* Top10 host cells (Invitrogen), as per the manufacturer's instructions. The sequences have been deposited in GenBank (**Table 1**).

Table 1. **List of *O. longicollis* (Oliver) individuals with Genebank Accession No. of COXI, COXII and tRNA^{Leu}**

Location (Individual No.)	Genbank Accession COXI	Genbank Accession COXII
Assam 1	KF856255	KJ446927
Assam 2	KJ446898	KJ446928
Assam 3	KJ446899	KJ446929
Assam 4	KJ446900	KJ446930
Assam 5	KJ446901	KJ446931
Jalgaon 1	KJ446902	KJ446932
Jalgaon 2	KJ446903	KJ446933
Jalgaon 3	KJ446904	KJ446934
Jalgaon 4	KJ446905	KJ446935
Jalgaon 5	KJ446906	KJ446936
Narayangaon 1	KJ446907	KJ446937
Narayangaon 2	KJ446908	KJ446938
Narayangaon 3	KJ446909	KJ446939
Narayangaon 4	KJ446910	KJ446940
Narayangaon 5	KJ446911	KJ446941
Trichy 1	KJ446912	KJ446942
Trichy 2	KJ446913	KJ446943
Trichy 3	KJ446914	KJ446944
Trichy 4	KJ446915	KJ446945
Trichy 5	KJ446916	KJ446946
Vaishali 1	KJ446917	KJ446947
Vaishali 2	KJ446918	KJ446948
Vaishali 3	KJ446919	KJ446949
Vaishali 4	KJ446920	KJ446950
Vaishali 5	KJ446921	KJ446951

Wayanad 1	KJ446922	KJ446952
Wayanad 2	KJ446923	KJ446953
Wayanad 3	KJ446924	KJ446954
Wayanad 4	KJ446925	KJ446955
Wayanad 5	KJ446926	KJ446956
Assam 1	tRNA Leucine	KJ446957

Detection of Wolbachia infection

Presence of *Wolbachia* infection was detected by PCR using *Wolbachia* 16S rDNA specific primers (WspecF and WspecR) and primers specific to *Wolbachia* surface protein gene (wspA-81F and wspA-691R) (**Table 2**). PCR was performed in a total reaction volume of 25 µl consisting of 25 ng DNA, 0.1 mM dNTPs, 100 pmol primer and 0.5 U Taq DNA Polymerase. PCR conditions were as follows: pre-denaturation at 94^0C for 2 min followed by thirty cycles of denaturation at 94^0C for 30 sec, annealing at 50.7^0C for 1 min, extension at 72^0C for 1 min; followed by a final extension at 72^0C for 4 min.

Table 2. Primers used to detect the presence of *Wolbachia* infection

Name of the primer	**Sequence of the primer (5'-3')**	**Tm (^{O}C)**	**Expected size of the amplicon (bp)**	**Reference**
WspecF	CATACCTATTCGAAAGGGATAG	50.7	438	Werren & Windsor (2000)
WspecR	AGCTTCGAGTGAAACCAATTC	50.7	438	Werren & Windsor (2000)
wspA-81F	TGGTCCAATAACTGATGAAGAAC	50.7	632	Zhou *et al.* (1998)
wspA-691R	AAAAATTAACGCTACTCCA	50.7	632	Zhou *et al.* (1998)

Sequence analysis

DNA sequence chromatograms were read and discrepancies between forward and reverse sequences were resolved using the Chromas software version 2.01 (http:// www.technelysium.com.au/chromas.html). CLUSTALW was used to generate the alignments (Thompson *et al.* 1994). The sequences were imported into MEGA v 5 for analysis of nucleotide diversity, genetic distances, substitution pattern and phylogenetic analysis (Tamura *et al.* 2007). Statistical support for the inferred nodes was obtained by bootstrapping in MEGA version 5 (Tamura *et al.* 2007). All bootstrap values were based on the performance of 1000 replicates. DnaSP v 5.10 was used for haplotype and indel polymorphism analysis. DAMBE v 5.3.27 was used to calculate Ts/Tv and divergence ratio. Correlation between genetic distance and geographical distance was done by Mantel test using the GenAlEx v.6.501 software. tRNAscan-SE Search Server (http://lowelab.ucsc.edu/tRNAscan-SE/) was used to scan the whole fragment to locate the tRNA leucine region. Vienna RNA secondary structure web server was used to get the tRNA leucine structure file. ORF finder tool (http://www.ncbi.nlm.nih.gov/projects/gorf/) was used to find the open reading frames of the COI and COII partial coding regions and these were translated to functional peptide sequences using Nucleotide Sequence Translation tool available online (http://www.ebi.ac.uk/Tools/st/).

RESULTS

Sequence Analysis of the COI-tRNA^{Leu}-COII region

The COI-tRNA^{Leu}-COII stretch included 835 bp (positions 1 to 835) of the partial coding sequence of cytochrome oxidase subunit I gene, 65 bp (position 836 to 901) of the complete coding sequence of tRNA^{Leu} and 602 bp (position 902 to 1502) of the partial coding sequence of cytochrome oxidase subunit II. The stretch of 1502 bp corresponded to positions 2068 to 3569 of *Spenophorus* sp. BYU-CO246 (Accession number, GU176342; Song *et al.* 2010). The aligned mitochondrial nucleotide sequences of the thirty O*doiporus longicollis* (Oliver) individuals yielded 1,502 characters of which 1463

(97.40%) were conserved and 23 (1.5%) were parsimony informative **(Table 3 and Fig. 1).**

Table 3. Conserved, variable and parsimony informative sites within the COI-tRNA^{Leu}- COII region

Region	Conseved	Variable	Parsimony
COI-tRNA^{leu}-COII	1463	39	23
COI	812	23	12
COII	590	12	10

```
TRICHY1          TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
TRICHY3          TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
TRICHY2          TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
TRICHY4          TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
TRICHY5          TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
NARAYANGAON2     TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
VAISHALI1        TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
JALGAON1         TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
JALGAON5         TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
JALGAON4         TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
VAISHALI4        TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
NARAYANGAON3     TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
VAISHALI3        TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
VAISHALI2        TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
JALGAON2         TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
JALGAON3         TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
NARAYANGAON1     TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
NARAYANGAON5     TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
VAISHALI5        TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
NARAYANGAON4     TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
Assam1           TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
WAYANAD1         TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
WAYANAD2         TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
WAYANAD3         TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
WAYANAD4         TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
WAYANAD5         TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
ASSAM2           TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
ASSAM3           TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
ASSAM4           TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
ASSAM5           TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
                 ************************************************************

TRICHY1          CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
TRICHY3          CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
TRICHY2          CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
TRICHY4          CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
TRICHY5          CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
NARAYANGAON2     CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
VAISHALI1        CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
JALGAON1         CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
JALGAON5         CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
JALGAON4         CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
VAISHALI4        CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
NARAYANGAON3     CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
```

```
VAISHALI3         CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
VAISHALI2         CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
JALGAON2          CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
JALGAON3          CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
NARAYANGAON1      CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
NARAYANGAON5      CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
VAISHALI5         CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
NARAYANGAON4      CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
Assam1            CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
WAYANAD1          CATATTATTAGTCAAGAAAGAGGAAAGAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
WAYANAD2          CATATTATTAGTCAAGAAAGAGGAAAGAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
WAYANAD3          CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
WAYANAD4          CATATTATTAGTCAAGAAAGAGGAAAAAAAGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
WAYANAD5          CATATTATTAGTCAAGAAAGAGGAAAAAAAGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
ASSAM2            CATATTATTAGTCAAGAAAGAGGAAAAAAAGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
ASSAM3            CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
ASSAM4            CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
ASSAM5            CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
                  ************************** ** ******************************

TRICHY1           GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
TRICHY3           GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
TRICHY2           GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
TRICHY4           GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
TRICHY5           GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
NARAYANGAON2      GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
VAISHALI1         GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
JALGAON1          GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
JALGAON5          GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
JALGAON4          GCCATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
VAISHALI4         GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
NARAYANGAON3      GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
VAISHALI3         GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
VAISHALI2         GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
JALGAON2          GCTATAATAGCAATTGGGCTTTAGGATTTTGTAGTCTGAGCTCACCATATATTCACAGTA 180
JALGAON3          GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTCACAGTA 180
NARAYANGAON1      GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
NARAYANGAON5      GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
VAISHALI5         GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
NARAYANGAON4      GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
Assam1            GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
WAYANAD1          GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
WAYANAD2          GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
WAYANAD3          GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
WAYANAD4          GCTATAATAGCAATTGGGCTTTTAGGATTCGTAGTCTGAGCTCACCATATATTTACAGTA 180
WAYANAD5          GCTATAATAGCAATTGGGCTTTTAGGATTCGTAGTCTGAGCTCACCATATATTTACAGTA 180
ASSAM2            GCTATAATAGCAATTGGGCTTTTAGGATTCGTAGTCTGAGCTCACCATATATTTACAGTA 180
ASSAM3            GCTATAATAGCAATTGGGCTTTTAGGATTCGTAGTCTGAGCTCACCATATATTTACAGTA 180
ASSAM4            GCTATAATAGCAATTGGGCTTTTAGGATTCGTAGTCTGAGCTCACCATATATTTACAGTA 180
ASSAM5            GCTATAATAGCAATTGGGCTTTTAGGATTCGTAGTCTGAGCTCACCATATATTTACAGTA 180
                  ** ******************  *  ** ********************** ******

TRICHY1           GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
TRICHY3           GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
TRICHY2           GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
TRICHY4           GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
TRICHY5           GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
NARAYANGAON2      GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
VAISHALI1         GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
JALGAON1          GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
JALGAON5          GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
JALGAON4          GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
VAISHALI4         GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
NARAYANGAON3      GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
VAISHALI3         GGGATAGATGTAGATACTCGGGCATATTCTACTTCAGCTACAATAATTATTGCTGTTCCT 240
VAISHALI2         GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
JALGAON2          GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
JALGAON3          GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
NARAYANGAON1      GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
NARAYANGAON5      GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
VAISHALI5         GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
```

```
NARAYANGAON4      GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
Assam1            GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
WAYANAD1          GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTACTGCTGTTCCT 240
WAYANAD2          GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTACTGCTGTTCCT 240
WAYANAD3          GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
WAYANAD4          GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
WAYANAD5          GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
ASSAM2            GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
ASSAM3            GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
ASSAM4            GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
ASSAM5            GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
                  *************************** ******************* **********

TRICHY1           ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
TRICHY3           ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
TRICHY2           ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCC 300
TRICHY4           ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCC 300
TRICHY5           ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
NARAYANGAON2      ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
VAISHALI1         ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
JALGAON1          ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
JALGAON5          ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
JALGAON4          ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
VAISHALI4         ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
NARAYANGAON3      ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
VAISHALI3         ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
VAISHALI2         ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
JALGAON2          ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
JALGAON3          ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
NARAYANGAON1      ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
NARAYANGAON5      ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
VAISHALI5         ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
NARAYANGAON4      ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
Assam1            ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
WAYANAD1          ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
WAYANAD2          ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
WAYANAD3          ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
WAYANAD4          ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
WAYANAD5          ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
ASSAM2            ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
ASSAM3            ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
ASSAM4            ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
ASSAM5            ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
                  ************************************************************

TRICHY1           CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
TRICHY3           CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
TRICHY2           CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
TRICHY4           CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
TRICHY5           CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
NARAYANGAON2      CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
VAISHALI1         CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
JALGAON1          CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
JALGAON5          CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
JALGAON4          CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
VAISHALI4         CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
NARAYANGAON3      CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
VAISHALI3         CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
VAISHALI2         CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
JALGAON2          CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
JALGAON3          CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
NARAYANGAON1      CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
NARAYANGAON5      CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
VAISHALI5         CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
NARAYANGAON4      CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
Assam1            CCTACTTCCTTATGAAGCTTAGGCTTTATCTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
WAYANAD1          CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
WAYANAD2          CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
WAYANAD3          CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCCTTTTACTATAGGAGGTCTAACTGGA 360
WAYANAD4          CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
WAYANAD5          CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
```

```
ASSAM2          CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
ASSAM3          CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
ASSAM4          CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
ASSAM5          CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
                **************************** **** ************************

TRICHY1         GTAGTATTAGCAAATGCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
TRICHY3         GTAGTATTAGCAAATGCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
TRICHY2         GTAGTATTAGCAAATGCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
TRICHY4         GTAGTATTAGCAAATGCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
TRICHY5         GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
NARAYANGAON2    GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
VAISHALI1       GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
JALGAON1        GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
JALGAON5        GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
JALGAON4        GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
VAISHALI4       GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
NARAYANGAON3    GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
VAISHALI3       GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
VAISHALI2       GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
JALGAON2        GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
JALGAON3        GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
NARAYANGAON1    GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
NARAYANGAON5    GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
VAISHALI5       GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
NARAYANGAON4    GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
Assam1          GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
WAYANAD1        GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
WAYANAD2        GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
WAYANAD3        GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
WAYANAD4        GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
WAYANAD5        GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
ASSAM2          GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
ASSAM3          GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
ASSAM4          GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
ASSAM5          GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
                *************** ****************************************

TRICHY1         CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
TRICHY3         CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
TRICHY2         CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
TRICHY4         CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
TRICHY5         CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
NARAYANGAON2    CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
VAISHALI1       CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
JALGAON1        CATTTCCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
JALGAON5        CATTTCCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
JALGAON4        CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
VAISHALI4       CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
NARAYANGAON3    CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
VAISHALI3       CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
VAISHALI2       CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
JALGAON2        CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
JALGAON3        CATTTTCATTATGTACTGTCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
NARAYANGAON1    CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
NARAYANGAON5    CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
VAISHALI5       CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
NARAYANGAON4    CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
Assam1          CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
WAYANAD1        CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
WAYANAD2        CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
WAYANAD3        CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
WAYANAD4        CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
WAYANAD5        CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
ASSAM2          CATTTTCATTATGTACTATCAATAGGAGCTGTACTTGCTATTATTGCAGGATTTATTCAA 480
ASSAM3          CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
ASSAM4          CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
ASSAM5          CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
                ***** *********** *************** *************************

TRICHY1         TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
```

```
TRICHY3          TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
TRICHY2          TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
TRICHY4          TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
TRICHY5          TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
NARAYANGAON2     TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
VAISHALI1        TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
JALGAON1         TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
JALGAON5         TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
JALGAON4         TGATTTCCTCTTCTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
VAISHALI4        TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
NARAYANGAON3     TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
VAISHALI3        TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
VAISHALI2        TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
JALGAON2         TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
JALGAON3         TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
NARAYANGAON1     TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
NARAYANGAON5     TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
VAISHALI5        TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
NARAYANGAON4     TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
Assam1           TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
WAYANAD1         TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
WAYANAD2         TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
WAYANAD3         TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
WAYANAD4         TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
WAYANAD5         TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
ASSAM2           TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
ASSAM3           TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
ASSAM4           TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
ASSAM5           TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
                 ************ ********************************************** 

TRICHY1          TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
TRICHY3          TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
TRICHY2          TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
TRICHY4          TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
TRICHY5          TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
NARAYANGAON2     TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
VAISHALI1        TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
JALGAON1         TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
JALGAON5         TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
JALGAON4         TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
VAISHALI4        TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
NARAYANGAON3     TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
VAISHALI3        TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
VAISHALI2        TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
JALGAON2         TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
JALGAON3         TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCCCAACATTTTCTAGGATTAAGTGGT 600
NARAYANGAON1     TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
NARAYANGAON5     TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
VAISHALI5        TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
NARAYANGAON4     TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
Assam1           TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
WAYANAD1         TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
WAYANAD2         TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
WAYANAD3         TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
WAYANAD4         TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
WAYANAD5         TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
ASSAM2           TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
ASSAM3           TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
ASSAM4           TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
ASSAM5           TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
                 ********************************** ********************** 

TRICHY1          ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
TRICHY3          ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
TRICHY2          ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
TRICHY4          ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
TRICHY5          ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
NARAYANGAON2     ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
VAISHALI1        ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
JALGAON1         ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCT 660
```

```
JALGAON5          ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCT 660
JALGAON4          ATACCACGACGTTATTCTGACTATCCGGATGCATATTATATATGAAACTCAATTTCTTCA 660
VAISHALI4         ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
NARAYANGAON3      ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
VAISHALI3         ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
VAISHALI2         ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
JALGAON2          ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
JALGAON3          ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
NARAYANGAON1      ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
NARAYANGAON5      ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
VAISHALI5         ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
NARAYANGAON4      ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
Assam1            ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
WAYANAD1          ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
WAYANAD2          ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
WAYANAD3          ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
WAYANAD4          ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
WAYANAD5          ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
ASSAM2            ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
ASSAM3            ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
ASSAM4            ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
ASSAM5            ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
                  ************************** ******************************

TRICHY1           ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTTATTATCTGAGAAGCT 720
TRICHY3           ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTTATTATCTGAGAAGCT 720
TRICHY2           ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTTATTATCTGAGAAGCT 720
TRICHY4           ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTTATTATCTGAGAAGCT 720
TRICHY5           ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTTATTATCTGAGAAGCT 720
NARAYANGAON2      ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTTATTATCTGAGAAGCT 720
VAISHALI1         ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTTATTATCTGAGAAGCT 720
JALGAON1          ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTTACTATCTGAGAAGCT 720
JALGAON5          ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTTACTATCTGAGAAGCT 720
JALGAON4          ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTTACTATCTGAGAAGCT 720
VAISHALI4         ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTTATTATCTGAGAAGCT 720
NARAYANGAON3      ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTTATTATCTGAGAAGCT 720
VAISHALI3         ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTTATTATCTGAGAAGCT 720
VAISHALI2         ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTTATTATCTGAGAAGCT 720
JALGAON2          GTTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTTATTATCTGAGAAGCT 720
JALGAON3          ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTTATTATCTGAGAAGCT 720
NARAYANGAON1      ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTTATTATCTGAGAAGCT 720
NARAYANGAON5      ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTTATTATCTGAGAAGCT 720
VAISHALI5         ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTTATTATCTGAGAAGCT 720
NARAYANGAON4      ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTTATTATCTGAGAAGCT 720
Assam1            ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTTATTATTTGAGAAGCT 720
WAYANAD1          ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTTATTATCTGAGAAGCT 720
WAYANAD2          ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTTATTATCTGAGAAGCT 720
WAYANAD3          ATTGGTAGACTAATCTCTTTAATTAGTGTTTTATACCTAATTTTTTATTATCTGAGAAGCT 720
WAYANAD4          ATTGGTAGACTAATCTCTTTAATTAGTGTTTTATACCTAATTTTTTATTATCTGAGAAGCT 720
WAYANAD5          ATTGGTAGACTAATCTCTTTAATTAGTGTTTTATACCTAATTTTTTATTATCTGAGAAGCT 720
ASSAM2            ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTTATTATTTGAGAAGCT 720
ASSAM3            ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTTATTATTTGAGAAGCT 720
ASSAM4            ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTTATTATTTGAGAAGCT 720
ASSAM5            ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTTATTATTTGAGAAGCT 720
                   ************************** ****************** *** *********

TRICHY1           TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
TRICHY3           TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
TRICHY2           TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
TRICHY4           TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
TRICHY5           TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
NARAYANGAON2      TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
VAISHALI1         TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
JALGAON1          TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
JALGAON5          TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
JALGAON4          TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
VAISHALI4         TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
NARAYANGAON3      TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
VAISHALI3         TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
VAISHALI2         TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
JALGAON2          TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
```

```
JALGAON3          TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
NARAYANGAON1      TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
NARAYANGAON5      TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
VAISHALI5         TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
NARAYANGAON4      TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
Assam1            TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTACCATCTTCCATTGAATGACTC 780
WAYANAD1          TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
WAYANAD2          TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
WAYANAD3          TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
WAYANAD4          TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTACCATCTTCCATTGAATGACTC 780
WAYANAD5          TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTACCATCTTCCATTGAATGACTC 780
ASSAM2            TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTACCATCTTCCATTGAATGACTC 780
ASSAM3            TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTACCATCTTCCATTGAATGACTC 780
ASSAM4            TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTACCATCTTCCATTGAATGACTC 780
ASSAM5            TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTACCATCTTCCATTGAATGACTC 780
                  ************************************ ********************

                                        COI
TRICHY1           CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATTTCTAA 840
TRICHY3           CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATTTCTAA 840
TRICHY2           CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATTTCTAA 840
TRICHY4           CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATTTCTAA 840
TRICHY5           CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATTTCTAA 840
NARAYANGAON2      CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATTTCTAA 840
VAISHALI1         CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATTTCTAA 840
JALGAON1          CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATTTCTAA 840
JALGAON5          CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATTTCTAA 840
JALGAON4          CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATTTCTAA 840
VAISHALI4         CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATTTCTAA 840
NARAYANGAON3      CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATTTCTAA 840
VAISHALI3         CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATTTCTAA 840
VAISHALI2         CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATTTCTAA 840
JALGAON2          CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATTTCTAA 840
JALGAON3          CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATTTCTAA 840
NARAYANGAON1      CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATTTCTAA 840
NARAYANGAON5      CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATTTCTAA 840
VAISHALI5         CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATTTCTAA 840
NARAYANGAON4      CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATTTCTAA 840
Assam1            CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATTTCTAA 840
WAYANAD1          CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATTTCTAA 840
WAYANAD2          CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATTTCTAA 840
WAYANAD3          CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATTTCTAA 840
WAYANAD4          CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATTTCTAA 840
WAYANAD5          CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATTTCTAA 840
ASSAM2            CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATTTCTAA 840
ASSAM3            CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATTTCTAA 840
ASSAM4            CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATTTCTAA 840
ASSAM5            CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATTTCTAA 840
                  ************************************************************

                  <-------------------- tRNA leucine -------------------->
TRICHY1           AATGGCAGAATAGTGCATTGGACTTAAACCCCAAAAATAAAGTTTAAACTTTTTTTAGAA 900
TRICHY3           AATGGCAGAATAGTGCATTGGACTTAAACCCCAAAAATAAAGTTTAAACTTTTTTTAGAA 900
TRICHY2           AATGGCAGAATAGTGCATTGGACTTAAACCCCAAAAATAAAGTTTAAACTTTTTTTAGAA 900
TRICHY4           AATGGCAGAATAGTGCATTGGACTTAAACCCCAAAAATAAAGTTTAAACTTTTTTTAGAA 900
TRICHY5           AATGGCAGAATAGTGCATTGGACTTAAACCCCAAAAATAAAGTTTAAACTTTTTTTAGAA 900
NARAYANGAON2      AATGGCAGAATAGTGCATTGGACTTAAACCCCAAAAATAAAGTTTAAACTTTTTTTAGAA 900
VAISHALI1         AATGGCAGAATAGTGCATTGGACTTAAACCCCAAAAATAAAGTTTAAACTTTTTTTAGAA 900
JALGAON1          AATGGCAGAATAGTGCATTGGACTTAAACCCCAAAAATAAAGTTTAAACTTTTTTTAGAA 900
JALGAON5          AATGGCAGAATAGTGCATTGGACTTAAACCCCAAAAATAAAGTTTAAACTTTTTTTAGAA 900
JALGAON4          AATGGCAGAATAGTGCATTGGACTTAAACCCCAAAAATAAAGTTTAAACTTTTTTTAGAA 900
VAISHALI4         AATGGCAGAATAGTGCATTGGACTTAAACCCCAAAAATAAAGTTTAAACTTTTTTTAGAA 900
NARAYANGAON3      AATGGCAGAATAGTGCATTGGACTTAAACCCCAAAAATAAAGTTTAAACTTTTTTTAGAA 900
VAISHALI3         AATGGCAGAATAGTGCATTGGACTTAAACCCCAAAAATAAAGTTTAAACTTTTTTTAGAA 900
VAISHALI2         AATGGCAGAATAGTGCATTGGACTTAAACCCCAAAAATAAAGTTTAAACTTTTTTTAGAA 900
JALGAON2          AATGGCAGAATAGTGCATTGGACTTAAACCCCAAAAATAAAGTTTAAACTTTTTTTAGAA 900
JALGAON3          AATGGCAGAATAGTGCATTGGACTTAAACCCCAAAAATAAAGTTTAAACTTTTTTTAGAA 900
NARAYANGAON1      AATGGCAGAATAGTGCATTGGACTTAAACCCCAAAAATAAAGTTTAAACTTTTTTTAGAA 900
NARAYANGAON5      AATGGCAGAATAGTGCATTGGACTTAAACCCCAAAAATAAAGTTTAAACTTTTTTTAGAA 900
VAISHALI5         AATGGCAGAATAGTGCATTGGACTTAAACCCCAAAAATAAAGTTTAAACTTTTTTTAGAA 900
NARAYANGAON4      AATGGCAGAATAGTGCATTGGACTTAAACCCCAAAAATAAAGTTTAAACTTTTTTTAGAA 900
```

```
Assam1          AATGGCAGAATAGTGCATTGGACTTAAACCCCAAAAATAAAGTTTAAACTTTTTTTAGAA 900
WAYANAD1        AATGGCAGAATAGTGCATTGGACTTAAACCCCAAAAATAAAGTTTAAACTTTTTTTAGAA 900
WAYANAD2        AATGGCAGAATAGTGCATTGGACTTAAACCCCAAAAATAAAGTTTAAACTTTTTTTAGAA 900
WAYANAD3        AATGGCAGAATAGTGCATTGGACTTAAACCCCAAAAATAAAGTTTAAACTTTTTTTAGAA 900
WAYANAD4        AATGGCAGAATAGTGCATTGGACTTAAACCCCAAAAATAAAGTTTAAACTTTTTTTAGAA 900
WAYANAD5        AATGGCAGAATAGTGCATTGGACTTAAACCCCAAAAATAAAGTTTAAACTTTTTTTAGAA 900
ASSAM2          AATGGCAGAATAGTGCATTGGACTTAAACCCCAAAAATAAAGTTTAAACTTTTTTTAGAA 900
ASSAM3          AATGGCAGAATAGTGCATTGGACTTAAACCCCAAAAATAAAGTTTAAACTTTTTTTAGAA 900
ASSAM4          AATGGCAGAATAGTGCATTGGACTTAAACCCCAAAAATAAAGTTTAAACTTTTTTTAGAA 900
ASSAM5          AATGGCAGAATAGTGCATTGGACTTAAACCCCAAAAATAAAGTTTAAACTTTTTTTAGAA 900
                ************************************************************

                          COII
TRICHY1         ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 960
TRICHY3         ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 960
TRICHY2         ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCTTCTTATAGAACAACTT 960
TRICHY4         ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCTTCTTATAGAACAACTT 960
TRICHY5         ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCTTCTTATAGAACAACTT 960
NARAYANGAON2    ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 960
VAISHALI1       ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 960
JALGAON1        ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 960
JALGAON5        ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 960
JALGAON4        ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 960
VAISHALI4       ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 960
NARAYANGAON3    ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 960
VAISHALI3       ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 960
VAISHALI2       ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 960
JALGAON2        ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 960
JALGAON3        ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 960
NARAYANGAON1    ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 960
NARAYANGAON5    ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 960
VAISHALI5       ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 960
NARAYANGAON4    ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 960
Assam1          ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 960
WAYANAD1        ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 960
WAYANAD2        ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 960
WAYANAD3        ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 960
WAYANAD4        ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 960
WAYANAD5        ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 960
ASSAM2          ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 960
ASSAM3          ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 960
ASSAM4          ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 960
ASSAM5          ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 960
                ****************************************** ****************

TRICHY1         CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 1020
TRICHY3         CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 1020
TRICHY2         CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 1020
TRICHY4         CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 1020
TRICHY5         CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 1020
NARAYANGAON2    CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 1020
VAISHALI1       CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 1020
JALGAON1        CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 1020
JALGAON5        CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 1020
JALGAON4        CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 1020
VAISHALI4       CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 1020
NARAYANGAON3    CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 1020
VAISHALI3       CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 1020
VAISHALI2       CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 1020
JALGAON2        CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 1020
JALGAON3        CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 1020
NARAYANGAON1    CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 1020
NARAYANGAON5    CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 1020
VAISHALI5       CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 1020
NARAYANGAON4    CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 1020
Assam1          CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 1020
WAYANAD1        CTTTTCTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 1020
WAYANAD2        CTTTTCTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 1020
WAYANAD3        CTTTTCTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 1020
WAYANAD4        CTTTTCTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 1020
```

```
WAYANAD5          CTTTTCTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 1020
ASSAM2            CTTTTCTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 1020
ASSAM3            CTTTTCTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 1020
ASSAM4            CTTTTCTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 1020
ASSAM5            CTTTTCTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 1020
                  ***** **************************************************

TRICHY1           ATATTACTAAGAATATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 1080
TRICHY3           ATATTACTAAGAATATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 1080
TRICHY2           ATATTACTAAGAATATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 1080
TRICHY4           ATATTACTAAGAATATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 1080
TRICHY5           ATATTACTAAGAATATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 1080
NARAYANGAON2      ATATTACTAAGAATATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 1080
VAISHALI1         ATATTACTAAGAATATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 1080
JALGAON1          ATATTACTAAGAATATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 1080
JALGAON5          ATATTACTAAGAATATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 1080
JALGAON4          ATATTACTAAGAGTATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 1080
VAISHALI4         ATATTACTAAGAATATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 1080
NARAYANGAON3      ATATTACTAAGAATATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 1080
VAISHALI3         ATATTACTAAGAATATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 1080
VAISHALI2         ATATTACTAAGAATATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 1080
JALGAON2          ATATTACTAAGAATATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 1080
JALGAON3          ATATTACTAAGAATATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 1080
NARAYANGAON1      ATATTACTAAGAATATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 1080
NARAYANGAON5      ATATTACTAAGAATATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 1080
VAISHALI5         ATATTACTAAGAATATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 1080
NARAYANGAON4      ATATTACTAAGAATATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 1080
Assam1            ATATTACTAAGAATATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 1080
WAYANAD1          ATATTACTAAGAATATTATTCAATAAATTCACTCACCGTTTTCTTCTAGAAGGACAAACA 1080
WAYANAD2          ATATTACTAAGAATATTATTCAATAAATTCACTCACCGTTTTCTTCTAGAAGGACAAACA 1080
WAYANAD3          ATATTACTAAGAATATTATTCAATAAATTCACTCACCGTTTTCTTCTAGAAGGACAAATA 1080
WAYANAD4          ATATTACTAAGAATATTATTCAATAAATTCACTCACCGTTTTCTTCTAGAAGGACAAATA 1080
WAYANAD5          ATATTACTAAGAATATTATTCAATAAATTCACTCACCGTTTTCTTCTAGAAGGACAAATA 1080
ASSAM2            ATATTACTAAGAATATTATTCAATAAATTCACTCACCGTTTTCTTCTAGAAGGACAAATA 1080
ASSAM3            ATATTACTAAGAATATTATTCAATAAATTCACTCACCGTTTTCTTCTAGAAGGACAAATA 1080
ASSAM4            ATATTACTAAGAATATTATTCAATAAATTCACTCACCGTTTTCTTCTAGAAGGACAAATA 1080
ASSAM5            ATATTACTAAGAATATTATTCAATAAATTCACTCACCGTTTTCTTCTAGAAGGACAAATA 1080
                  ************ ********************** ********************** *

TRICHY1           ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 1140
TRICHY3           ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 1140
TRICHY2           ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 1140
TRICHY4           ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 1140
TRICHY5           ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 1140
NARAYANGAON2      ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 1140
VAISHALI1         ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 1140
JALGAON1          ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 1140
JALGAON5          ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 1140
JALGAON4          ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 1140
VAISHALI4         ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 1140
NARAYANGAON3      ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 1140
VAISHALI3         ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 1140
VAISHALI2         ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 1140
JALGAON2          ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 1140
JALGAON3          ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 1140
NARAYANGAON1      ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 1140
NARAYANGAON5      ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 1140
VAISHALI5         ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 1140
NARAYANGAON4      ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 1140
Assam1            ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 1140
WAYANAD1          ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTTCCATCA 1140
WAYANAD2          ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTTCCATCA 1140
WAYANAD3          ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTTCCATCA 1140
WAYANAD4          ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTTCCATCA 1140
WAYANAD5          ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTTCCATCA 1140
ASSAM2            ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTTCCATCA 1140
ASSAM3            ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTTCCATCA 1140
ASSAM4            ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTTCCATCA 1140
ASSAM5            ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTTCCATCA 1140
                  ************************************************** ******
```

```
TRICHY1          CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 1200
TRICHY3          CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 1200
TRICHY2          CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 1200
TRICHY4          CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 1200
TRICHY5          CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 1200
NARAYANGAON2     CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 1200
VAISHALI1        CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 1200
JALGAON1         CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 1200
JALGAON5         CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 1200
JALGAON4         CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 1200
VAISHALI4        CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 1200
NARAYANGAON3     CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 1200
VAISHALI3        CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 1200
VAISHALI2        CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 1200
JALGAON2         CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 1200
JALGAON3         CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 1200
NARAYANGAON1     CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 1200
NARAYANGAON5     CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 1200
VAISHALI5        CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 1200
NARAYANGAON4     CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 1200
Assam1           CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 1200
WAYANAD1         CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 1200
WAYANAD2         CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 1200
WAYANAD3         CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 1200
WAYANAD4         CTTCGTTTATTATATATTTTAGATGAAATTAATAATCCCTCTATTACTATTAAAGCTATT 1200
WAYANAD5         CTTCGTTTATTATATATTTTAGATGAAATTAATAATCCCTCTATTACTATTAAAGCTATT 1200
ASSAM2           CTTCGTTTATTATATATTTTAGATGAAATTAATAATCCCTCTATTACTATTAAAGCTATT 1200
ASSAM3           CTTCGTTTATTATATATTTTAGATGAAATTAATAATCCCTCTATTACTATTAAAGCTATT 1200
ASSAM4           CTTCGTTTATTATATATTTTAGATGAAATTAATAATCCCTCTATTACTATTAAAGCTATT 1200
ASSAM5           CTTCGTTTATTATATATTTTAGATGAAATTAATAATCCCTCTATTACTATTAAAGCTATT 1200
                 ********************************** ************************

TRICHY1          GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 1260
TRICHY3          GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 1260
TRICHY2          GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 1260
TRICHY4          GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 1260
TRICHY5          GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 1260
NARAYANGAON2     GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 1260
VAISHALI1        GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 1260
JALGAON1         GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 1260
JALGAON5         GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 1260
JALGAON4         GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 1260
VAISHALI4        GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 1260
NARAYANGAON3     GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 1260
VAISHALI3        GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 1260
VAISHALI2        GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 1260
JALGAON2         GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 1260
JALGAON3         GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 1260
NARAYANGAON1     GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 1260
NARAYANGAON5     GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 1260
VAISHALI5        GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 1260
NARAYANGAON4     GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 1260
Assam1           GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 1260
WAYANAD1         GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 1260
WAYANAD2         GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 1260
WAYANAD3         GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 1260
WAYANAD4         GGCCACCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 1260
WAYANAD5         GGCCACCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 1260
ASSAM2           GGCCACCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 1260
ASSAM3           GGCCACCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 1260
ASSAM4           GGCCACCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 1260
ASSAM5           GGCCACCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 1260
                 ***** ****************************************************

TRICHY1          TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 1320
TRICHY3          TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 1320
TRICHY2          TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 1320
TRICHY4          TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 1320
TRICHY5          TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 1320
NARAYANGAON2     TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 1320
VAISHALI1        TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 1320
```

```
JALGAON1          TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 1320
JALGAON5          TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 1320
JALGAON4          TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 1320
VAISHALI4         TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 1320
NARAYANGAON3      TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 1320
VAISHALI3         TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 1320
VAISHALI2         TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 1320
JALGAON2          TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 1320
JALGAON3          TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 1320
NARAYANGAON1      TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 1320
NARAYANGAON5      TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 1320
VAISHALI5         TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 1320
NARAYANGAON4      TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 1320
Assam1            TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 1320
WAYANAD1          TATATAATTCCTACTAATGACTTAAATACCTATAATTTCCGTTTATTAGATGTAGATAAC 1320
WAYANAD2          TATATAATTCCTACTAATGACTTAAATACCTATAATTTCCGTTTATTAGATGTAGATAAC 1320
WAYANAD3          TATATAATTCCTACTAATGACTTAAATACCTATAATTTCCGTTTATTAGATGTAGATAAC 1320
WAYANAD4          TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTATTAGATGTAGATAAC 1320
WAYANAD5          TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTATTAGATGTAGATAAC 1320
ASSAM2            TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTATTAGATGTAGATAAC 1320
ASSAM3            TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTATTAGATGTAGATAAC 1320
ASSAM4            TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTATTAGATGTAGATAAC 1320
ASSAM5            TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTATTAGATGTAGATAAC 1320
                  ***************** ************************** ***************

TRICHY1           CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 1380
TRICHY3           CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 1380
TRICHY2           CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 1380
TRICHY4           CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 1380
TRICHY5           CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 1380
NARAYANGAON2      CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 1380
VAISHALI1         CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 1380
JALGAON1          CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 1380
JALGAON5          CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 1380
JALGAON4          CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATTTATAGATGTAATT 1380
VAISHALI4         CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 1380
NARAYANGAON3      CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 1380
VAISHALI3         CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 1380
VAISHALI2         CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 1380
JALGAON2          CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 1380
JALGAON3          CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 1380
NARAYANGAON1      CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 1380
NARAYANGAON5      CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 1380
VAISHALI5         CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 1380
NARAYANGAON4      CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 1380
Assam1            CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 1380
WAYANAD1          CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 1380
WAYANAD2          CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 1380
WAYANAD3          CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 1380
WAYANAD4          CGATTAATTGTCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 1380
WAYANAD5          CGATTAATTGTCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 1380
ASSAM2            CGATTAATTGTCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 1380
ASSAM3            CGATTAATTGTCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 1380
ASSAM4            CGATTAATTGTCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 1380
ASSAM5            CGATTAATTGTCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 1380
                  ********* ********************************** *************

TRICHY1           CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 1440
TRICHY3           CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 1440
TRICHY2           CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 1440
TRICHY4           CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 1440
TRICHY5           CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 1440
NARAYANGAON2      CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 1440
VAISHALI1         CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 1440
JALGAON1          CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 1440
JALGAON5          CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 1440
JALGAON4          CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 1440
VAISHALI4         CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 1440
NARAYANGAON3      CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 1440
VAISHALI3         CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 1440
VAISHALI2         CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 1440
```

```
JALGAON2          CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 1440
JALGAON3          CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 1440
NARAYANGAON1      CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 1440
NARAYANGAON5      CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 1440
VAISHALI5         CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 1440
NARAYANGAON4      CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 1440
Assam1            CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 1440
WAYANAD1          CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 1440
WAYANAD2          CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 1440
WAYANAD3          CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 1440
WAYANAD4          CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 1440
WAYANAD5          CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 1440
ASSAM2            CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 1440
ASSAM3            CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 1440
ASSAM4            CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 1440
ASSAM5            CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 1440
                  ************************************************************

TRICHY1           CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 1500
TRICHY3           CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 1500
TRICHY2           CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 1500
TRICHY4           CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 1500
TRICHY5           CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 1500
NARAYANGAON2      CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 1500
VAISHALI1         CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 1500
JALGAON1          CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 1500
JALGAON5          CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 1500
JALGAON4          CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 1500
VAISHALI4         CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 1500
NARAYANGAON3      CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 1500
VAISHALI3         CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 1500
VAISHALI2         CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 1500
JALGAON2          CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 1500
JALGAON3          CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 1500
NARAYANGAON1      CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 1500
NARAYANGAON5      CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 1500
VAISHALI5         CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 1500
NARAYANGAON4      CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 1500
Assam1            CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 1500
WAYANAD1          CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 1500
WAYANAD2          CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 1500
WAYANAD3          CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 1500
WAYANAD4          CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 1500
WAYANAD5          CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 1500
ASSAM2            CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 1500
ASSAM3            CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 1500
ASSAM4            CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 1500
ASSAM5            CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 1500
                  ************************************************************

TRICHY1           GG 1502
TRICHY3           GG 1502
TRICHY2           GG 1502
TRICHY4           GG 1502
TRICHY5           GG 1502
NARAYANGAON2      GG 1502
VAISHALI1         GG 1502
JALGAON1          GG 1502
JALGAON5          GG 1502
JALGAON4          GG 1502
VAISHALI4         GG 1502
NARAYANGAON3      GG 1502
VAISHALI3         GG 1502
VAISHALI2         GG 1502
JALGAON2          GG 1502
JALGAON3          GG 1502
NARAYANGAON1      GG 1502
NARAYANGAON5      GG 1502
VAISHALI5         GG 1502
NARAYANGAON4      GG 1502
Assam1            GG 1502
```

```
WAYANAD1          GG 1502
WAYANAD2          GG 1502
WAYANAD3          GG 1502
WAYANAD4          GG 1502
WAYANAD5          GG 1502
ASSAM2            GG 1502
ASSAM3            GG 1502
ASSAM4            GG 1502
ASSAM5            GG 1502
                  **
```

Fig. 1. ClustalW alignment of the COI-tRNA^{Leu}-COII region from the thirty weevils of *Odoiporus longicollis* (Oliver)

The base composition was AT rich with A = 31.9%, T = 38.1%, C = 17.7% and G = 12.9% with almost 70% AT in the coding regions **(Table 4).** This strong AT bias agrees with published information for insect mitochondrial DNA (Sheffield *et al.* 2008; Simons *et al.* 1994).

Table 4. Base composition of COI/COII, COI and COII at the three codon positions

Codon nucleotide position	**Base composition (%)**				**A+T%**
	A	**C**	**G**	**T**	
All COI/COII sites	31.9	17.7	12.9	38.1	70
COI Pos 1	39.0	11.9	3.2	46.0	85.0
COI Pos 2	31.5	1501	22.3	31	62.5
COI Pos 3	20.9	22.1	16.5	40	62.1
All COI sites	30.5	16.4	14.1	39.1	69.6
COII Pos 1	31.7	22.9	20.0	25.0	56.7
COII Pos 2	26.9	18.4	12.9	42.0	68.9
COII Pos 3	40.0	13.8	0.4	46.0	86.0
All COII sites	32.9	18.3	11.1	37.7	70.6

Haplotype analysis

The total number of sites selected for haplotype analysis (excluding sites with gaps / missing data) was 1502. Number of variable sites was found to be 39. Fifteen haplotypes (gaps excluded) were identified amongst these 30 individuals and the mean haplotype diversity (Hd) was 0.8989 (**Table 5).** Excepting for haplotype 8 which was present in all the individuals from Narayangaon and individuals 2, 4 and 5 from Vaishali, none of the other haplotypes were common between locations

Table 5. Distribution of haplotype among the thirty weevils based on the COI-tRNA^{Leu}-COII sequence

Haplotype	Location	Individual	Total no
Hap 1	Assam	Assam1	1
Hap 2	Assam	Assam2	1
Hap 3	Assam	Assam3, 4, 5	3
Hap 4	Jalgaon	Jalgaon1, 5	2
Hap 5	Jalgaon	Jalgaon2	1
Hap 6	Jalgaon	Jalgaon3	1
Hap 7	Jalgaon	Jalgaon4	1
Hap 8	Narayangaon Vaishali	Narayangaon 1, 2, 3, 4, 5 and Vaishali1, 2, 4, 5	9
Hap 9	Trichy	Trichy1, 3	2
Hap 10	Trichy	Trichy2, 4	2
Hap 11	Trichy	Trichy5	1
Hap 12	Vaishali	Vaishali3	1
Hap 13	Wayanad	Wayanad1, 2	2

Hap 14	Wayanad	Wayanad3	1
Hap 15	Wayanad	Wayanad4, 5	2

Ts/Tv ratio of COI-tRNA^{Leu}-COII region

The extent of saturation in the COI-tRNA^{Leu}-COII region data set was determined by the Ts/Tv ratio and the Iss values (Xia *et al.* 2003). A Ts/Tv ratio of 9.3736, the increase in the Ts/Tv ratio with genetic distance and the I_{ss} value of 0.0062 which was significantly less than the Iss.c value of 0.7747 (p= 0.0022 , DF= 1501) suggested that substitutions in this data set have not reached saturation (**Fig. 2 and Table 6)**.

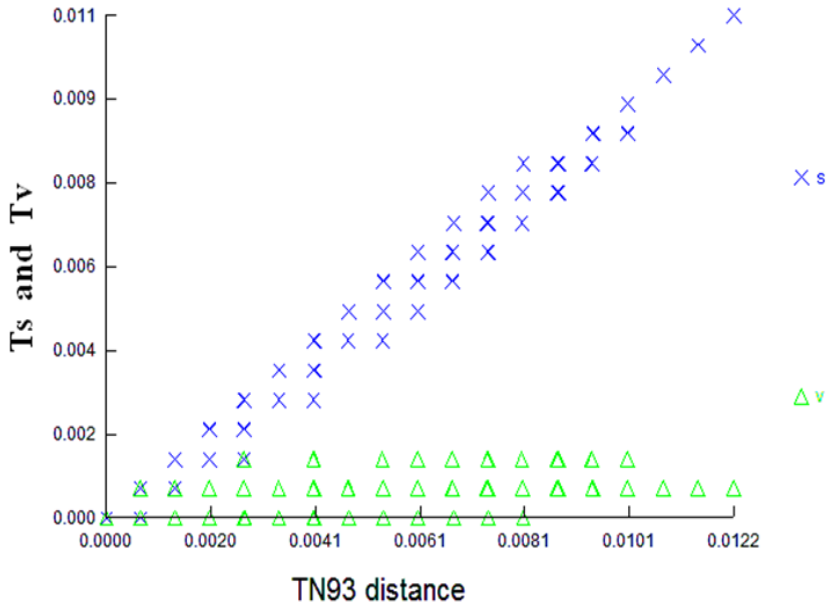


Fig. 2. Graph showing relationship between Ts and Tv and divergence of COI-tRNA^{Leu}-COII region

Table 6. Xia's test and nucleotide diversity (Synonymous and Non synonymous) for COI-tRNA^{leu}-COII, COI and COII at the three codon positions

Codon position	Iss	Iss.c	PC (two tailed)	Nucleotide diversity (π) in synonymous sites	Nucleotide diversity (π) in non synonymous sites
COI-tRNAL-COII (1st)	0.0062	0.7747	0.0000 (DF 1501)	0.01323	0.00206
COI-tRNAL-COII (2nd)	0.0030	0.7047	0.0000 (DF 1500)	0.00699	0.00483
COI-tRNAL-COII (3rd)	0.0146	0.7045	0.0000 (DF 1500)	0.00568	0.00431
COI (1st)	0.0055	0.7415	0.0000 (DF 834)	0.00941	0.00232
COI (2nd)	0.0047	0.6640	0.0000 (DF 833)	0.00635	0.00326
COI (3rd)	0.0114	0.6638	0.0000 (DF 833)	0.00517	0.00365
COII (1st)	0.0069	0.7188	0.0000 (DF 601)	0.01997	0.00172
COII (2nd)	0.0029	0.6464	0.0000 (DF 600)	0.00302	0.00650
COII (3rd)	0.0194	0.6462	0.0000 (DF 600)	0.00412	0.00614

For estimating the phylogenetic relationships amongst the thirty COI-tRNA^{Leu}-COII sequences, the T92+G model was selected as the best fit model using Mega v5, as it gave the lowest BIC score of 5171.4 with lnL = -2264.3. Substitution pattern and rates which were estimated using this model indicated that the most frequent substitutions were transitions of the G-A and C-T(U) types (**Table 7**). Total no of transitions was found to be thirty-one (A-C, T-C) and total no of transversions was three (T-G, A-T).

Table 7. Maximum Likelihood Estimate of Substitution Matrix based on T92 + G model

	A	T/U	C	G
A	-	*1.07*	*0.46*	**14.04**
T/U	*1.07*	-	**14.04**	*0.46*
C	*1.07*	**32.90**	-	*0.46*
G	**32.90**	*1.07*	*0.46*	-

Each entry is the probability of substitution (*r*) from one base (row) to another base (column). Substitution pattern and rates were estimated under the Tamura (1992) model (+G) [1]. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories, [+*G*], parameter = 0.0500). For simplicity, sum of *r* values is made equal to 100, the nucleotide frequencies are A = 35.05%, T/U = 35.05%, C = 14.95%, and G = 14.95%. For estimating ML values, a tree topology was automatically computed. The maximum Log likelihood for this computation was -2263.342. The analysis involved 30 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. There were a total of 1504 positions in the final dataset. Evolutionary analyses were conducted in MEGA]. All positions containing gaps and missing data were eliminated. There were a total of 458 positions in the final dataset. Rates of different transitional substitutions are shown in **bold** and those of transversionsal substitutions are shown in *italics*.

Genetic distance based on the COI-tRNA^{Leu}-COII region

The intra-population genetic distance ranged from 0.000 (Narayangaon and Vaishali) to 0.005 (Jalgaon and Wayanad) with mean diversity within populations of 0.002 (S.E. 0.000). The inter-population genetic distance ranged from 0.000 (between Narayangaon and Vaishali) to 0.009 (between Assam and Jalgaon) with a mean inter-population diversity of 0.005 (S.E. 0.001).

Phylogenetic analyses based on the COI-tRNA^{Leu}-COII region

Phylogenetic analysis was done using the thirty COI-tRNA^{Leu}-COII gene sequences as well as the haplotype data set with the maximum likelihood algorithm in MEGA 5 using the Tamura-3 parameter model with Gamma distribution and no invariant sites (**Figs. 3a and b**).

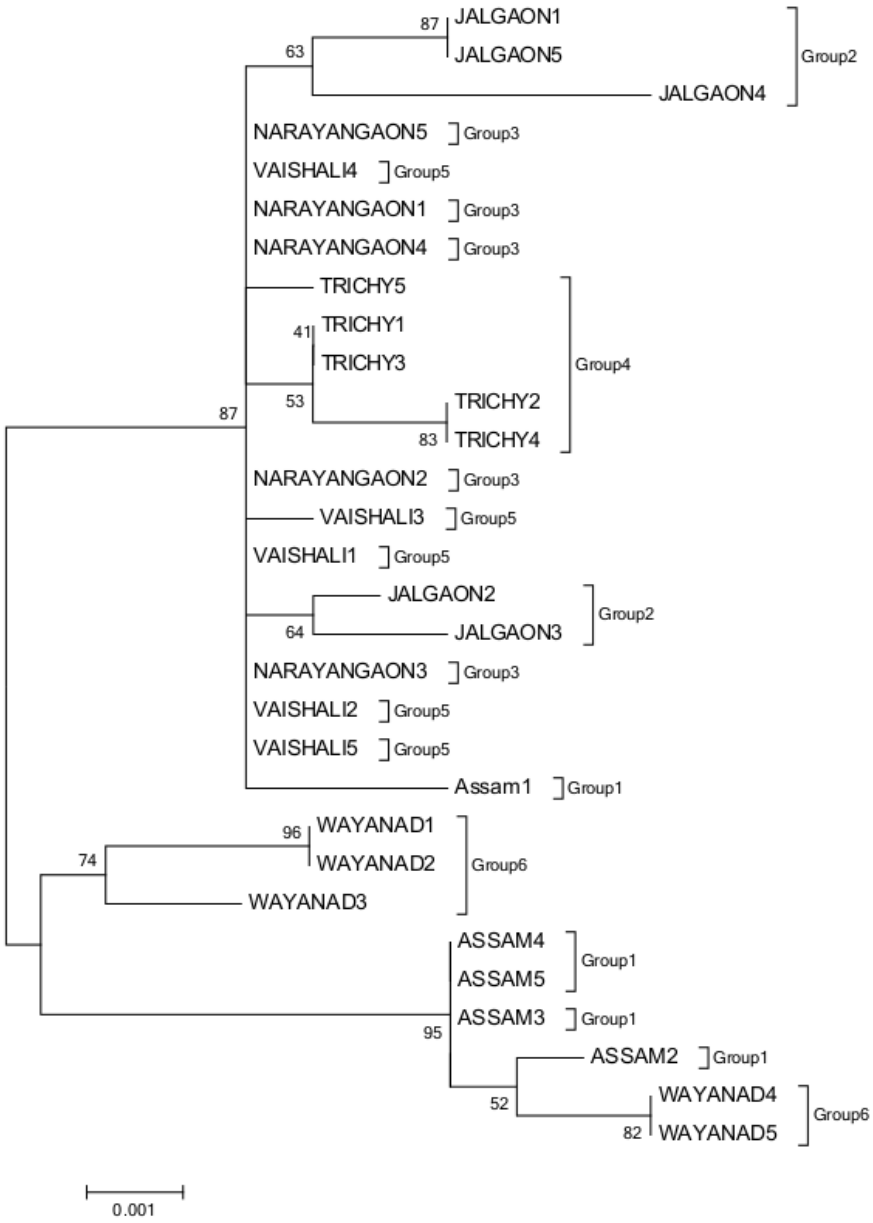


Fig. 3a. Maximum likelihood tree based on the COI-tRNA^{Leu}-COII sequences of the thirty *O. longicollis* (Oliver) individuals

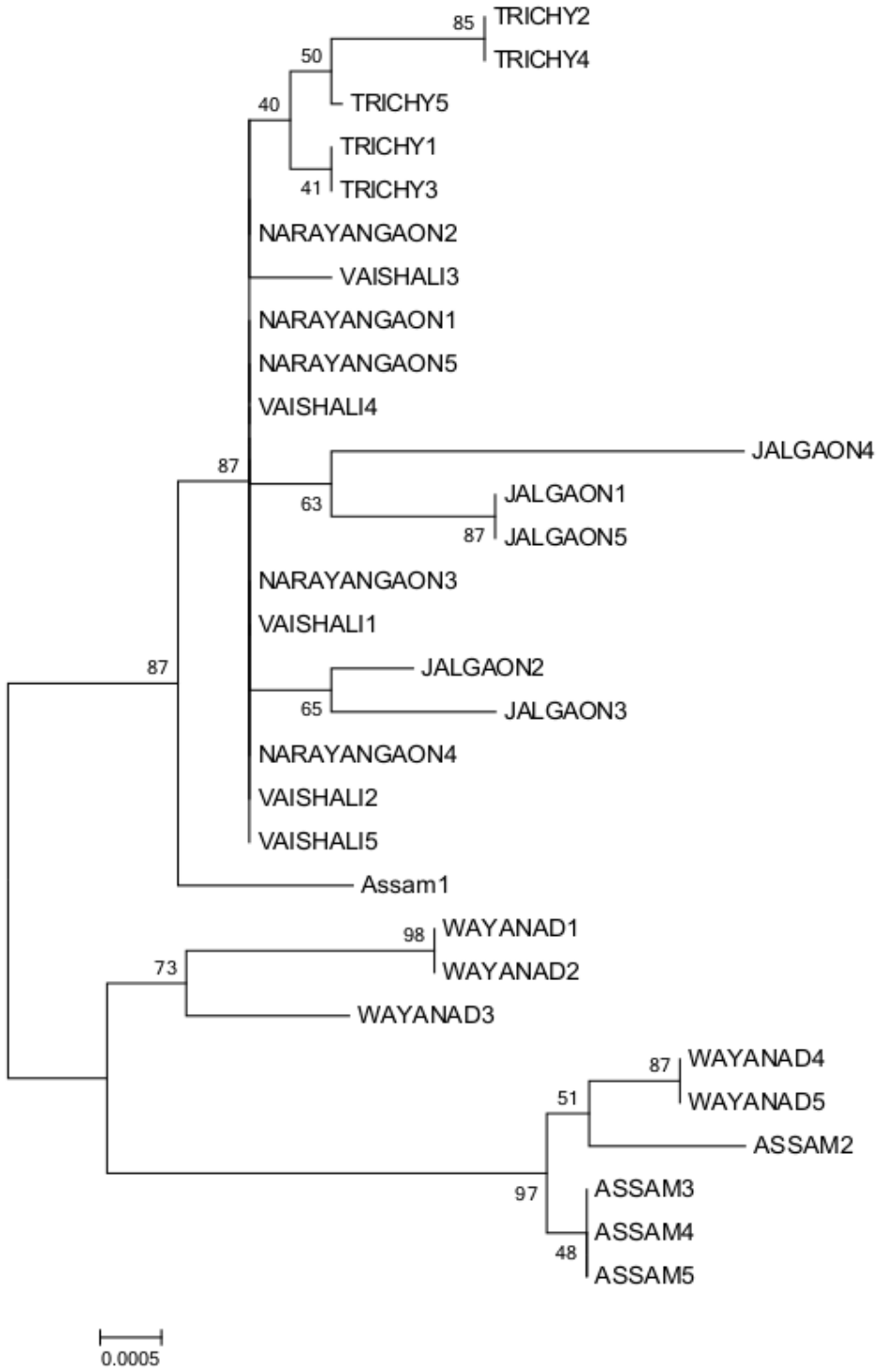


Fig. 3b. Maximum likelihood tree based on the COI-tRNA^{Leu}-COII haplotype data of the thirty *O. longicollis* (Oliver) individuals

In both the trees, individuals 2, 3, 4 and 5 from Assam and all the five individuals from Wayanad cluster into one group. No haplotype based grouping or phylogeographic distribution of the populations. There was no strong correlation between geographic distance and genetic distance [Mantel test (1967)] (p= 0.030) (**Fig. 4**).

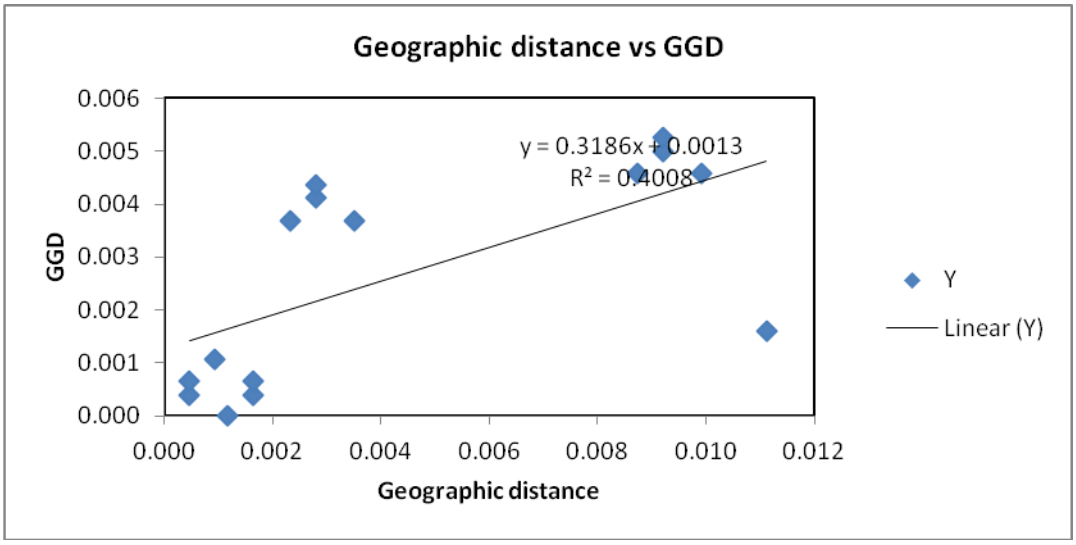


Fig. 4. Correlation between genetic distance (GGD) and geographic distance based on the COI-tRNA^{Leu}-COII region (Mantel test) (p=0.030)

Estimates of Population differentiation

Pairwise Fst values varied from 0.75556 to 0.999 and pair wise Gst values varied from 0.000 to 0.48448 with seven of the fifteen comparisons giving significant p values (**Table 8**).

Table 8. F_{ST}, G_{ST}, Nm and χ^2 values between the six *O. longicollis* (Oliver) populations based on the COI-tRNA^{Leu}-COII sequences

Populatin	Assam	Jalgaon	Narayagaon	Trichy	Vaishali	Wayanad
Assam	000					
Jalgaon	F_{ST} 0.56061 (Nm 0.20) G_{ST} 0.11111 (Nm 2.00) χ^2:0.1247 ns df=6	000				
Narayagan	F_{ST} 0.75556 (Nm 0.08) G_{ST} 0.48448 (Nm 0.27) χ^2:0.0186* (*0.01<P<0.05) df =3	F_{ST} 0.14286 (Nm 1.50) G_{ST} 0.37931 (Nm 0.41) χ^2:0.0404* (*0.01<P<0.05) df= 4	000			
Trichy	F_{ST} 0.72222 (Nm 0.10) G_{ST} 0.14286 (Nm 1.50) χ^2:0.0752 (ns) df= 5	F_{ST} 0.26667 (Nm 0.69) G_{ST} 0.081808 (Nm 2.83) χ^2: 0.1247 (ns) df=6	F_{ST} 0.55556 (Nm 0.20) G_{ST} 0.42857 (Nm 0.33) χ^2:0.0186* (*0.01<P<0.05) df=3	000		
Vaishali	F_{ST} 0.73913 (Nm 0.09) G_{ST} 0.29032 (Nm 0.61) χ^2:0.0404* (*0.01<P<0.05) df = 4	F_{ST} 0.13636 (Nm 1.58) G_{ST} 0.21212 (Nm 0.93) χ^2:0.0752 (ns) df =5	F_{ST} 0.0000 (Nm 000) G_{ST} 0.0000 (Nm 1801) χ^2:0.2918 (ns) df=1	F_{ST} 0.50000 (Nm 0.25) G_{ST} 0.25000 (Nm 0.75) χ^2:0.0404* (*0.01<P<0.05) df=4	000	
Wayanad	F_{ST} 0.29455 (Nm 0.60) G_{ST} 0.14286 (Nm 1.50) χ^2:0.0752 (ns) df = 5	F_{ST} 0.43212 (Nm 0.29) G_{ST} 0.08108 (Nm 2.83) χ^2: 0.1247 (ns) df=6	F_{ST} 0.61111 (Nm 0.16) G_{ST} 0.42857 (Nm 0.33) χ^2:0.0186* (* 0.01<P<0.05) df=3	F_{ST} 0.60185 (Nm 0.17) G_{ST} 0.1111 (Nm 2.00) χ^2:0.0752 (ns) df=5	F_{ST} 0.59783 (Nm 0.17) G_{ST} 0.25000 (Nm 0.75) χ^2:0.0404* (*0.01<P<0.05) df=4	000

Overall estimates for whole population of F_{ST} [0.53637 (Nm 0.22)], G_{ST} [0.32500 (Nm 0.52)], χ^2 (p values) [123.333, 0.0001 *** (***, P<0.001)

The overall genetic differentiation estimate between the populations gave a highly significant p value of 0.0001 ($p<0.001$; $\chi^2 = 123.333$; df=75), suggesting restricted gene flow between the populations. Non-significant values of -1.00348 and -1.949 respectively with Tajma's D test and Fu's *Fs* test support the neutrality hypothesis indicating that the populations under study are normal evolving populations without any bottleneck or genetic drift which could result in genetic differentiation (**Table 9**).

Table 9. Results for Tajima's Neutrality Test

Codon position	***m***	***S***	p_s	$\theta^{\#}$	Π^{*}	D^{**}
All three positions (1st, 2nd and 3rd)	30	35	0.023318	0.005886	0.004391	-0.927790
1st codon	30	16	0.031936	0.008061	0.004318	-1.577113
2nd codon	30	12	0.023952	0.006046	0.006410	0.196550
3rd codon	30	7	0.014028	0.003541	0.002437	-0.921107

Test for Wolbachia infection

Wolbachia specific primers did not give any amplification products.

Sliding window analysis of the COI-tRNA^{Leu}-COII region for nucleotide diversity

Nucleotide diversity (π) of the COI-tRNA^{Leu}-COII region was estimated using sliding window analysis with DnaSP v 5.10 (Nei, 1987 equation 10.5; Rozas *et al.* 2003). As evident from the graph, localized nucleotide divergence was observed, which is an indicator of the total divergence (**Figs. 5 and 6**) (Roe & Sperling 2007). Localized divergence was quite variable among windows across the full COI-tRNA^{leu}-COI sequence length. Non-random regional variations has been reported for mt DNA (Lin & Danforth 2004; Broughton & Reneau 2006; Galtier *et al.* 2006).

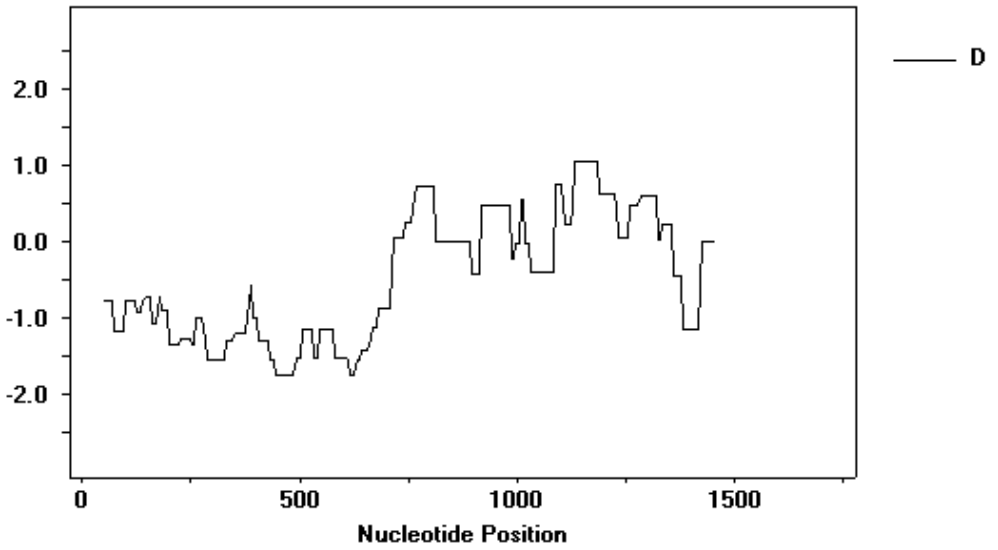


Fig. 5. Sliding window analysis of the COI-tRNA^{Leu}-COI region for Tajma's D

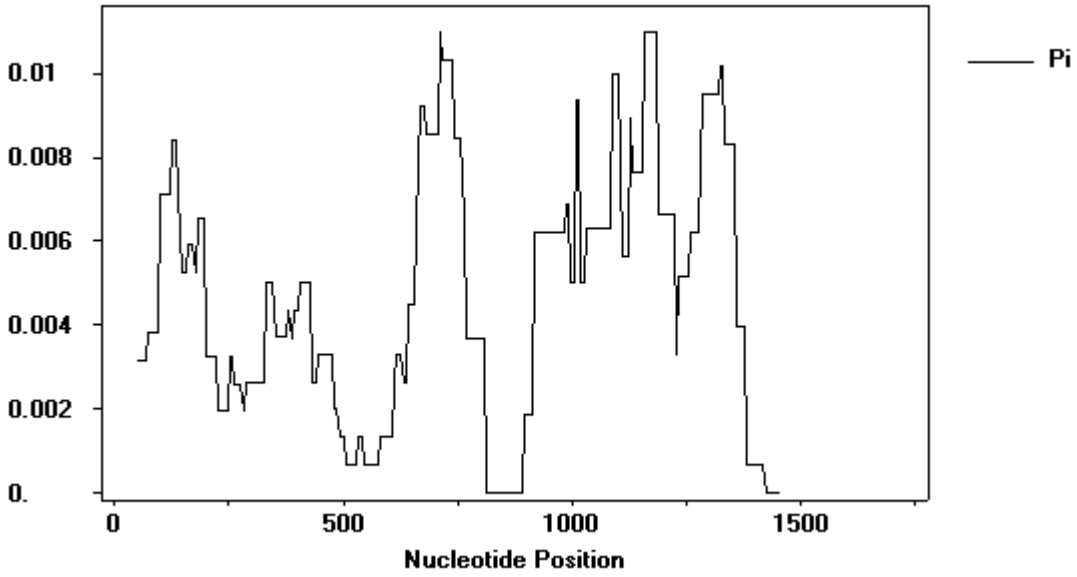


Fig. 6. Sliding window analysis of the COI-tRNA^{Leu}-COII region for nucleotide diversity

Analysis of the COI region

The M1-M2 primers amplified 835 bp of the 3' region of the COI gene. This region included 812 conserved, 23 variable and 12 parsimony informative sites (**Tables 1 and 3**). ClustalW alignment of the nucleotide sequences of COI of all the thirty individuals shows that of the 21 base substitutions, six, five and ten occur in the first, second and third codon positions respectively (**Table 10 and Fig. 7**).

Table 10. Details of nucleotide changes in the partial COI region of the thirty individuals of *O. longicollis* (Oliver)

Serial no	Position of nucleotide	Position of amino acid	Type of nucleotide change	Change in codon	Change in codon position	Amino acid change	Type of change	Individuals	Structural region of the COI protein
1	87	29	Transition	AAA→AAG	3rd	K	Synonymous	Wayanad 1, 2	I3
2	90	30	Transition	AAG→AAA	3rd	K	Synonymous	Assam 2, Wayanad 4, 5	I3
3	123	41	Transition	GCT→GCC	3rd	A	Synonymous	Jalgaon 4	I3
4	150	50	Transition	TTT→TTC	3rd	F	Synonymous	Wayanad 4, 5, Assam 2, 3, 4, 5	M7
5	174	58	Transition	TTT→TTC	3rd	F	Synonymous	Jalgaon 3	M7
6	209	70	Transition	TTT→TCT	2nd	F→S	Non synonymous	Vaishali 3	I4
7	230	77	Transition	ATT→ACT	2nd	I→T	Non synonymous	Wayanad 1, 2	I4
8	330	110	Transition	ATT→ATC	3rd	I	Synonymous	Assam 1	E5
9	335	112	Transition	CTT→CCT	2nd	L→P	Non synonymous	Wayanad 3	E5
10	376	126	Transversion	TCA→GCA	1st	S→A	Conserved substitution	Trichy 1, 2, 3, 4	M10
11	426	142	Transition	TTT→TTC	3rd	F	Synonymous	Jalgaon 1, 5	M10
12	438	146	Transition	CTA→CTG	3rd	L	Synonymous	Jalgaon 3	M10
13	454	152	Transition	TTT→CTT	1st	F→L	Conserved substitution	Assam 2	M10
14	493	165	Transition	TTT→CTT	1st	F→L	Conserved substitution	Jalgaon 4	I5
15	576	192	Transition	CCT→CCC	3rd	P	Synonymous	Jalgaon 3	M11
16	626	209	Transition	CAG→CGG	2nd	Y	Synonymous	Jalgaon 4	E6
17	661	221	Transition	ATT→GTT	1st	I→V	Conserved substitution	Jalgaon 2	M12
18	688	230	Transition	ATT→GTT	1st	I→V	Conserved substitution	Wayanad 3, 4, 5	M12
19	707	236	Transition	ATT→ACT	2nd	I→T	Non synonymous	Jalgaon 1, 4, 5	COOH
20	711	237	Transition	ATC→ATT	3rd	I	Synonymous	Assam 1, 2, 3, 4, 5	COOH
21	760	254	Transition	TCA→CCA	1st	S→P	Semi conservative	Wayanad 4, 5, Assam 1, 2, 3, 4, 5	COOH

```
WAYANAD1          TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
WAYANAD2          TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
NARAYANGAON1      TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
NARAYANGAON2      TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
TRICHY5           TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
TRICHY2           TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
TRICHY4           TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
TRICHY1           TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
TRICHY3           TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
NARAYANGAON3      TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
VAISHALI1         TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
NARAYANGAON4      TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
VAISHALI3         TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
VAISHALI2         TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
VAISHALI4         TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
VAISHALI5         TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
NARAYANGAON5      TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
JALGAON1          TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
JALGAON5          TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
JALGAON4          TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
JALGAON2          TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
JALGAON3          TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
WAYANAD3          TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
WAYANAD4          TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
WAYANAD5          TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
ASSAM2            TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
ASSAM3            TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
ASSAM4            TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
ASSAM5            TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
Assam1            TGATTTTTTGGTCATCCAGAAGTTTATATTTTAATTCTACCAGGATTTGGTATAATTTCT 60
                  ************************************************************

WAYANAD1          CATATTATTAGTCAAGAAAGAGGAAAGAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
WAYANAD2          CATATTATTAGTCAAGAAAGAGGAAAGAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
NARAYANGAON1      CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
NARAYANGAON2      CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
TRICHY5           CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
TRICHY2           CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
TRICHY4           CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
TRICHY1           CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
TRICHY3           CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
NARAYANGAON3      CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
VAISHALI1         CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
NARAYANGAON4      CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
VAISHALI3         CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
VAISHALI2         CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
VAISHALI4         CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
VAISHALI5         CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
NARAYANGAON5      CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
JALGAON1          CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
JALGAON5          CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
JALGAON4          CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
JALGAON2          CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
JALGAON3          CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
WAYANAD3          CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
WAYANAD4          CATATTATTAGTCAAGAAAGAGGAAAAAAAGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
WAYANAD5          CATATTATTAGTCAAGAAAGAGGAAAAAAAGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
ASSAM2            CATATTATTAGTCAAGAAAGAGGAAAAAAAGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
ASSAM3            CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
ASSAM4            CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
ASSAM5            CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
Assam1            CATATTATTAGTCAAGAAAGAGGAAAAAAGGAAGCCTTTGGAGTCTTAGGAATAATCTAT 120
                  ************************* ** ******************************

WAYANAD1          GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
WAYANAD2          GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
NARAYANGAON1      GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
NARAYANGAON2      GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
TRICHY5           GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
TRICHY2           GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
TRICHY4           GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
```

```
TRICHY1          GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
TRICHY3          GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
NARAYANGAON3     GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
VAISHALI1        GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
NARAYANGAON4     GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
VAISHALI3        GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
VAISHALI2        GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
VAISHALI4        GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
VAISHALI5        GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
NARAYANGAON5     GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
JALGAON1         GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
JALGAON5         GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
JALGAON4         GCCATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
JALGAON2         GCTATAATAGCAATTGGGCTTTAGGATTTTGTAGTCTGAGCTCACCATATATTCACAGTA 180
JALGAON3         GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTCACAGTA 180
WAYANAD3         GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
WAYANAD4         GCTATAATAGCAATTGGGCTTTTAGGATTCGTAGTCTGAGCTCACCATATATTTACAGTA 180
WAYANAD5         GCTATAATAGCAATTGGGCTTTTAGGATTCGTAGTCTGAGCTCACCATATATTTACAGTA 180
ASSAM2           GCTATAATAGCAATTGGGCTTTTAGGATTCGTAGTCTGAGCTCACCATATATTTACAGTA 180
ASSAM3           GCTATAATAGCAATTGGGCTTTTAGGATTCGTAGTCTGAGCTCACCATATATTTACAGTA 180
ASSAM4           GCTATAATAGCAATTGGGCTTTTAGGATTCGTAGTCTGAGCTCACCATATATTTACAGTA 180
ASSAM5           GCTATAATAGCAATTGGGCTTTTAGGATTCGTAGTCTGAGCTCACCATATATTTACAGTA 180
Assam1           GCTATAATAGCAATTGGGCTTTTAGGATTTGTAGTCTGAGCTCACCATATATTTACAGTA 180
                 ** ******************  *  ** ********************** ******

WAYANAD1         GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTACTGCTGTTCCT 240
WAYANAD2         GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTACTGCTGTTCCT 240
NARAYANGAON1     GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
NARAYANGAON2     GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
TRICHY5          GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
TRICHY2          GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
TRICHY4          GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
TRICHY1          GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
TRICHY3          GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
NARAYANGAON3     GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
VAISHALI1        GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
NARAYANGAON4     GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
VAISHALI3        GGGATAGATGTAGATACTCGGGCATATTCTACTTCAGCTACAATAATTATTGCTGTTCCT 240
VAISHALI2        GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
VAISHALI4        GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
VAISHALI5        GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
NARAYANGAON5     GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
JALGAON1         GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
JALGAON5         GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
JALGAON4         GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
JALGAON2         GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
JALGAON3         GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
WAYANAD3         GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
WAYANAD4         GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
WAYANAD5         GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
ASSAM2           GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
ASSAM3           GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
ASSAM4           GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
ASSAM5           GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
Assam1           GGGATAGATGTAGATACTCGGGCATATTTTACTTCAGCTACAATAATTATTGCTGTTCCT 240
                 ************************** ******************** **********

WAYANAD1         ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
WAYANAD2         ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
NARAYANGAON1     ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
NARAYANGAON2     ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
TRICHY5          ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
TRICHY2          ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCC 300
TRICHY4          ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCC 300
TRICHY1          ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
TRICHY3          ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
NARAYANGAON3     ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
VAISHALI1        ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
NARAYANGAON4     ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
VAISHALI3        ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
VAISHALI2        ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
```

```
VAISHALI4          ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
VAISHALI5          ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
NARAYANGAON5       ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
JALGAON1           ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
JALGAON5           ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
JALGAON4           ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
JALGAON2           ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
JALGAON3           ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
WAYANAD3           ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
WAYANAD4           ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
WAYANAD5           ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
ASSAM2             ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
ASSAM3             ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
ASSAM4             ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
ASSAM5             ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
Assam1             ACTGGAATCAAAATTTTTAGATGATTAGCAACTTACCATGGCGCCCGAATTACTTATTCT 300
                   ************************************************************

WAYANAD1           CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
WAYANAD2           CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
NARAYANGAON1       CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
NARAYANGAON2       CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
TRICHY5            CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
TRICHY2            CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
TRICHY4            CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
TRICHY1            CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
TRICHY3            CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
NARAYANGAON3       CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
VAISHALI1          CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
NARAYANGAON4       CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
VAISHALI3          CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
VAISHALI2          CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
VAISHALI4          CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
VAISHALI5          CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
NARAYANGAON5       CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
JALGAON1           CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
JALGAON5           CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
JALGAON4           CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
JALGAON2           CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
JALGAON3           CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
WAYANAD3           CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCCTTTTACTATAGGAGGTCTAACTGGA 360
WAYANAD4           CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
WAYANAD5           CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
ASSAM2             CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
ASSAM3             CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
ASSAM4             CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
ASSAM5             CCTACTTCCTTATGAAGCTTAGGCTTTATTTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
Assam1             CCTACTTCCTTATGAAGCTTAGGCTTTATCTTTCTTTTTACTATAGGAGGTCTAACTGGA 360
                   *************************** **** ************************

WAYANAD1           GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
WAYANAD2           GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
NARAYANGAON1       GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
NARAYANGAON2       GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
TRICHY5            GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
TRICHY2            GTAGTATTAGCAAATGCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
TRICHY4            GTAGTATTAGCAAATGCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
TRICHY1            GTAGTATTAGCAAATGCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
TRICHY3            GTAGTATTAGCAAATGCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
NARAYANGAON3       GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
VAISHALI1          GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
NARAYANGAON4       GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
VAISHALI3          GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
VAISHALI2          GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
VAISHALI4          GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
VAISHALI5          GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
NARAYANGAON5       GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
JALGAON1           GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
JALGAON5           GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
JALGAON4           GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
JALGAON2           GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
```

```
JALGAON3          GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
WAYANAD3          GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
WAYANAD4          GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
WAYANAD5          GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
ASSAM2            GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
ASSAM3            GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
ASSAM4            GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
ASSAM5            GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
Assam1            GTAGTATTAGCAAATTCATCAATTGATATCATCTTACATGATACTTATTATGTAGTAGCC 420
                  *************** ******************************************** 

WAYANAD1          CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
WAYANAD2          CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
NARAYANGAON1      CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
NARAYANGAON2      CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
TRICHY5           CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
TRICHY2           CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
TRICHY4           CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
TRICHY1           CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
TRICHY3           CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
NARAYANGAON3      CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
VAISHALI1         CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
NARAYANGAON4      CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
VAISHALI3         CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
VAISHALI2         CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
VAISHALI4         CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
VAISHALI5         CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
NARAYANGAON5      CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
JALGAON1          CATTTCCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
JALGAON5          CATTTCCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
JALGAON4          CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
JALGAON2          CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
JALGAON3          CATTTTCATTATGTACTGTCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
WAYANAD3          CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
WAYANAD4          CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
WAYANAD5          CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
ASSAM2            CATTTTCATTATGTACTATCAATAGGAGCTGTACTTGCTATTATTGCAGGATTTATTCAA 480
ASSAM3            CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
ASSAM4            CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
ASSAM5            CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
Assam1            CATTTTCATTATGTACTATCAATAGGAGCTGTATTTGCTATTATTGCAGGATTTATTCAA 480
                  ***** *********** *************** ************************** 

WAYANAD1          TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
WAYANAD2          TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
NARAYANGAON1      TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
NARAYANGAON2      TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
TRICHY5           TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
TRICHY2           TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
TRICHY4           TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
TRICHY1           TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
TRICHY3           TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
NARAYANGAON3      TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
VAISHALI1         TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
NARAYANGAON4      TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
VAISHALI3         TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
VAISHALI2         TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
VAISHALI4         TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
VAISHALI5         TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
NARAYANGAON5      TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
JALGAON1          TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
JALGAON5          TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
JALGAON4          TGATTTCCTCTTCTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
JALGAON2          TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
JALGAON3          TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
WAYANAD3          TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
WAYANAD4          TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
WAYANAD5          TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
ASSAM2            TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
ASSAM3            TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
ASSAM4            TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
```

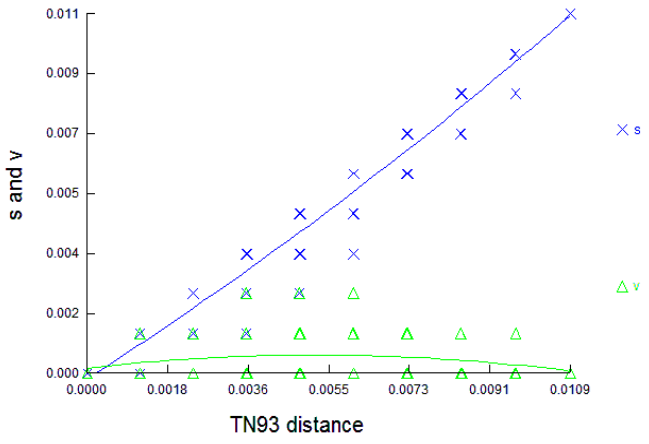
```
ASSAM5          TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
Assam1          TGATTTCCTCTTTTTACAGGATTAACTTTAAACCAAAAATATTTAAAAATCCAGTTTTTT 540
                ************ ********************************************

WAYANAD1        TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
WAYANAD2        TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
NARAYANGAON1    TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
NARAYANGAON2    TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
TRICHY5         TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
TRICHY2         TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
TRICHY4         TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
TRICHY1         TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
TRICHY3         TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
NARAYANGAON3    TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
VAISHALI1       TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
NARAYANGAON4    TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
VAISHALI3       TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
VAISHALI2       TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
VAISHALI4       TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
VAISHALI5       TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
NARAYANGAON5    TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
JALGAON1        TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
JALGAON5        TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
JALGAON4        TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
JALGAON2        TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
JALGAON3        TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCCCAACATTTTCTAGGATTAAGTGGT 600
WAYANAD3        TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
WAYANAD4        TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
WAYANAD5        TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
ASSAM2          TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
ASSAM3          TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
ASSAM4          TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
ASSAM5          TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
Assam1          TCTATATTTTTAGGAGTAAATTTAACTTTCTTTCCTCAACATTTTCTAGGATTAAGTGGT 600
                ********************************** ************************

WAYANAD1        ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
WAYANAD2        ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
NARAYANGAON1    ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
NARAYANGAON2    ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
TRICHY5         ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
TRICHY2         ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
TRICHY4         ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
TRICHY1         ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
TRICHY3         ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
NARAYANGAON3    ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
VAISHALI1       ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
NARAYANGAON4    ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
VAISHALI3       ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
VAISHALI2       ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
VAISHALI4       ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
VAISHALI5       ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
NARAYANGAON5    ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
JALGAON1        ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCT 660
JALGAON5        ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCT 660
JALGAON4        ATACCACGACGTTATTCTGACTATCCGGATGCATATTATATATGAAACTCAATTTCTTCA 660
JALGAON2        ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
JALGAON3        ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
WAYANAD3        ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
WAYANAD4        ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
WAYANAD5        ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
ASSAM2          ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
ASSAM3          ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
ASSAM4          ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
ASSAM5          ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
Assam1          ATACCACGACGTTATTCTGACTATCCAGATGCATATTATATATGAAACTCAATTTCTTCA 660
                ************************* ******************************

WAYANAD1        ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTATTATCTGAGAAGCT 720
WAYANAD2        ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTATTATCTGAGAAGCT 720
NARAYANGAON1    ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTATTATCTGAGAAGCT 720
```

```
NARAYANGAON2      ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTATTATCTGAGAAGCT 720
TRICHY5           ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTATTATCTGAGAAGCT 720
TRICHY2           ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTATTATCTGAGAAGCT 720
TRICHY4           ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTATTATCTGAGAAGCT 720
TRICHY1           ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTATTATCTGAGAAGCT 720
TRICHY3           ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTATTATCTGAGAAGCT 720
NARAYANGAON3      ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTATTATCTGAGAAGCT 720
VAISHALI1         ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTATTATCTGAGAAGCT 720
NARAYANGAON4      ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTATTATCTGAGAAGCT 720
VAISHALI3         ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTATTATCTGAGAAGCT 720
VAISHALI2         ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTATTATCTGAGAAGCT 720
VAISHALI4         ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTATTATCTGAGAAGCT 720
VAISHALI5         ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTATTATCTGAGAAGCT 720
NARAYANGAON5      ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTATTATCTGAGAAGCT 720
JALGAON1          ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTACTATCTGAGAAGCT 720
JALGAON5          ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTACTATCTGAGAAGCT 720
JALGAON4          ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTACTATCTGAGAAGCT 720
JALGAON2          GTTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTATTATCTGAGAAGCT 720
JALGAON3          ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTATTATCTGAGAAGCT 720
WAYANAD3          ATTGGTAGACTAATCTCTTTAATTAGTGTTTTATACCTAATTTTTATTATCTGAGAAGCT 720
WAYANAD4          ATTGGTAGACTAATCTCTTTAATTAGTGTTTTATACCTAATTTTTATTATCTGAGAAGCT 720
WAYANAD5          ATTGGTAGACTAATCTCTTTAATTAGTGTTTTATACCTAATTTTTATTATCTGAGAAGCT 720
ASSAM2            ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTATTATTTGAGAAGCT 720
ASSAM3            ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTATTATTTGAGAAGCT 720
ASSAM4            ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTATTATTTGAGAAGCT 720
ASSAM5            ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTATTATTTGAGAAGCT 720
Assam1            ATTGGTAGACTAATCTCTTTAATTAGTATTTTATACCTAATTTTTATTATTTGAGAAGCT 720
                   ************************** ****************** *** *********

WAYANAD1          TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
WAYANAD2          TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
NARAYANGAON1      TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
NARAYANGAON2      TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
TRICHY5           TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
TRICHY2           TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
TRICHY4           TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
TRICHY1           TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
TRICHY3           TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
NARAYANGAON3      TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
VAISHALI1         TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
NARAYANGAON4      TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
VAISHALI3         TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
VAISHALI2         TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
VAISHALI4         TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
VAISHALI5         TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
NARAYANGAON5      TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
JALGAON1          TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
JALGAON5          TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
JALGAON4          TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
JALGAON2          TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
JALGAON3          TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
WAYANAD3          TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTATCATCTTCCATTGAATGACTC 780
WAYANAD4          TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTACCATCTTCCATTGAATGACTC 780
WAYANAD5          TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTACCATCTTCCATTGAATGACTC 780
ASSAM2            TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTACCATCTTCCATTGAATGACTC 780
ASSAM3            TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTACCATCTTCCATTGAATGACTC 780
ASSAM4            TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTACCATCTTCCATTGAATGACTC 780
ASSAM5            TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTACCATCTTCCATTGAATGACTC 780
Assam1            TTTTCTGTAAAACGAATAAATCTTTCAGGAATAAACCTACCATCTTCCATTGAATGACTC 780
                  ************************************ ******************

WAYANAD1          CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATT 835
WAYANAD2          CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATT 835
NARAYANGAON1      CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATT 835
NARAYANGAON2      CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATT 835
TRICHY5           CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATT 835
TRICHY2           CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATT 835
TRICHY4           CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATT 835
TRICHY1           CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATT 835
TRICHY3           CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATT 835
NARAYANGAON3      CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATT 835
```

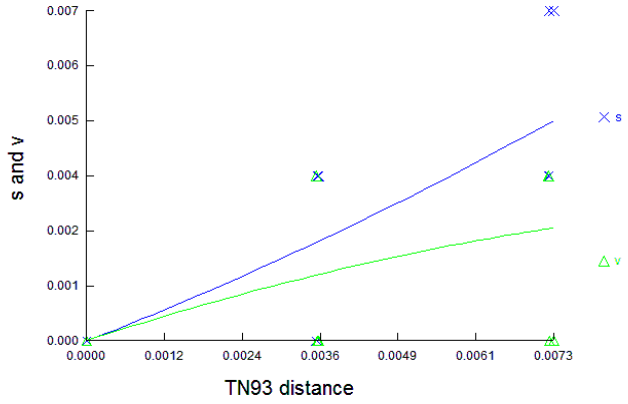
```
VAISHALI1          CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATT 835
NARAYANGAON4       CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATT 835
VAISHALI3          CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATT 835
VAISHALI2          CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATT 835
VAISHALI4          CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATT 835
VAISHALI5          CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATT 835
NARAYANGAON5       CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATT 835
JALGAON1           CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATT 835
JALGAON5           CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATT 835
JALGAON4           CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATT 835
JALGAON2           CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATT 835
JALGAON3           CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATT 835
WAYANAD3           CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATT 835
WAYANAD4           CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATT 835
WAYANAD5           CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATT 835
ASSAM2             CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATT 835
ASSAM3             CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATT 835
ASSAM4             CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATT 835
ASSAM5             CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATT 835
Assam1             CAATTTTTCCCTCCTGCAGACCATAGATTCCTTGAATTACCAATTATCACAAATT 835
                   ********************************************************
```

Fig 7. ClustalW alignment of the partial COI nucleotide sequences of the thirty *O. longicollis* (Oliver) individuals

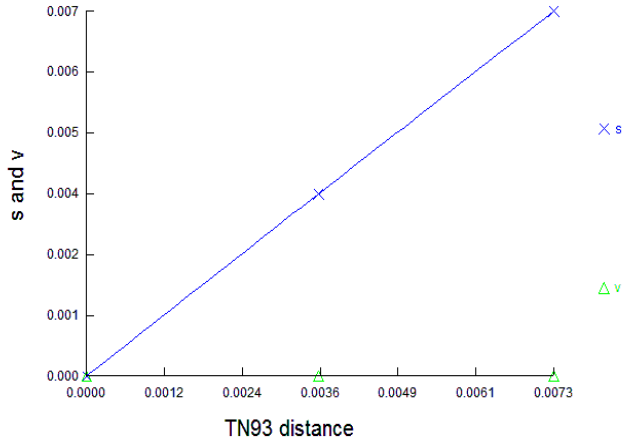
Nucleotide diversity (π) for COI was 0.005098 (Std Dev 0.002797) and no indels were present in the alignment. The overall base composition was A+T rich with the first position of the codons showing almost 85% AT (**Table 3**). The substitution saturation at each position of the codons was determined by Xia's test (Xia *et al.* 2003; Xia *et al.* 2009) and by plotting the Ts and Tv versus genetic distance (**Figs. 8 a-c and Table 6**). The Iss values in each case were significantly less than the Iss.c value suggesting that the substitutions have not reached saturation.



(a)



(b)



(c)

Figs. 8. Saturation analysis of the three codon positions of the partial COI gene from the thirty *O. longicollis (Oliver)* individuals. Ts and TV are plotted against the pairwise sequence divergences. a - 1st codon position; b - 2nd codon position; c - 3rd codon position

The partial COI gene translated into a protein of 278 amino acids using the invertebrate translation code. The ClustalW protein alignment of the thirty individuals is shown in **Fig. 9**.

```
WAYANAD1          WFFGHPEVYILILPGFGMISHIISQESGKKEAFGVLGMIYAMMAIGLLGFVVWAHHMFTV 60
WAYANAD2          WFFGHPEVYILILPGFGMISHIISQESGKKEAFGVLGMIYAMMAIGLLGFVVWAHHMFTV 60
VAISHALI3         WFFGHPEVYILILPGFGMISHIISQESGKKEAFGVLGMIYAMMAIGLLGFVVWAHHMFTV 60
TRICHY1           WFFGHPEVYILILPGFGMISHIISQESGKKEAFGVLGMIYAMMAIGLLGFVVWAHHMFTV 60
TRICHY2           WFFGHPEVYILILPGFGMISHIISQESGKKEAFGVLGMIYAMMAIGLLGFVVWAHHMFTV 60
TRICHY3           WFFGHPEVYILILPGFGMISHIISQESGKKEAFGVLGMIYAMMAIGLLGFVVWAHHMFTV 60
TRICHY4           WFFGHPEVYILILPGFGMISHIISQESGKKEAFGVLGMIYAMMAIGLLGFVVWAHHMFTV 60
JALGAON2          WFFGHPEVYILILPGFGMISHIISQESGKKEAFGVLGMIYAMMAIGLLGFVVWAHHMFTV 60
VAISHALI5         WFFGHPEVYILILPGFGMISHIISQESGKKEAFGVLGMIYAMMAIGLLGFVVWAHHMFTV 60
VAISHALI4         WFFGHPEVYILILPGFGMISHIISQESGKKEAFGVLGMIYAMMAIGLLGFVVWAHHMFTV 60
VAISHALI2         WFFGHPEVYILILPGFGMISHIISQESGKKEAFGVLGMIYAMMAIGLLGFVVWAHHMFTV 60
VAISHALI1         WFFGHPEVYILILPGFGMISHIISQESGKKEAFGVLGMIYAMMAIGLLGFVVWAHHMFTV 60
TRICHY5           WFFGHPEVYILILPGFGMISHIISQESGKKEAFGVLGMIYAMMAIGLLGFVVWAHHMFTV 60
NARAYANGAON5      WFFGHPEVYILILPGFGMISHIISQESGKKEAFGVLGMIYAMMAIGLLGFVVWAHHMFTV 60
JALGAON1          WFFGHPEVYILILPGFGMISHIISQESGKKEAFGVLGMIYAMMAIGLLGFVVWAHHMFTV 60
JALGAON4          WFFGHPEVYILILPGFGMISHIISQESGKKEAFGVLGMIYAMMAIGLLGFVVWAHHMFTV 60
JALGAON5          WFFGHPEVYILILPGFGMISHIISQESGKKEAFGVLGMIYAMMAIGLLGFVVWAHHMFTV 60
NARAYANGAON2      WFFGHPEVYILILPGFGMISHIISQESGKKEAFGVLGMIYAMMAIGLLGFVVWAHHMFTV 60
NARAYANGAON4      WFFGHPEVYILILPGFGMISHIISQESGKKEAFGVLGMIYAMMAIGLLGFVVWAHHMFTV 60
NARAYANGAON3      WFFGHPEVYILILPGFGMISHIISQESGKKEAFGVLGMIYAMMAIGLLGFVVWAHHMFTV 60
NARAYANGAON1      WFFGHPEVYILILPGFGMISHIISQESGKKEAFGVLGMIYAMMAIGLLGFVVWAHHMFTV 60
JALGAON3          WFFGHPEVYILILPGFGMISHIISQESGKKEAFGVLGMIYAMMAIGLLGFVVWAHHMFTV 60
Assam1            WFFGHPEVYILILPGFGMISHIISQESGKKEAFGVLGMIYAMMAIGLLGFVVWAHHMFTV 60
ASSAM2            WFFGHPEVYILILPGFGMISHIISQESGKKEAFGVLGMIYAMMAIGLLGFVVWAHHMFTV 60
ASSAM3            WFFGHPEVYILILPGFGMISHIISQESGKKEAFGVLGMIYAMMAIGLLGFVVWAHHMFTV 60
ASSAM4            WFFGHPEVYILILPGFGMISHIISQESGKKEAFGVLGMIYAMMAIGLLGFVVWAHHMFTV 60
ASSAM5            WFFGHPEVYILILPGFGMISHIISQESGKKEAFGVLGMIYAMMAIGLLGFVVWAHHMFTV 60
WAYANAD4          WFFGHPEVYILILPGFGMISHIISQESGKKEAFGVLGMIYAMMAIGLLGFVVWAHHMFTV 60
WAYANAD5          WFFGHPEVYILILPGFGMISHIISQESGKKEAFGVLGMIYAMMAIGLLGFVVWAHHMFTV 60
WAYANAD3          WFFGHPEVYILILPGFGMISHIISQESGKKEAFGVLGMIYAMMAIGLLGFVVWAHHMFTV 60
                  ************************************************************

WAYANAD1          GMDVDTRAYFTSATMITAVPTGIKIFSWLATYHGARITYSPTSLWSLGFIFLFTMGGLTG 120
WAYANAD2          GMDVDTRAYFTSATMITAVPTGIKIFSWLATYHGARITYSPTSLWSLGFIFLFTMGGLTG 120
VAISHALI3         GMDVDTRAYSTSATMIIAVPTGIKIFSWLATYHGARITYSPTSLWSLGFIFLFTMGGLTG 120
TRICHY1           GMDVDTRAYFTSATMIIAVPTGIKIFSWLATYHGARITYSPTSLWSLGFIFLFTMGGLTG 120
TRICHY2           GMDVDTRAYFTSATMIIAVPTGIKIFSWLATYHGARITYSPTSLWSLGFIFLFTMGGLTG 120
TRICHY3           GMDVDTRAYFTSATMIIAVPTGIKIFSWLATYHGARITYSPTSLWSLGFIFLFTMGGLTG 120
TRICHY4           GMDVDTRAYFTSATMIIAVPTGIKIFSWLATYHGARITYSPTSLWSLGFIFLFTMGGLTG 120
JALGAON2          GMDVDTRAYFTSATMIIAVPTGIKIFSWLATYHGARITYSPTSLWSLGFIFLFTMGGLTG 120
VAISHALI5         GMDVDTRAYFTSATMIIAVPTGIKIFSWLATYHGARITYSPTSLWSLGFIFLFTMGGLTG 120
VAISHALI4         GMDVDTRAYFTSATMIIAVPTGIKIFSWLATYHGARITYSPTSLWSLGFIFLFTMGGLTG 120
VAISHALI2         GMDVDTRAYFTSATMIIAVPTGIKIFSWLATYHGARITYSPTSLWSLGFIFLFTMGGLTG 120
VAISHALI1         GMDVDTRAYFTSATMIIAVPTGIKIFSWLATYHGARITYSPTSLWSLGFIFLFTMGGLTG 120
TRICHY5           GMDVDTRAYFTSATMIIAVPTGIKIFSWLATYHGARITYSPTSLWSLGFIFLFTMGGLTG 120
NARAYANGAON5      GMDVDTRAYFTSATMIIAVPTGIKIFSWLATYHGARITYSPTSLWSLGFIFLFTMGGLTG 120
JALGAON1          GMDVDTRAYFTSATMIIAVPTGIKIFSWLATYHGARITYSPTSLWSLGFIFLFTMGGLTG 120
JALGAON4          GMDVDTRAYFTSATMIIAVPTGIKIFSWLATYHGARITYSPTSLWSLGFIFLFTMGGLTG 120
JALGAON5          GMDVDTRAYFTSATMIIAVPTGIKIFSWLATYHGARITYSPTSLWSLGFIFLFTMGGLTG 120
NARAYANGAON2      GMDVDTRAYFTSATMIIAVPTGIKIFSWLATYHGARITYSPTSLWSLGFIFLFTMGGLTG 120
NARAYANGAON4      GMDVDTRAYFTSATMIIAVPTGIKIFSWLATYHGARITYSPTSLWSLGFIFLFTMGGLTG 120
NARAYANGAON3      GMDVDTRAYFTSATMIIAVPTGIKIFSWLATYHGARITYSPTSLWSLGFIFLFTMGGLTG 120
NARAYANGAON1      GMDVDTRAYFTSATMIIAVPTGIKIFSWLATYHGARITYSPTSLWSLGFIFLFTMGGLTG 120
JALGAON3          GMDVDTRAYFTSATMIIAVPTGIKIFSWLATYHGARITYSPTSLWSLGFIFLFTMGGLTG 120
Assam1            GMDVDTRAYFTSATMIIAVPTGIKIFSWLATYHGARITYSPTSLWSLGFIFLFTMGGLTG 120
ASSAM2            GMDVDTRAYFTSATMIIAVPTGIKIFSWLATYHGARITYSPTSLWSLGFIFLFTMGGLTG 120
ASSAM3            GMDVDTRAYFTSATMIIAVPTGIKIFSWLATYHGARITYSPTSLWSLGFIFLFTMGGLTG 120
ASSAM4            GMDVDTRAYFTSATMIIAVPTGIKIFSWLATYHGARITYSPTSLWSLGFIFLFTMGGLTG 120
ASSAM5            GMDVDTRAYFTSATMIIAVPTGIKIFSWLATYHGARITYSPTSLWSLGFIFLFTMGGLTG 120
WAYANAD4          GMDVDTRAYFTSATMIIAVPTGIKIFSWLATYHGARITYSPTSLWSLGFIFLFTMGGLTG 120
WAYANAD5          GMDVDTRAYFTSATMIIAVPTGIKIFSWLATYHGARITYSPTSLWSLGFIFLFTMGGLTG 120
WAYANAD3          GMDVDTRAYFTSATMIIAVPTGIKIFSWLATYHGARITYSPTSLWSLGFIFPFTMGGLTG 120
```

```
                 ********* ****** ********************************** ********

WAYANAD1         VVLANSSIDIILHDTYYVVAHFHYVLSMGAVFAIIAGFIQWFPLFTGLTLNQKYLKIQFF 180
WAYANAD2         VVLANSSIDIILHDTYYVVAHFHYVLSMGAVFAIIAGFIQWFPLFTGLTLNQKYLKIQFF 180
VAISHALI3        VVLANSSIDIILHDTYYVVAHFHYVLSMGAVFAIIAGFIQWFPLFTGLTLNQKYLKIQFF 180
TRICHY1          VVLANASIDIILHDTYYVVAHFHYVLSMGAVFAIIAGFIQWFPLFTGLTLNQKYLKIQFF 180
TRICHY2          VVLANASIDIILHDTYYVVAHFHYVLSMGAVFAIIAGFIQWFPLFTGLTLNQKYLKIQFF 180
TRICHY3          VVLANASIDIILHDTYYVVAHFHYVLSMGAVFAIIAGFIQWFPLFTGLTLNQKYLKIQFF 180
TRICHY4          VVLANASIDIILHDTYYVVAHFHYVLSMGAVFAIIAGFIQWFPLFTGLTLNQKYLKIQFF 180
JALGAON2         VVLANSSIDIILHDTYYVVAHFHYVLSMGAVFAIIAGFIQWFPLFTGLTLNQKYLKIQFF 180
VAISHALI5        VVLANSSIDIILHDTYYVVAHFHYVLSMGAVFAIIAGFIQWFPLFTGLTLNQKYLKIQFF 180
VAISHALI4        VVLANSSIDIILHDTYYVVAHFHYVLSMGAVFAIIAGFIQWFPLFTGLTLNQKYLKIQFF 180
VAISHALI2        VVLANSSIDIILHDTYYVVAHFHYVLSMGAVFAIIAGFIQWFPLFTGLTLNQKYLKIQFF 180
VAISHALI1        VVLANSSIDIILHDTYYVVAHFHYVLSMGAVFAIIAGFIQWFPLFTGLTLNQKYLKIQFF 180
TRICHY5          VVLANSSIDIILHDTYYVVAHFHYVLSMGAVFAIIAGFIQWFPLFTGLTLNQKYLKIQFF 180
NARAYANGAON5     VVLANSSIDIILHDTYYVVAHFHYVLSMGAVFAIIAGFIQWFPLFTGLTLNQKYLKIQFF 180
JALGAON1         VVLANSSIDIILHDTYYVVAHFHYVLSMGAVFAIIAGFIQWFPLFTGLTLNQKYLKIQFF 180
JALGAON4         VVLANSSIDIILHDTYYVVAHFHYVLSMGAVFAIIAGFIQWFPLLTGLTLNQKYLKIQFF 180
JALGAON5         VVLANSSIDIILHDTYYVVAHFHYVLSMGAVFAIIAGFIQWFPLFTGLTLNQKYLKIQFF 180
NARAYANGAON2     VVLANSSIDIILHDTYYVVAHFHYVLSMGAVFAIIAGFIQWFPLFTGLTLNQKYLKIQFF 180
NARAYANGAON4     VVLANSSIDIILHDTYYVVAHFHYVLSMGAVFAIIAGFIQWFPLFTGLTLNQKYLKIQFF 180
NARAYANGAON3     VVLANSSIDIILHDTYYVVAHFHYVLSMGAVFAIIAGFIQWFPLFTGLTLNQKYLKIQFF 180
NARAYANGAON1     VVLANSSIDIILHDTYYVVAHFHYVLSMGAVFAIIAGFIQWFPLFTGLTLNQKYLKIQFF 180
JALGAON3         VVLANSSIDIILHDTYYVVAHFHYVLSMGAVFAIIAGFIQWFPLFTGLTLNQKYLKIQFF 180
Assam1           VVLANSSIDIILHDTYYVVAHFHYVLSMGAVFAIIAGFIQWFPLFTGLTLNQKYLKIQFF 180
ASSAM2           VVLANSSIDIILHDTYYVVAHFHYVLSMGAVLAIIAGFIQWFPLFTGLTLNQKYLKIQFF 180
ASSAM3           VVLANSSIDIILHDTYYVVAHFHYVLSMGAVFAIIAGFIQWFPLFTGLTLNQKYLKIQFF 180
ASSAM4           VVLANSSIDIILHDTYYVVAHFHYVLSMGAVFAIIAGFIQWFPLFTGLTLNQKYLKIQFF 180
ASSAM5           VVLANSSIDIILHDTYYVVAHFHYVLSMGAVFAIIAGFIQWFPLFTGLTLNQKYLKIQFF 180
WAYANAD4         VVLANSSIDIILHDTYYVVAHFHYVLSMGAVFAIIAGFIQWFPLFTGLTLNQKYLKIQFF 180
WAYANAD5         VVLANSSIDIILHDTYYVVAHFHYVLSMGAVFAIIAGFIQWFPLFTGLTLNQKYLKIQFF 180
WAYANAD3         VVLANSSIDIILHDTYYVVAHFHYVLSMGAVFAIIAGFIQWFPLFTGLTLNQKYLKIQFF 180
                 *****:************************:************:***************

WAYANAD1         SMFLGVNLTFFPQHFLGLSGMPRRYSDYPDAYYMWNSISSIGSLISLISILYLIFIIWEA 240
WAYANAD2         SMFLGVNLTFFPQHFLGLSGMPRRYSDYPDAYYMWNSISSIGSLISLISILYLIFIIWEA 240
VAISHALI3        SMFLGVNLTFFPQHFLGLSGMPRRYSDYPDAYYMWNSISSIGSLISLISILYLIFIIWEA 240
TRICHY1          SMFLGVNLTFFPQHFLGLSGMPRRYSDYPDAYYMWNSISSIGSLISLISILYLIFIIWEA 240
TRICHY2          SMFLGVNLTFFPQHFLGLSGMPRRYSDYPDAYYMWNSISSIGSLISLISILYLIFIIWEA 240
TRICHY3          SMFLGVNLTFFPQHFLGLSGMPRRYSDYPDAYYMWNSISSIGSLISLISILYLIFIIWEA 240
TRICHY4          SMFLGVNLTFFPQHFLGLSGMPRRYSDYPDAYYMWNSISSIGSLISLISILYLIFIIWEA 240
JALGAON2         SMFLGVNLTFFPQHFLGLSGMPRRYSDYPDAYYMWNSISSVGSLISLISILYLIFIIWEA 240
VAISHALI5        SMFLGVNLTFFPQHFLGLSGMPRRYSDYPDAYYMWNSISSIGSLISLISILYLIFIIWEA 240
VAISHALI4        SMFLGVNLTFFPQHFLGLSGMPRRYSDYPDAYYMWNSISSIGSLISLISILYLIFIIWEA 240
VAISHALI2        SMFLGVNLTFFPQHFLGLSGMPRRYSDYPDAYYMWNSISSIGSLISLISILYLIFIIWEA 240
VAISHALI1        SMFLGVNLTFFPQHFLGLSGMPRRYSDYPDAYYMWNSISSIGSLISLISILYLIFIIWEA 240
TRICHY5          SMFLGVNLTFFPQHFLGLSGMPRRYSDYPDAYYMWNSISSIGSLISLISILYLIFIIWEA 240
NARAYANGAON5     SMFLGVNLTFFPQHFLGLSGMPRRYSDYPDAYYMWNSISSIGSLISLISILYLIFIIWEA 240
JALGAON1         SMFLGVNLTFFPQHFLGLSGMPRRYSDYPDAYYMWNSISSIGSLISLISILYLIFTIWEA 240
JALGAON4         SMFLGVNLTFFPQHFLGLSGMPRRYSDYPDAYYMWNSISSIGSLISLISILYLIFTIWEA 240
JALGAON5         SMFLGVNLTFFPQHFLGLSGMPRRYSDYPDAYYMWNSISSIGSLISLISILYLIFTIWEA 240
NARAYANGAON2     SMFLGVNLTFFPQHFLGLSGMPRRYSDYPDAYYMWNSISSIGSLISLISILYLIFIIWEA 240
NARAYANGAON4     SMFLGVNLTFFPQHFLGLSGMPRRYSDYPDAYYMWNSISSIGSLISLISILYLIFIIWEA 240
NARAYANGAON3     SMFLGVNLTFFPQHFLGLSGMPRRYSDYPDAYYMWNSISSIGSLISLISILYLIFIIWEA 240
NARAYANGAON1     SMFLGVNLTFFPQHFLGLSGMPRRYSDYPDAYYMWNSISSIGSLISLISILYLIFIIWEA 240
JALGAON3         SMFLGVNLTFFPQHFLGLSGMPRRYSDYPDAYYMWNSISSIGSLISLISILYLIFIIWEA 240
Assam1           SMFLGVNLTFFPQHFLGLSGMPRRYSDYPDAYYMWNSISSIGSLISLISILYLIFIIWEA 240
ASSAM2           SMFLGVNLTFFPQHFLGLSGMPRRYSDYPDAYYMWNSISSIGSLISLISILYLIFIIWEA 240
ASSAM3           SMFLGVNLTFFPQHFLGLSGMPRRYSDYPDAYYMWNSISSIGSLISLISILYLIFIIWEA 240
ASSAM4           SMFLGVNLTFFPQHFLGLSGMPRRYSDYPDAYYMWNSISSIGSLISLISILYLIFIIWEA 240
ASSAM5           SMFLGVNLTFFPQHFLGLSGMPRRYSDYPDAYYMWNSISSIGSLISLISILYLIFIIWEA 240
WAYANAD4         SMFLGVNLTFFPQHFLGLSGMPRRYSDYPDAYYMWNSISSIGSLISLISVLYLIFIIWEA 240
WAYANAD5         SMFLGVNLTFFPQHFLGLSGMPRRYSDYPDAYYMWNSISSIGSLISLISVLYLIFIIWEA 240
WAYANAD3         SMFLGVNLTFFPQHFLGLSGMPRRYSDYPDAYYMWNSISSIGSLISLISVLYLIFIIWEA 240
                 ****************************************:********:***** ****

WAYANAD1         FSVKRMNLSGMNLSSSIEWLQFFPPADHSFLELPIITN 278
WAYANAD2         FSVKRMNLSGMNLSSSIEWLQFFPPADHSFLELPIITN 278
VAISHALI3        FSVKRMNLSGMNLSSSIEWLQFFPPADHSFLELPIITN 278
TRICHY1          FSVKRMNLSGMNLSSSIEWLQFFPPADHSFLELPIITN 278
TRICHY2          FSVKRMNLSGMNLSSSIEWLQFFPPADHSFLELPIITN 278
```

```
TRICHY3           FSVKRMNLSGMNLSSSIEWLQFFPPADHSFLELPIITN 278
TRICHY4           FSVKRMNLSGMNLSSSIEWLQFFPPADHSFLELPIITN 278
JALGAON2          FSVKRMNLSGMNLSSSIEWLQFFPPADHSFLELPIITN 278
VAISHALI5         FSVKRMNLSGMNLSSSIEWLQFFPPADHSFLELPIITN 278
VAISHALI4         FSVKRMNLSGMNLSSSIEWLQFFPPADHSFLELPIITN 278
VAISHALI2         FSVKRMNLSGMNLSSSIEWLQFFPPADHSFLELPIITN 278
VAISHALI1         FSVKRMNLSGMNLSSSIEWLQFFPPADHSFLELPIITN 278
TRICHY5           FSVKRMNLSGMNLSSSIEWLQFFPPADHSFLELPIITN 278
NARAYANGAON5      FSVKRMNLSGMNLSSSIEWLQFFPPADHSFLELPIITN 278
JALGAON1          FSVKRMNLSGMNLSSSIEWLQFFPPADHSFLELPIITN 278
JALGAON4          FSVKRMNLSGMNLSSSIEWLQFFPPADHSFLELPIITN 278
JALGAON5          FSVKRMNLSGMNLSSSIEWLQFFPPADHSFLELPIITN 278
NARAYANGAON2      FSVKRMNLSGMNLSSSIEWLQFFPPADHSFLELPIITN 278
NARAYANGAON4      FSVKRMNLSGMNLSSSIEWLQFFPPADHSFLELPIITN 278
NARAYANGAON3      FSVKRMNLSGMNLSSSIEWLQFFPPADHSFLELPIITN 278
NARAYANGAON1      FSVKRMNLSGMNLSSSIEWLQFFPPADHSFLELPIITN 278
JALGAON3          FSVKRMNLSGMNLSSSIEWLQFFPPADHSFLELPIITN 278
Assam1            FSVKRMNLSGMNLPSSIEWLQFFPPADHSFLELPIITN 278
ASSAM2            FSVKRMNLSGMNLPSSIEWLQFFPPADHSFLELPIITN 278
ASSAM3            FSVKRMNLSGMNLPSSIEWLQFFPPADHSFLELPIITN 278
ASSAM4            FSVKRMNLSGMNLPSSIEWLQFFPPADHSFLELPIITN 278
ASSAM5            FSVKRMNLSGMNLPSSIEWLQFFPPADHSFLELPIITN 278
WAYANAD4          FSVKRMNLSGMNLPSSIEWLQFFPPADHSFLELPIITN 278
WAYANAD5          FSVKRMNLSGMNLPSSIEWLQFFPPADHSFLELPIITN 278
WAYANAD3          FSVKRMNLSGMNLSSSIEWLQFFPPADHSFLELPIITN 278
                  ************.**********************
```

Fig. 9. ClustalW alignment of the partial translated COI region of the thirty *O. longicollis* (Oliver) individuals

The ClustalW alignment of the partial COI protein of one representative individual *i.e.* Assam 1 with the COI protein of a few insects *i.e. Sphenophorus* sp. BYU-CO246m, *Chaetosoma scaritides*, *Tribolium castaneum*, *Priasilpha obscura, Cyphon* sp. BT0012, *Locusta migratoria*, *Drosophila yakuba*, *A.gambiae* and *Apis mellifera ligustica* identified thirteen highly conserved structurally important regions ie external loops E5 and E6, M6-M12 and internal loops I3-15 in the partial COI protein sequences of the thirty individuals (**Fig. 10**).

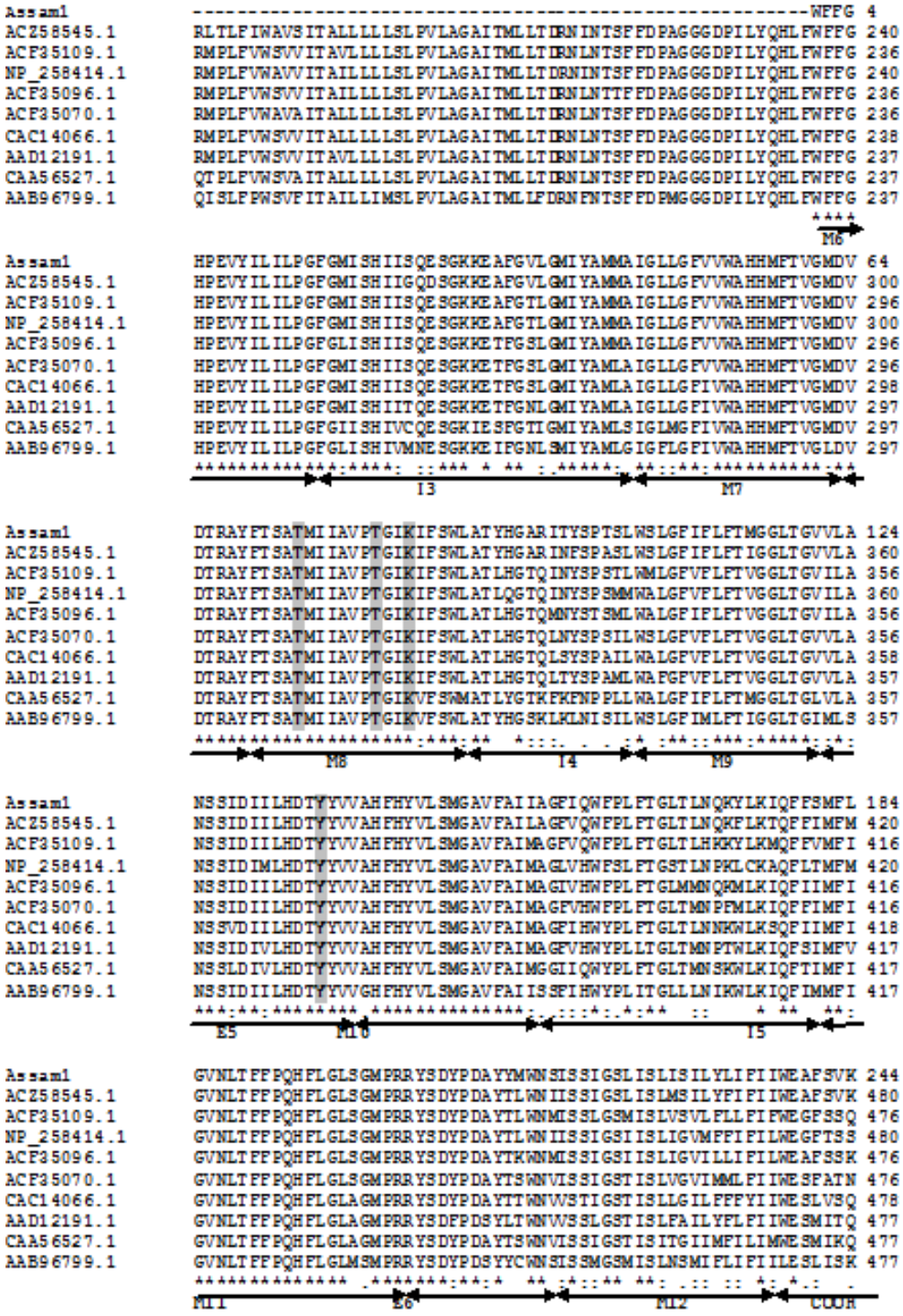


Fig. 10. ClustalW alignment of the partial COI sequence of Assam (individual 1) with the COI protein of a few insects. *Sphenophorus* sp. BYU-CO246 (ACZ58545.1), *Chaetosoma scaritides* (ACF35096.1), *Tribolium castaneum* (NP_258414.1), *Priasilpha obscura* (ACF35109.1), *Cyphon* sp. BT0012 (ACF35070.1), *Locusta migratoria* (CAA56527.1), *Drosophila yakuba* (CAC14066.1), *A. gambiae* (AAD12191.1) and *Apis*

mellifera ligustica (AAB96799.1). Protein ID of the COI protein is indicated in parenthesis. The different structural regions are shown by double headed arrows. Functionally significant residues are in bold and shaded in grey. and are described in the text. The numbering of the residues as described in the text are with respect to the COI protein of *D. yakuba*.)

A Codon Bias Index (CBI) of 0.362 was observed for COI and the RSCU values are shown in **Table 11.**

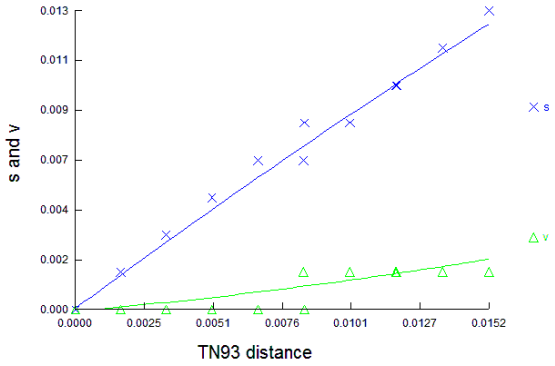
Table 11. Codon usage bias in the partial COI protein of the thirty weevils of *O. longicollis* (Oliver)

Codon	Count	RSCU	Codon	Count	RSCU	Codon	Count	RSCU	Codon	Count	RSCU
UUU(F)	22	1.42	UCU(S)	5	0.68	UAU(Y)	19.5	1.26	UGU(C)	5.9	1.07
UUC(F)	8.9	0.58	UCC(S)	9	1.22	UAC(Y)	11.4	0.74	UGC(C)	5.1	0.93
UUA(L)	9	1.86	UCA(S)	7.9	1.08	UAA(*)	3	0.86	UGA(W)	3	0.5
UUG(L)	3.2	0.65	UCG(S)	1	0.14	UAG(*)	4	1.14	UGG(W)	9	1.5
CUU(L)	8	1.66	CCU(P)	1	0.49	CAU(H)	2.1	1.97	CGU(R)	0.2	0.67
CUC(L)	3	0.62	CCC(P)	0.1	0.03	CAC(H)	0	0.03	CGC(R)	1	**3.33**
CUA(L)	2	0.41	CCA(P)	6.1	**2.98**	CAA(Q)	3	1.2	CGA(R)	0	0
CUG(L)	3.8	0.79	CCG(P)	1	0.49	CAG(Q)	2	0.8	CGG(R)	0	0
AUU(I)	13	1.63	ACU(T)	2	0.81	AAU(N)	21.9	1.63	AGU(S)	8	1.09
AUC(I)	3	0.37	ACC(T)	4	1.59	AAC(N)	5	0.37	AGC(S)	9	1.22
AUA(M)	5	1.11	ACA(T)	2	0.8	AAA(K)	7.9	1.45	AGA(S)	6	0.82
AUG(M)	4	0.89	ACG(T)	2	0.8	AAG(K)	3	0.55	AGG(S)	13	1.76
GUU(V)	1	**3.87**	GCU(A)	1	4	GAU(D)	2	1.33	GGU(G)	0	0
GUC(V)	0	0.13	GCC(A)	0	0	GAC(D)	1	0.67	GGC(G)	1	**2.03**
GUA(V)	0	0	GCA(A)	0	0	GAA(E)	2.1	2	GGA(G)	1	1.97
GUG(V)	0	0	GCG(A)	0	0	GAG(E)	0	0	GGG(G)	0	0

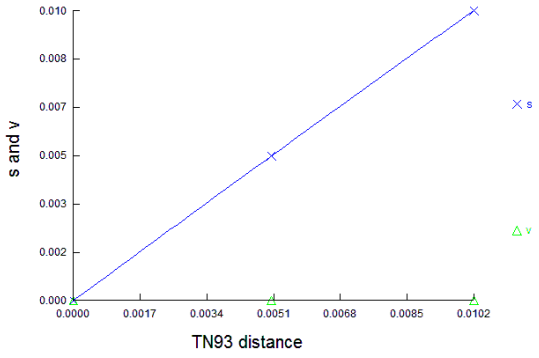
(Relatively higher values of RSCU are shown in bold)

Analysis of the COII region

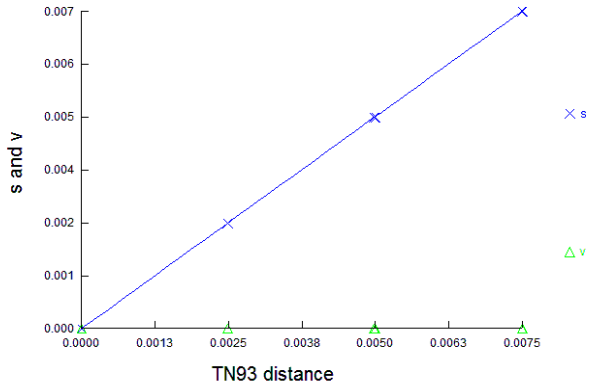
602 bp of the 5' region of the COII gene was amplified by the M1-M2 primers. Nucleotide diversity (π) for COII was 0.005591 (Std Dev 0.003174). Indels were absent in the aligned sequences**.** The sequence has a high proportion (70.6%) of A+T, with the third codon position showing almost 86% A+T **(Table 3).** The plots of Ts and Tv versus genetic distance for each codon position and Xia's test for saturation at each codon position indicated that the nucleotide substitutions at each of the codon positions has not reached saturation. (**Figs. 11 a-c and Table 6**).



(a)



(b)



(c)

Fig. 11. Saturation analysis of the three codon positions of the partial COII gene from the thirty *O. longicollis (Oliver)* individuals. Ts and TV are plotted against the pairwise sequence divergences. a - 1st codon position; b - 2nd codon position; c - 3rd codon position

The ClustalW alignment of the partial COII gene of the thirty individuals shows twelve base substitutions of which three, three and six occur in the first, second and third codon positions (**Fig. 12 and Table 12).**

Table 12. Details of nucleotide changes in the partial COII region of the thirty individuals of *O. longicollis* (Oliver)

Serial no	Position of nucleotide	Position of amino acid	Type of nucleotide change	Change in codon	Change in codon position	Amino acid change	Type of change	Individuals	Structural region of the COII protein
1	44	15	Transition	CCT→CTT	2nd	P→L	Non synonymous	Trichy 2, 4, 5	Intermembrane
2	66	23	Transition	TTT→TTC	3rd	F	Synonymous	Assam 2, 3, 4, 5, Wayanad 1, 2, 3, 4, 5	Intermembrane
3	132	45	Transition	ATA→GTA	1st	M→V	Conserved substitution	Jalgaon 4	M1
4	155	52	Transition	CAT→CAC	3rd	H	Synonymous	Assam 2, 3, 4, 5, Wayanad 1, 2, 3, 4, 5	M1
5	179	60	Transition	ATA→ACA	2nd	M→T	Non synonymous	Wayanad 1, 2	I1
6	234	78	Transversion	CTA→CTT	3rd	L	Synonymous	Assam 2, 3, 4, 5, Wayanad 1, 2, 3, 4, 5	M2
7	276	92	Transition	CCC→TCC	1st	P	Synonymous	Assam 2, 3, 4, 5,	Intermembrane

								Wayanad 1, 2, 3, 4, 5	
8	306	103	Transition	CAT→CAC	3rd	H	Synonymous	Assam 2, 3, 4, 5, Wayanad 1, 2, 3, 4, 5	Periplasmic region
9	378	126	Transition	AAC→AAT	3rd	N	Synonymous	Wayanad 1, 2, 3	Periplasmic region
10	405	135	Transition	TTA→TTG	3rd	L	Synonymous	Assam 2, 3, 4, 5, Wayanad 1, 2, 3, 4, 5	Periplasmic region
11	430	144	Transition	ATC→GTC	1st	L→V	Non synonymous	Assam 2, 3, 4, 5, Wayanad 4, 5	Periplasmic region
12	467	156	Transition	TCT→TTT	2nd	S→F	Non synonymous	Jalgaon 4	Periplasmic region

```
VAISHALI4         ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 60
VAISHALI5         ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 60
VAISHALI3         ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 60
VAISHALI2         ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 60
VAISHALI1         ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 60
TRICHY3           ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 60
TRICHY1           ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 60
NARAYANGAON5      ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 60
NARAYANGAON4      ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 60
NARAYANGAON3      ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 60
NARAYANGAON2      ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 60
NARAYANGAON1      ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 60
JALGAON5          ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 60
JALGAON3          ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 60
JALGAON2          ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 60
JALGAON1          ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 60
Assam1            ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 60
TRICHY2           ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCTTCTTATAGAACAACTT 60
TRICHY4           ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCTTCTTATAGAACAACTT 60
TRICHY5           ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCTTCTTATAGAACAACTT 60
JALGAON4          ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 60
ASSAM2            ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 60
ASSAM3            ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 60
ASSAM4            ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 60
ASSAM5            ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 60
WAYANAD4          ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 60
WAYANAD5          ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 60
WAYANAD1          ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 60
WAYANAD2          ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 60
WAYANAD3          ATTGCAACATGAAAAACCTTAATACTCCAAGACAGCGCTTCTCCTCTTATAGAACAACTT 60
                  ************************************** ***************

VAISHALI4         CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 120
VAISHALI5         CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 120
VAISHALI3         CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 120
VAISHALI2         CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 120
VAISHALI1         CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 120
TRICHY3           CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 120
TRICHY1           CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 120
NARAYANGAON5      CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 120
NARAYANGAON4      CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 120
NARAYANGAON3      CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 120
NARAYANGAON2      CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 120
NARAYANGAON1      CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 120
JALGAON5          CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 120
JALGAON3          CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 120
JALGAON2          CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 120
JALGAON1          CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 120
Assam1            CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 120
```

```
TRICHY2          CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 120
TRICHY4          CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 120
TRICHY5          CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 120
JALGAON4         CTTTTTTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 120
ASSAM2           CTTTTCTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 120
ASSAM3           CTTTTCTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 120
ASSAM4           CTTTTCTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 120
ASSAM5           CTTTTCTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 120
WAYANAD4         CTTTTCTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 120
WAYANAD5         CTTTTCTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 120
WAYANAD1         CTTTTCTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 120
WAYANAD2         CTTTTCTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 120
WAYANAD3         CTTTTCTTCCATGATCACGCCTTATTAATTTTAGTTCTTATTACTATTCTAGTAGGACAA 120
                 ***** **************************************************

VAISHALI4        ATATTACTAAGAATATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 180
VAISHALI5        ATATTACTAAGAATATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 180
VAISHALI3        ATATTACTAAGAATATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 180
VAISHALI2        ATATTACTAAGAATATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 180
VAISHALI1        ATATTACTAAGAATATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 180
TRICHY3          ATATTACTAAGAATATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 180
TRICHY1          ATATTACTAAGAATATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 180
NARAYANGAON5     ATATTACTAAGAATATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 180
NARAYANGAON4     ATATTACTAAGAATATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 180
NARAYANGAON3     ATATTACTAAGAATATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 180
NARAYANGAON2     ATATTACTAAGAATATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 180
NARAYANGAON1     ATATTACTAAGAATATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 180
JALGAON5         ATATTACTAAGAATATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 180
JALGAON3         ATATTACTAAGAATATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 180
JALGAON2         ATATTACTAAGAATATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 180
JALGAON1         ATATTACTAAGAATATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 180
Assam1           ATATTACTAAGAATATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 180
TRICHY2          ATATTACTAAGAATATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 180
TRICHY4          ATATTACTAAGAATATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 180
TRICHY5          ATATTACTAAGAATATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 180
JALGAON4         ATATTACTAAGAGTATTATTCAATAAATTCACTCATCGTTTTCTTCTAGAAGGACAAATA 180
ASSAM2           ATATTACTAAGAATATTATTCAATAAATTCACTCACCGTTTTCTTCTAGAAGGACAAATA 180
ASSAM3           ATATTACTAAGAATATTATTCAATAAATTCACTCACCGTTTTCTTCTAGAAGGACAAATA 180
ASSAM4           ATATTACTAAGAATATTATTCAATAAATTCACTCACCGTTTTCTTCTAGAAGGACAAATA 180
ASSAM5           ATATTACTAAGAATATTATTCAATAAATTCACTCACCGTTTTCTTCTAGAAGGACAAATA 180
WAYANAD4         ATATTACTAAGAATATTATTCAATAAATTCACTCACCGTTTTCTTCTAGAAGGACAAATA 180
WAYANAD5         ATATTACTAAGAATATTATTCAATAAATTCACTCACCGTTTTCTTCTAGAAGGACAAATA 180
WAYANAD1         ATATTACTAAGAATATTATTCAATAAATTCACTCACCGTTTTCTTCTAGAAGGACAAACA 180
WAYANAD2         ATATTACTAAGAATATTATTCAATAAATTCACTCACCGTTTTCTTCTAGAAGGACAAACA 180
WAYANAD3         ATATTACTAAGAATATTATTCAATAAATTCACTCACCGTTTTCTTCTAGAAGGACAAATA 180
                 ************ ********************** ********************** *

VAISHALI4        ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 240
VAISHALI5        ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 240
VAISHALI3        ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 240
VAISHALI2        ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 240
VAISHALI1        ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 240
TRICHY3          ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 240
TRICHY1          ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 240
NARAYANGAON5     ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 240
NARAYANGAON4     ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 240
NARAYANGAON3     ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 240
NARAYANGAON2     ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 240
NARAYANGAON1     ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 240
JALGAON5         ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 240
JALGAON3         ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 240
JALGAON2         ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 240
JALGAON1         ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 240
Assam1           ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 240
TRICHY2          ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 240
TRICHY4          ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 240
TRICHY5          ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 240
JALGAON4         ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTACCATCA 240
ASSAM2           ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTTCCATCA 240
ASSAM3           ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTTCCATCA 240
ASSAM4           ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTTCCATCA 240
```

```
ASSAM5          ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTTCCATCA 240
WAYANAD4        ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTTCCATCA 240
WAYANAD5        ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTTCCATCA 240
WAYANAD1        ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTTCCATCA 240
WAYANAD2        ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTTCCATCA 240
WAYANAD3        ATCGAATTAATTTGAACTATTTTACCTGCAATTATTCTTATTCTAATTGCTCTTCCATCA 240
                ************************************************** ******

VAISHALI4       CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 300
VAISHALI5       CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 300
VAISHALI3       CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 300
VAISHALI2       CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 300
VAISHALI1       CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 300
TRICHY3         CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 300
TRICHY1         CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 300
NARAYANGAON5    CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 300
NARAYANGAON4    CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 300
NARAYANGAON3    CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 300
NARAYANGAON2    CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 300
NARAYANGAON1    CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 300
JALGAON5        CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 300
JALGAON3        CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 300
JALGAON2        CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 300
JALGAON1        CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 300
Assam1          CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 300
TRICHY2         CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 300
TRICHY4         CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 300
TRICHY5         CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 300
JALGAON4        CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 300
ASSAM2          CTTCGTTTATTATATATTTTAGATGAAATTAATAATCCCTCTATTACTATTAAAGCTATT 300
ASSAM3          CTTCGTTTATTATATATTTTAGATGAAATTAATAATCCCTCTATTACTATTAAAGCTATT 300
ASSAM4          CTTCGTTTATTATATATTTTAGATGAAATTAATAATCCCTCTATTACTATTAAAGCTATT 300
ASSAM5          CTTCGTTTATTATATATTTTAGATGAAATTAATAATCCCTCTATTACTATTAAAGCTATT 300
WAYANAD4        CTTCGTTTATTATATATTTTAGATGAAATTAATAATCCCTCTATTACTATTAAAGCTATT 300
WAYANAD5        CTTCGTTTATTATATATTTTAGATGAAATTAATAATCCCTCTATTACTATTAAAGCTATT 300
WAYANAD1        CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 300
WAYANAD2        CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 300
WAYANAD3        CTTCGTTTATTATATATTTTAGATGAAATTAATAACCCCTCTATTACTATTAAAGCTATT 300
                ********************************** ************************

VAISHALI4       GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 360
VAISHALI5       GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 360
VAISHALI3       GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 360
VAISHALI2       GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 360
VAISHALI1       GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 360
TRICHY3         GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 360
TRICHY1         GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 360
NARAYANGAON5    GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 360
NARAYANGAON4    GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 360
NARAYANGAON3    GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 360
NARAYANGAON2    GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 360
NARAYANGAON1    GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 360
JALGAON5        GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 360
JALGAON3        GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 360
JALGAON2        GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 360
JALGAON1        GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 360
Assam1          GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 360
TRICHY2         GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 360
TRICHY4         GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 360
TRICHY5         GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 360
JALGAON4        GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 360
ASSAM2          GGCCACCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 360
ASSAM3          GGCCACCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 360
ASSAM4          GGCCACCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 360
ASSAM5          GGCCACCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 360
WAYANAD4        GGCCACCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 360
WAYANAD5        GGCCACCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 360
WAYANAD1        GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 360
WAYANAD2        GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 360
WAYANAD3        GGCCATCAATGATACTGATCATATGAATATTCTGATTACAAAAATATTGAATTTGATTCT 360
                ***** ************************************************** 
```

```
VAISHALI4         TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 420
VAISHALI5         TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 420
VAISHALI3         TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 420
VAISHALI2         TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 420
VAISHALI1         TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 420
TRICHY3           TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 420
TRICHY1           TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 420
NARAYANGAON5      TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 420
NARAYANGAON4      TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 420
NARAYANGAON3      TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 420
NARAYANGAON2      TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 420
NARAYANGAON1      TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 420
JALGAON5          TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 420
JALGAON3          TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 420
JALGAON2          TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 420
JALGAON1          TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 420
Assam1            TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 420
TRICHY2           TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 420
TRICHY4           TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 420
TRICHY5           TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 420
JALGAON4          TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTGTTAGATGTAGATAAC 420
ASSAM2            TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTATTAGATGTAGATAAC 420
ASSAM3            TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTATTAGATGTAGATAAC 420
ASSAM4            TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTATTAGATGTAGATAAC 420
ASSAM5            TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTATTAGATGTAGATAAC 420
WAYANAD4          TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTATTAGATGTAGATAAC 420
WAYANAD5          TATATAATTCCTACTAACGACTTAAATACCTATAATTTCCGTTTATTAGATGTAGATAAC 420
WAYANAD1          TATATAATTCCTACTAATGACTTAAATACCTATAATTTCCGTTTATTAGATGTAGATAAC 420
WAYANAD2          TATATAATTCCTACTAATGACTTAAATACCTATAATTTCCGTTTATTAGATGTAGATAAC 420
WAYANAD3          TATATAATTCCTACTAATGACTTAAATACCTATAATTTCCGTTTATTAGATGTAGATAAC 420
                  ********** ********** ********** ********** **********

VAISHALI4         CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 480
VAISHALI5         CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 480
VAISHALI3         CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 480
VAISHALI2         CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 480
VAISHALI1         CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 480
TRICHY3           CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 480
TRICHY1           CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 480
NARAYANGAON5      CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 480
NARAYANGAON4      CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 480
NARAYANGAON3      CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 480
NARAYANGAON2      CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 480
NARAYANGAON1      CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 480
JALGAON5          CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 480
JALGAON3          CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 480
JALGAON2          CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 480
JALGAON1          CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 480
Assam1            CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 480
TRICHY2           CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 480
TRICHY4           CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 480
TRICHY5           CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 480
JALGAON4          CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATTTATAGATGTAATT 480
ASSAM2            CGATTAATTGTCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 480
ASSAM3            CGATTAATTGTCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 480
ASSAM4            CGATTAATTGTCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 480
ASSAM5            CGATTAATTGTCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 480
WAYANAD4          CGATTAATTGTCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 480
WAYANAD5          CGATTAATTGTCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 480
WAYANAD1          CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 480
WAYANAD2          CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 480
WAYANAD3          CGATTAATTATCCCATTTGAAACCCAAATCCGCCTACTAGTTACATCTATAGATGTAATT 480
                  ********* ********** ********** ********** **********

VAISHALI4         CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 540
VAISHALI5         CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 540
VAISHALI3         CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 540
VAISHALI2         CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 540
VAISHALI1         CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 540
TRICHY3           CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 540
```

```
TRICHY1          CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 540
NARAYANGAON5     CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 540
NARAYANGAON4     CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 540
NARAYANGAON3     CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 540
NARAYANGAON2     CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 540
NARAYANGAON1     CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 540
JALGAON5         CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 540
JALGAON3         CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 540
JALGAON2         CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 540
JALGAON1         CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 540
Assam1           CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 540
TRICHY2          CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 540
TRICHY4          CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 540
TRICHY5          CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 540
JALGAON4         CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 540
ASSAM2           CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 540
ASSAM3           CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 540
ASSAM4           CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 540
ASSAM5           CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 540
WAYANAD4         CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 540
WAYANAD5         CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 540
WAYANAD1         CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 540
WAYANAD2         CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 540
WAYANAD3         CACTCTTGAACAATTCCTTCTTTAGGTATCAAAATTGATGGAACTCCTGGACGTCTAAAC 540
                 ************************************************************

VAISHALI4        CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 600
VAISHALI5        CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 600
VAISHALI3        CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 600
VAISHALI2        CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 600
VAISHALI1        CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 600
TRICHY3          CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 600
TRICHY1          CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 600
NARAYANGAON5     CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 600
NARAYANGAON4     CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 600
NARAYANGAON3     CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 600
NARAYANGAON2     CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 600
NARAYANGAON1     CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 600
JALGAON5         CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 600
JALGAON3         CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 600
JALGAON2         CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 600
JALGAON1         CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 600
Assam1           CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 600
TRICHY2          CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 600
TRICHY4          CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 600
TRICHY5          CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 600
JALGAON4         CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 600
ASSAM2           CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 600
ASSAM3           CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 600
ASSAM4           CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 600
ASSAM5           CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 600
WAYANAD4         CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 600
WAYANAD5         CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 600
WAYANAD1         CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 600
WAYANAD2         CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 600
WAYANAD3         CAAGTAGGACTTATAATTAATCGTCCTGGATTATTTTATGGTCAATGTTCAGAAATTTGT 600
                 ************************************************************

VAISHALI4          GG 602
VAISHALI5          GG 602
VAISHALI3          GG 602
VAISHALI2          GG 602
VAISHALI1          GG 602
TRICHY3            GG 602
TRICHY1            GG 602
NARAYANGAON5       GG 602
NARAYANGAON4       GG 602
NARAYANGAON3       GG 602
NARAYANGAON2       GG 602
```

```
NARAYANGAON1      GG 602
JALGAON5          GG 602
JALGAON3          GG 602
JALGAON2          GG 602
JALGAON1          GG 602
Assam1            GG 602
TRICHY2           GG 602
TRICHY4           GG 602
TRICHY5           GG 602
JALGAON4          GG 602
ASSAM2            GG 602
ASSAM3            GG 602
ASSAM4            GG 602
ASSAM5            GG 602
WAYANAD4          GG 602
WAYANAD5          GG 602
WAYANAD1          GG 602
WAYANAD2          GG 602
WAYANAD3          GG 602
                  **
```

Fig. 12. ClustalW alignment of the partial COII nucleotide sequences of the thirty *O. longicollis* (Oliver) individuals

The partial COII gene translated into a protein of 200 amino acids using the invertebrate translation code. The ClustalW alignment of the partial amino acid sequences of COII of all the thirty individuals is shown in **Fig. 13.**

```
ASSAM2            IATWKTLMLQDSASPLMEQLLFFHDHALLILVLITILVGQMLLSMLFNKFTHRFLLEGQM 60
ASSAM3            IATWKTLMLQDSASPLMEQLLFFHDHALLILVLITILVGQMLLSMLFNKFTHRFLLEGQM 60
ASSAM4            IATWKTLMLQDSASPLMEQLLFFHDHALLILVLITILVGQMLLSMLFNKFTHRFLLEGQM 60
ASSAM5            IATWKTLMLQDSASPLMEQLLFFHDHALLILVLITILVGQMLLSMLFNKFTHRFLLEGQM 60
WAYANAD4          IATWKTLMLQDSASPLMEQLLFFHDHALLILVLITILVGQMLLSMLFNKFTHRFLLEGQM 60
WAYANAD5          IATWKTLMLQDSASPLMEQLLFFHDHALLILVLITILVGQMLLSMLFNKFTHRFLLEGQM 60
Assam1            IATWKTLMLQDSASPLMEQLLFFHDHALLILVLITILVGQMLLSMLFNKFTHRFLLEGQM 60
JALGAON1          IATWKTLMLQDSASPLMEQLLFFHDHALLILVLITILVGQMLLSMLFNKFTHRFLLEGQM 60
JALGAON2          IATWKTLMLQDSASPLMEQLLFFHDHALLILVLITILVGQMLLSMLFNKFTHRFLLEGQM 60
JALGAON3          IATWKTLMLQDSASPLMEQLLFFHDHALLILVLITILVGQMLLSMLFNKFTHRFLLEGQM 60
WAYANAD1          IATWKTLMLQDSASPLMEQLLFFHDHALLILVLITILVGQMLLSMLFNKFTHRFLLEGQT 60
WAYANAD2          IATWKTLMLQDSASPLMEQLLFFHDHALLILVLITILVGQMLLSMLFNKFTHRFLLEGQT 60
WAYANAD3          IATWKTLMLQDSASPLMEQLLFFHDHALLILVLITILVGQMLLSMLFNKFTHRFLLEGQM 60
VAISHALI5         IATWKTLMLQDSASPLMEQLLFFHDHALLILVLITILVGQMLLSMLFNKFTHRFLLEGQM 60
VAISHALI4         IATWKTLMLQDSASPLMEQLLFFHDHALLILVLITILVGQMLLSMLFNKFTHRFLLEGQM 60
VAISHALI3         IATWKTLMLQDSASPLMEQLLFFHDHALLILVLITILVGQMLLSMLFNKFTHRFLLEGQM 60
VAISHALI2         IATWKTLMLQDSASPLMEQLLFFHDHALLILVLITILVGQMLLSMLFNKFTHRFLLEGQM 60
VAISHALI1         IATWKTLMLQDSASPLMEQLLFFHDHALLILVLITILVGQMLLSMLFNKFTHRFLLEGQM 60
TRICHY3           IATWKTLMLQDSASPLMEQLLFFHDHALLILVLITILVGQMLLSMLFNKFTHRFLLEGQM 60
TRICHY2           IATWKTLMLQDSASLLMEQLLFFHDHALLILVLITILVGQMLLSMLFNKFTHRFLLEGQM 60
TRICHY4           IATWKTLMLQDSASLLMEQLLFFHDHALLILVLITILVGQMLLSMLFNKFTHRFLLEGQM 60
TRICHY5           IATWKTLMLQDSASLLMEQLLFFHDHALLILVLITILVGQMLLSMLFNKFTHRFLLEGQM 60
TRICHY1           IATWKTLMLQDSASPLMEQLLFFHDHALLILVLITILVGQMLLSMLFNKFTHRFLLEGQM 60
NARAYANGAON5      IATWKTLMLQDSASPLMEQLLFFHDHALLILVLITILVGQMLLSMLFNKFTHRFLLEGQM 60
NARAYANGAON4      IATWKTLMLQDSASPLMEQLLFFHDHALLILVLITILVGQMLLSMLFNKFTHRFLLEGQM 60
NARAYANGAON3      IATWKTLMLQDSASPLMEQLLFFHDHALLILVLITILVGQMLLSMLFNKFTHRFLLEGQM 60
NARAYANGAON2      IATWKTLMLQDSASPLMEQLLFFHDHALLILVLITILVGQMLLSMLFNKFTHRFLLEGQM 60
NARAYANGAON1      IATWKTLMLQDSASPLMEQLLFFHDHALLILVLITILVGQMLLSMLFNKFTHRFLLEGQM 60
JALGAON5          IATWKTLMLQDSASPLMEQLLFFHDHALLILVLITILVGQMLLSMLFNKFTHRFLLEGQM 60
JALGAON4          IATWKTLMLQDSASPLMEQLLFFHDHALLILVLITILVGQMLLSVLFNKFTHRFLLEGQM 60
                  ************** ****************************:**************

ASSAM2            IELIWTILPAIILILIALPSLRLLYILDEINNPSITIKAIGHQWYWSYEYSDYKNIEFDS 120
ASSAM3            IELIWTILPAIILILIALPSLRLLYILDEINNPSITIKAIGHQWYWSYEYSDYKNIEFDS 120
ASSAM4            IELIWTILPAIILILIALPSLRLLYILDEINNPSITIKAIGHQWYWSYEYSDYKNIEFDS 120
ASSAM5            IELIWTILPAIILILIALPSLRLLYILDEINNPSITIKAIGHQWYWSYEYSDYKNIEFDS 120
```

```
WAYANAD4          IELIWTILPAIILILIALPSLRLLYILDEINNPSITIKAIGHQWYWSYEYSDYKNIEFDS 120
WAYANAD5          IELIWTILPAIILILIALPSLRLLYILDEINNPSITIKAIGHQWYWSYEYSDYKNIEFDS 120
Assam1            IELIWTILPAIILILIALPSLRLLYILDEINNPSITIKAIGHQWYWSYEYSDYKNIEFDS 120
JALGAON1          IELIWTILPAIILILIALPSLRLLYILDEINNPSITIKAIGHQWYWSYEYSDYKNIEFDS 120
JALGAON2          IELIWTILPAIILILIALPSLRLLYILDEINNPSITIKAIGHQWYWSYEYSDYKNIEFDS 120
JALGAON3          IELIWTILPAIILILIALPSLRLLYILDEINNPSITIKAIGHQWYWSYEYSDYKNIEFDS 120
WAYANAD1          IELIWTILPAIILILIALPSLRLLYILDEINNPSITIKAIGHQWYWSYEYSDYKNIEFDS 120
WAYANAD2          IELIWTILPAIILILIALPSLRLLYILDEINNPSITIKAIGHQWYWSYEYSDYKNIEFDS 120
WAYANAD3          IELIWTILPAIILILIALPSLRLLYILDEINNPSITIKAIGHQWYWSYEYSDYKNIEFDS 120
VAISHALI5         IELIWTILPAIILILIALPSLRLLYILDEINNPSITIKAIGHQWYWSYEYSDYKNIEFDS 120
VAISHALI4         IELIWTILPAIILILIALPSLRLLYILDEINNPSITIKAIGHQWYWSYEYSDYKNIEFDS 120
VAISHALI3         IELIWTILPAIILILIALPSLRLLYILDEINNPSITIKAIGHQWYWSYEYSDYKNIEFDS 120
VAISHALI2         IELIWTILPAIILILIALPSLRLLYILDEINNPSITIKAIGHQWYWSYEYSDYKNIEFDS 120
VAISHALI1         IELIWTILPAIILILIALPSLRLLYILDEINNPSITIKAIGHQWYWSYEYSDYKNIEFDS 120
TRICHY3           IELIWTILPAIILILIALPSLRLLYILDEINNPSITIKAIGHQWYWSYEYSDYKNIEFDS 120
TRICHY2           IELIWTILPAIILILIALPSLRLLYILDEINNPSITIKAIGHQWYWSYEYSDYKNIEFDS 120
TRICHY4           IELIWTILPAIILILIALPSLRLLYILDEINNPSITIKAIGHQWYWSYEYSDYKNIEFDS 120
TRICHY5           IELIWTILPAIILILIALPSLRLLYILDEINNPSITIKAIGHQWYWSYEYSDYKNIEFDS 120
TRICHY1           IELIWTILPAIILILIALPSLRLLYILDEINNPSITIKAIGHQWYWSYEYSDYKNIEFDS 120
NARAYANGAON5      IELIWTILPAIILILIALPSLRLLYILDEINNPSITIKAIGHQWYWSYEYSDYKNIEFDS 120
NARAYANGAON4      IELIWTILPAIILILIALPSLRLLYILDEINNPSITIKAIGHQWYWSYEYSDYKNIEFDS 120
NARAYANGAON3      IELIWTILPAIILILIALPSLRLLYILDEINNPSITIKAIGHQWYWSYEYSDYKNIEFDS 120
NARAYANGAON2      IELIWTILPAIILILIALPSLRLLYILDEINNPSITIKAIGHQWYWSYEYSDYKNIEFDS 120
NARAYANGAON1      IELIWTILPAIILILIALPSLRLLYILDEINNPSITIKAIGHQWYWSYEYSDYKNIEFDS 120
JALGAON5          IELIWTILPAIILILIALPSLRLLYILDEINNPSITIKAIGHQWYWSYEYSDYKNIEFDS 120
JALGAON4          IELIWTILPAIILILIALPSLRLLYILDEINNPSITIKAIGHQWYWSYEYSDYKNIEFDS 120
                  ************************************************************

ASSAM2            YMIPTNDLNTYNFRLLDVDNRLIVPFETQIRLLVTSMDVIHSWTIPSLGIKIDGTPGRLN 180
ASSAM3            YMIPTNDLNTYNFRLLDVDNRLIVPFETQIRLLVTSMDVIHSWTIPSLGIKIDGTPGRLN 180
ASSAM4            YMIPTNDLNTYNFRLLDVDNRLIVPFETQIRLLVTSMDVIHSWTIPSLGIKIDGTPGRLN 180
ASSAM5            YMIPTNDLNTYNFRLLDVDNRLIVPFETQIRLLVTSMDVIHSWTIPSLGIKIDGTPGRLN 180
WAYANAD4          YMIPTNDLNTYNFRLLDVDNRLIVPFETQIRLLVTSMDVIHSWTIPSLGIKIDGTPGRLN 180
WAYANAD5          YMIPTNDLNTYNFRLLDVDNRLIVPFETQIRLLVTSMDVIHSWTIPSLGIKIDGTPGRLN 180
Assam1            YMIPTNDLNTYNFRLLDVDNRLIIPFETQIRLLVTSMDVIHSWTIPSLGIKIDGTPGRLN 180
JALGAON1          YMIPTNDLNTYNFRLLDVDNRLIIPFETQIRLLVTSMDVIHSWTIPSLGIKIDGTPGRLN 180
JALGAON2          YMIPTNDLNTYNFRLLDVDNRLIIPFETQIRLLVTSMDVIHSWTIPSLGIKIDGTPGRLN 180
JALGAON3          YMIPTNDLNTYNFRLLDVDNRLIIPFETQIRLLVTSMDVIHSWTIPSLGIKIDGTPGRLN 180
WAYANAD1          YMIPTNDLNTYNFRLLDVDNRLIIPFETQIRLLVTSMDVIHSWTIPSLGIKIDGTPGRLN 180
WAYANAD2          YMIPTNDLNTYNFRLLDVDNRLIIPFETQIRLLVTSMDVIHSWTIPSLGIKIDGTPGRLN 180
WAYANAD3          YMIPTNDLNTYNFRLLDVDNRLIIPFETQIRLLVTSMDVIHSWTIPSLGIKIDGTPGRLN 180
VAISHALI5         YMIPTNDLNTYNFRLLDVDNRLIIPFETQIRLLVTSMDVIHSWTIPSLGIKIDGTPGRLN 180
VAISHALI4         YMIPTNDLNTYNFRLLDVDNRLIIPFETQIRLLVTSMDVIHSWTIPSLGIKIDGTPGRLN 180
VAISHALI3         YMIPTNDLNTYNFRLLDVDNRLIIPFETQIRLLVTSMDVIHSWTIPSLGIKIDGTPGRLN 180
VAISHALI2         YMIPTNDLNTYNFRLLDVDNRLIIPFETQIRLLVTSMDVIHSWTIPSLGIKIDGTPGRLN 180
VAISHALI1         YMIPTNDLNTYNFRLLDVDNRLIIPFETQIRLLVTSMDVIHSWTIPSLGIKIDGTPGRLN 180
TRICHY3           YMIPTNDLNTYNFRLLDVDNRLIIPFETQIRLLVTSMDVIHSWTIPSLGIKIDGTPGRLN 180
TRICHY2           YMIPTNDLNTYNFRLLDVDNRLIIPFETQIRLLVTSMDVIHSWTIPSLGIKIDGTPGRLN 180
TRICHY4           YMIPTNDLNTYNFRLLDVDNRLIIPFETQIRLLVTSMDVIHSWTIPSLGIKIDGTPGRLN 180
TRICHY5           YMIPTNDLNTYNFRLLDVDNRLIIPFETQIRLLVTSMDVIHSWTIPSLGIKIDGTPGRLN 180
TRICHY1           YMIPTNDLNTYNFRLLDVDNRLIIPFETQIRLLVTSMDVIHSWTIPSLGIKIDGTPGRLN 180
NARAYANGAON5      YMIPTNDLNTYNFRLLDVDNRLIIPFETQIRLLVTSMDVIHSWTIPSLGIKIDGTPGRLN 180
NARAYANGAON4      YMIPTNDLNTYNFRLLDVDNRLIIPFETQIRLLVTSMDVIHSWTIPSLGIKIDGTPGRLN 180
NARAYANGAON3      YMIPTNDLNTYNFRLLDVDNRLIIPFETQIRLLVTSMDVIHSWTIPSLGIKIDGTPGRLN 180
NARAYANGAON2      YMIPTNDLNTYNFRLLDVDNRLIIPFETQIRLLVTSMDVIHSWTIPSLGIKIDGTPGRLN 180
NARAYANGAON1      YMIPTNDLNTYNFRLLDVDNRLIIPFETQIRLLVTSMDVIHSWTIPSLGIKIDGTPGRLN 180
JALGAON5          YMIPTNDLNTYNFRLLDVDNRLIIPFETQIRLLVTSMDVIHSWTIPSLGIKIDGTPGRLN 180
JALGAON4          YMIPTNDLNTYNFRLLDVDNRLIIPFETQIRLLVTFMDVIHSWTIPSLGIKIDGTPGRLN 180
                  ***********************:*********** ************************

ASSAM2              QVGLMINRPGLFYGQCSEIC 200
ASSAM3              QVGLMINRPGLFYGQCSEIC 200
ASSAM4              QVGLMINRPGLFYGQCSEIC 200
ASSAM5              QVGLMINRPGLFYGQCSEIC 200
WAYANAD4            QVGLMINRPGLFYGQCSEIC 200
WAYANAD5            QVGLMINRPGLFYGQCSEIC 200
Assam1              QVGLMINRPGLFYGQCSEIC 200
JALGAON1            QVGLMINRPGLFYGQCSEIC 200
JALGAON2            QVGLMINRPGLFYGQCSEIC 200
JALGAON3            QVGLMINRPGLFYGQCSEIC 200
```

```
WAYANAD1          QVGLMINRPGLFYGQCSEIC 200
WAYANAD2          QVGLMINRPGLFYGQCSEIC 200
WAYANAD3          QVGLMINRPGLFYGQCSEIC 200
VAISHALI5         QVGLMINRPGLFYGQCSEIC 200
VAISHALI4         QVGLMINRPGLFYGQCSEIC 200
VAISHALI3         QVGLMINRPGLFYGQCSEIC 200
VAISHALI2         QVGLMINRPGLFYGQCSEIC 200
VAISHALI1         QVGLMINRPGLFYGQCSEIC 200
TRICHY3           QVGLMINRPGLFYGQCSEIC 200
TRICHY2           QVGLMINRPGLFYGQCSEIC 200
TRICHY4           QVGLMINRPGLFYGQCSEIC 200
TRICHY5           QVGLMINRPGLFYGQCSEIC 200
TRICHY1           QVGLMINRPGLFYGQCSEIC 200
NARAYANGAON5      QVGLMINRPGLFYGQCSEIC 200
NARAYANGAON4      QVGLMINRPGLFYGQCSEIC 200
NARAYANGAON3      QVGLMINRPGLFYGQCSEIC 200
NARAYANGAON2      QVGLMINRPGLFYGQCSEIC 200
NARAYANGAON1      QVGLMINRPGLFYGQCSEIC 200
JALGAON5          QVGLMINRPGLFYGQCSEIC 200
JALGAON4          QVGLMINRPGLFYGQCSEIC 200
                  ********************
```

Fig. 13. ClustalW alignment of the partial translated COII region of the thirty *O. longicollis* (Oliver) individuals

The ClustalW alignment of the partial COII of one representative individual ie Assam 1 with the COII protein sequences of a few insects ie *D.betulae, B. Formaneki, P. Marginatus and O. rugosostriatus* identified the membrane spanning helices M1 and M2, the separating tract I1 and the periplasmic domain (**Fig. 14).**

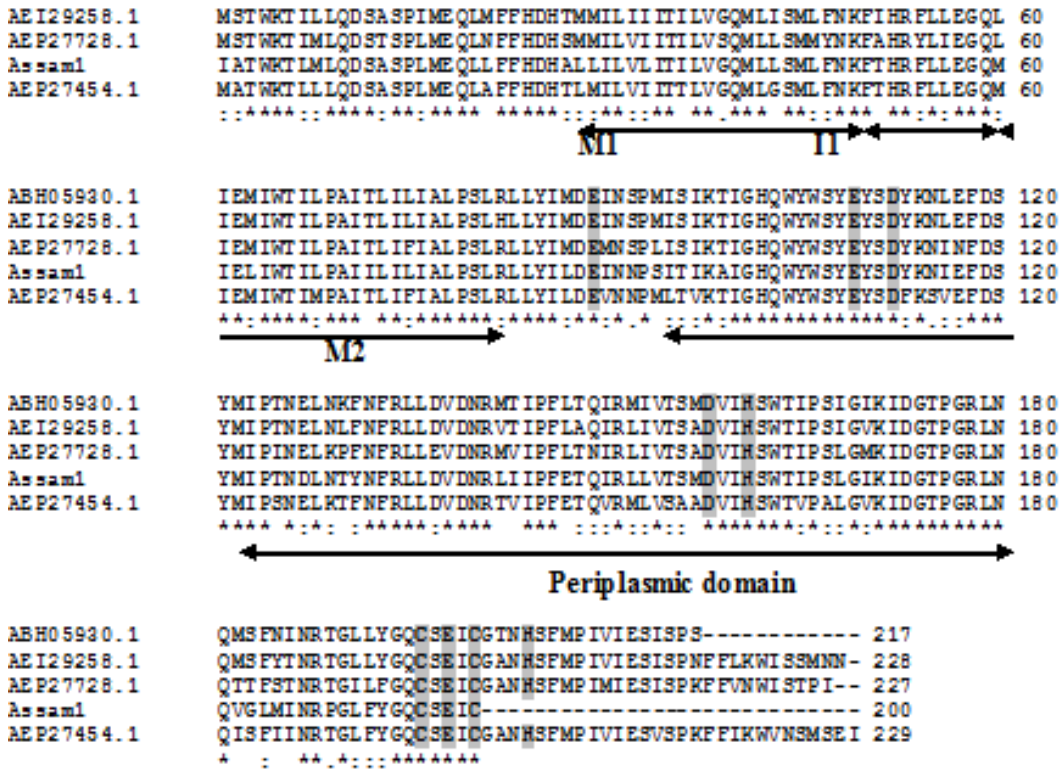


Fig. 14. ClustalW alignment of the partial COII sequence of Assam (individual 1) with the COII proteins of a few insects. *D.betulae* (AEP27454), *B. formaneki* (ABH 05930), *P. marginatus* (AEI29258) and *O. rugosostriatus* (AEP27728). Protein IDs of the COII proteins are indicated in parenthesis. The different structural regions are shown by double headed arrows. Functionally significant residues are in bold and shaded in grey and are described in the text.

A Codon Bias Index (CBI) of 0.610 was observed for COII and the RSCU values are shown in **Table 13.**

Table 13. Codon usage bias in the partial COII protein of the thirty weevils of *O. longicollis* (Oliver)

Codon	Count	RSCU	Codon	Count	RSCU	Codon	Count	RSCU	Codon	Count	RSCU
UUU(F)	4.7	1.05	UCU(S)	7	**4.66**	UAU(Y)	6	1.5	UGU(C)	2	2
UUC(F)	4.3	0.95	UCC(S)	0	0	UAC(Y)	2	0.5	UGC(C)	0	0
UUA(L)	16.3	**2.87**	UCA(S)	3	2.01	UAA(*)	0	0	UGA(W)	5	2

UUG(L)	0.7	0.12	UCG(S)	0	0	UAG(*)	0	0	UGG(W)	0	0
CUU(L)	8.4	1.48	CCU(P)	5.9	**2.65**	CAU(H)	2.5	1	CGU(R)	5	2.86
CUC(L)	1	0.18	CCC(P)	1	0.45	CAC(H)	2.5	1	CGC(R)	1	0.57
CUA(L)	7.7	1.35	CCA(P)	2	0.9	CAA(Q)	8	2	CGA(R)	1	0.57
CUG(L)	0	0	CCG(P)	0	0	CAG(Q)	0	0	CGG(R)	0	0
AUU(I)	23	1.72	ACU(T)	6	1.99	AAU(N)	6.3	1.26	AGU(S)	0	0
AUC(I)	3.8	0.28	ACC(T)	3	0.99	AAC(N)	3.7	0.74	AGC(S)	1	0.67
AUA(M)	7.9	2	ACA(T)	3.1	1.02	AAA(K)	5	2	AGA(S)	1	0.67
AUG(M)	0	0	ACG(T)	0	0	AAG(K)	0	0	AGG(S)	0	0
GUU(V)	2	1.28	GCU(A)	3	2	GAU(D)	8	1.6	GGU(G)	2	0.89
GUC(V)	0.2	0.13	GCC(A)	1	0.67	GAC(D)	2	0.4	GGC(G)	1	0.44
GUA(V)	4	2.59	GCA(A)	2	1.33	GAA(E)	8	2	GGA(G)	6	**2.67**
GUG(V)	0	0	GCG(A)	0	0	GAG(E)	0	0	GGG(G)	0	0

Protein Hydrophobicity Plots of COI and COII

Figs 15 and 16 show the hydrophobicity plots of COI and COII which are typical of membrane helical regions of membrane bound protein. Kyte and Doolittle (1982) is a widely applied scale for delineating hydrophobic character of a protein. Regions with values above 0 are hydrophobic in character.

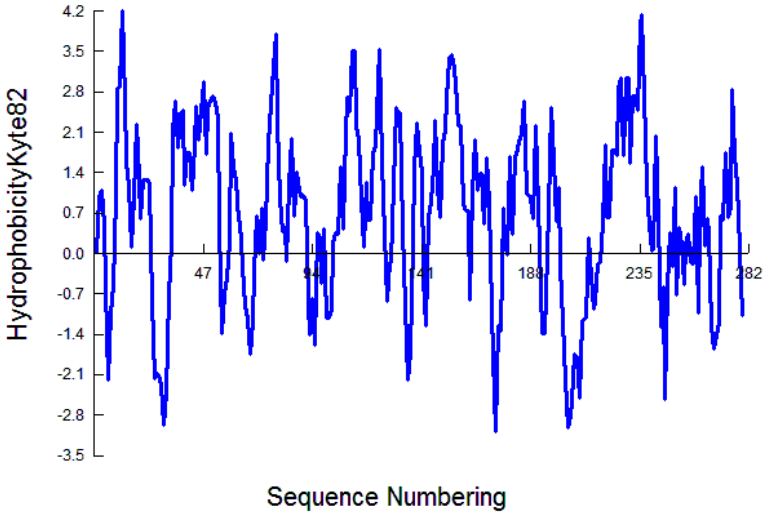


Fig. 15. Hydrophobicity plot of COI

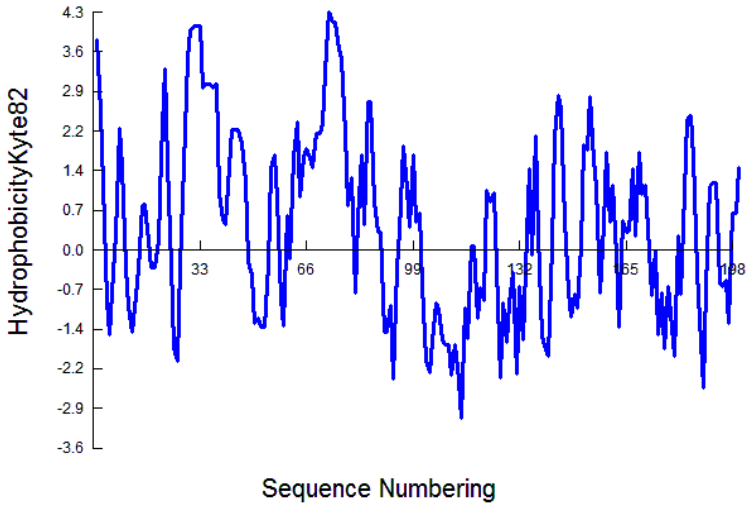


Fig. 16. Hydrophobicity plot of COII

Analysis of tRNA^{Leu}

The tRNA^{Leu} sequences of all the thirty individuals are identical and could be folded into the usual two dimensional structures using tRNAscan-SE (Lowe & Eddy 1997) (**Fig. 17**).

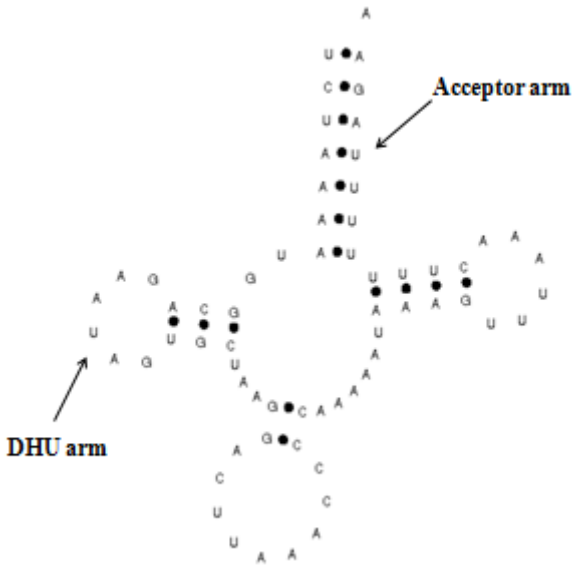


Fig. 17. 2D structure of the tRNA^{Leu} of the Assam1 individual

The conserved regions in the tRNA^{Leu} gene are shown in the ClustalW alignment of Assam1 tRNA^{Leu} with four other Coleopteran insects (Nardi *et al.* 2003) (**Fig. 18**).

```
DQ768215.1      TCTAATATGGCAGATTAGTGCATTGGATTTAAACCCCAAAAATAAAGTTT 50
EF398270.1      TCTAATATGGCAGATTAGTGCATTGGATTTAAGATCCAAATATAAAGTCA 50
Assam1          TCTAAAATGGCAGAATAGTGCATTGGACTTAAACCCCAAAAATAAAGTTT 50
AF467886.1      TCTAATATGGCAGAATAGTGCAATGGATTTAAACCCCATATATAAAGATT 50
15626344        TCTATTATGGCAGATTAGTGCAATGGACTTAAACCCCATTTATAAAGTAA 50
                 ****  ******** ******* **** ****   ***   ******
                 aa-arm       DHU-arm       Anticodon arm TψC-arm


DQ768215.1       ACCCTTTTTTTAGAA 65
EF398270.1       --TCTTTTTTTAGAA 63
                 AAACTTTTTTTAGAA 65
          1      AAACTTTTTTTAGAA 65
                 ATACTTTTTTTAGAA 65
                    ************
                         aa-arm
```

Fig. 18. ClustalW of Assam1 tRNA^{Leu} with four coleopteran insects. *Anoplophora glabripennis* (DQ768215.1), *Pyrophorus divergens* (EF398270.1), *Crioceris duodecimpunctata* (AF467886.1) and *Tribolium castaneum* (15626344)

DISCUSSION

In the present work, the sequence of the COI-tRNA^{Leu}-COII region was used to assess the genetic diversity of thirty weevils representing six Indian populations of O*doiporus longicollis* (Oliver). As observed in insects and crustaceans, the L2 gene encoding the two codon families UUR, lies between the COI and COII genes.

The COI-tRNA^{Leu}-COII sequenced region showed the typical AT bias as observed in insect mitochondrial DNA (Frati *et al.* 1997). The G-A and C-T(U) transitions were more frequent than the T(U)-C and A-G transitions. The nucleotide composition of the partial COI and COII genes is AT rich as observed in insect mitochondrial genes (Frati *et al.* 1997). Such a transition bias has been suggested to be caused due to the intrinsic property of the mitochondrial DNA where the spontaneous deamination of C to T results in an increase in the GC→AT transitions (Belle *et al.* 2005)**.** The AT bias was highest in the 1st codon position of COI and in the 3rd codon position of COII. The COI and COII genes are under functional constraint because they encode proteins essential for the electron transport system. High rates of transitions and low rates of transversions are observed at all the three codon positions in the partial COI and COII genes, a characteristic of mitochondrial genes (de Bruin 1983). The amino acid composition of the resulting protein can be changed by both transitions and transversions but transitions, by and large, do not alter the amino-acid of the resulting protein, whereas the biochemical difference between the original and mutated protein is more pronounced for a mutation caused by a transversion resulting in favouring against transversions (de Bruin 1983; Zhang 2000). Bias towards transitions is most evident when conspecific populations are being compared, as is observed in the present study (Frati *et al.* 1997).

The partial protein sequence of the COI gene of the thirty *O. longicollis* individuals deduced using the invertebrate genetic code has 238 amino acids which on a protein BLAST showed identity with the COI of other insects. The COI gene in all the thirty individuals terminates with the amino acid N encoded by the codon AAA (invertebrate mitochondrial code) followed by TT before the start of the tRNA^{Leu} gene. The common termination codons identified in other Coleopterans include TAA, ATT, TTT, ACT, TAG and T (Sheffield *et al.* 2008). However, there are no reports of any

insect mitochondrial gene with a TT termination codon. It is very likely that in the present case also, T may be serving as a termination codon and the additional T could represent a non-coding intergenic spacer between COI and $tRNA^{Leu}$. Non-coding intergenic spacer regions between coding genes have been reported in the mitochondrial genome of several insects (Sheffield *et al.* 2008). In insects, the termination codon is not complete and complete termination codons are produced by post-trancriptional polyadenylation (Ojala *et al.* 1981).

The translated partial sequence of COI includes thirteen structurally important regions *i.e.* external loops E5 and E6, M6-M12, internal loops I3-15 and the COOH terminal. ClustalW alignment of the partial COI protein of the thirty individuals exhibits high conservation with a few substitutions (**Table 10**). Substitutions in the third codon position resulted in synonymous amino acid substitutions; those in the first codon position resulted in conservative or semi-conservative replacements whereas second codon position substitutions resulted by and large in amino acid replacements. These observations are as expected for mitochondrial DNA substitutions (de Bruin 1983; Zhang 2000). Transmembrane helices M1, M6, M7 and M10 have been implicated in providing the sites for binding of metal ligands which interact with two heam groups and copper atom, found to be essential for the activity of COI (Lunt *et al.* 1996). The conserved helix M8, which is the region of cytochrome oxidase proton conduction channel, has three conserved polar residues (T308, T315 and K318) which are involved with the COI proton-conduction channel and these three residues are found to be completely conserved across all taxa of organisms. TY370 in E5 is suggested to be closely associated with heme A. Transmembrane helices M5 and M11 are also found to be highly conserved and have a structural role (Lunt *et al.* 1996).

The COOH region of COI is the least conserved and highly variable when the protein alignment is compared with a few other insect COI (**Fig. 10**) (Lunt *et al.* 1996). The COII gene begins immediately after the $tRNA^{Leu}$ gene with the start codon ATT. ATG, ATT, ATC and ATA are common initiation codons in insects (Frati *et al.* 1997). Fearnley & Walker (1987) showed that in the human mitochondrial ND2 gene, ATT is translated as methionine and the same could be true in case of insect mitochondrial genes.

COII transfers electrons from cytochrome c to the catalytic subunit I. This subunit provides the substrate-binding site of the complex and harbours the copper centre called as Cu(A). Ligation of two copper atoms to each molecule of COII involves residues H161, C196, C200 and H204 on the COII molecule (Frati *et al.*1997). COII also provides a cleft for the binding of cytochrome c. D112, D158 and E198 on the COII molecule which are highly conserved in pterygote insects interact with a ring of lysine residues on the cytochrome c facilitating this binding. In bovines in addition to D112, D158 and E198, Glu 114 is also involved in this function. As Glu114 is not present in insects, Glu 109 is proposed to play a similar role. Glu 89 is conserved in all insects and is suggested to play an important role in the structure and function of COII. Amino acid residues 26–48 and 60–80 appear to form hydrophobic regions and are probably membrane spanning helices 1 and 2 (M1, M2). These regions and the separating tract (I1) are variable. Region 95-170 represents the periplasmic domain (Frati *et al.*1997).

Sixteen and nineteen amino acids in COI and COII respectively reflect a bias against codons ending in a C or G (**Tables 11 and 13**). In the COII gene, CAU and CAC both coding for histidine gave a RSCU value of 1. Majority of the anticodons have U as the wobble base which allows efficient pairing with the synonymous codons as compared with anticodons which have a G or C as the wobble base (Sun *et al.* 2009). A similar observation was reported for pterygote insects and primates (Frati *et al.* 1997). Xia (1996) suggested that in insects, transcriptional efficiency can be increased by maximising the use of A in the third codon position of protein coding genes. In the COI gene, highest RSCU values are observed for the codons UGG, CGC, AGG and GGC which code for W, R, S and G respectively. Considering the general A+T bias observed in the third codon position, these codons should have been less favoured. However in case of these four amino acids, codon usage does not appear to be governed by A+T bias, but probably is influenced by selection for the corresponding tRNA molecules.

The populations did not show any phylogeographic distribution and no strong correlateion was observed between genetic distance and geographic distance. However, the overall genetic differentiation among the populations was highly significant. The Fu & Li's (1993) D and F tests were non-significant for this mitochondrial region leading to

the rejection of background selection occurring in the data set being analysed. The overall genetic differentiation observed using the mitochondrial fragment suggested that the populations were isolated and this result did not correlate with the genetic differentiation observed with the rDNA markers ITS1 and ITS2 that have been described in Chapter 4 of this thesis. Analysis using each of these nuclear markers indicated gene flow between the populations which is expected considering the strong flying ability of the weevils. Such a discrepancy in the results obtained using a mitochondrial marker and nuclear markers can be attributed to sex biased dispersal (Goudet *et al.* 2002). A simultaneous study of two genes which differ in their mode of inheritance makes it possible to assess whether the observed genetic differentiation is due to a bias in dispersal of one of the genders, because male and female insects of several species have been reported to differ in their migratory and colonization behaviour. (Roderick 1996; Sielezniew *et al.* 2011). For markers such as the mitochondrial markers which are uni-parentally inherited, the male sex does not contribute to the mitochondrial genome of the off-spring, whereas for nuclear genes such as the rDNA genes which are biparentally inherited, both sexes contribute to the diversity in the progeny. Hence a difference in the genetic structure between these two markers is expected when sex biased dispersal occurs. For species in which the males disperse and the females are philopatric, the estimated genetic differentiation between populations is expected to be higher when estimated using mitochondrial DNA or any other maternal marker, than using a biparental marker which is observed in the present study. The two key factors which have been suggested to favour sex biased dispersal are avoidance of inbreeding and asymmetry in the intensity in local competition about reproductive resources (Gros *et al.* 2008). Sex biased dispersal has been reported in a number of insects ie damselflies (Beirinckx *et al.* 2006), cactophilic *Drosophila* (Markow & Castrezana 2000), mayfly *Callibaetis ferrugineus hageni* (Caudill 2003), ant *Formica exsecta* (Sundstrom *et al.*2003), bark beetle *Ips typographicus* (Salle *et al.* 2007), carrabid beetles (Lagisz *et al.*2010), queenless ant *Diacamma cyaneiventre* (Doums *et al.*2002), butterfly *Plebejus argus* (Sielezniew *et al.*2011) and grasshoppers (Kindler *et al.*2012). The results of the present study support a male biased dispersal of *Odoiporus longicollis* (Oliver) weevils. *Odoiporus longicollis* (Oliver) weevils are fast fliers and the by flight.

Cytoplasmic incompatibility (CI) causing *Wolbachia* could replace the original mitochondrial haplotypes by hitchhiking through the population and such a selective sweep could be mistaken for a population bottle neck or a founder effect. *Wolbachia* infections within a population would thus cause mitochondrial isolation inspite of nuclear gene flow. Hence, phylogenies based on such mitochondrial data would contradict those derived from nuclear DNA markers such as ITS1 and ITS2. Therefore it is necessary to check for *Wolbachia* infections in the populations being studied when mitochondrial DNA data is being used to derive phylogenies (Arthofer *et al.*2010). In the present work, *Wolbachia* specific primers were used to check for the presence of *Wolbachia* in the six populations. However, no amplification products were observed confirming that the observed significant overall genetic differentiation was caused due to male biased gene flow.

Knowledge about pest dispersal and movement and their relationship to pest management practices, environmental factors and ecosystem can be used to predict pest infestations. An understanding of pest dispersal in important for the success of IPM programmes (Aylor & Irwin 1999; Byrne 1999). Knowledge about the prevalence of male biased gene flow in *Odoiporus longicollis* (Oliver) could aid in the IPM of bananas.

REFERENCES

Ahern RG, Hawthorne DJ and Raupp MJ (2009) *Molecular Ecology*, ***18*****(2)**, 343-356.

Aylor DE and Irwin ME (1999) *Agricultural and Forest Meteorology*, **97,** 233-234.

Arthofer W, Avtzis DN, Riegler M and Stauffer C (2010) *ZooKeys*, **(56)**, 269.

Baskaran M (2011) *Populations from Tamil Nadu, South India*, **11(05)**, 12-15.

Beirinckx K, Gossum HV, Lajeunesse MJ and Forbes MR (2006) *Oikos*, ***113*****(3),** 539-547.

Belle E, Piganeau G, Gardner M and Eyre-Walker A (2005) *Gene*, **355**, 58-66.

Broughton RE and Reneau PC (2006) *Molecular biology and evolution*, **23(8)**, 1516-1524.

Brown JK (1994) *FAO Plant Protection Bulletin*, **42(1/2)**, 3-32.

Byrne DN (1999) *Agricultural and Forest Meteorology*, **97,** 309-316.

Caudill CC (2003) *Oikos*, **101(3)**, 624-630.

de Bruin MH (1983) *Nature*, **304**, 234-240.

Dellaporta SL, Wood J and Hicks JB (1983) *Plant Molecular Biology Reporter*, **1(4)**, 19-21.

Crozier RH, Crozier YC and Mackinlay AG (1989) *Molecular Biology and Evolution*, **6(4)**, 399-411.

de Bruin M (1983) *Nature* **304**, 234–241.

Dellaporta SL, Wood J and Hicks JB (1983) *Plant Molecular Biology Reporter*, **1(4)**, 19-21.

Doums C, Cabrera H and Peeters C (2002) *Molecular Ecology*, ***11*(11)**, 2251-2264.

Fearnley IM and Walker JE (1987) *Biochemistry*, **26(25)**, 8247-8251.

Frati F, Simon C, Sullivan J and Swofford DL (1997) *Journal of Molecular Evolution*, ***44*(2)**, 145-158.

Fu YX and Li WH (1993) *Genetics* **133**, 693–709.

Galtier N, Enard D, Radondy Y, Bazin E and Belkhir K (2006) *Genome Research*, **16(2)**, 215-222.

Goudet J, Perrin N and Waser P (2002) *Molecular Ecology*, ***11*(6)**, 1103-1114.

Gros A, Hovestadt T and Poethke HJ (2008) *Ecological Modelling*, ***219*(1)**, 226-233.

Hepburn HR, Smith DR, Radloff SE and Otis GW (2001) *Apidologie*, ***32*(1)**, 3-24.

Hepburn HR, Smith DR, Radloff SE and Otis GW (2001) *Apidologie*, **32(1)**, 3-24.

Horn A, Roux-Morabito G, Lieutier F and Kerdelhue C (2006) *Molecular Ecology*, **15(6)**, 1603-1615.

Hufbauer RA, Bogdanowicz SM and Harrison RG (2004) *Molecular Ecology*, ***13*(2)**, 337-348.

Kindler E, Arlettaz R and Heckel G (2012) *Molecular Phylogenetics and Evolution*, **65**, 695–704.

Kyte J and Doolittle RF (1982) *Journal of Molecular Biology*, **157(1)**, 105-132.

Lagisz M, Wolff K, Sanderson RA and Laskowski R (2010) *Journal of Insect Science*, ***10***.

Lin CP and Danforth BN (2004) *Molecular Phylogenetics and Evolution*, **30(3)**, 686-702.

Lowe TM and Eddy SR (1997) *Nucleic acids research*, **25(5)**, 0955-964.

Lunt DH, Zhang DX, Szymura JM and Hewltt OM (1996) *Insect Molecular Biology*, **5(3)**, 153-165.

Markow TA and Castrezana S (2000) *Oikos*, **89(2)**, 378–386.

Memon N, Meier R, Manan A and Su KFY (2006) *Systematic Entomology*, **31(4)**, 703-710.

Nardi F, Carapelli A, Dallai R and Frati F (2003) *Insect Molecular Biology*, **12**, 605–611.

Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Ojala D, Montoya J and Attardi G (1981) *Nature*, **290**, 470–474.

Orsini L, Koivulehto H and Hanski I (2007) *Cladistics*, **23**, 145-168.

Roderick GK (1996) *Annual Review of Entomology*, ***41*(1)**, 325-352.

Roe AD and Sperling FA (2007) *Molecular Phylogenetics and Evolution*, **44(1)**, 325-345.

Rozas J, Sanchez-DelBarrio JC, Messeguer X and Rozas R (2003) *Bioinformatics*, **19(18)**, 2496-2497.

Salle A, Arthofer W, Lieutier F, Stauffer C and Kerdelhue C (2007) *Biological Journal of the Linnean Society*, ***90*(2)**, 239-246.

Sheffield NC, Song H, Cameron SL and Whiting MF (2008) *Molecular Biology and Evolution*, **25(11)**, 2499-2509.

Sielezniew M, Ponikwicka D, Ratkiewicz M, Rutkowski R, Dziekanska I and Kostro-Ambroziak A (2011) *European Journal Entomology*, **108**, 537-545.

Simons C, Frati F, Beckenbach A, Crespi B, Liu H and Floors P (1994) *Annals of the Entomological Society of America*, **87(6)**, 651-701.

Song H, Sheffield NC, Cameron SL, Miller KB and Whiting MF (2010) *Systematic Entomology*, **35(3)**, 429-448.

Sun Z, Wan DG, Murphy RW, Ma L, Zhang XS and Huang DW (2009) *Genes & Genomics*, ***31*(1)**, 65-71.

Sundstrom L, Keller L and Chapuisat M (2003) *Evolution*, ***57*(7)**, 1552-1561.

Szalanski AL, Roehrdanz RL and Taylor DB (2000) *Florida Entomologist*, 262-267.

Tamura K (1992) *Molecular Biology and Evolution*, **9(5)**, 814-825.

Tamura K, Dudley J, Nei M and Kumar S (2007) *Molecular Biology and Evolution*, **24(8)**, 1596-1599.

Thompson JD, Desmond GH and Toby JG (1994) *Nucleic Acids Research*, **22.22**, 4673-4680.

Wan Q-H, Wu H, Fujihara T and Fang S-G (2004) *Electrophoresis*, **25**, 2165-2176.

Werren JH and Windsor DM (2000) *Proceedings of the Royal Society B: Biological Sciences*, **267**, 1277-1285.

Xia X (1996) *Genetics*, **144(3)**, 1309-1320.

Xia X, Xie Z and Li WH (2003) *Journal of Molecular Evolution*, **56(3)**, 362-370.

Xia X and Lemey P (2009) The Phylogenetic Handbook: A Practical Approach to DNA and Protein Phylogeny. 2nd edition Cambridge University Press.

Zhang J (2000) *Journal of Molecular Evolution*, ***50*(1)**, 56-68.

Zhou WF, Rousset F and O'Neill S (1998) *Proceedings of the Royal Society B: Biological Sciences,***265,** 509-515

CHAPTER 6

Characterization of potential genes for the control of banana stem weevil: An insight into the study of the specificity of interaction between the amylase of *O. longicollis* (Oliver) with the wheat monomeric and dimeric α-amylase inhibitors, by homology modeling

SUMMARY

α-amylases and proteinases are important digestive enzymes which play a central role in the digestive metabolism in those insects that live on seeds and plant parts. The larvae and adults of banana pseudostem weevil feed voraciously on the stem parts and complete their development within the stem. Hence they depend to a large extent on their proteinases and α-amylases for their survival. The purified monomeric and dimeric α-amylase inhibitor fractions from a local variety of wheat seeds completely inhibit the α amylase activity from this pest and do not inhibit porcine pancreatic amylase and human salivary amylase. These characteristics make these inhibitors attractive candidates for genetic engineering to develop banana transgenics with improved tolerance towards banana pseudostem weevil. This chapter describes **(i)** the cloning of the partial gene encoding the amylase (the cloned segment includes all the conserved regions and the catalytic domain) **(ii)** cloning of the genes encoding the monomeric and dimeric α-amylase inhibitors from a local variety of wheat and **(iii)** study of the molecular basis of the specificity of inhibition between the α-amylase and the inhibitors by homology modeling. Such a study is the first step towards developing banana transgenics using such genes.

INTRODUCTION

Proteinaceous enzyme inhibitors from plants have been extensively studied, because they play a role in host-plant resistance to insect and microbial pests (Yetter *et al.* 1979; Gatehouse *et al.* 1986; Silano 1987; Ryan 1984, 1990; Richardson 1991). α-amylases are essential enzymes for insect growth and development, and hence inhibitors of α-amylases may have detrimental effects on the insect's life cycle when present in the diet. Wheat kernels are particularly rich in inhibitors of α-amylases of insects and mammals (Kneen & Sandsted 1943, 1946; Buonocore & Silano 1986; Baker 1988a; Baker & Lum 1989; Gomez *et al.* 1989; Feng *et al.* 1991a). Some of these are selective for insect enzymes, in that they inhibit α-amylases from insects strongly but inhibit mammalian salivary or pancreatic a-amylases only weakly or not at all. Insect-selective α-amylase inhibitors in wheat are primarily monomeric proteins (Silano *et al.* 1975), and, based on an electrophoretic mobility of 0.28, these proteins have been called the 0.28 group of inhibitors (Silano 1987). The insect-selectivity, monomeric architecture, and small size (approx. 14 kDa) of these inhibitors make their genes attractive candidates for introduction into plants to increase resistance to attack by insect and microbial pests. As these α-amylase inhibitors could function as insect control proteins, it is important to elucidate structure-function relationships in these proteins, and also understand the physico-chemical basis for the selectivity exhibited by individual α-amylase inhibitors. Such studies would be able to throw light on how the structures of these proteins determine their patterns of inhibition against different α-amylases. The selectivity exhibited by a particular inhibitor can be much more subtle than simply distinguishing between insect and mammalian α-amylases,

α-amylases from different insect species differ substantially in their sensitivity to particular inhibitor fractions (Feng *et al.* 1991a, b). For example, the two α-amylases from the rice weevil exhibited marked differences in their sensitivities to inhibitor fractions or purified inhibitors (Baker 1987, 1988b, 1989; Feng *et al.* 1991a; Chen *et al.* 1992). Another example is the well characterized inhibitor α AI-1 which inhibits porcine pancreatic amylase (PPA), human salivary amylase (HAS) and the α-amylases of several bruchids but not the α-amylase of the Mexican bean weevil *Zabrotes subfasciatus* (ZSA)

(Powers & Whitaker 1977; Powers & Culbertson 1983; Ishimoto & Kitamura 1989). Similarly, the α-amylase inhibitor isolated from amaranthus seeds, AAI, strongly inhibits the larvae of red flour beetle (*Tribolium castaneum*), the yellow meal worm (*Tenebrio molitor*) and the grain borer (*Prostephanus truncatus*) but does not inhibit mammalian α-amylases or proteases (Chagolla-Lopez *et al.* 1994). There is a much greater abundance of the inhibitors in wheat than in rice (Baker 1988a; Feng *et al.* 1991a, b)

Because of such a specificity of interaction between the proteinaceous inhibitors and the amylases, use of a genetic engineering strategy requires purification and characterization of the α-amylases of the target insects as well as of the potential inhibitor as well as an understanding of the structure to function relationship. Such studies would require assembling a group of inhibitors whose structures are known and whose patterns of inhibition against a repertoire of insect and mammalian enzymes is known. An improved understanding of the structural basis of the interaction between the α-amylases from *O. longicollis* and the potential inhibitor will help to design inhibitors with more desirable characteristics.

In the present work, the purified monomeric and dimeric α-amylase inhibitors from a local variety of wheat seeds are shown to completely inhibits the α amylase activity of this pest. The dimeric α-amylase inhibitor from this source is also specific for the insect amylase and does not inhibit mammalian amylases unlike the reported dimeric α-amylase inhibitors (Sharma *et al*. 2009). These characteristics would make these inhibitors ideal candidates for use of genetic engineering approaches for the development of transgenic bananas with improved tolerance towards this pest. As a first step towards this long term objective, the partial gene encoding the amylase of this pest as well as the genes encoding the monomeric and dimeric α-amylase inhibitors have been cloned and the interaction between the amylase and these inhibitors has been studied using homology modeling.

MATERIALS AND METHODS

Collection of samples

Live larvae of *O. longicollis* (Oliver) were collected from infested banana pseudostems from banana fields in Trichy, Tamil-Nadu, India, and stored at −70°C till further use.

Preparation of crude α-amylase from *O. longicollis* (Oliver) larvae

One gram of adult insect larvae was crushed in liquid nitrogen using a mortar and pestle. The powdered material was suspended in five volumes of 150 mM NaCl or 150 mM Tris-HCl (pH 7.5) and centrifuged at 10,000 rpm for 10 min at 4°C. The pellets were discarded and the supernatant was assayed for enzyme activity. It was observed that extraction with 150 mM NaCl yielded 37% more enzyme activity as compared with 150 mM Tris-HCl pH 7.5. So, all further extractions were carried out in 150 mM NaCl. The NaCl extract containing the crude amylase enzyme was dialyzed against 10 mM sodium-phosphate buffer (pH 7.0) and then concentrated by lyophilization to one fourth its volume and then re-dialyzed against the same buffer. The recovery on lyophilization was 88%.

Assay for amylase activity

The amylase activity was determined by the dinitrosalicylic acid (DNS) method (Bernfield *et al.* 1955). The reaction mixture contained 0.25 ml of suitably diluted enzyme in sodium-phosphate buffer (50 mM, pH 8.0) and 0.25 ml of (1% w/v) starch solution for amylase activity. The reaction was terminated by the addition of 0.5 ml DNSA. The tubes were heated in a boiling water bath for 5 min and the color intensity was read at 540 nm after dilution with 5 ml of distilled water. A standard glucose curve was used for calculating enzyme activities. The linearity curve of the enzyme was

determined for calculation of unit activity. One IU of amylase activity is defined as the amount of enzyme required to produce 1 μmole of glucose/min.

Isolation and purification of the monomeric and dimeric α-amylase inhibitors from wheat (Variety: MP Sehore)

Extraction of wheat α-amylase inhibitors was carried out as described by Petrucci *et al.* (1974). In a preliminary experiment it was observed that a 150 mM NaCl solution was more effective in extraction of the α-amylase inhibitors from wheat flour than 100 mM Tris/HCl buffer pH 7.6. Hence, 150 mM NaCl was used for extraction of the α-amylase inhibitors. 100 g of finely ground whole wheat flour (Variety MP Sehore) was stirred in 300 ml of 150 mM NaCl solutions at room temperature for 3 h. The suspension was then centrifuged for 10,000 rpm for 20 min. at 4°C. The supernatant liquid was subjected to ammonium sulphate fractionation to isolate the α-amylase inhibitors.

Ammonium sulphate precipitation of α-amylase inhibitors

The crude α-amylase inhibitor solution was first precipitated with 0–0.4 M $(NH_4)_2SO4$ to remove non-proteinaceous components. The resultant supernatant liquid after centrifugation was salted out with 0.4–2.0 M $(NH_4)_2SO4$. The precipitate was collected by centrifugation at 10,000 rpm for 10 min and dissolved in one tenth the volume of double distilled water and dialyzed against water for 24 h with four changes. The inhibitor fraction was concentrated by lyophilization. Since the crude preparation had endogenous α-amylase activity, the inhibitor solution was boiled for 1 min at 100°C to inactivate the endogenous α-amylase activity.

Purification of α-amylase inhibitors by FPLC Superose-12 Gel permeation column chromatography

The crude ammonium sulphate fraction of the wheat α-amylase inhibitor containing 6 mg protein and 7000 units of inhibitor activity in 300 μl was loaded on a

FPLC-Superose 12 gel permeation column (10 × 300 mM). Phosphate buffer pH 7.4 of composition 137 mM-NaCl, 3 mM-KCl, 1 mM-KH_2PO_4, 8 mM Na_2HPO_4 was used for elution. The flow rate of the column was maintained at 18.5 ml/h. A total of twenty-five fractions of one ml each were collected and the individual fractions were assayed for inhibitor activity. The FPLC profile revealed the presence of five distinct peaks (**Fig. 1a**). Inhibitor activity was detected only in peak three and no activity was detected in any of the other peaks. In order to separate the dimer and monomer, the individual FPLC fractions of peak three were collected from several runs, dialyzed, lyophilized and loaded on Sephadex G-50 gel permeation column chromatography. The buffer system used was the same as described for FPLC. The flow rate of the column was maintained at 15 ml/h. Two peaks were obtained as shown in **Fig. 1b**. The two peak fractions were analyzed by native PAGE. The two samples were found to be homogenous and corresponded to the dimer and monomer (**Fig. 1c**). Other FPLC fractions were processed similarly.

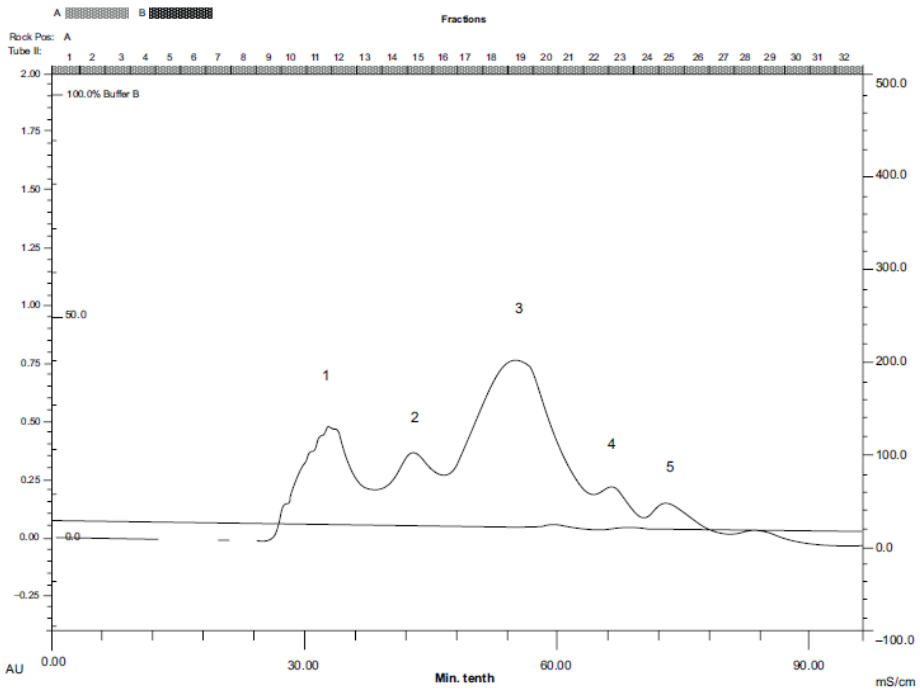


Fig. 1a. Profile of FPLC-gel permeation column chromatography (Superose 12 prep. Grade)

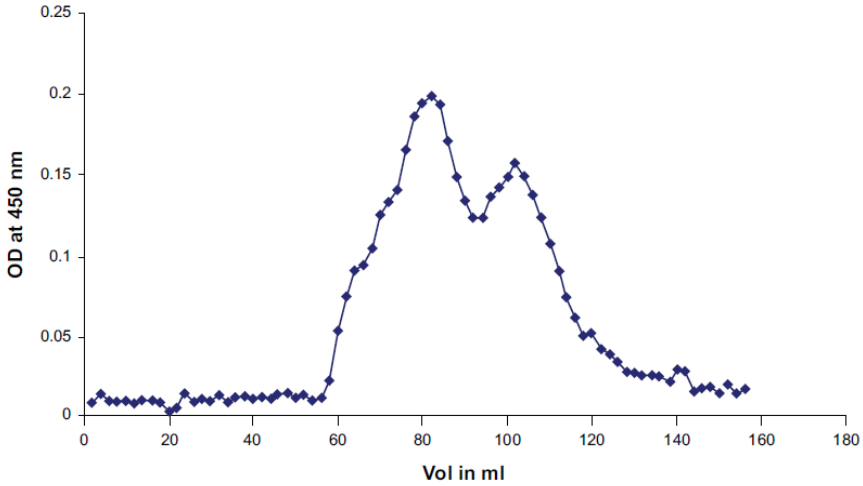


Fig. 1b. Profile of G 50 gel permeation column chromatography

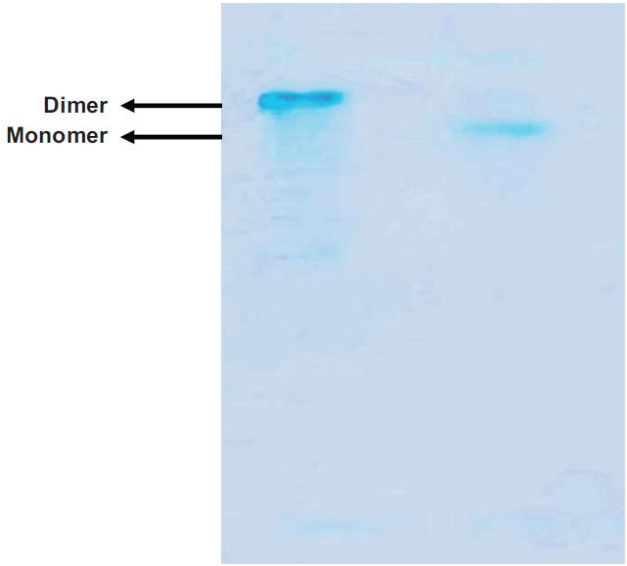


Fig. 1c. Native PAGE of purified monomeric and dimeric α-amylase inhibitors

Assay for inhibitor activity

The assay for determining inhibitor activity was carried out essentially as described for amylase activity excepting that the reaction mixture contained suitable aliquots of inhibitor. One unit of inhibitor activity is defined as the amount of inhibitor required to inhibit 50% of enzyme activity. Commercial dimeric α-amylase inhibitor from

wheat (Sigma Catalogue No. A1520) was used as a control as it inhibits human salivary α-amylase (Sigma Catalogue No. A0521) and porcine pancreatic α-amylase (Sigma Catalogue No. A3176).

PAGE electrophoresis

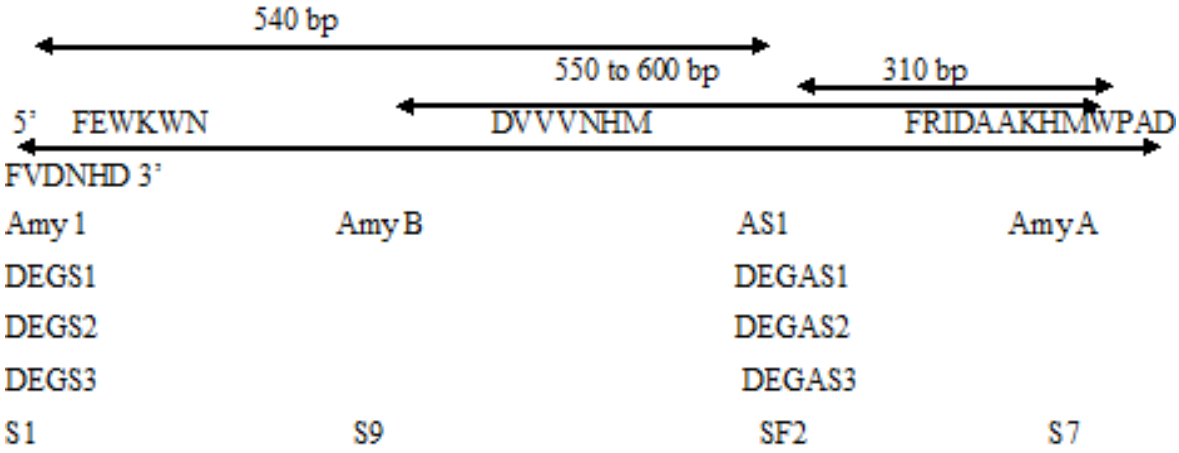
Non-denaturing gel electrophoresis was performed according to the procedure described by Davis *et al.* (1964) using a 10% separating gel at pH 8.9. Bromophenol blue was used as the tracking dye.

Molecular weight determination of the inhibitors by MALDI-TOF

After electrophoresis the protein bands were excised from the gel and transferred into two separate vials. Sufficient extraction solution (100 µl) consisting of a mixture of formic acid/glass distilled water/2-propanol (1:3:2) was added to the crushed gel to completely cover the gel pieces. The tubes were closed and vigorously shaken at room temperature for 4 to 8 h. After vortexing, the sample was microfuged. The supernatant was retrieved with a micropipette. The crushed gel was washed once with an equal volume of fresh extraction solution and the washes were combined with the supernatant. The two supernatants were concentrated and subjected to MALDI-TOF (Cohen & Chait 1997). MALDI-TOF analysis showed that the molecular weights of the dimer and monomer were 13,330 and 13,159 Daltons respectively, suggesting that band 1 was a dimer and band 2 was a monomer.

Cloning of the gene encoding the α-amylase from *O. longicollis* (Oliver)

The schematic representation of the conserved regions in α-amylases and the expected size of the amplicons is shown below:



For cloning the gene encoding the α-amylase from *O. longicollis* (Oliver), primers were designed based on the conserved region sequences of α-amylase genes (**Table 1**).

Table 1. Primers used in the cloning of the α-amylase gene from *O. longicollis* (Oliver), based on conserved regions of α-amylases

Primer	Forward (5'-3')	Reverse (5'-3')	Expected Size of the amplicon (bp) (Tm ^{o}C)
SF2-S7	TCCGGATCGACGCCGCCAA	TCCGGATCGACGCCGCCAA	310 (66^{o}C)
Amy1-AS1	ATCGTGCAGCTGTTCGAGTGG	CATGTGGCCGGCCGAC	540 (55^{o}C)

Total RNA was isolated using RNeasy kit (Qiagen, Catalogue number 74104) as per the manufactures' instructions. cDNA was made using Reverse transcriptase from Invitrogen (SuperScript^{TM} III One-Step RT-PCR System with Platinum^{R} Taq DNA

Polymerase Kit, Catalogue number 12574-018). Primers SF2/S7 and Amy1/AS1 amplified fragments of expected size ~310 bp and ~540 bp (**Figs 2a and b**).

PCR program

cDNA synthesis
1 cycle: 45–60°C for 15–30 minutes
Denaturation
1 cycle: 94°C for 2 minutes

PCR amplification
40 cycles: 94°C for 15 seconds (denature)
55-65°C for 30 seconds (anneal)
68°C for 1 minute per kilobase (extend)

Final extension (optional)
1 cycle: 68°C for 5 minutes

These fragments were cloned in pCR4 TOPO vector (TOPO TA Cloning^{R} Kit for sequencing, Catalogue number K4575-01) and plasmid DNA was isolated from positive clones using the plasmid DNA isolation kit (QIAprep^{R} Spin Miniprep Kit (50), Catalogue number 27104) as per the manufacturer's instructions. The derived sequence matched with α amylase gene sequences from insects. The stretch of the amylase gene delimited by primers SF2-S7 and Amy1-AS1 was 310 bp and 540 bp long respectively (**Figs. 2a and b**). The gene sequence has been submitted to GenBank with accession number KJ146681) (**Fig. 3**).

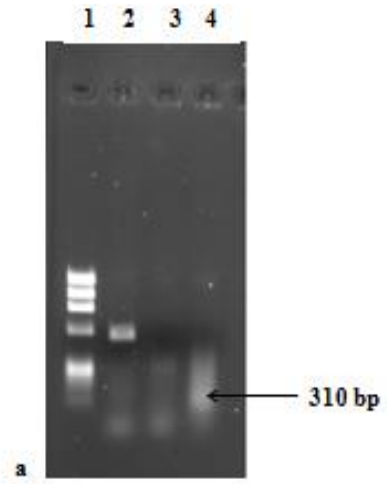


Figure. a) Lane 1- ϕX17 marker
Lane 4- 310 bp fragment of *O. longicollis*
α-amylase gene obtained after RT reaction

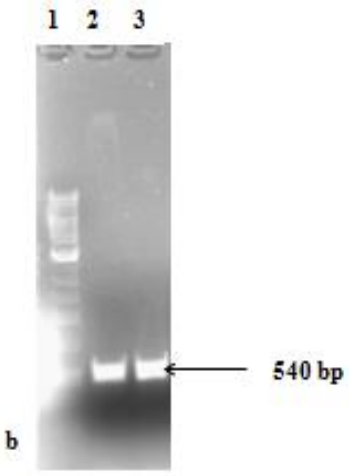


Figure. a) Lane 1- ϕX17 marke
Lane 2 and 3- 540 bp fragment of *O. longicollis*
α-amylase gene obtained after RT reaction

Fig. 2a and b. PCR-amplified and cloned fragment of banana pseudostem weevil amylase

```
 10 ctgttcgagtggaaacacgtcgatgtagcggcggaatgtgagaat
    L  F  E  W  K  H  V  D  V  A  A  E  C  E  N
 55 ttcctcagtcaaaaaggatatggaggagttcaagtttcaccgcca
    F  L  S  Q  K  G  Y  G  G  V  Q  V  S  P  P
100 agtgaaaacgctatagtcgaaggtcgtccatggtgggaaaaatat
    S  E  N  A  I  V  E  G  R  P  W  W  E  K  Y
145 caaccagtcagctatgttttaaataaccgagctggagatgaagac
    Q  P  V  S  Y  V  L  N  N  R  A  G  D  E  D
190 gcttttgccggcatggttaaacgttgtaatgctgttggaattaga
    A  F  A  G  M  V  K  R  C  N  A  V  G  I  R
235 atttatatagatctggttgctaatcatatggcaacttcaagcggt
    I  Y  I  D  L  V  A  N  H  M  A  T  S  S  G
280 gaaggtactgctggagatacctgcgatcctggcagcaaatcatat
    E  G  T  A  G  D  T  C  D  P  G  S  K  S  Y
325 ccagctgtatcatattcaagtgaaaacttccatccttcttgtgac
    P  A  V  S  Y  S  S  E  N  F  H  P  S  C  D
370 attgactacaacgacgcagcttccatcagaaattgtgaactctct
    I  D  Y  N  D  A  A  S  I  R  N  C  E  L  S
415 ggcttgaaggacttggatcaaagtcaggattatgtcaggggtaaa
    G  L  K  D  L  D  Q  S  Q  D  Y  V  R  G  K
460 atcattgattacctaaaccatcttgtcgaccttggtgtagccgga
    I  I  D  Y  L  N  H  L  V  D  L  G  V  A  G
505 ttccgtgtggacgcggcgaaacacatcccggcccagcacctggag
    F  R  V  D  A  A  K  H  I  P  A  Q  H  L  E
550 cagatccgcgcccgggtcggcgacggatcggtctactggaagcag
    Q  I  R  A  R  V  G  D  G  S  V  Y  W  K  Q
595 gagatgatctacggggcgggcgaggcgatccaccccgacgagtac
    E  M  I  Y  G  A  G  E  A  I  H  P  D  E  Y
640 acgggcgcgggcgacgtgcaggagttccgccacgccttcgacatc
    T  G  A  G  D  V  Q  E  F  R  H  A  F  D  I
```

```
685 aagcgcatctcccagaacgaacgcctcgcgtacctggacgacttc
    K  R  I  S  Q  N  E  R  L  A  Y  L  D  D  F
730 ggcccagagctgggggctacctgcccgtcccgcccggcgcccggc
    G  P  E  L  G  A  T  C  P  S  R  P  A  P  G
775 gtcttcgtcgacaaccacgaa 795
    V  F  V  D  N  H  E
```

Fig. 3. Partial gene and translated protein sequence of banana pseudostem weevil amylase (Genbank accession no. KJ146681**)**

Cloning of the genes encoding the monomeric and dimeric inhibitor from wheat (Variety: MP Sehore)

Forward and reverse primers were designed according to Wang *et al.* (2006) for amplifying the genes encoding the monomeric and dimeric α-amylase inhibitors from wheat (variety MP Sehore),) (**Table 2**).

Table 2. Primers used to amplify the monomeric and dimeric α-amylase inhibitor genes

Primer	Forward (5'-3')	Reverse (5'-3')	Expected size of the amplicon (bp) (Tm^oC)
WMAIF2-WMAIR2	ATGTGGATGAAGACCGTG	CACGCACCGCACCATTACTT	532 (58^oC)
WDAIF1-WDAIR1	CTATGTATGCTCGTGGCGAC	GCTACTCATTCGCTTGACTAGAC	436 (60^oC)

Total DNA was isolated from wheat seedlings using the protocol of Dellaporta (1983) and was used as a template for the PCR reaction, the PCR program as follows:

PCR program

Initial denaturation 94^0C– 5 mins

Denaturation 94^0C – 30 secs

Annealing 58^0C and 60^0C - 45 secs } 25 cycles

Extensions 72^0C – 1 min

Final Extension 72^0C – 10 min

The size of the amplicons obtained were 535 bp and 436 bp for the wheat monomeric and dimeric α-amylase inhibitors (**Figs. 4a and b**). These fragments were cloned in pCR4 TOPO vector (Invitrogen) and transformed into Top10 *E. coli* cells as per the manufacturer's instructions. Plasmid DNA was isolated from positive clones using the plasmid DNA isolation kit (QIAprep^{R} Spin Miniprep Kit (50), Catalogue number 27104) as per the manufacturer's instructions. The derived sequence matched with reported genes sequences of monomeric and dimeric wheat α-amylase inhibitors (**Figs. 5a and b**). These sequences have been submitted to GenBank with Accession numbers KJ146680 and GQ374443 respectively.

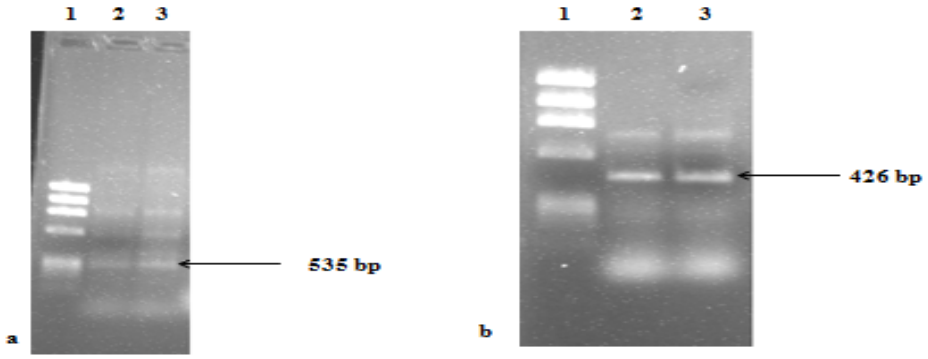


Figure. a) PCR products with WMAIF2 and WMAIR2 primers
Lane 1- ϕX174 marker
Lane 3- 535 bp WMAI PCR product.

Figure. b) PCR products with WDAIF1 and WDAIR1 primers
Lane 1- ϕX174 marker
Lane 2, 3- 426 bp amplification product

Fig. 4 a and b. PCR-amplified products of WMAI and WDAI

```
  1 atgtggatgaagaccgtgttctgggggctcctagtattcatgctc
    M  W  M  K  T  V  F  W  G  L  L  V  F  M  L
 46 gtggcgacaacaatggcggtcgagtatggtgcaaggagccataac
    V  A  T  T  M  A  V  E  Y  G  A  R  S  H  N
 91 agtggtccttggagttggtgcgatccggcgacgggctacaaggtg
    S  G  P  W  S  W  C  D  P  A  T  G  Y  K  V
136 agcgcactcacgggctgccgggcaatggtgaagctccagtgtgtg
    S  A  L  T  G  C  R  A  M  V  K  L  Q  C  V
181 ggcagtcaggtgcccgaggctgtcctaagagattgctgccagcag
    G  S  Q  V  P  E  A  V  L  R  D  C  C  Q  Q
226 ctggccgacatcaacaacgaatggtgcaggtgcggggacctcagc
    L  A  D  I  N  N  E  W  C  R  C  G  D  L  S
271 agcatgttgcgtagtgtttatcaggagctcggcgtgcgtgagggg
    S  M  L  R  S  V  Y  Q  E  L  G  V  R  E  G
316 aaggaggtgctcccaggttgccggaaggaggtgatgaagctcacg
    K  E  V  L  P  G  C  R  K  E  V  M  K  L  T
361 gcggcgagcgtgcctgaggtctgcaaggtgcccattcccaacccg
    A  A  S  V  P  E  V  C  K  V  P  I  P  N  P
```

```
406 tcgggagacagagcaggtgtctgctactgggccgcgtatcctgac
    S  G  D  R  A  G  V  C  Y  W  A  A  Y  P  D
451 gtctag 456
    V  *
```

Fig. 5a. Sequence of WMAI (Genbank accession no.KJ146680**)**
(pdb id AHN92451.1)

```
  1 atgctcgtggcgacacccatagcggccgagtacgacgcatggagc
    M  L  V  A  T  P  I  A  A  E  Y  D  A  W  S
 46 gttaacagtggtccctggatgtgctatccagggtatgcctttaag
    V  N  S  G  P  W  M  C  Y  P  G  Y  A  F  K
 91 gtgccagcgctccctggctgtcgtccagtgctgaagctccagtgc
    V  P  A  L  P  G  C  R  P  V  L  K  L  Q  C
136 aatggcagccaggtgcccgaggctgtcctaagggactgctgccag
    N  G  S  Q  V  P  E  A  V  L  R  D  C  C  Q
181 cagctcgcccacatcagcgagtggtgcaggtgcggggccctctac
    Q  L  A  H  I  S  E  W  C  R  C  G  A  L  Y
226 agcatgttggacagcatgtataaggagcatggcgcgcaggaggga
    S  M  L  D  S  M  Y  K  E  H  G  A  Q  E  G
271 caggcagggacaggagcgttcccacgctgccggagggaggtggtg
    Q  A  G  T  G  A  F  P  R  C  R  R  E  V  V
316 aagctgacggcggcgagcatcacggcggtctgcaagctacccatc
    K  L  T  A  A  S  I  T  A  V  C  K  L  P  I
361 gtcattgatgcgtctggagatggagcgtatgtctgcaagggtgtg
    V  I  D  A  S  G  D  G  A  Y  V  C  K  G  V
406 gccgcatacccggacgtctag 426
    A  A  Y  P  D  V  *
```

Fig. 5b. Sequence of WDAI (Genbank accession no.GQ374443**)**

(pdb id ACU51027.1)

Homology Modeling

Homology models were built for the partial amylase, wheat monomeric α-amylase inhibitor (WMAI), wheat dimeric α-amylase inhibitor (WDAI), amylase-WMAI complex and amylase-WDAI complex. For the identification of template for building the homology model of the partial sequence of *O. longicollis* amylase and its monomeric and dimeric inhibitors from wheat, BLAST search was performed against the PDB database. The software Prime was used for building the homology model for both the proteins. The model obtained was refined by carrying out energy minimization using Desmond.

The cloned partial gene sequence of *O. longicollis* amylase, wheat monomeric and dimeric α amylase inhibitors were translated into the protein sequence using the 'ORF Finder' program at the NCBI site (http://www.ncbi.nlm.nih.gov/projects/gorf). The signal sequence in WMAI and WDAI was determined by SignalP version 3.0 (http://www.cbs.dtu.dk/services/SignalP/). Sequence analysis was performed with the

BLAST program of the National Centre for Biotechnology Information (NCBI), NIH, Bethesda, MD, USA. Sequence alignments were obtained using the online tools of ClustalW server (http://www.ebi.ac.uk/Tools/msa/clustalw2/). The phylogenetic analysis was performed using the the online BLAST online tool (http://www.ncbi.nlm.nih.gov/blast/treeview/treeView.cgi) and Newick Tree format was downloaded and tree was viewed and saved in MEGA v 5.2. The three-dimensional structure of the amylase of *O. longicollis* Oliver (Copleoptera: Curculionidea) was predicted from its amino acid sequence by homology modeling (PDBSUM). The structural model which was found to be the closest to this amylase was that of the amylase from *T. molitor* L. (pdb: 1TMQ A) Chain A. The geometry of the predicted model was checked using the program PROCHECK.

RESULTS AND DISCUSSION

Sequence analysis of the partial amylase gene from *O. longicollis* (Oliver)

The cloned partial sequence of *O. longicollis* (Oliver) amylase gene was 795 bp long and has been submitted to GenBank with accession number KJ146681 (**Fig. 3**). The derived sequence translated into a protein of 265 amino acids.

ClustalW alignment of the partial amylase sequence of *O. longicollis* showed significant homology with the α-amylases from other insects (*Helicoverpa armigera, Ostrinia nubilalis, Scirpophaga incertulas, Diatraea saccharalis, Tenebrio molitor)* as well as with porcine pancreatic amylase (PPA) and human salivary amylase (HSA) (**Fig. 6**).

```
gi|156968285|gb|ABU98614.1|         --MFRLILLLAAVSLALAYKNPHYASGRTTMVHLFEWKWDDIAAECERFL 48
gi|438387|gb|AAA03715.1|            --MLRFVVLLAGLALTLAFKNPHYSGDRTTMVHLFEWKWDDIADECERFL 48
gi|157154545|gb|ABV24963.1|         MIPLILLAVAVFARAANDYKNPYYVPGRSVNVHLFEWKWEDIAEECERFL 50
gi|33151028|gb|AAP97393.1|          MVKLLIVALVALAVFANAYKNPHYAPGRSVNVHLFEWKWDDIATECENFL 50
gi|5902775|sp|P56634.1|AMY_TEN      ------------------QKDANFASGRNSIVHLFEWKWNDIADECERFL 32
gi|6056338|gb|AAF02828.1|AF064      ---MKLFLLLSAFGFCWAQYAPQTQSGRTSIVHLFEWRWVDIALECERYL 47
gi|15988376_10-396                  ---------------------------RTSIVHLFEWRWVDIALECERYL 23
O.longicollis                       -------------------------------IVQLFEWKHVDVAAECENFL 20
                                                                  *:****:  *:* ***.:*

gi|156968285|gb|ABU98614.1|         GPRGYGGIQISPPNENLAIWSANRPWWERYQPISYRLVTRSGNEQQFASM 98
gi|438387|gb|AAA03715.1|            GPNGFGGIQISPPNENLIIRAHNRPWWERYQPMSYRLITRSGNEQQFTNM 98
gi|157154545|gb|ABV24963.1|         GPNGFGGVQISPPNENVIIWSANRPWWERYQPISYKLTTRSGDNTQLANM 100
gi|33151028|gb|AAP97393.1|          GPRGFGGIQISPPNENVVLWTYNRPWWERYQPMSYLLDTRSGDEAQFADM 100
```

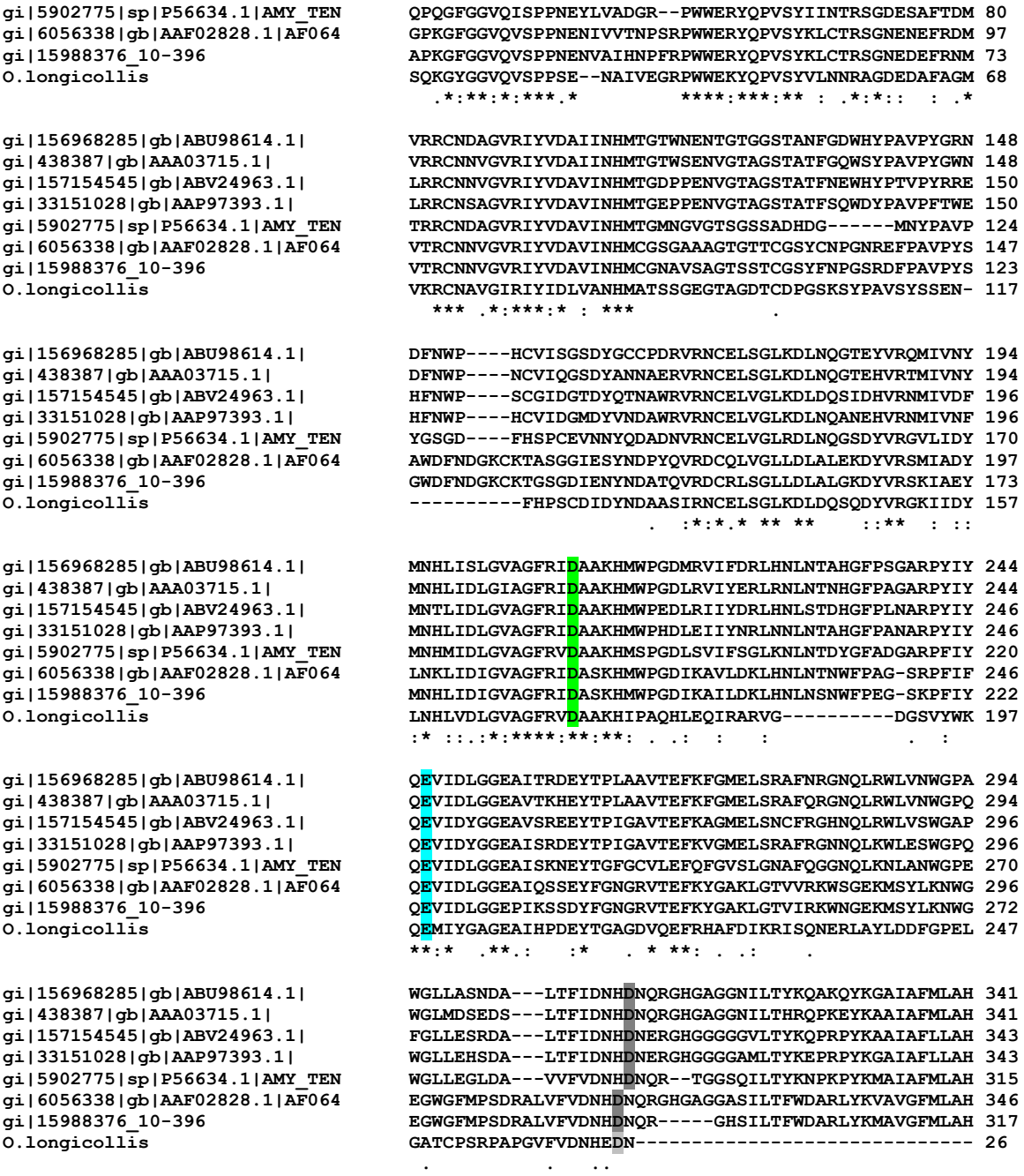


Fig. 6. Clustal of *O. longicollis* (Oliver) amylase with amylase sequences from other organisms

Gene bank accession numbers of organisms are indicated in the clustal: ABU98614.1: *Helicoverpa armigera;* AAA03715.1: *Ostrinia nubilalis;* ABV24963.1: *Scirpophaga incertulas;* AAP97393.1: *Diatraea saccharalis;* P56634.1: *Tenebrio molitor;* AAF02828.1: Porcine pancreatic amylase; 15988376: Human salivary amylase

From this Fig it is seen that the functionally significant residues present in the other amylases are also present in the *O. longicollis* amylase. These include Asp 171 and Glu 199 which are involved in catalysis (Machius *et al.* 1996); His 87 and His 176 are involved in substrate binding; Asn 85, Arg 133, Asp 142 and His 176 are involved in Ca2+ binding and Arg 170 is involved in Cl- binding. The role of chloride in α-amylase catalytic activity has been established by various researchers (Levitzki & Steer 1974; Feller *et al.* 1996; Aghazari *et al.* 1998).The chloride binding site of mammalian pancreatic and salivary α-amylases includes Arg-195, Asn–298, the side-chain amines of Arg-337 and a water molecule (Larson *et al.* 1994; Brayer *et al.* 1995; Qian *et al.* 1995; Ramasubbu *et al.* 1996). It has been observed that the basic residue Arg/Lys-337 is substituted by Glu in chloride-independent a-amylases which is active without chloride. This suggests that the chloride ion was not essential to the catalytic mechanism. In vertebrates, beetles and flies, the Cl- ion is coordinated by six ligands i.e. Arg 210 (PPA) and Arg 351(PPA) in a bidentate mode and Asn 312 (PPA) in a unidentate mode (Strobl *et al.* 1998a, b). This Arg 210 of PPA is present as Arg 170 in *O. longicollis* (Oliver) amylase (**Fig. 6**). Arg 351 of PPA has been found to be replaced by a lysine residue in bacterial a-amylase (Aghajari *et al.* 1998). The α-amylases of beetles and flies have either Arg or Lys in this position, which is suitable for activity in an acidic gut environment. This region is however not included in the derived partial sequence of *O. longicollis* amylase.

The amino acid composition of *O. longicollis* (Oliver) amylase is shown in **Table 3**. The percentage of hydrophobic residues (Ala, Val, Ile, Leu, Phe, Pro and Met) in the partial sequence is 37.8%. This value compares well with similar values of 37% for *Tenebrio molitor* α-amylase, 40% for *Helicoverpa armigera*, 35% for Human pancreatic α-amylase (HPA), 34% for Human salivary α-amylase (HSA) and 34% for Porcine pancreatic α-amylase (PPA).

Table 3. Amino acid composition of *O. longicollis* (Oliver) amylase

Amino acid		No. of Residues	
Alanine	Ala	A	25 (9.9%)
Arginine	Arg	R	13 (5.2%)
Aspargine	Asn	N	11 (4.4%)
Aspartic acid	Asp	D	21 (8.3%)
Cysteine	Cys	C	6 (2.4%)
Glutamine	Gln	Q	10 (4.0%)
Glutamic acid	Glu	E	16 (6.3%)
Glycine	Gly	G	24 (9.5%)
Histidine	His	H	8 (3.2%)
Isoleucine	Ile	I	14 (5.6%)
Leucine	Leu	L	13 (5.2%)
Lysine	Lys	K	10 (4.0%)
Methionine	Met	M	3 (1.2%)
Phenylalanine	Phe	F	7 (2.8%)
Proline	Pro	P	13(5.2%)
Serine	Ser	S	18(7.1%)
Threonine	Thr	T	5 (2.0%)
Tryptophan	Trp	W	3 (1.2%)
Tyrosine	Tyr	Y	13 (5.2%)
Valine	Val	V	19(7.5%)

Hydrophobic interactions have been suggested to contribute to the thermostability of proteins as evident by the value of the instability index (II) computed to be 34.68 (Aghajari *et al.* 1998). This value classifies the protein as stable. **Table 4** shows the codon usage statistics of the amylase from banana pseudostem weevil

Table 4. Codon usage statistics of the partial amylase from *O. longicollis*

Amino-acids represented by single letter code	Codon	Frequency of codon usage per 1000.	Amino-acids represented by single letter code	Codon	Frequency of codon usage per 1000
A	GCG	33.96	W	TGG	15.09
A	GCA	7.55	Y	TAT	26.42
A	GCT	33.96	Y	TAC	22.64
A	GCC	18.87	K	AAG	11.32
C	TGT	15.09	K	AAA	26.42

C	TGC	7.55	L	TTG	7.55
D	GAT	30.19	L	TTA	3.77
D	GAC	52.83	L	CTG	18.87
E	GAG	30.19	L	CTA	3.77
E	GAA	37.74	L	CTT	7.55
F	TTT	3.77	L	CTC	11.32
F	TTC	30.19	M	ATG	11.32
G	GGG	7.55	N	AAT	18.87
G	GGA	30.19	N	AAC	26.42
G	GGT	18.87	P	CCG	15.09
G	GGC	33.96	P	CCA	18.87
H	CAT	11.32	P	CCT	7.55
H	CAC	22.64	P	CCC	7.55
I	ATA	7.55	Q	CAG	26.42
I	ATT	15.09	Q	CAA	15.09
I	ATC	33.96	R	AGG	3.77
S	AGT	15.09	R	AGA	7.55
S	AGC	11.32	R	CGG	3.77
S	TCG	3.77	R	CGA	3.77
S	TCA	18.87	R	CGT	11.32
S	TCT	7.55	R	CGC	11.32
S	TCC	11.32	T	ACG	3.77
V	GTG	11.32	T	ACA	0.00
V	GTA	11.32	T	ACT	7.55
V	GTT	22.64	T	ACC	7.55
V	GTC	33.96			

Sequence analysis of the gene encoding WMAI

The derived WMAI gene sequence was 532 bp long which translated into a protein of 130 amino acids. Region 1-21 was predicted to be the signal peptide with SignalP 4.1 server with a prediction score of 0.945 (**Fig. 7**).

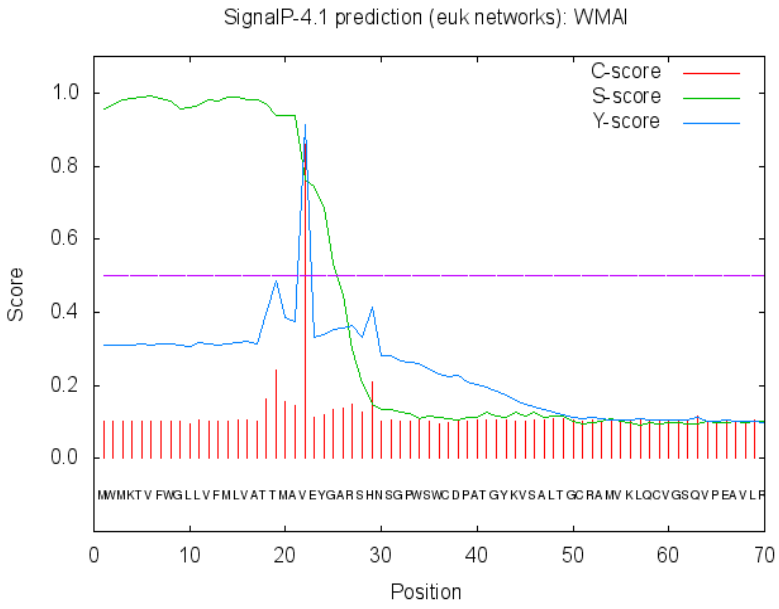


Fig. 7. Presence of signal peptide as predicted by SignalP-4.1 siftware

Fig. 8 shows the ClustalW alignment of WMAI (pdb id AHN92451.1) from different varieties of wheat. Various regions which are conserved across the varieties have been identified and highlighted. The residue highlighted in yellow is the residue where signal peptide is cut; residues highlighted in green are dimer interface (polypeptide binding) site; residues highlighted in blue are α-amylase binding (polypeptide binding) sites. In case of signal peptide sequence of *T. dicoccoides* (ACQ84008.1), it was observed that a hydrophobic amino acid Valine (V12) has been replaced with hydrophobic Leucine (L12) as compared with other sequences of wheat monomeric inhibitors compared here. Since this is a synonymous replacement, it does not affect the overall charge of the signal peptide. This comparison also points out the high homology among the wheat monomeric inhibitors reported from different sources.

Fig. 8. ClustalW of WMAI α-amylase inhibitor with monomeric α-amylase inhibitors form different varieties of wheat. Sequence (pdb id AHN92451.1) highlighted in gray is WMAI sequence used in this study. (Residue highlighted in yellow is the residue where signal peptide is cut; residues highlighted in green 69,119,123,129 are dimer interface (polypeptide binding) site; residues highlighted in blue 81..83,90 are α-amylase binding (polypeptide binding) site). GenBank accession numbers represents monomeric α-amylase inhibitor: ACQ83727.1 (*Triticum dicoccoides*); AHN92451.1 (*T. aestivum*); EMT18098.1 A-amylase inhibitor 0.28 (*Aegilops tauschii*); ACQ83969.1 (*T. dicoccoides*); ACQ83918.1 (*T. dicoccoides*); ACQ83859.1 (*T. dicoccoides);* CAA11410.1 (*T. aestivum*) ACQ84008.1 (*T. dicoccoides*).

Sequence analysis of the gene encoding WDAI

The cloned WDAI gene sequence was 426 bp long which translated into a protein of 141 amino acids. When the whole sequence of WDAI (pdb id ACU51027.1) was tested for the presence of signal peptide using SignalP software, the signal sequence could not be identified. Hence, the reported WDAI sequences from Chinese spring varieties *i.e.* AAY42613.1 and AAY42614.1 were used for comparison (Wang *et al.* 2006). It seems the SiganlP software has not been trained for identifying the signal sequence from this kind of proteins. But as reported in Wang *et al.* (2006), based on deduced amino acid sequences of the mature protein sequences of the dimeric α-amylase inhibitors 0.19 (ExPASy database: P01085) and 0.53 (ExPASy database: P01084) (Maeda *et al.* 1983, 1985), it was found that there were 51 and 372 bp sequences encoding the signal peptide and mature protein, respectively. On comparison of the nucleotide sequence of WDAI in the present work with the nucleotide sequences of wheat dimeric proteins reported by Wang et al., 2006 using Blast, the WDAI reported in the present work showed 98% identity with sequences reported by Wang *et al*, 2006. Based on these observations, it was concluded that in this case also there were 51 and 375 bp sequences encoding the signal peptide and mature protein, respectively (Wang *et al.* 2006). This dimeric wheat inhibitor is sequentially placed in the class 0.19 dimer type inhibitors of wheat. **Fig. 9** shows the ClustalW alignment of WDAI with other known inhibitors of this class. Sequence comparison showed that it differs from the dimeric α-amylase inhibitor from *Triticum turgidum* subsp. *dicoccoides* (GeneBank ACP40697.1) by just three amino acids. Similarly it differs by 8 residues with respect to sequence of 0.19 inhibitor structure reported in protein data bank (1HSS) and also by the presence of 17 extra residues at the N-terminus. The alignment for WDAI shows that residues 18 to 140 match with residues 1 to 123 of 1HSS_A (**Fig. 9**). In case of signal peptide sequence of wheat dimeric inhibitors from different sources there were two replacements Y11→H (ACP40828.1 and ACP40676.1) and V16→I (ACP40684.1) as compared with other sequences compared here. There are seven synonymous replacements in the dimeric inhibitor sequences: R_{33}→E (ACP40844.1), O_{57}→E (AAY42617.1), D_{64}→H (ACU51027.1), V_{87}→A (ACU51027.1), S_{99}→R (ACU51027.1), V_{115}→I (ACP40676.1), D_{123}→N (ACP40676.1)

and A_{141}→V (ACU51027.1). Since these are synonymous replacements, they do not affect the overall property of the amino acids. These replacements have retained the high homology among the wheat dimeric inhibitors reported from different sources.

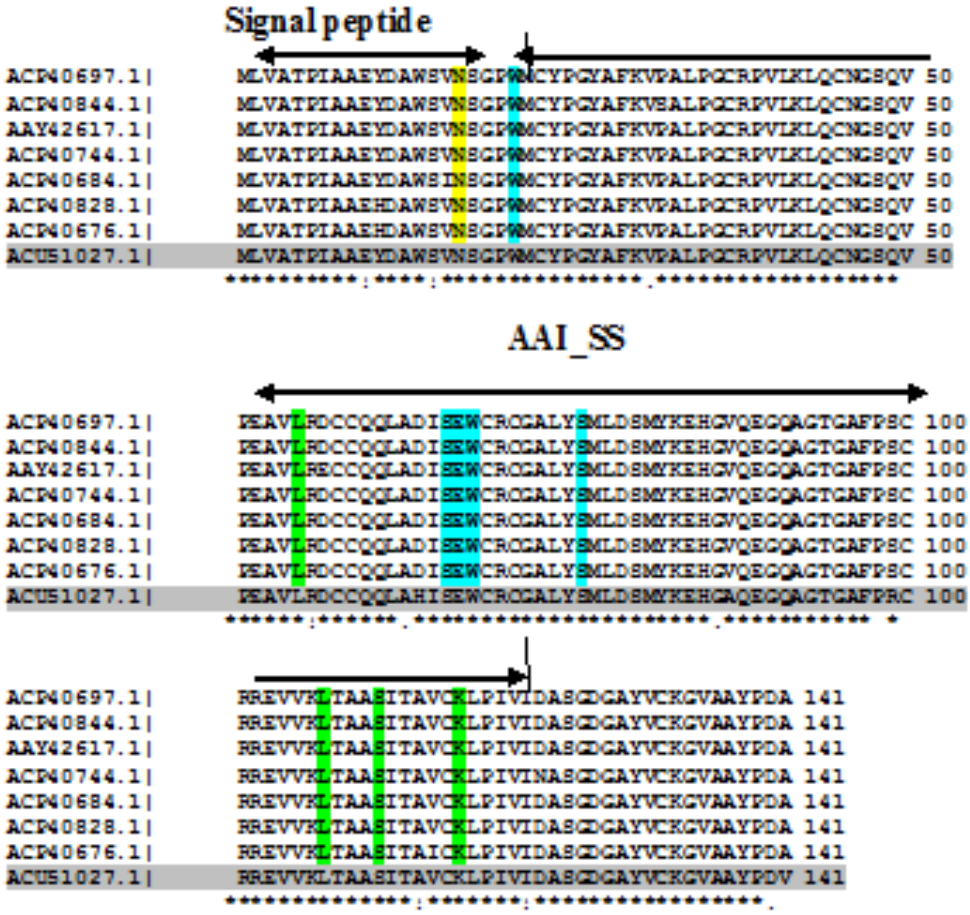


Fig. 9. ClustalW of WDAI α-amylase inhibitor with dimeric α-amylase inhibitors form different varieties of wheat. Sequence (pdb id ACU51027.1) highlighted in gray is WDAI sequence used in this study. (Residue highlighted in yellow is the residue where signal peptide is cut; residues highlighted in green 55, 107, 111, 117 are dimer interface (polypeptide binding) site; residues highlighted in blue 21, 66..68, 75 are α-amylase binding (polypeptide binding) site). GenBank accession numbers represents dimeric α-amylase inhibitor: ACP40697.1 (*Triticum dicoccoides*); ACP40844.1 (*T. dicoccoides*); AAY42617.1 (*T. aestivum*); ACP40744.1 (*T. dicoccoides*); ACP40684.1 (*T. dicoccoides*); ACP40828.1 (*T. dicoccoides*); ACP40676.1 (*T. dicoccoides*); ACU51027.1 (*T. aestivum*).

Bioinformatics and Homology modeling

Homology modeling of O. longicollis partial amylase

The derived protein sequence of the partial amylase of *O. longicollis* has all the conserved regions present in amylases and also the catalytic domain. An attempt was made to build a homology model of *O. longicollis* amylase using the partial translated amino-acid sequence. For homology modeling, the first step is to identify a suitable template. BLAST local alignment search against PDB database using the partial sequence of *O. longicollis* amylase protein as query resulted in identifying the crystal structure of an amylase protein from *Tenebrio molitor* larval α-amylase (PDB Id 1TMQ_A) which shares a sequence identity of 50% over 95% query coverage. The alignment was found to be significant with Blast e-value of 2e-79 (e-value of 0 means highly significant alignment). The structure of 1TMQ_A has been determined at a resolution of 1.64 Angstrom and this is the only available 3D structure of an insect amylase. The *Tenebrio molitor* larval α-amylase protein (1TMQ_A) is 471 residues long and includes two domains ie the catalytic domain from residues 1 to 378 and the C-terminal beta-sheet domain from 379 to 471. The *O. longicollis* amylase protein matches with residues 20-272 of 1TMQ_A which is the catalytic domain. The numbering of *O. longicollis* amylase sequence has been done according to the blast alignment with TMA where the conserved sequence LFEWK matches with the first conserved region of TMQ amylase as well as with the first conserved region of amylases from other organisms. Therefore it was possible to partially model the catalytic domain of *O. longicollis* amylase (**Fig. 10**). The superimposition of the partial *O. longicollis* amylase on the amylase from *Tenebrio molitor* L. (TMA) in the homology model is shown in **Fig. 11**.

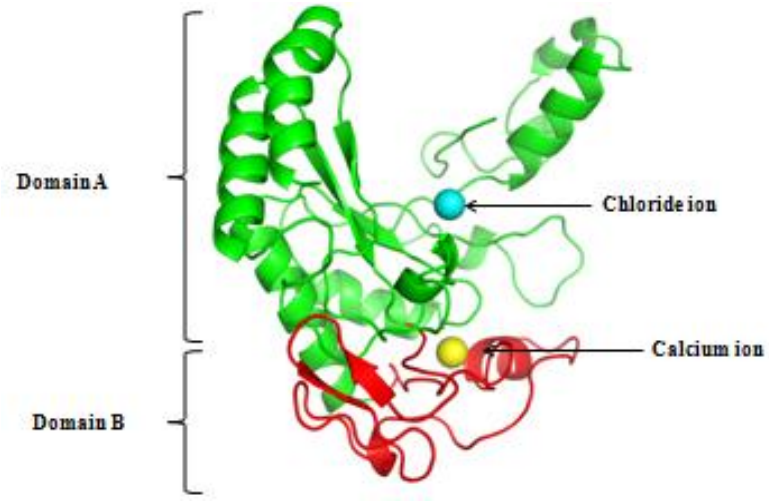


Fig. 10. Homology model of the partial *O. longicollis* amylase (Domain A is represented in green; blue ball is the chloride ion binding site; Domain B is represented in red; yellow ball is the calcium binding site)

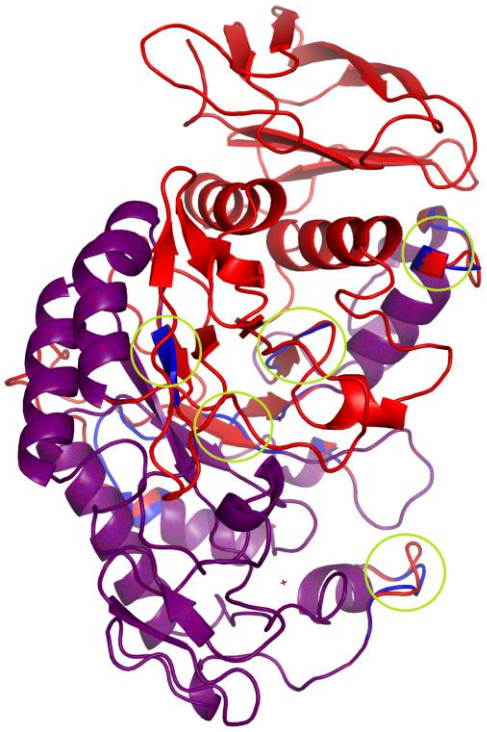


Fig. 11. Homology model of the partial *O. longicollis* amylase superimposed on TMQ amylase (The regions shown in magenta are the regions which overlap and red coloured regions are the portions which are not present in partial *O. longicollis* sequence.

The regions circled in green are the structural regions which do not match exactly with the template (TMA) but still have formed similar structural elements.)

Structural features of the partial O. longicollis amylase

Fig. 12 shows the topological diagram of helices and beta sheets in the *O. longicollis* partial amylase. There are thirteen helices in the partial amylase (**Fig. 12 and Table 5**).

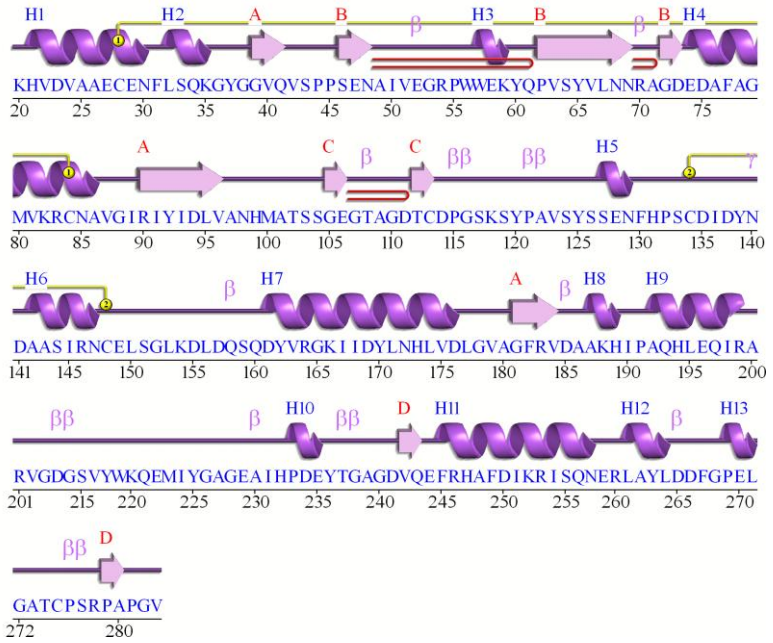


Fig. 12. Overall secondary structure of *O. longicollis* amylase showing helices, loops and β-sheets

Table 5. No of helices in *O. longicollis* amylase

No	Start	End	Type	No. residues	Sequence
1	His21	Asn30	H	10	HVDVAAECEN
2	Leu32	Lys35	H	4	LSQK
3	Trp57	Lys59	G	3	WEK
4	Glu74	Ala86	H	13	EDAFAGMVKRCNA
5	Ser127	Asn129	G	3	SEN
6	Ala142	Asn147	H	6	AASIRN
7	Asp161	Asp176	H	16	DYVRGKIIDYLNHLVD
8	Ala187	His189	G	3	AKH
9	Ala192	Arg199	H	8	AQHLEQIR

10	Pro233	Glu235	G	3	PDE
11	Phe245	Asn257	H	13	FRHAFDIKRISQN
12	Leu261	Leu264	G	4	LAYL
13	Pro269	Leu271	G	3	PEL

(The data displayed in the table for each helix includes the helix number (assigned sequentially starting with 1 at the N terminus of the protein), the residue numbers corresponding to the start and end of the helices, the helix type (H (α helix) or G (3,10) helix. This is followed by the number of residues in the helix. In the final column in the table the helix's amino acid sequence has been given.**)**

The 3D structure of *T. molitor* amylase complexed with various inhibitors is also well studied. Since *O. longicollis* is an important pest of banana and plantains it is important to have knowledge about the structure of enzymes such as amylase which can be a potential target for genetic engineering approaches. The 3D model developed for *O. longicollis* amylase will serve to compare coleopteran amylases more accurately, because there is paucity of information on the amylases of this particular group of insects.

Comparison of structural model of O. longicollis amylase with TMA

Tenebrio molitor is the only insect for which 3D structure of α-amylase (TMA) is known. This enzyme was found to be well adapted to the slightly acidic physiological environment of the larval midgut with a pH optimum of 5.8 for the cleavage of starch (Buonocore *et al.* 1976). The structure of TMA consists of a single polypeptide chain of 471 amino-acid residues, one calcium ion, one chloride ion and 261 water molecules (Strobl *et al.* 1998a). The protein folds into three distinct domains, named A, B and C. Domain A, the major structural unit of TMA is composed of two segments (residues 1-97 and 160-379) and forms a (b/a)8-barrel; an eight-stranded, parallel β-barrel embraced by a concentric circle of eight helical segments (seven a helices and one 310-helix). This domain contains the catalytic site and the ligand binding residues whereas domain B is globular and is inserted into domain A. It is formed by several extended segments and a short α-helix (residue 98-159). This domain forms a cavity against the β barrel of domain A in which the calcium ion is bound. This cation is of fundamental importance for the

structural integrity of a-amylases. Finally, domain C is located exactly opposite to domain B on the other side of domain A. The C domain comprises the C-terminal residues 380-471 and forms a separated folding unit, exclusively made of β sheet. Eight of the 10 strands fold into a β sandwich structure with Greek key topology. The conservation of the interface of A and C domains among a-amylases from different sources suggests an important role for enzyme activity, stability and folding. In the porcine pancreatic a-amylase (PPA), the interface between C and A domains contains a secondary starch-binding site, occupied by maltose in one crystal structure.

In comparison with TMA, the modeled structure of *O. longicollis* amylase shows similar structural features (**Fig. 13**). The partial polypeptide chain consists of 265 amino-acid residues, one calcium ion and one chloride ion. The partial protein has been modeled into two distinct domains, A and B (domain C is not present in the partial sequence). As compared to TMA Domain A which is the major structural unit is composed of two segments (residues 21-97 and 160-265) and forms a (b/a)8-barrel and thirteen helices (**Fig. 13**).

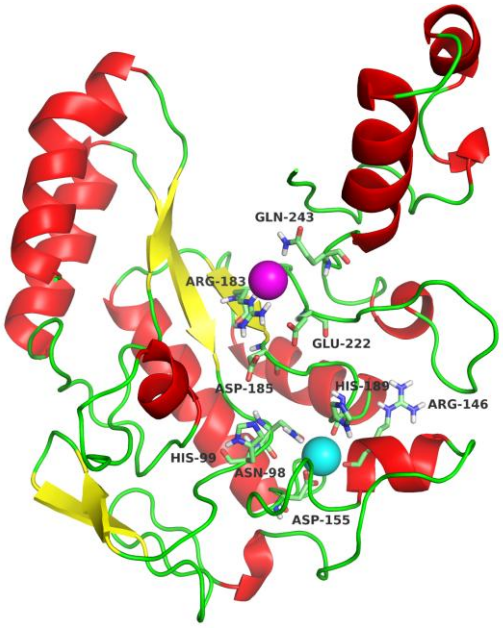


Fig. 13. Modeled structure of *O. longicollis* partial amylase showing catalytic residues (In the figure α helices are represented by red ribbons, beta sheets by yellow arrows, magenta colored ball is chloride binding site and cyan colored ball is calcium binding site).

As in TMA this domain seems to contain the catalytic site and the ligand binding residues whereas domain B is globular and is inserted into domain A. This domain forms a cavity against the β barrel of domain A in which the calcium ion is bound which bounds to Asn 98, Arg 146, Asp 155 and His 189 (**Fig. 13**). TMA, in common with almost all determined a-amylase structures contains a calcium ion at a conserved position. The calcium-binding site of TMA is located at the interface between domains A and B, near to the catalytic centre. The Ca2+ ion is important for activity due its contact with His189 which interacts with the fourth sugar of the substrate, bound in the active site, forming a hinge between the catalytic-site and the Ca2+-binding site (Strobl *et al.* 1998b). TMA crystal structures also contain a chloride anion. It has been proposed that the chloride may be capable of allosterically activating TMA (Buonocore *et al.* 1976) due to its proximity to a water molecule, which probably initiates substrate cleavage (Mazur *et al.* 1994). This modeling and comparative analysis has helped to identify the putative active site residues of *O. longicollis* α-amylase.

Insect and mammalian α-amylases display high homology in their primary and tertiary structures. For example, a-amylase from the yellow mealworm (TMA) and porcine pancreatic a-amylase (PPA) share 54% sequence identity (Strobl *et al.* 1997), and 410 structurally conserved Ca atoms superimposed with an r.m.s.d. of 2.2 A ° (Strobl *et al.* 1998b). It is likely that both mammalian and insect a-amylases have a similar mechanism of action.

Homology modeling of WMAI and WDAI

BLAST local alignment search against PDB database using the translated WMAI and WDAI protein sequences as query resulted in the 0.19 α-amylase inhibitor (Chain A) from wheat kernel (PDB Id 1HSS_A) as the best template.

Homology model of the WMAI

The monomeric wheat a-amylase inhibitor belongs to the category of 0.28 cereal type inhibitor and is one of the most fully characterized inhibitor of this kind. Inhibitors

of this type are specific for insect amylases and are highly active against α-amylase of *T. molitor* (TMA). They only weakly inhibit the α-amylases from human saliva and pancreas and the pancreas of some avians (Silano *et al.* 1975). The clustalW alignment of WMAI shows that residues from 31 to 150 of WMAI match with residues 1 to 123 of 1HSS_A. Hence the N-terminal 30 residues of WMAI could not be modeled. The homology model of the WMAI is shown in **Fig. 14**. According to the modeled structure, all 10 cysteine residues in the molecule form disulfide bridges Cys37–Cys84, Cys51–Cys72, Cys59–Cys112, Cys73 – Cys 128, and Cys86 – Cys143 (Poerio *et al*. 1991) (**Table 6 and Fig. 15).**

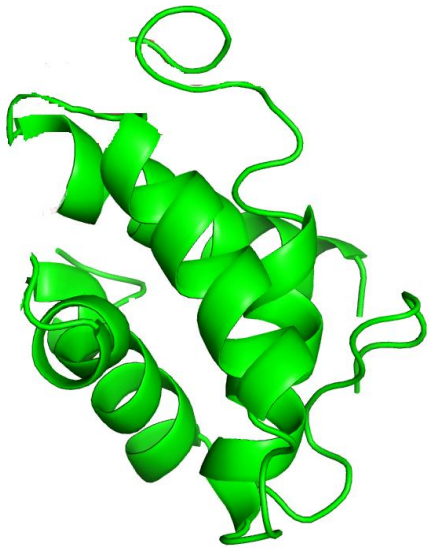


Fig. 14. Homology model of WMAI

Table 6. Disulfide bonds in WMAI

1st cysteine	**2nd cysteine**	**Type**	**Chi1**	**Chi2**	**Chi3**	**Chi2p**	**Chi1p**
C 37	C 84	LHS	-81.9	-78.3	-82.3	-38.0	-53.7
C 51	C 72	RHH	-68.7	-57.4	9406	48.4	170.8
C 59	C 112	RHH	65.7	66.0	71.7	-150.9	-59.7
C 73	C 128	LHS	-60.0	-51.7	-79.7	-62.8	-73.9
C 86	C 143	LHS	-63.5	-51.6	-78.9	-71.3	-74.2

Number of disulphides in chain B: 5

The data displayed for each disulphide bridge is as follows: The table shows the residue numbers of the two cysteines involved in the disulphide bridge and the type of disulphide. The type of bridge is abbreviated (RHH: right hand hook; SRH: short right hand hook; LHS: left handed spiral; RHS: right handed spiral). Chi1, chi2, chi3, chi2' and chi1' values and the distance between the C-α atoms of the residues involved are also recorded.

Homology model of the banana pseudostem weevil amylase-WMAI complex

Using the structure of *Tenebrio molitor* larval α-amylase in complex with ragi bifunctional inhibitor **(**PDB Id 1HSS_A) inhibitor complex as template, the complex of banana pseudostem amylase and WMAI was built (**Fig. 15**). The resulting model was refined by carrying out energy minimization using Desmond.

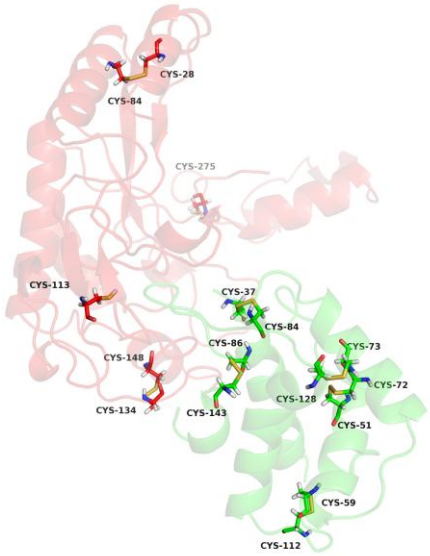


Fig. 15. Homology model of *O. longicollis* amylase complexed with WMAI. (Banana pseudostem amylase structure is shown in red and WMAI structure is shown in green Disulfide bonds of the amylase are shown in red and WMAI disulfide bonds are shown in green with the number of interacting cysteine residues)

The 3-D structure of the complex between TMA and 0.28 wheat inhibitor (Payan 2004) has been used for comparative analysis with the banana pseudostem weevil-WMAI homology model. Some of the structural characteristics of this complex in comparison with TMA-0.28 complexed model are as follows:

The overall 3-D modelled structure of the wheat monomeric inhibitor 0.28 (**Fig. 16**) is similar to that of 0.19 and RBI. Four helices (H1 to H4) with an ‘‘up and down’’ topology constitute the main body of the molecule. The helices are linked together by loop segments, L1 (residues 59–65), L2 (residues 78–86), L3 (residues 100–113) and L4 (residues 129-143) (**Fig. 16**). The C-terminal loop L5 (residue from 148 to COOH) follows after helix H5.

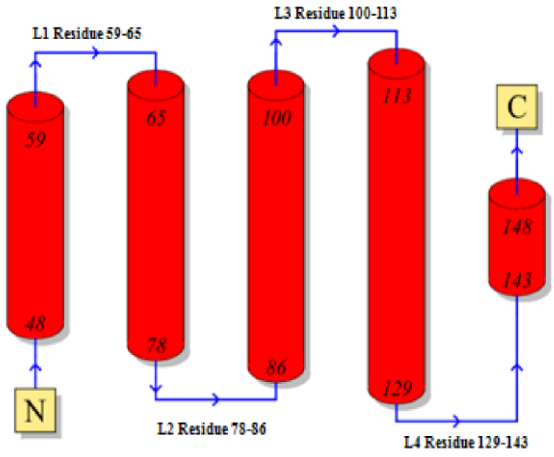
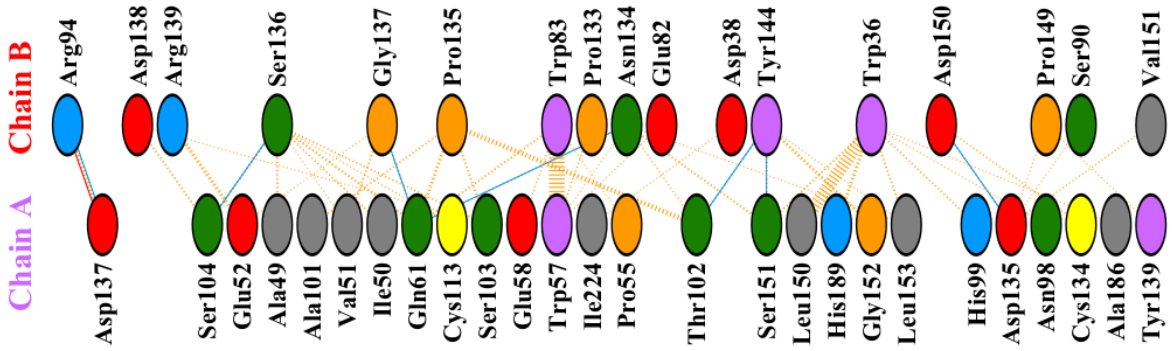


Fig. 16. Topological diagram of WMAI secondary structure

The 3-D structure of the complex of WMAI inhibitor – banana pseudostem weevil amylase (**Figs. 17a and b**) shows three regions of contact (Payan 2004). In the proposed model of WMAI and banana pseudostem weevil amylase, different kinds of interactions between the aminoacid residues of the two molecules are observed. Important interactions are: Arg94, Ser136, Gly137, Asn134, Tyr144 and Asp150 form hydrogen bond with Asp137, Ser104, Gln61, Thr102, Ser151 and Asp135 shown in **Fig. 17a** in the order of placement of helices.



(Chain A shows the interacting residues of *O. longicollis* amylase and Chain B shows the interacting residues of WMAI)

Fig. 17a. Residue interactions across interface coloured by residue type in WMAI-amylase complex

Key: Salt bridges (red); Disulphide bonds (yellow); Hydrogen bonds (blue); Non-bonded contacts (orange).

The number of H-bond lines between any two residues indicates the number of potential hydrogen bonds between them. For non-bonded contacts, which can be plentiful, the width of the striped line is proportional to the number of atomic contacts.

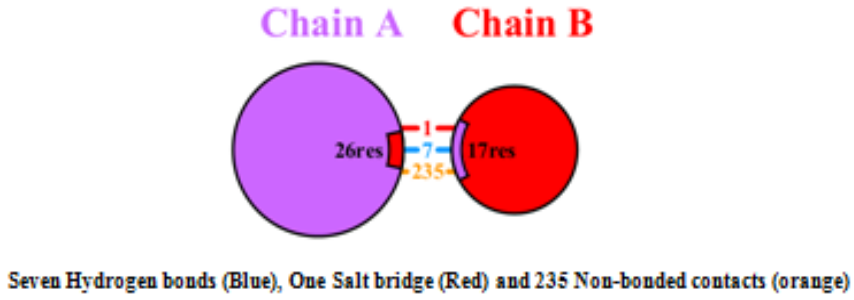


Fig. 17b. Schematic diagram of interactions between protein chains of WMAI-banana pseudostem weevil amylase

Interacting chains are joined by coloured lines, each representing a different type of interaction, as per the key above. The area of each circle is proportional to the surface area of the corresponding protein chain. The extent of the interface region on each chain is represented by the black wedge whose size signifies the interface surface area. Statistics for this interface are given below.

There are seven Hydrogen bonds between the residues of amylase shown in **Table 7,** One Salt bridge (Red) and 235 non-bonded contacts (orange). One salt bridge between Asp 137 (amylase) and Arg94 (WMAI) must be providing stability to the amylase inhibitor complex.

Table 7. Hydrogen bond interaction between the amino acid residues of *O. longicollis* amylase and WMAI

Sr. No of H-bond	Residue in *O. longicollis* amylase	Residue in WMAI amylase
1	Asp137	Arg94
2	Ser104	Ser136
3	Gln61	Gly137
4	Gln61	Pro133
5	Thr102	Tyr144
6	Ser151	Tyr144
7	Asp135	Asp150

There are many non-bonded contacts most prominent being Trp83-Trp57 and Trp36-Leu150 between WMAI and the amylase. Out of these non-bonded interactions, Trp57 is near Ile 224, and Leu150 is near His189 in the amylase which is very close to the catalytic residues His185 and Glu222. These structural interactions indicate some kind of steric interference responsible for the inhibitory activity of WMAI. In the case of TMA-inhibitor complex, role of glycine-rich loop (residues 304 to 310 in PPA) at the C-terminal end of TMA amylase has been proposed (Strobl *et al.* 1998a)**.** Since the presence and/or absence of the above-mentioned loops is the main structural difference between the α-amylases from mammals and from insects, it has been suggested that an inhibitor exclusively directed against mammalian α -amylases makes crucial contacts to these

loops, while an inhibitor solely directed against the insect enzymes might, for instance, be sterically hindered from binding by at least one of the loops. Consequently, an inhibitor active against enzymes from both sources should make only minor or no interactions with these areas. The majority of interactions of tendamistat and of the bean α -amylase inhibitor to PPA are actually made in other contact areas (Strobl *et al.* 1998b). Therefore, it is not surprising that both inhibitors are also active against beetle α -amylases (Schroeder *et al.* 1995). In the present case, since only 265 amino acid are available for comparison, from the modeling study it can be only proposed that the similar interaction in other areas of contact may be leading to inhibitory activity of this class of inhibitors against insect α -amylases. It may be noted that (**Fig. 17b**) the number of nonbonded interaction (235) is far more than hydrogen bond (7) and number of salt bridges is even lower (only one in case of WMAI). In the modeled interaction of 0.19 with TMA (Franco *et al.* 2000), the residues located at the C-terminal end of the inhibitor molecule, superimposed with the facing enzyme region. It has been proposed that in the RBI– TMA structural analysis (Strobl *et al.* 1998b), access of substrate may also be prevented by steric hindrance conferred by these C-terminal residues. Thus, it is can be suggested that this may be a general mechanism for these a-amylase inhibitors of the cereal a-amylase superfamily.

Homology model of WDAI

The wheat dimeric α amylase inhibitor referred to as 0.19 inhibitor inhibits α-amylases from human saliva, pig pancreas, chick pancreas, the yellow mealworm, and *Bacillus subtilis* (Franco *et al.* 2000). The ClustalW alignment of WDAI with a few inhibitors of this class is shown in **Fig. 9.** The wheat inhibitor characterized here has an extra C-terminal valine residue compared to others of this type.

The structured model of the WDAI is shown in **Fig. 18**. The overall 3-D modeled structure of the wheat dimeric inhibitor 0.19 is similar to that of 0.19 and RBI. All 10 cysteine residues in 0.19 form disulfide bonds, occurring in the pairs: Cys23–Cys69, Cys37–Cys58, Cys45–Cys100, Cys59–Cys116 and Cys71–Cys132 (**Fig. 18 and Table 8**).

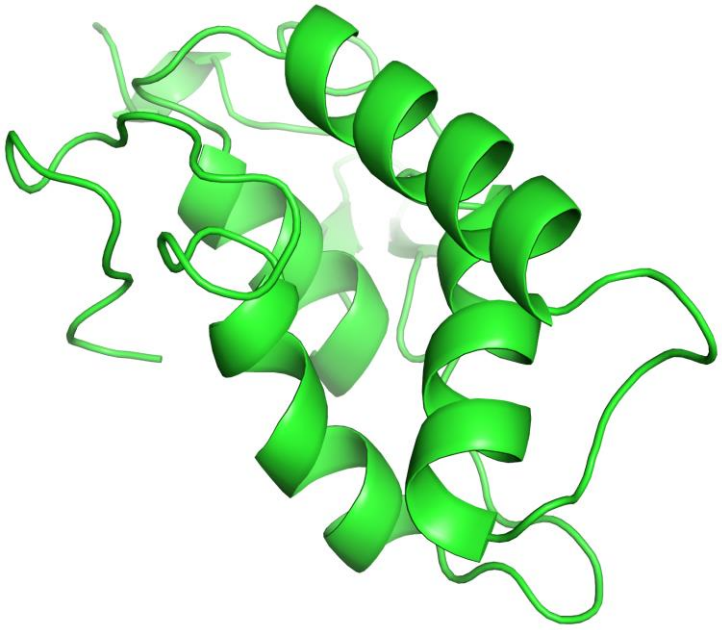


Fig. 18. Homology model of the WDAI

The overall topology and helical arrangement is shown in **Fig. 19.** The helices and beta sheets are linked together by loop segments, L1 (residues 45–51), L2 (residues 64–68), L3 (residues 85–87), L4 (residues 91-101), L5 (residues 110-111), L6 (residues 116-120), L7 (residues 122-127), L8 (residues 129-132) (**Fig. 19**). The C-terminal loop L9 (residues 137 to COOH) follows after helix H7.

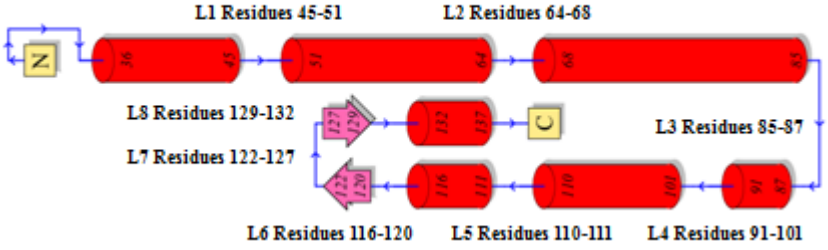


Fig. 19. Topological diagram of WDAI secondary structure

Table 8. Disulfide bonds in WDAI

1st cysteine	2nd cysteine	Type	Chi1	Chi2	Chi3	Chi2p	Chi1p
C 23	C 69	LHS	-66.9	-90.2	-76.1	-42.8	-56.1
C 37	C 58	RHH	-66.9	-59.8	95.9	53.8	169.4
C 45	C 100	RHH	67.1	66.8	67.9	-159.2	-57.7
C 59	C 116	LHS	-58.8	-54.9	-76.2	-66.7	-69.0
C 71	C 132	LHS	-65.8	-54.1	-78.0	-68.7	-64.1

Number of disulphides in chain B: 5

The 3-D structure of the complex between TMA and 0.28 (Payan 2004) has been used for the comparative analysis of the banana pseudostem weevil amylase complexed with WDAI homology model (**Fig. 20**). **Table 8** shows the amino acid residues involved in the formation of conserved disulfide bonds in the 3D structure of WDAI.

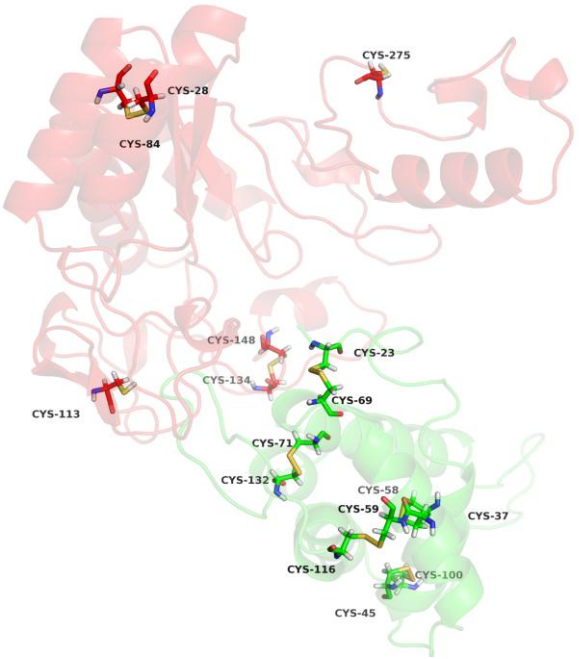


Fig. 20. The modeled structure of the banana pseudostem weevil amylase with the WDAI (The amylase structure is shown in red and WDAI structure is shown in green. The disulfide bonds of the amylase are shown in red and WDAI disulfide bonds are shown in green with the number of interacting cystine residues)

Important structural characteristics of this complex in comparison with TMA are described below:

The regions of contact and the interacting residues in the 3-D structure of the complex of WDAI with banana pseudostem weevil are shown in **Figs 21a and b**.

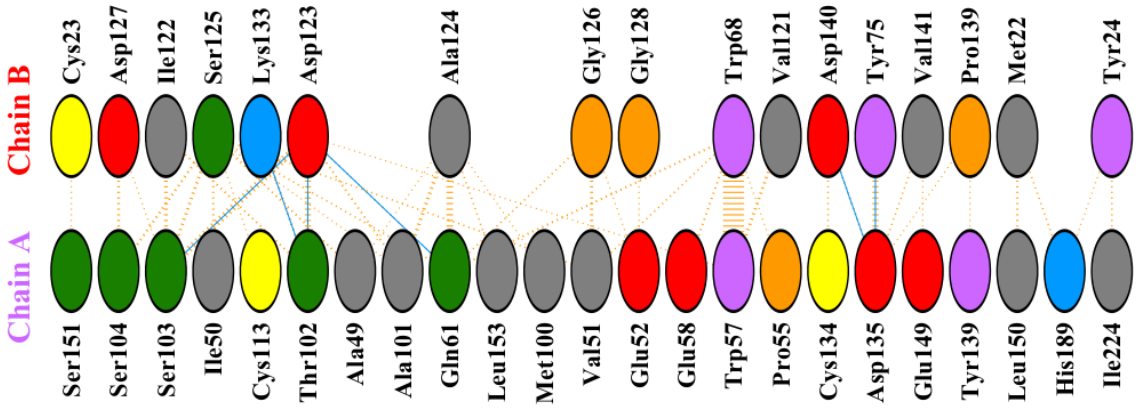


Fig. 21a. Residue interactions across interface coloured by residue type in WDAI-banana pseudostem weevil amylase complex (Chain A shows the interacting residues of the amylase and Chain B shows the interacting residues of WDAI).

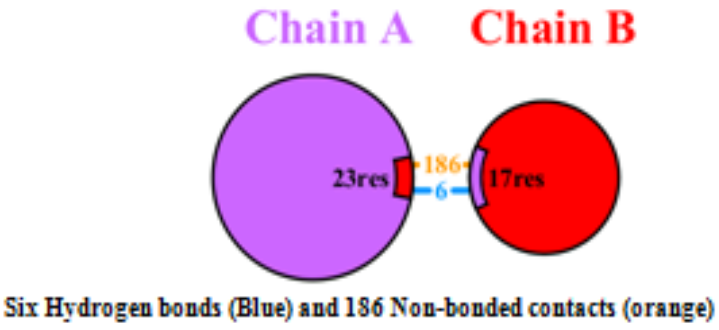


Fig. 21b. Schematic diagram of interactions between protein chains of WDAI and banana pseudostem weevil amylase.

There are six Hydrogen bonds between *O. longicollis* amylase as shown in **Table 9**. There are 186 Non-bonded contacts (orange) (**Fig. 21b**).

Table 9. Hydrogen bond interaction between the amino acid residues *O. longicollis* amylase and WDAI

S. No of H-bond	Residue in *O. longicollis amylase*	Residue in WDAI amylase
1	Ser103	Lys133
2	Thr102	Asp123
3	Thr102	Asp123
4	Gln61	Asp123
5	Asp135	Asp140
6	Asp135	Tyr75

In the proposed model of WDAI and the amylase several amino acids have been shown to interact in many ways. Few important interactions are described: Lys133, Asp123, Asp140 and Tyr75 form hydrogen bond with Lie50, Thr102, Ala101 and Asp135 shown in the **Fig. 21a** in the order of placement of helices. There are many nonbonded contacts most prominent being Ala124-Gln61 and Trp68-Trp57 between WDAI and the amylase. These nonbonded interactions are far from the catalytic residues His185 and Glu222 in the modeled structure. His189 of amylase forms a nonbonded interaction with Met22 and Tyr24 of WDAI where as Ile224 of the amylase, which is close to catalytic residues forms a nonbonded interaction with Tyr24 of WDAI. Although these interacting residues are different from WMAI and the amylase, it is clear from the WDAI-amylase structural model that in this case also there is an involvement of nonbonded interactions which are responsible for some kind of steric hindrance leading to inhibitory activity of WDAI. The reason may be the prevention of binding of starch substrate to the active site of enzyme. The reasons for this kind of inhibition has been already been discussed in the section describing the WMAI. In case of WDAI-amylase also it may be noted that (**Fig. 21b**) the number of nonbonded interaction (186) is far more than hydrogen bond (6) and there is no salt bridges involved. This suggests that this may be a general mechanism for these α-amylase inhibitors of the cereal α-amylase superfamily.

There are experimental evidences to show that the inhibitor protein isolated is specific for insect amylases and does not inhibit human salivary amylase (HAS). It is

reported that among the wheat amylase inhibitors, the 0.19 type class shows wider specificity for both insect and mammalian α-amylases whereas 0.28 class is more specific for insect amylases.

Comparison of WMAI and WDAI

Comparative modeling studies between 0.19 and 0.28 types inhibitors for insect and mammalian amylases have shown that the residues 104–107, VVDA of 0.19 (1HSS) are important for the global mechanism of inhibition which also includes specificity. VVDA of 0.19 (1HSS) represents Valine-Valine-Aspartate-Alanine (hydrophobic-hydrophobic-hydrophilic-hydrophobic) stretch. In the present comparative study, the clustalW alignment shows that VVDA of 0.19 (1HSS) has mutated to a shorter PNP in 0.28 WMAI (133-135) and VIDA (121-124) in WDAI (0.19) type for insect amylase (**Fig. 22**).

```
WMAI            MWMKTVFWGLLVFMLVATTMAVEYGARSHNSGPWSWCDPATGYKVSALTGCRAMVKLQCV 60
WDAI            --------------MLVATPIAAEYDAWSVNSGPWMCYPGYAFKVPALPGCRPVLKLQCN 46
                          ::.:*  .        ..    * *. .:**.**.***.::****

WMAI            GSQVPEAVLRDCCQQLADINNEWCRCGDLSSMLRSVYQELGVREGKE---VLPGCRKEVM 117
WDAI            GSQVPEAVLRDCCQQLAHIS-EWCRCGALYSMLDSMYKEHGAQEGQAGTGAFPRCRREVV 105
                ****************.*. ****** * *** *:*:* *.:**:    .:* **:**:
WMAI            KLTAASVPEVCKVPIPNPSGDRAGVCYWAAYPDV-- 151
WDAI            KLTAASITAVCKLPIVIDASGDGAYVCKGVAAYPDV 141
                ******:. ***:**   :.. ..    .. .
```

Fig. 22. Clustal alignment of WMAI and WDAI proteins

In case of WMAI, PNP represents Proline-Asparagine-Proline (hydrophobic-hydrophilic-hydrophobic) and for WDAI, VIDA represents Valine-Isoleucine-Aspartate-Alanine (hydrophobic- hydrophobic-hydrophilic-hydrophobic). This clearly shows that although there is a change in amino acids but the specific combination of hydrophobic-hydrophilic-hydrophobic amino acid has been maintained. Similar conclusion was drawn in a different modeling study using four different inhibitors from wheat (Franco *et al.* 2000). Further structural studies only can provide an exact description of the inhibitor-amylase interactions and specificity. An ideal inhibitor should be able to specifically inhibit the insect amylases. We have identified and purified a monmeric dimeric inhibitor from a local variety of wheat which unlike other amylase inhibitors reported to date, does

not inhibit porcine a-amylase or human salivary α-amylase but specifically inhibits the *O. longicollis* amylase. Elucidation of the molecular basis of specificity of interaction between the banana pseudostem weevil amylase and the wheat inhibitor will not only help understand the specificity of the wheat inhibitor but can help to design inhibitors specific for other coleopteran amylases.

Description of Ramachandran plots

The Ramachandran plot shows the phi-psi torsion angles for all residues in the structure (except those at the chain termini). Glycine residues are separately identified by triangles as these are not restricted to the regions of the plot appropriate to the other sidechain types. The colouring/shading on the plot represents the different regions (see below) described in Morris *et al.* (1992): the darkest areas (here shown in *red*) correspond to the "core" regions representing the most favourable combinations of phi-psi values. Ideally, one would hope to have over 90% of the residues in these "core" regions. The percentage of residues in the "core"regions is one of the better guides to stereochemical quality.

The models of the banana pseudostem weevil amylase with WMAI and WDAI were evaluated for their correct geometry and stereo chemistry by using software program PROCHECK (Laskowski *et al.* 1993). PROCHECK checks the overall quality of model by plotting Ramachandran map, which is a two-dimensional scatter plot between the two dihedral angles (Phi and Psi) of residues in a polypeptide chain. The overall percentage of residues that are lying in the allowed and disallowed region of this plot is used to assess the overall quality of model. A good model is one having less percentage of residues lying in the disallowed region of plot, hence with less stearic clashes.

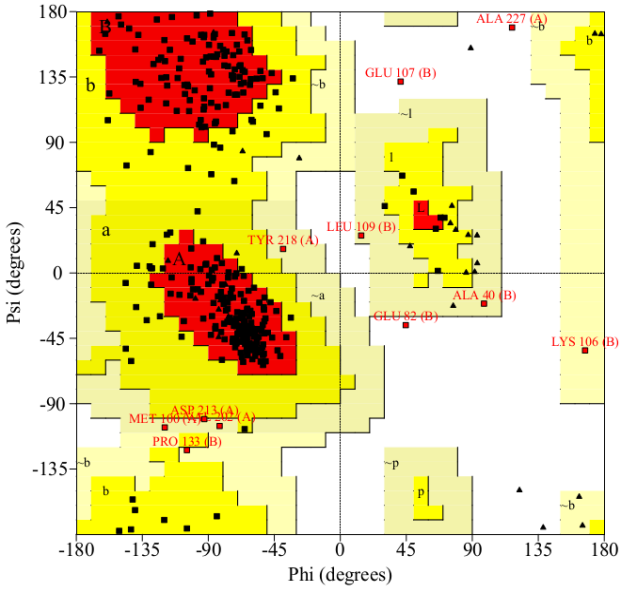


Fig. 23a. Main Ramachandran plot for the amylase-WMAI model

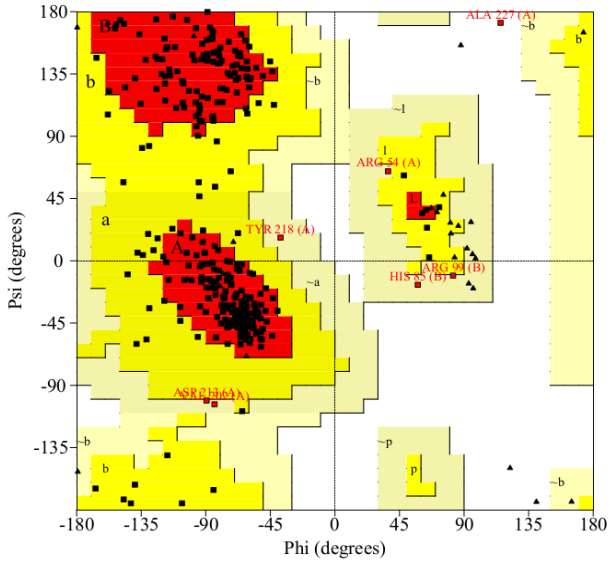


Fig. 23b. Main Ramachandran plot for the amylase-WDAI model (The different regions on the Ram achandran plot are as described in Morris *et al.* (1992). The regions are labelled as follows: A - Core α; L - Core left-handed α; a - Allowed α; l - Allowed left-handed α ; ~a - Generous α; ~l - Generous left-handed α; B - Core beta p - Allowed epsilon; b - Allowed beta; ~p - Generous epsilon; ~b - Generous beta.)

The different regions were taken from the observed phi-psi distribution for 121,870 residues from 463 known X-ray protein structures. The two most favoured

regions are the "core" and "allowed" regions which correspond to 10° x 10° pixels having more than 100 and 8 residues in them, respectively. The "generous" regions were defined by Morris *et al.* (1992) by extending out by 20° (two pixels) all round the "allowed" regions. In fact, very few residues are found in these "generous" regions, so they can probably be treated much like the "disallowed" region.

The predicted model for protein-protein interaction checked against ideal geometry parameters and it was showing that more than 99% of the residues were falling in the allowed regions of the Ramachandran plot. Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions. More than 99% of bond lengths and more than 91% of bond angles were within limits. G-factors provide a measure of how unusual, or out-of-the-ordinary, a property is. The overall G factor output by the program shows that the model has acceptable geometry (**Figs 23a and b, Table 10**).

Table 10. Properties associated with Ramachandran plot

Protein-protein interaction	**Residues in most favoured regions**	**Residues in additional allowed regions**	**Residues in generously allowed regions**	**Residues in disallowed regions**	**G-Factors**		
					Dihedral angle*	**Main-chain covalent force***	**Overall**
Monomer-Amylase	255 (82.5%)	44 (14.2%)	7 (2.3%)	3 (1.0%)	-0.50	-0.3	-0.39
Dimer-Amylase	262 (84.0%)	43 (13.8%)	6 (1.9%)	1 (0.3%)	-0.44	-0.17	-0.30

Values below -0.5* - unusual

Thus it can be concluded that the predicted models for the banana pseudostem amylase with the WMAI and WDAI are suitable for understanding the probable reason of inhibitory activity of WMAI and WDAI against insect α amylases.

REFERENCES

Aghajari N, Haser R, Feller G and Gerday C (1998) *Protein Science*, **7(3)**, 564-572.

Baker JE (1987) *Journal of Stored Products Research*, **23(2)**, 125-131.

Baker JE (1988a) *Entomologia Experimentalis et Applicata*, **46(1)**, 47-54.

Baker JE (1988b) *Insect Biochemistry*, **18**, 107-116.

Baker JE (1989) *Comparative Biochemistry and Physiology*, **92B**, 389-393.

Baker JE and Lum PTM (1989) *Journal of Economic Entomology*, **82(6)**, 1548-1553.

Bernfeld P (1955) *Methods in Enzymology*, **1**, 149-158.

Brayer GD, Luo Y and Withers SG (1995) *Protein Science*, **4(9)**, 1730-1742.

Buonocore V, Poerio E, Pace W, Petrucci T, Silano V and Tomasi M (1976) *FEBS Letters*, **67(2)**, 202-206.

Buonocore V and Silano V (1986) In *Nutritional and toxicological significance of enzyme inhibitors in foods* (pp. 483-507). Springer US.

Chagolla-Lopez A, Blanco-Labra A, Patthy A, Sanchez R and Pongor S (1994) *Journal of Biological Chemistry*, **269(38)**, 23675-23680.

Chen W, Hoy JW and Schneider RW (1992) *Experimental Mycology*, **16**, 22–34.

Cohen SL and Chait BT (1997) *Analytical Biochemistry*, **247(2)**, 257-267.

Davis BJ (1964) *Annals of the New York Academy of Sciences*, **121(2)**, 404-427.

Dellaporta SL, Wood J and Hicks JB (1983) *Plant Molecular Biology Reporter*, **1(4)**, 19-21.

Feller G, le Bussy O, Houssier C and Gerday C (1996) *Journal of Biological Chemistry*, **271(39)**, 23836-23841.

Feng G, Chen M, Kramer KJ and Reeck GR (1991a) *Cereal Chemistry*, **68**, 95-99.

Feng G, Chen M, Kramer KJ and Reeck GR (1991b) *Cereal Chemistry*, **68**, 516-521.

Franco OL, Rigden DJ, Melo RF, Bloch C, Silva CP and Grossi de Sa MF (2000) *European Journal of Biochemistry*, **267(8)**, 2166-2173.

Gatehouse AMR, Fenton KA, Jepson and Pavey DJ (1986) *Journal of the Science of Food and Agriculture*, **37**, 727-734.

Gomez L, Sanchez-Monge R, Garcia-Olmedo F and Salcedo G (1989) *Proceedings of the National Academy of Sciences*, **86(9)**, 3242-3246.

Ishimoto M and Kitamura K (1989) *Applied Entomology and Zoo1ogy*, **24(3)**, 28l-286.

Kneen E and Sandstedt RM (1943) *Journal of the American Chemical Society*, **65(6)**, 1247-1247.

Kneen E and Sandstedt RM (1946) *Archives of Biochemistry*, **9**, 235-249.

Larson SB, Greenwood A, Cascio D, Day J and McPherson A (1994) *Journal of Molecular Biology*, **235(5)**, 1560-1584.

Laskowski RA, MacArthur MW, Moss DS and Thornton JM (1993) *Journal of Applied Crystallography*, **26(2)**, 283-291.

Levitzki A and Steer ML (1974) *European Journal of Biochemistry*, **41(1)**, 171-180.

Machius M, Vertesy L, Huber R and Wiegand G (1996) *Journal of Molecular Biology*, **260(3)**, 409-421.

Maeda K, Hase T and Matsubara H (1983) *Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology*, **743(1)**, 52-57.

Maeda K, Kakabayashi S and Matsubara H (1985) *Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology*, **828(3)**, 213-221.

Mazur AK, Haser R and Payan F (1994) *Biochemical and Biophysical Research Communications*, 204(1), 297-302.

Morris AL, MacArthur MW, Hutchinson EG and Thornton JM (1992) *Proteins: Structure, Function, and Bioinformatics*, **12(4)**, 345-364.

Payan F (2004) *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, **1696(2)**, 171-180.

Petrucci T, Tomasi M, Cantagalli P and Silano V (1974) *Phytochemistry*, 13(11), 2487-2495.

Powers JR and Whitaker JR (1977) *Journal of Food Biochemistry*, 1, 217–238.

Powers JR and Culberston JD (1983) *Cereal Chemistry*, 60, 427-429.

Qian M, Haser R and Payan F (1995) *Protein Science*, **4(4)**, 747-755.

Ramasubbu N, Paloth V, Luo Y, Brayer GD and Levine MJ (1996) *Acta Crystallographica Section D: Biological Crystallography*, **52(3)**, 435-446.

Richardson M (1991) *Methods in Plant Biochemistry*, **5**, 259-305.

Ryan CA (1984) In: Verma, D.P.S. and Hohn, T.H. (eds.) Plant gene research: genes involved in microbe plant interactions, Springer Verlag, Wien,New York, pp. 375-386.

Schroeder HE, Gollasch S, Moore A, Tabe LM, Craig S, Hardie DC and Higgins TJ (1995) *Plant Physiology*, **107(4)**, 1233-1239.

Sharma P, Shankar PR, Subramaniam G, Kumar A, Tandon A, Suresh CG, Meenakshi V and Kumar LS (2009) *International Journal of Insect Science*, 1, 29–44.

Silano V (1987) In: Enzymes and Their Roles in Cereal Technology (edited by Kruger J. E., Lineback D. and Stauffer C. E.), pp. 141-199.

Silano V, Furia M, Gianfreda L, Macri A, Palescandolo R, Rab A and Valfre F (1975) *Biochimica et Biophysica Acta (BBA)-Enzymology*, **391(1)**, 170-178.

Strobl S, Gomis-Ruth FX, Maskos K, Frank G, Huber R and Glockshuber R (1997) *FEBS Letters*, **409(1)**, 109-114.

Strobl S, Maskos K, Betz M, Wiegand G, Huber R, Gomis-Rueth FX and Glockshuber R (1998a) *Journal of Molecular Biology*, **278(3)**, 617-628.

Strobl S, Maskos K, Wiegand G, Huber R, Gomis-Ruth FX and Glockshuber R (1998b) *Structure*, **6(7)**, 911-921.

Wang JR, Yan ZH, Wei YM, Nevo E, Baum BR and Zheng YL (2006) *Journal of Cereal Science*, **43(3)**, 360-368.

Yetter MA, Saunders RM and Boles HP (1979) *Cereal Chemistry*. 56, 243-244.

CHAPTER 7

Conclusions

The important conclusions of this work are summarised below:

1) In the present work, nuclear markers (RAPD, ISSR, AFLP, rDNA- ITS1, ITS2) and mitochondrial markers (COI-tRNA^{Leu}-COII) were used to study genetic diversity of *Odoiporus longicollis* (Oliver) populations collected from different hot spot locations of banana growing regions of India. Theses studies are the first molecular studies on this important banana pest. Amongst these nuclear markers, AFLPs were able to describe the genetic variation precisely as **(i)** AFLPs revealed a strong correlation between genetic distance and geographic distance and **(ii)** the UPGMA dendrogram based on the AFLP data showed a phylogeographic distribution of the populations.

2) Among the six populations, the Assam population is the oldest as **(i)** it shows the highest genetic diversity based on the ITS1 and ITS2 sequence analysis **(ii)** it is separated from the other five populations by a longer branch length in the phylogenetic trees derived by AFLPs, ITS1 and ITS2 data. This observation supports the initial migration of this pest into India via the north-east, from the centre of origin of bananas in south-east Asia which is considered to be the primary centre of diversification and earliest domestication of this crop (Simmonds 1962).

3) Secondary structure of the ITS1 region of *O. longicollis* (Oliver) presented in this work is the first report from Curculionidea. The ITS2 secondary structure of this pest conforms to the pan-eukaryotic model. In the ITS1 secondary structure phylogenetic tree, the individuals grouped according to their secondary structures such that each group showed specific structural characteristics. The phylogenetic trees based on the secondary structures of ITS1 and ITS2 of the thirty *O. longicollis* (Oliver) individuals were congruent, though the secondary structure of these two regions bears no resemblance to each other, suggesting that secondary structures of these two regions are important `markers' in phylogenetic analysis.

4) The present analysis based on nine RAPD (113 loci) and nine ISSR markers (71 loci) suggests that there is restricted gene flow between the populations (For RAPDs : G_{ST} = 0.545, Nm = 0.4174; For ISSRs : G_{ST} = 0.4461, Nm = 0.6207). However, these results are in contrast to those obtained using the rDNA markers *i.e.* ITS1 and ITS2 (For ITS1 : Gst = 0.12704, Nm =1.72 ; For ITS2 : GST = 0.15094, Nm = 1.41).

Possible explanations for this discrepancy are as follows : Gene flow is determined from the distribution of genetic variation among populations. Gene flow is inferred from F_{ST} values which gives estimates of the amount of differentiation among the populations, and F_{ST} values and related parameters are based on correct estimates of allele frequencies (Wright 1978). While estimating gene flow with genetic markers, the distribution of alleles is determined. Molecular markers, with some exception, mainly detect variation in non-coding DNA regions and hence the variation expressed by these markers is assumed to be selectively neutral. Different molecular markers also differ in the amount of variability that they display. Hence, for determining gene flow, it is necessary to choose the appropriate marker which would display the right amount of variability for the spatial scale under study. Markers such as RFLPs are co-dominant and hence are able to differentiate the heterozygotes from homozygotes and such markers allow easy estimation of allele frequency. On the other hand, markers such as RAPDs, ISSRs and AFLPs are dominant and with such markers it is not possible to differentiate the heterozygotes from homozygotes, making it difficult to calculate right estimates of allele frequencies (Charlesworth 2003). Such dominant markers allow the estimation of genotype frequencies. In addition, the occurrence of null alleles would also give wrong estimates of allele frequencies at particular loci and hence wrong estimates of F_{ST}. As seen in **Tables 1and 2,** Nm values for the loci amplified with RAPDs vary from 0.0649 to 6.0, while similar values for the loci amplified with ISSRs vary from 0.0727 to 4.0, and the values reported in this work are an average (**Tables 1 and 2)**. Hence, these values are subjected to change depending upon the number of markers/loci used in the analysis. A large number of polymorphic as well as monomorphic loci representing a random sample of the genome are required to be included in the study to have better estimates of genetic differentiation among the populations. (Nei & Roychoudhury 1972; Fauvelot *et al*. 2007). Analysis of DNA sequence data and variability in the sequences has several advantages *i.e.* **(i)** DNA variants are often neutral or close to neutral **(ii)** DNA sequence variants that are shared between populations can be detected and **(iii)** The number of haplotypes in a sample at a locus can be identified (Charlesworth 2003). In the light of the above discussion, it can be concluded that the

DNA sequence data is more reliable and hence on the basis of the DNA sequence analysis of the ITS1 and ITS2 regions it is concluded that there is gene flow between the populations. This can be attributed to the strong flying ability of this pest.

A large number of polymorphic as well as monomorphic loci representing a random sample of the genome are required to be included in the study to have better estimates of genetic differentiation among the populations. (Nei & Roychoudhury 1972).

Table 1. POPGENE Output for RAPD Data

Locus	Sample Size	Ht	Hc	Hs	Gst	Gcs	Nm(Gst)* Nm(Gcs)*
L1	30	0.2778	0.0000	0.0000	1.0000	****	0.0000 ****
L2	30	0.4444	0.2933	0.2933	0.3400	0.0000	0.9706 2000.0000
L3	30	0.3200	0.2400	0.2400	0.2500	0.0000	1.5000 2000.0000
L4	30	0.4444	0.1867	0.1867	0.5800	0.0000	0.3621 2000.0000
L5	30	0.4444	0.1867	0.1867	0.5800	0.0000	0.3621 2000.0000
L6	30	0.4444	0.1067	0.1067	0.7600	0.0000	0.1579 2000.0000
L7	30	0.3911	0.1867	0.1867	0.5227	0.0000	0.4565 2000.0000
L8	30	0.3578	0.2133	0.2133	0.4037	0.0000	0.7385 2000.0000
L9	30	0.0000	0.0000	0.0000	****	****	**** ****
L10	30	0.1244	0.1067	0.1067	0.1429	0.0000	3.0000 2000.0000
L11	30	0.3911	0.0800	0.0800	0.7955	0.0000	0.1286 2000.0000
L12	30	0.1800	0.0800	0.0800	0.5556	0.0000	0.4000 2000.0000
L13	30	0.0644	0.0533	0.0533	0.1724	0.0000	2.4000 2000.0000
L14	30	0.2778	0.2400	0.2400	0.1360	0.0000	3.1765 2000.0000
L15	30	0.4800	0.2933	0.2933	0.3889	0.0000	0.7857 2000.0000
L16	30	0.1244	0.0800	0.0800	0.3571	0.0000	0.9000 2000.0000
L17	30	0.0644	0.0533	0.0533	0.1724	0.0000	2.4000 2000.0000
L18	30	0.2778	0.2133	0.2133	0.2320	0.0000	1.6552 2000.0000
L19	30	0.1244	0.0800	0.0800	0.3571	0.0000	0.9000 2000.0000
L20	30	0.2778	0.1600	0.1600	0.4240	0.0000	0.6792 2000.0000
L21	30	0.3911	0.1067	0.1067	0.7273	0.0000	0.1875 2000.0000
L22	30	0.4444	0.0000	0.0000	1.0000	****	0.0000 ****
L23	30	0.1244	0.1067	0.1067	0.1429	0.0000	3.0000 2000.0000
L24	30	0.4644	0.0533	0.0533	0.8852	0.0000	0.0649 2000.0000
L25	30	0.3200	0.2400	0.2400	0.2500	0.0000	1.5000 2000.0000
L26	30	0.2778	0.1600	0.1600	0.4240	0.0000	0.6792 2000.0000
L27	30	0.2778	0.1067	0.1067	0.6160	0.0000	0.3117 2000.0000
L28	30	0.4200	0.0533	0.0533	0.8730	0.0000	0.0727 2000.0000
L29	30	0.0644	0.0533	0.0533	0.1724	0.0000	2.4000 2000.0000
L30	30	0.0644	0.0533	0.0533	0.1724	0.0000	2.4000 2000.0000
L31	30	0.0644	0.0533	0.0533	0.1724	0.0000	2.4000 2000.0000
L32	30	0.3200	0.1600	0.1600	0.5000	0.0000	0.5000 2000.0000
L33	30	0.1244	0.0800	0.0800	0.3571	0.0000	0.9000 2000.0000
L34	30	0.0644	0.0533	0.0533	0.1724	0.0000	2.4000 2000.0000
L35	30	0.1800	0.1600	0.1600	0.1111	0.0000	4.0000 2000.0000
L36	30	0.1244	0.1067	0.1067	0.1429	0.0000	3.0000 2000.0000
L37	30	0.3200	0.2133	0.2133	0.3333	0.0000	1.0000 2000.0000
L38	30	0.1800	0.0800	0.0800	0.5556	0.0000	0.4000 2000.0000
L39	30	0.3578	0.2667	0.2667	0.2547	0.0000	1.4634 2000.0000
L40	30	0.1244	0.0800	0.0800	0.3571	0.0000	0.9000 2000.0000
L41	30	0.3578	0.0800	0.0800	0.7764	0.0000	0.1440 2000.0000
L42	30	0.2311	0.0533	0.0533	0.7692	0.0000	0.1500 2000.0000
L43	30	0.1244	0.0800	0.0800	0.3571	0.0000	0.9000 2000.0000
L44	30	0.4644	0.1600	0.1600	0.6555	0.0000	0.2628 2000.0000
L45	30	0.2311	0.1333	0.1333	0.4231	0.0000	0.6818 2000.0000
L46	30	0.1800	0.0800	0.0800	0.5556	0.0000	0.4000 2000.0000
L47	30	0.1244	0.1067	0.1067	0.1429	0.0000	3.0000 2000.0000

L48	30	0.1800	0.1333	0.1333	0.2593	0.0000	1.4286 2000.0000
L49	30	0.0644	0.0533	0.0533	0.1724	0.0000	2.4000 2000.0000
L50	30	0.1244	0.1067	0.1067	0.1429	0.0000	3.0000 2000.0000
L51	30	0.2311	0.2133	0.2133	0.0769	0.0000	6.0000 2000.0000
L52	30	0.4644	0.0533	0.0533	0.8852	0.0000	0.0649 2000.0000
L53	30	0.3578	0.1867	0.1867	0.4783	0.0000	0.5455 2000.0000
L54	30	0.4978	0.2667	0.2667	0.4643	0.0000	0.5769 2000.0000
L55	30	0.0644	0.0533	0.0533	0.1724	0.0000	2.4000 2000.0000
L56	30	0.3578	0.0800	0.0800	0.7764	0.0000	0.1440 2000.0000
L57	30	0.0644	0.0533	0.0533	0.1724	0.0000	2.4000 2000.0000
L58	30	0.4444	0.1067	0.1067	0.7600	0.0000	0.1579 2000.0000
L59	30	0.4644	0.2133	0.2133	0.5407	0.0000	0.4248 2000.0000
L60	30	0.4911	0.0800	0.0800	0.8371	0.0000	0.0973 2000.0000
L61	30	0.5000	0.2400	0.2400	0.5200	0.0000	0.4615 2000.0000
L62	30	0.4644	0.0533	0.0533	0.8852	0.0000	0.0649 2000.0000
L63	30	0.4200	0.2400	0.2400	0.4286	0.0000	0.6667 2000.0000
L64	30	0.1800	0.1333	0.1333	0.2593	0.0000	1.4286 2000.0000
L65	30	0.3911	0.1600	0.1600	0.5909	0.0000	0.3462 2000.0000
L66	30	0.3578	0.1333	0.1333	0.6273	0.0000	0.2970 2000.0000
L67	30	0.4800	0.1067	0.1067	0.7778	0.0000	0.1429 2000.0000
L68	30	0.4200	0.0533	0.0533	0.8730	0.0000	0.0727 2000.0000
L69	30	0.4444	0.1067	0.1067	0.7600	0.0000	0.1579 2000.0000
L70	30	0.4800	0.1333	0.1333	0.7222	0.0000	0.1923 2000.0000
L71	30	0.2311	0.0533	0.0533	0.7692	0.0000	0.1500 2000.0000
L72	30	0.1800	0.1333	0.1333	0.2593	0.0000	1.4286 2000.0000
L73	30	0.1244	0.1067	0.1067	0.1429	0.0000	3.0000 2000.0000
L74	30	0.1244	0.1067	0.1067	0.1429	0.0000	3.0000 2000.0000
L75	30	0.2311	0.1600	0.1600	0.3077	0.0000	1.1250 2000.0000
L76	30	0.4800	0.2400	0.2400	0.5000	0.0000	0.5000 2000.0000
L77	30	0.4200	0.2133	0.2133	0.4921	0.0000	0.5161 2000.0000
L78	30	0.1244	0.0800	0.0800	0.3571	0.0000	0.9000 2000.0000
L79	30	0.3911	0.2667	0.2667	0.3182	0.0000	1.0714 2000.0000
L80	30	0.2778	0.2133	0.2133	0.2320	0.0000	1.6552 2000.0000
L81	30	0.3911	0.1867	0.1867	0.5227	0.0000	0.4565 2000.0000
L82	30	0.3578	0.1067	0.1067	0.7019	0.0000	0.2124 2000.0000
L83	30	0.1244	0.0800	0.0800	0.3571	0.0000	0.9000 2000.0000
L84	30	0.1244	0.0800	0.0800	0.3571	0.0000	0.9000 2000.0000
L85	30	0.1800	0.1600	0.1600	0.1111	0.0000	4.0000 2000.0000
L86	30	0.1244	0.1067	0.1067	0.1429	0.0000	3.0000 2000.0000
L87	30	0.3911	0.0800	0.0800	0.7955	0.0000	0.1286 2000.0000
L88	30	0.3911	0.0800	0.0800	0.7955	0.0000	0.1286 2000.0000
L89	30	0.3200	0.0533	0.0533	0.8333	0.0000	0.1000 2000.0000
L90	30	0.3200	0.1600	0.1600	0.5000	0.0000	0.5000 2000.0000
L91	30	0.4444	0.1067	0.1067	0.7600	0.0000	0.1579 2000.0000
L92	30	0.0644	0.0533	0.0533	0.1724	0.0000	2.4000 2000.0000
L93	30	0.4200	0.2933	0.2933	0.3016	0.0000	1.1579 2000.0000
L94	30	0.1244	0.1067	0.1067	0.1429	0.0000	3.0000 2000.0000
L95	30	0.0644	0.0533	0.0533	0.1724	0.0000	2.4000 2000.0000
L96	30	0.1244	0.0800	0.0800	0.3571	0.0000	0.9000 2000.0000
L97	30	0.1244	0.0800	0.0800	0.3571	0.0000	0.9000 2000.0000
L98	30	0.1244	0.0800	0.0800	0.3571	0.0000	0.9000 2000.0000
L99	30	0.4978	0.1333	0.1333	0.7321	0.0000	0.1829 2000.0000
L100	30	0.4444	0.1867	0.1867	0.5800	0.0000	0.3621 2000.0000
L101	30	0.4444	0.0000	0.0000	1.0000	****	0.0000 ****
L102	30	0.4911	0.1333	0.1333	0.7285	0.0000	0.1863 2000.0000
L103	30	0.3578	0.1867	0.1867	0.4783	0.0000	0.5455 2000.0000
L104	30	0.0644	0.0533	0.0533	0.1724	0.0000	2.4000 2000.0000
L105	30	0.1244	0.0800	0.0800	0.3571	0.0000	0.9000 2000.0000
L106	30	0.0644	0.0533	0.0533	0.1724	0.0000	2.4000 2000.0000
L107	30	0.1800	0.0800	0.0800	0.5556	0.0000	0.4000 2000.0000
L108	30	0.1244	0.0800	0.0800	0.3571	0.0000	0.9000 2000.0000
L109	30	0.2311	0.1600	0.1600	0.3077	0.0000	1.1250 2000.0000
L110	30	0.1244	0.0800	0.0800	0.3571	0.0000	0.9000 2000.0000
L111	30	0.2311	0.1600	0.1600	0.3077	0.0000	1.1250 2000.0000
L112	30	0.0644	0.0533	0.0533	0.1724	0.0000	2.4000 2000.0000
L113	30	0.1800	0.1333	0.1333	0.2593	0.0000	1.4286 2000.0000
Mean	30	0.2682	0.1220	0.1220	0.5450	0.0000	0.4174 2000.0000
St. Dev		0.0221	0.0049	0.0049			

```
* Nm = estimate of gene flow from Gst or Gcs. E.g., Nm = 0.5(1 - Gst)/Gst;
  See McDermott and McDonald, Ann. Rev. Phytopathol. 31:353-373 (1993).
The number of polymorphic loci is :     112
The percentage of polymorphic loci is : 99.12
```

Table 2. POPGENE Output for ISSR Data

Locus	Sample Size	Ht	Hc	Hs	Gst	Gcs	Nm(Gst)*	Nm(Gcs)*
L1	30	0.2778	0.2400	0.2400	0.1360	0.0000	3.1765	2000.0000
L2	30	0.2778	0.2133	0.2133	0.2320	0.0000	1.6552	2000.0000
L3	30	0.2311	0.1867	0.1867	0.1923	0.0000	2.1000	2000.0000
L4	30	0.4978	0.4000	0.4000	0.1964	0.0000	2.0455	2000.0000
L5	30	0.4444	0.0000	0.0000	1.0000	****	0.0000	****
L6	30	0.0000	0.0000	0.0000	****	****	****	****
L7	30	0.1800	0.1600	0.1600	0.1111	0.0000	4.0000	2000.0000
L8	30	0.0644	0.0533	0.0533	0.1724	0.0000	2.4000	2000.0000
L9	30	0.2311	0.1333	0.1333	0.4231	0.0000	0.6818	2000.0000
L10	30	0.3200	0.2933	0.2933	0.0833	0.0000	5.5000	2000.0000
L11	30	0.4444	0.3467	0.3467	0.2200	0.0000	1.7727	2000.0000
L12	30	0.1800	0.1333	0.1333	0.2593	0.0000	1.4286	2000.0000
L13	30	0.0000	0.0000	0.0000	****	****	****	****
L14	30	0.4200	0.2133	0.2133	0.4921	0.0000	0.5161	2000.0000
L15	30	0.1800	0.0800	0.0800	0.5556	0.0000	0.4000	2000.0000
L16	30	0.4200	0.2400	0.2400	0.4286	0.0000	0.6667	2000.0000
L17	30	0.4911	0.4267	0.4267	0.1312	0.0000	3.3103	2000.0000
L18	30	0.3578	0.1867	0.1867	0.4783	0.0000	0.5455	2000.0000
L19	30	0.3578	0.1867	0.1867	0.4783	0.0000	0.5455	2000.0000
L20	30	0.4200	0.3200	0.3200	0.2381	0.0000	1.6000	2000.0000
L21	30	0.0644	0.0533	0.0533	0.1724	0.0000	2.4000	2000.0000
L22	30	0.4978	0.4533	0.4533	0.0893	0.0000	5.1000	2000.0000
L23	30	0.1800	0.1600	0.1600	0.1111	0.0000	4.0000	2000.0000
L24	30	0.0000	0.0000	0.0000	****	****	****	****
L25	30	0.4911	0.3467	0.3467	0.2941	0.0000	1.2000	2000.0000
L26	30	0.2311	0.1600	0.1600	0.3077	0.0000	1.1250	2000.0000
L27	30	0.2311	0.1600	0.1600	0.3077	0.0000	1.1250	2000.0000
L28	30	0.0000	0.0000	0.0000	****	****	****	****
L29	30	0.0644	0.0533	0.0533	0.1724	0.0000	2.4000	2000.0000
L30	30	0.3911	0.1333	0.1333	0.6591	0.0000	0.2586	2000.0000
L31	30	0.4644	0.1333	0.1333	0.7129	0.0000	0.2013	2000.0000
L32	30	0.4444	0.1600	0.1600	0.6400	0.0000	0.2813	2000.0000
L33	30	0.4911	0.1867	0.1867	0.6199	0.0000	0.3066	2000.0000
L34	30	0.4978	0.1333	0.1333	0.7321	0.0000	0.1829	2000.0000
L35	30	0.4978	0.2400	0.2400	0.5179	0.0000	0.4655	2000.0000
L36	30	0.4644	0.1333	0.1333	0.7129	0.0000	0.2013	2000.0000
L37	30	0.4800	0.2667	0.2667	0.4444	0.0000	0.6250	2000.0000
L38	30	0.4444	0.2933	0.2933	0.3400	0.0000	0.9706	2000.0000
L39	30	0.0000	0.0000	0.0000	****	****	****	****
L40	30	0.4978	0.3200	0.3200	0.3571	0.0000	0.9000	2000.0000
L41	30	0.4800	0.1333	0.1333	0.7222	0.0000	0.1923	2000.0000
L42	30	0.4800	0.2400	0.2400	0.5000	0.0000	0.5000	2000.0000
L43	30	0.0000	0.0000	0.0000	****	****	****	****
L44	30	0.0644	0.0533	0.0533	0.1724	0.0000	2.4000	2000.0000
L45	30	0.4200	0.2667	0.2667	0.3651	0.0000	0.8696	2000.0000
L46	30	0.3200	0.2133	0.2133	0.3333	0.0000	1.0000	2000.0000
L47	30	0.0000	0.0000	0.0000	****	****	****	****
L48	30	0.1244	0.1067	0.1067	0.1429	0.0000	3.0000	2000.0000
L49	30	0.4978	0.2933	0.2933	0.4107	0.0000	0.7174	2000.0000
L50	30	0.0000	0.0000	0.0000	****	****	****	****
L51	30	0.4911	0.2400	0.2400	0.5113	0.0000	0.4779	2000.0000
L52	30	0.3578	0.2667	0.2667	0.2547	0.0000	1.4634	2000.0000
L53	30	0.4200	0.2667	0.2667	0.3651	0.0000	0.8696	2000.0000
L54	30	0.4800	0.2400	0.2400	0.5000	0.0000	0.5000	2000.0000
L55	30	0.4978	0.1600	0.1600	0.6786	0.0000	0.2368	2000.0000
L56	30	0.4911	0.2933	0.2933	0.4027	0.0000	0.7416	2000.0000
L57	30	0.4200	0.3200	0.3200	0.2381	0.0000	1.6000	2000.0000

L58	30	0.4200	0.2133	0.2133	0.4921	0.0000	0.5161 2000.0000
L59	30	0.4800	0.3467	0.3467	0.2778	0.0000	1.3000 2000.0000
L60	30	0.4978	0.1600	0.1600	0.6786	0.0000	0.2368 2000.0000
L61	30	0.2778	0.2133	0.2133	0.2320	0.0000	1.6552 2000.0000
L62	30	0.4200	0.1600	0.1600	0.6190	0.0000	0.3077 2000.0000
L63	30	0.5000	0.2400	0.2400	0.5200	0.0000	0.4615 2000.0000
L64	30	0.4911	0.1600	0.1600	0.6742	0.0000	0.2416 2000.0000
L65	30	0.3911	0.2400	0.2400	0.3864	0.0000	0.7941 2000.0000
L66	30	0.0000	0.0000	0.0000	****	****	**** ****
L67	30	0.0000	0.0000	0.0000	****	****	**** ****
L68	30	0.4978	0.3733	0.3733	0.2500	0.0000	1.5000 2000.0000
L69	30	0.4200	0.0533	0.0533	0.8730	0.0000	0.0727 2000.0000
L70	30	0.4800	0.0800	0.0800	0.8333	0.0000	0.1000 2000.0000
L71	30	0.1244	0.0800	0.0800	0.3571	0.0000	0.9000 2000.0000
Mean	30	0.3208	0.1777	0.1777	0.4461	0.0000	0.6207 2000.0000
St. Dev		0.0333	0.0139	0.0139			

* Nm = estimate of gene flow from Gst or Gcs. E.g., Nm = 0.5(1 - Gst)/Gst;
  See McDermott and McDonald, Ann. Rev. Phytopathol. 31:353-373 (1993).
The number of polymorphic loci is : 61
The percentage of polymorphic loci is : 85.92

5) The results based on the mitochondrial gene fragment suggests a male-biased gene flow in the pest. The advantages of sex-biased dispersal are avoidance of inbreeding and reduced competition for reproductive resources.

The results presented in this thesis are important in terms of pest management, considering the status of *O. longicollis* (Oliver) in India (Dutt & Maiti 1972). Since banana is the fourth most important crop of the developing world, it is important to have a cumulative approach towards advancing banana yields in major banana producing regions of world. In this regard, several research organizations and scientists have formed a network towards the improvement of several cultivars of banana and development of IPM approaches for the control of pests and diseases which affect banana plantations. One such organization is INIBAP-ASPNET (Advancing banana and plantains R & D in Asia and the Pacific), which has been very active around the world, particularly in Asia and the Pacific Region. The INIBAP member countries and institutions participate in various programmes such as germplasm utilization, disease control, organic cultivation techniques and so on.

Integrated pest management (IPM) has been defined in a context of the associated environment and population dynamics of the pest species and the practices utilized in a compatible manner to maintain the pest population levels below those causing ‘economic injury’ (Smith & Reynolds 1966). Biological control of insect pests is one of the approaches to control pest populations (Hall & Ehler 1979; Greathead & Greathead1992).

Genetic characterization is crucial to ensure that there is a single species and not a species complex involved in a biological control release program designed for IPM. But it is also important that the insect pest under the study itself be genetically characterized because populations of the same species may differ in their biological characteristics important for biological control (Narang *et al.* 1993). Genetic diversity studies of naturally occurring insect populations are important for identifying genetic variations in populations of different geographic origins. Hence the recognition of intraspecific variation can also be crucial for the success of biological control programs (Powell & Walton 1989; Narang *et al.* 1993, Unruh & Woolley 1999).

For understanding the biology of invading insect species and designing control strategies for them, a genetic approach to geographic or population structure is necessary (Carey 1991; Unruh & Woolley 1999). The present work is the first study of the population genetic structure of *O. longicollis* (Oliver) using various nuclear and mitochondrial markers. The phenomenon of gene flow has been observed in the populations of *O. longicollis* (Oliver) which would facilitate the rapid spread of traits that could compromise IPM control programmes, such as insecticide resistance or behavioural modifications. Such information on molecular markers based studies can be used to enhance the efficacy of pest management practices and serve as an impetus for wider studies on pest biology, ecology and behavior.

To understand the invasions of insect pest and for designing the control methods using IPM approach it is important to study the geographic structure of the pest (Carey 1991). The pest insects can evolve to develop resistance to chemical insecticides as well as transgenic plants producing insecticides (Tabashnik 1994). Study of gene flow and the geographic structure of pest populations can help in the prediction of estimating the rate of resistance development (Caprio & Tabashnik 1992; Gould 1994; Mallet & Porter 1992; Tabashnik 1994). There are reports that migration contributes to the development of insecticide resistance (Daly & Gregg 1985; Korman *et al.* 1993; Caprio & Tabashnik 1992; Dunley & Croft 1994)**.** Gene flow estimates in resistance management is related to selection and adaptation in different environments (Roderick 1996). The patterns of resistance variation within and among insect populations depend on insect movement and strategy of insecticide application (Daly 1993).

Sex-pheromone based control is one of the control strategies being adopted for the control of *O. longicollis* (Oliver) (discussed in Chapter 1). Studies have indicated the presence of female-produced sex pheromone system in the pest, but the olfactory activity of both male and female is mediated by the male-produced aggregation pheromone (Prasuna et al. 2008). Since male-biased gene flow has been observed in the present study (Chapter 5), it is possible to utilize synthetic aggregation pheromone in integrated pest management of this pest as it would attract both male and female weevils for trapping, leading to higher efficiency due to specificity. This technique can be an alternative eco-friendly technique for the integrated pest management of these weevils. Knowledge about pest dispersal and movement and their relationship to pest management practices, environmental factors and ecosystem can be used to predict pest infestations. An understanding of pest dispersal in important for the success of IPM programmes (Aylor & Irwin 1999; Byrne 1999). Knowledge about the prevalence of male biased gene flow in *Odoiporus longicollis* (Oliver) could aid in the IPM of bananas.

Due to the ever increasing demands for increased food production it is necessary to develop and apply novel biotechnology methods towards improving crop varieties in reasonably short time in a cost-effective way. The major objective of IPM approach is to enhance the resistance of plants to pests and pathogens by using a combination of control strategies including the judicious use of pesticides, crop rotation, field sanitation, and above all exploitation of inherently resistant plant varieties. These IPM based farming practices include several approaches such as chemical control, biological control, cultural control, mechanical control, and genetic control. Genetic control or the use of transgenic crops (Estruch *et al*. 1997) involves expression of foreign insecticidal genes and they have made a significant contribution to sustainable agriculture and is an important component of IPM. Haq *et al.* (2004) has reviewed the use and advantages of plant protease inhibitor genes to combat against insects, pests, and pathogens by developing resistant crop varieties. The development and use of transgenic crops expressing *Bacillus thuringiensis* (Bt) proteins has many success stories. Another insecticidal gene from the bacterium *Bacillus cereus*, known as the vegetative insecticidal protein (vip) gene has been found to be effective against corn rootworms (Chrispeels & Sadava 2003). Presently the main strategy adopted to develop insect-resistant plants via genetic engineering is

based on the use of Bt toxin genes. This has helped in the development of transgenic crops with a high resistance to specific insects. These Cry protein genes (crystal proteins or endotoxins) have the advantages of high specificity, a short life in the environment, and a high and fast toxic activity over conventional control techniques (Chrispeels & Sadava 2003; Jouanin *et al.* 1998).

However, limitations of transgenic Bt plants include increased persistence of the Bt toxin within the plant leading to development of insect resistance (Moar *et al.* 1995), narrow range of target insects, and a complex mode of action because they need to be proteolytically activated after solubilization in the guts of susceptible insects (Knowles 1994; Schnepf *et al.* 1998). So there is a need to find alternative candidate genes for such purpose (Oppert *et al.* 1997). Genes encoding proteinase and α-amylase inhibitors for control of insects have also found wide application (Discussed in Chapter 1). These transgenic systems are sustainable and have decreased inputs of energy and chemicals, and at the end will not generate harmful outputs such as pesticide residues (Boulter 1993).

α-amylase inhibitors play important roles in plant defense mechanisms, particularly against insects, and several of these inhibitors have been expressed in different crops to increase their resistance to particular insects (discussed in chapter 1). In Chapter 6, the interacting residues between inhibitor and pest amylase have been identified as well as conserved cysteine residues have been identified which form disulfide bonds which are important in establishing the active conformation of these inhibitors (Laskowski & Kato 1980). α-amylases are ubiquitous proteins and play a key role in carbohydrate metabolism in various insects, especially the ones which feed on starchy seeds during their development and thus depend on their a-amylases for survival (Franco *et al.* 2000; MacGregor *et al*. 2000). The Uganda Banana Biotechnology Project (UBBP) in collaboration with the Forestry and Agricultural Biotechnology Institute (FABI) is currently investigating the potential of cysteine proteinase inhibitors from rice and papaya in GM banana. *In vitro* studies at FABI have shown that these cysteine proteinases are strongly inhibited by both a purified recombinant rice cystatin, oryzacystatin-I (OC-I) (Abe *et al.* 1987) and papaya cystatin (Kiggundu *et al*. 2003). These recombinant papaya cystatin significantly reduce the early growth and development of weevil larvae (Kiggundu *et al.* 2003). Additionally, site-directed

mutagenesis is being carried out at FABI to improve inhibition and stability of the inhibitor for weevil control. To confer resistance for nematodes, a transgenic banana variety expressing a modified OC-I gene has been produced by the John Innes Centre (JIC, UK) and the University of Leeds (UL, UK) (Philip Vain, personal communication). The field testing of these plants is proposed to be done in Uganda.

In addition, lectin gene isolated form snowdrop, *Galanthus nivalis* agglutinin (GNA), is toxic to several insect pests in the orders Homotera, Coleoptera and Lepidoptera (Tinjuangjun 2002). Currently the effect of GNA and the *Aegopodium podagraria* lectin (APA) among others, on three nematode species pathogenic to banana is being tested in-vivo at Katholieke Universiteit Leuven (KUL) in Belgium in collaboration with the UBBP (Carlens 2002). All such genes are potential candidates for weevil control.

In the present work the genes encoding the monomeric and dimeric inhibitors from wheat which inhibit the amylase from this pest have been cloned and the specificity of interaction between these inhibitors and the amylases has been studied using homology modelling. This study is the first step towards developing transgenic banana plants with improved tolerance to this pest,

The way forward

There is no doubt that plant biotechnology, in recent years has played a key role in the sustainable production of several crop varieties (for example green revolution in the 1950s and 1960s). Dr. Norman E. Borlaug received the 1970 Nobel Prize, for his work in the 'Green Revolution,' which saved millions of lives from famine in India, Mexico, and the Middle East. In the 1940s, he began research in Mexico and developed new disease resistance high-yield varieties of wheat. The success of the Green Revolution in Mexico, led to the spread of these technologies worldwide in the 1950s and 1960s. Similarly, the above mentioned efforts for the development of transgenic banana varieties can revolutionize the banana production in the developing world. In order to achieve these objectives, it is important to understand resistance mechanisms and the development of rapid screening methods Kehe *et al.* (2000). Further work on the identification of

insecticidal proteins, other than Bt and proteinase inhibitors against the banana weevil is needed. This also emphasizes the development of a reliable and efficient bioassay system. Work was done at CIRAD in this direction (Kiggundu *et al*. 2010) At FABI, a system to rear weevil larvae on banana stems vacuum infiltrated with inhibitors has been developed to test *in vivo* effects on larval development of the inhibitors which will help in rapid discovery of proteins with insecticidal properties to banana weevil (Kiggundu *et al*. 2010) For a sustainable solution to pest problems for banana crops, high yielding and pest resistant cultivars can be developed by using genetic modification techniques.

The successful transformation of bananas may depend on factors such as: the consistent expression of desired traits, at the required level and no interference of the transgene with other metabolic processes (Van der Vyver *et al.* 2003a; Foyer *et al.* 2003). Since banana plants are polyploid, GM banana needs to be vegetatively propagated. So there are serious concerns about the genomic integrity of the off-spring plant which may led to mutations and epigenetic changes as a result of the plant transformation process (Karp 1993; Phillips *et al.* 1994; Labra *et al.* 2001). Future research must also focus on the long-term performance of GM bananas under various field conditions, including the long-term response of weevils and effects to their natural predators.

The information obtained from the present work will help in the development of transgenic banana plants with resistance against banana pseudostem weevils. Losses due to this pest can be minimized by the development of transgenic banana and thus these amylase inhibitor genes could make a promising contribution along with other methods of control towards maximizing yields in the third world countries. Although lot of progress has been done in other crops there is no work done for developing bananas with improved tolerance towards *O. longicollis* (Oliver). The information obtained from this study can provide building blocks for developing full fledged IPM approaches based on banana transgenics. Although, there is a plethora of information available, this small contribution has a lot of promising possibilities of pest control through plant-derived inhibitor genes, but this work needs to be further explored and prudently tapped for implementation in IPM programmes.

References

Abe K, Kondo H and Arai S (1987) *Agricultural and Biological Chemistry*, **51**, 2763-2768.

Aylor DE and Irwin ME (1999) *Agricultural and Forest Meteorology*, *97*(4), 233-234.

Boulter D (1993) *Biochemistry*, **34(6)**, 1453-1466.

Byrne DN (1999) Special issue: Aerial dispersal of pests and pathogens. *Agricultural and Forest Meteorology*, **97**, 309–316.

Caprio MA and Tabashnik BE (1992) *Journal of Economic Entomology*, **85(3)**, 611-620.

Carlens K (2002) In: Novel Approaches to the improvement of banana production in Eastern Africa-the application of biotechnological methodologies Project Report (Annexes) INIBAP, Montpellier France.

Carey JR (1991) *Science*, **253(5026)**, 1369-1373.

Charlesworth (2003) *Phiosophical Transactions of the Royal Society, London.* B, **358**, 1051–1070

Chrispeels MJ and Sadava DE (2003) In: Plants, Genes and Crop Biotechnology. Edited by Chrispeels MJ, Sadava DE. Jones and Barlett Publishers; 2003: 52-75

Daly JC and Gregg P (1985) *Bulletin of Entomological Research*, **75(01)**, 169-184.

Daly JC (1993) *Genetica*, **90(2-3)**, 217-226.

Dunley JE and Croft BA (1994) *Experimental & Applied Acarology*, **18(4)**, 201-211.

Dutt N and Maiti BB (1972) *Indian Journal of Entomology*, **34(1)**, 20-30.

Estruch JJ, Carozzi NB, Desai N, Duck NB, Warren GW and Koziel MG (1997) *Nature Biotechnology*, **15(2)**, 137-141.

Fauvelot C, Lemaire C, Planes S and Bonhomme F (2007) *Heredity*, **99,** 331-339.

Foyer CH, Groten K and Kunert K (2003) In: Encyclopaedia of Applied Plant Sciences, edited by Thomas B, Murphy DJ, and Murray B. New York, London: Academic Press. 419–430, 2003.

Franco OL, Rigden DJ, Melo FR, Bloch Jr C, Silva CP and Grossi de Sa MF (2000) *European Journal of Biochemistry*, **267(8)**, 2166-2173.

Gould F (1994) *Biocontrol Science and Technology*, **4(4)**, 451-461.

Greathead DJ and Greathead AH (1992) *Biocontrol News and Information*, **13(4)**, 61N-68N.

Hall R and Ehler LE (1979) *Bulletin of the ESA*, **25(4)**, 280-283.

Haq SK, Atif SM and Khan RH (2004) *Archives of Biochemistry and Biophysics*, **431(1)**, 145-159.

Jouanin L, Bonade-Bottino M, Girard C, Morrot G and Giband M (1998) *Plant Science*, **131(1)**, 1-11.

Karp A (1993) In: Genetic instability in regenerated and transgenic plants. Agro-Food-Industry Hi-Tech May/June, 7-12.

Kehe M, Toto HC and Gnonhouri P (2000) In: Craenen, K., Ortiz, R., Karamura, E. B., Vuylsteke, D. R., (Eds.), *Acta Horticulturae*, **540**, 497-503.

Kiggundu A, Kunert K, Viljoen A and Muchwezi JM (2003) In: Genetic transformation strategies to address the major constraints to banana and plantain production in Africa, 119-130, *Montpellier, France: INIBAP*, 74-107.

Kiggundu A, Muchwezi J, Van der Vyver C, Viljoen A, Vorster J, Schlüter U, Kunert K and Michaud D (2010) *Archives of Insect Biochemistry and Physiology*, **73(2)**, 87-105.

Knowles BH (1994) *Advances in Insect Physiology*, **24**, 275-308.

Korman AK, Mallet J, Goodenough JL, Graves JB, Hayes JL, Hendricks DE and Wall M (1993) *Annals of the Entomological Society of America*, **86(2)**, 182-188.

Labra M, Savini C, Bracale M, Pelucchi N, Colombo L, Bardini M and Sala F (2001) *Plant Cell Reports*, **20**, 325-330.

Laskowski Jr M and Kato I (1980) *Annual Review of Biochemistry*, **49(1)**, 593-626.

MacGregor EA, Bazin SL, Ens EW, Lahnstein J, Macri LJ, Shirley NJ and MacGregor AW (2000) *Journal of Cereal Science*, **31(1)**, 79-90.

Mallet J and Porter P (1992) *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **250(1328)**, 165-169.

Moar WJ, Pusztai-Carey M, Van Faassen HENK, Bosch D, Frutos R, Rang C and Adang MJ (1995) *Applied and Environmental Microbiology*, **61(6)**, 2086-2092.

Narang SK, Tabachnick WJ and Faust RM (1993) In: Narang SK, Barlett AC, Faust RM, editors. *Applications of Genetics to Arthropods of Biological Control Significance*, 19–52. Boca Raton, Florida: CRC Press, Inc.

Nei M and Roychoudhury A K (1972) *Science* **177**, 434-436.

Oppert B, Kramer KJ, Beeman RW, Johnson D and McGaughey WH (1997) *Journal of Biological Chemistry*, **272(38)**, 23473-23476.

Phillips RL, Kaeppler SM and Olhoft P (1994) *Proceedings of the National Academy of Science USA*, **91**, 5222-5226.

Powell W and Walton MP (1989) In: Loxdale HD, den Hollander J, editors. *Electrophoretic Studies on Agricultural Pests*, 443–65. Oxford: Clarendon.

Prasuna AL, Jyothi KN, Prasad AR, Yadav JS and Padmanaban B (2008) *Current Science (00113891)*, **94(7)**.

Roderick GK (1996) *Annual Review of Entomology*, **41(1)**, 325-352.

Schnepf E, Crickmore N, Van Rie J, Lereclus D, Baum J, Feitelson J and Dean DH (1998) *Microbiology and molecular biology reviews*, **62(3)**, 775-806.

Simmonds NW (1962) The Evolution of the Bananas. Longmans, London, Great Britain.

Smith RF and Reynolds HT (1966) In: *Proceedings of FAO (United Nations Food and Agriculture Organisation) Symposium on Integrated Pest Control*, Vol. 1, FAO, Rome, pp. 11–17.

Tabashnik BE (1994) *Annual Review of Entomology*, **39(1)**, 47-79.

Tinjuangjun P (2002) Snowdrop lectin gene in transgenic plants:its potential for Asian agriculture. Agbiotechnet.com (ABN091).

Unruh TR and Woolley JB (1999) Molecular *Biological Control*, 57-85.

Van der Vyver C, Schneidereit J, Driscoll S, Turner J, Kunert KJ and Foyer CH (2003a) *Plant Biotechnology Journal*, **1**, 101-112.

Wright S (1978) Evolution and the Genetics of Populations, Vol. 4. University of Chicago Press, Chicago, Ill.

Following are the List of Manuscripts which are based on the thesis:

1. "Cloning and sequence analysis of the amylase gene from the rice pest *Scirpophaga incertulas* Walker and its inhibitor from wheat (variety MP sehore)."*International Journal of Insect Science* 1 (2009): 29-44.

2. "Molecular characterization and utility of the internal transcribed spacers I and II in genetic diversity analysis of Indian populations of *Odoiporus longicollis* (Oliver), a serious pest of bananas".--manuscript communicated.

3. "Male biased gene flow in banana pseudostem weevil as revealed by analysis of the COI-tRNALeu- COII region." --manuscript communicated.

4. "Genetic diversity among Indian populations of the banana pseudostem weevil (*O. longicollis* Oliver) as revealed by RAPDs, ISSRs and AFLPs."--Manuscript under preparation.