RESEARCH REPORT No. 287 **ISSN 0970-8103**

TRACING THE ORIGIN AND SPREAD OF TEAK DEFOLIATOR OUTBREAKS THROUGH A MOLECULAR APPROACH

(KFRI 366/01 - DBT Project File No. BT/PR/1656/AGR/05/105/99 dated 15.2.2001-
(March 2001- August 2004)

V.V.Sudheendrakumar
T.V.Sajeev
R.V.Varma
Forest Protection Division, KFRI

Moinak Banerjee
N. Chandrasekhar
Human Molecular Genetics Laboratory
Rajiv Gandhi Centre for Biotechnology
Thiruvananthapuram

Kerala Forest Research Institute
Peechi- 680 653, Kerala, India

2. CONTENTS

ABSTRACT

The teak defoliator, *Hyblaea puera*, is recognized as the most important pest of teak in India. Outbreak of this insect is a regular annual phenomenon in teak plantations, which results in substantial reduction in volume increment of teak. An understanding of the origin and spread of the outbreaks of this pest, which erupt all of a sudden following the pre-monsoon rain every year, is an important prerequisite for developing appropriate control strategies. In large teak plantation areas, *H. puera* outbreaks begin in small epicentres and later spread to larger areas. The present study was undertaken in about 8,500 ha of teak plantations (latitudes 11^{o} 10'N and 11^{o} 25'N and longitudes 76^{o} 10'E and 76^{o} 25'E) in Nilambur, Kerala during the pest incidence season in 2001-2002 to elucidate the relationship between the different populations of the teak defoliator. The objectives of the study were: 1. To gather information on the spatial and temporal pattern of distribution of teak defoliator populations during the early stage of outbreak in teak plantations. 2. To examine the genetic relationship between discrete populations of teak defoliator during the build-up of an outbreak using molecular techniques and thereby illuminate the origin of outbreak populations.

The area was divided into 19 blocks and 189 Observation Units (Ous).The average area of one OU was about 50 ha. The outbreaks were monitored at fortnightly intervals with the help of twenty trained field workers. The populations were classified into 'endemic', 'epicenter' and 'epidemic' populations based on the time of occurrence and size of infestation. Populations were classified as 'endemic', 'epicenter' and 'epidemic', based on their time of occurrence and the density of the population as represented by the area it infests. Endemics are insects belonging to the low-density population level; epicenters are patchy, medium density outbreaks that occur during the pre-monsoon season, whilst epidemic represents large area, high-density outbreak populations. Using the duration of each instar (egg- 1 day; instars I and II -2 days each, instars III to V- 3 days each; pre-pupa- 1 day and pupa- 4 days), the temporal data on outbreaks was examined to see whether each subsequent outbreak

Deleted: plantations which

Deleted: in substantial

Deleted: 1

Deleted: ¶
¶

Deleted: The plantation maps were prepared by interpreting the aerial photographs. Only the teak plantations were interpreted and mapped. The map was brought to the scale 1:50,000 and major features like drainage and roads were marked by superimposing it on Survey of India (SOI) topographic sheets of the same scale

could be explained on the basis of previous outbreak. The larval samples of the population suspected to be related were subjected to molecular analysis to confirm genetic relatedness.

A novel method of screening for nuclear and mitochondrial DNA polymorphism using (RAGEPs) was standardized for estimating the genetic variation within and between populations. This method is based on PCR technique using single gene specific primers with nil to moderate level of degeneracy. Based on the criteria of polymorphic content and specificity to the teak defoliator genome, 11 nuclear Random Amplified Gene Encoded Primers (RAGEPs) and 11 mitochondrial RAGEPS were selected from a batch of 57 n-RAGEPs and 37 m-RAGEPs. Using this method, the relationship between different populations was traced out.

The ecological data generated suggested a relationship between endemic populations and some of the epicenter populations and similarly between epicenter populations and some of the outbreak populations. However, the molecular studies did not reveal any relationship between endemic and epicenter populations. The study thus gave little evidence to show that the aggregation of moths belonging to the endemic populations causes the epicenter populations. The study reconfirmed the relationship between epicenter populations and some of the epidemic populations (Outbreaks) as revealed through the ecological studies. This finding is relevant from the point of view of management of the teak defoliator to some extent through managing the epicenter populations which occupy comparatively a small area in large scale plantations.

1. INTRODUCTION

The teak defoliator, *Hyblaea puera*, is recognized as the most important pest of teak in India. Outbreak of this insect is a regular annual phenomenon in teak plantations, which results in substantial reduction in volume increment of teak (Nair *et al* 1985). Repeated outbreaks occur every year during the early part of the growth season. Although biological control using parasitoids was tested in the past it did not succeed due to the unique population dynamics of the pest involving shifting foci of infestations. Chemical control was also tested experimentally but again it has its limitations.

Deleted: plantations which

Deleted: in substantial

Deleted: involving shifting

An understanding of the origin and spread of the outbreaks of this pest, which erupt all of a sudden following the pre-monsoon rain every year, is an important prerequisite for developing appropriate control strategies. Evidence gathered from the past decade on the population dynamics of *H. puera* indicates habitual, short range movements of emerging moth populations, suggesting that these spread to larger areas, generation after generation, affecting entire teak plantations (Nair and Sudheendrakumar, 1986). Nair and Mohandas, 1996 reported that in large teak plantation areas, *H. puera* outbreaks began in small epicentres and later spread to larger areas. It was suspected that population build-up in the early outbreak epicentres may account for the subsequent widespread outbreaks. It appeared that the populations could be classified as 'endemic', 'epicenter' and 'epidemic', based on their time of occurrence and the density of the population as represented by the area it infests. Endemics are insects belonging to the low-density population level; epicenters are patchy, medium density outbreaks that occur during the pre-monsoon season, whilst epidemic represents large area, high-density outbreak populations. It was felt that an understanding of the cause-effect relationship between the initial small outbreaks and the large outbreaks that occur later is crucial for practical control of the pest.

A study carried out in a 10,000 ha plantation area at Nilambur, Kerala, showed that some of the epicenters do contribute to development of outbreaks (Nair *et al.,* 1988). However, all such outbreaks could not be attributed to local build up, indicating the possibility of the

involvement of jmmigrant populations. The above study using the time lapse (developmental time) between two outbreaks to determine whether an earlier outbreak was responsible for causing the subsequent outbreak was not fully reliable even when coincidence occurred. It becomes therefore necessary to apply advanced methods to address this problem. Hence the study was undertaken to verify the results of the previous observations.

Deleted: immmigrant

Recently, molecular markers have been used to enhance understanding of insect displacements, especially including estimates of movement of particular genotypes and/or biotypes, reproductive strategy and success. Such approaches have also been used to study founder events (Taberner *et al*., 1997), geographical invasions (Stone and Sunnucks, 1993), small and large scale displacements (Daly and Gregg, 1985; Loxdale and Lushai, 2001) including movement of entire population (Loxdale and Lushai, 1999), and even altitudinal movements related to habitat patchiness and persistence (Liebherr,1988). Molecular data can yield valuable information when integrated with information from ethology, field ecology, comparative morphology, systematics and paleontology (Avise, 1994). Use of direct and indirect methods of tracking insects along with description of the role and utility of various molecular markers – protein and DNA – in monitoring insect dispersal, has been extensively reviewed (Osborne *et al*., 2001).

Arbitrarily-primed DNA markers involving the polymerase chain reaction (PCR), have proved very useful for genetic fingerprinting and for facilitating positional cloning of genes. This class of markers are particularly important for less studied species, for which genome sequence information is generally not known. These technologies include randomly amplified polymorphic DNA (RAPDs) (Welsh and McClelland, 1990; Williams *et al*., 1990), DNA amplification fingerprinting (DAF) (Caetano-Anolles *et al*.,1991) and amplified fragment length polymorphisms (AFLPs) (Vos *et al*., 1995).

RAPD is one of the most powerful and efficient method for analysing any species whose genome content has not been worked out previously. RAPD markers are produced by PCR using short oligonucleotide primers of randomly chosen sequence. More than 400 different 10-base primers are commercially available to identify RAPD variation. The RAPD process

typically reveals several polymorphic genetic segments per primer within population. Other segments per primer may appear as monomorphic bands within or across populations. The degree of variability observed for many primers suggests that this technique will be useful for a variety of questions, including individual identification, paternity analysis, strain identification and phylogenetic analysis.

In this study, it was envisaged to generate ecological data to interpret the relationship between different populations of the teak defoliator and to verify the same using molecular tools and techniques.

The objectives of the study were:

1. To gather information on the spatial and temporal pattern of distribution of teak defoliator populations during the early stage of outbreak in teak plantations.

2. To examine the genetic relationship between discrete populations of teak defoliator during the build-up of an outbreak using molecular techniques and thereby illuminate the origin of outbreak populations.

Deleted: ¶
The present project is the logical continuation of an earlier work carried out by Nair et al (1998) in which an attempt was made to trace the Hyblaea outbreaks

Deleted: epicentres of

Deleted: the teak plantation areas, H. puera outbreaks begin in small epicentres and later spread to larger areas. It was suspected that population build-up in early outbreak epicentres may account for the subsequent widespread outbreaks. However, the study using the time lapse (developmental time) between the two outbreaks to determine whether an outbreak is related to an earlier out break showed that all the subsequent outbreaks could not be attributed to previous outbreaks indicating the possibility of migrant populations.¶
¶
The present study uses Molecular methods, based on PCR and Random Amplified Gene Encoded Primers to understand the relationship between different populations.¶

2. MATERIALS AND METHODS

2.1. Ecological studies

2.1.1. Outbreak monitoring and collection of insect samples for studies

The study was carried out in about 8,500 ha of teak plantations (latitudes 11^{o} 10'N and 11^{o} 25'N and longitudes 76^{o} 10'E and 76^{o} 25'E) in Nilambur, Kerala during the pest incidence season in 2001-2002. The area was divided into 19 blocks and 189 Observation Units (OUs) (Fig. 1) based on natural boundaries like streams, roads and footpaths. The average area of one OU was about 50 ha.

Twenty trained field workers were employed to carry out teak defoliator outbreak monitoring. Observations were made at fortnightly intervals. The details of location of pest incidence and extent of infestation were later transferred to the field map to understand the spatial pattern of infestation.

Larval samples were collected from the infestation sites. Whenever fifth instar larvae were available, ten larvae were preserved in 70 % alcohol and stored in deep freezer (-20^{o}C). If only lower stages like third or fourth instar were available they were reared up to 5th instar in the laboratory. A total of ten 5th instar larvae were preserved from each sample. Balance of the larvae were reared out to the next generation. Each sample was assigned code number containing the details of Year/Month/Date/Block/Grid/ Generation for further reference.

A separate culture was maintained in the laboratory for each distinct outbreak patch. It was labeled as "P" (Parent generation) and from each lot 10 larvae were preserved with code as the above for the analysis. These larvae were reared into next generations to get the F1 population. A sample from F1 was also reared into the next generation. Samples of P, F1 and F2 generations were preserved for molecular characterization.

Deleted: 1

Deleted: ¶
¶

Deleted: The plantation maps were prepared by interpreting the aerial photographs. Only the teak plantations were interpreted and mapped. The map was brought to the scale 1:50,000 and major features like drainage and roads were marked by superimposing it on Survey of India (SOI) topographic sheets of the same scale

Comment [KN1]: Not relevant here.

Deleted: Data on flushing level of teak, intensity of defoliation due to Hyblaea , stages of the pest, nature and extent of incidence

Deleted: .

Comment [KN2]: No need for a separate subheading

Deleted: Rearing and preservation of subsequent generations:¶
¶

Using the duration of each instar (egg-one day; instars 1 and 2 –two days each, instars 3 to 5- three days each; pre-pupa- one day and pupa—four days) the temporal data on outbreaks was examined to see whether each subsequent outbreak could be explained on the basis of previous outbreak. Based on the ecological data, possibly related populations were identified (Nair, *et al*., 1988) which were subjected to molecular studies.

Comment [KN3]: This much detail is not necessary. Will confuse the readers who are not familiar with Hyblaea.

Deleted: To examine this, the generation time from egg to egg was used. Based on the duration of each instar (given above) and pre-oviposition period of 2 days, the egg-to-egg period works out to 21 days. Unpublished laboratory rearing data had shown that this period was sometimes completed in 19 days, when the larvae were maintained on teak leaf at ambient temperature. Therefore 19 days was taken as the minimum generation time. The maximum generation time was taken as 26 days by adding the oviposition period of 5 days

Deleted: (Sudheendrakumar

Deleted: ,

Deleted: 1994)

Deleted: to the normal egg-to-egg period of 21 days. Thus, if an infestation (egg laying) was observed between the 20th and 26th day of an earlier infestation, it could be argued that the second represented the F1 generation of the first. While there is no certainty about this, the converse is always true, i.e., if the second outbreak did not occur during that interval, it was not caused by moths produced in the first outbreak, although the borderline cases may be suspect because of possible variation in the generation time under natural conditions. ¶

2.2. Molecular studies

2.2.1. DNA Isolation:

The total DNA from the individual insect larva was extracted using the protocol as described by Andrew and Gary (1995) with a minor modification of the method of isolation and purification. The protocol is as described below:

1) Powder the larval sample using liquid Nitrogen.
2) Add 500µl Lifton buffer Proteinase-K (0.2 M sucrose, 0.05 M EDTA, 0.1 M Tris, 0.5% SDS, (pH 9.0) Proteinase K should be 50ug/ml.
3) Incubate for 55°C for 2 hours invert the tubes for every 30 mins or 37°C for overnight.
4) Add 75µl of 8M Potassium acetate vortex and incubate in ice for 30 mins and Spin at 14,000 RPM for 15 Mins
5) Add 1µl of 10mg/ml RNAse, Invert the mix and incubate for 42°C for 1 to 2 hours
6) Add 300µl of Tris saturated Phenol with pH 7.5 and 300µl of 24:1 chloroform isoamyl alcohol and gently invert the tubes
7) Spin the tube at12, 000 RPM for 5mins
8) Remove the Aqueous phase to another new tube
9) Add 400µl of 24:1 chloroform isoamyl alcohol and invert the mix gently and spin it 12, 000 RPM for 5 mins
10) Remove the aqueous phase to another tube add 50µl of 3M sodium acetate and add 95% ethanol and incubate for overnight in ice or -20°C
11) Spin for 14,000 RPM for 15mins and decant the supernatant
12) Wash the Pellet twice in 70% ethanol at 14,000 RPM for 5mins
13) Vacuum or air dry the DNA pellet and resuspend in TE buffer for longer storage

The DNA extracted was quantified using a spectrophotometer at 260nm (Shimadzu). The quality of the DNA was checked by taking the absorbance ratios of 260/280 nm.

2.2.2. Polymerase Chain Reaction

The endemic, epicenter and outbreak populations of *H. puera* were discriminated using novel method of screening of nuclear and mitochondrial DNA polymorphism using Random Amplified Gene Encoded Primers (RAGEPs). This method was based on PCR using single

gene specific primers with nil to moderate level degeneracy. RAGEP-PCR reactions were performed at low stringency using gene specific primers obtained from University of British Columbia (UBC). RAGEP protocol is based on the principle of RAPD PCR.

Both nuclear and mitochondrial DNA RAGEP amplifications were performed in a total volume of 30µl. The reaction mixture that gave the best reproducible results was used for routine RAPD- PCR's. (Table 1).

Each reaction consisted of 1x Taq buffer with 1.5mM $MgCl_2$, 1.2 U of *Taq* polymerase (BG), 0.25mM of dNTPs (Amersham) and 12 pM of primer per reaction. Primers were initially screened for polymorphism and repeatability. Amplifications were performed in similar cycling conditions in a Thermocycler (Biorad) programmed as follows: initial denaturation at 95°C for 5 min followed by 45 cycles of cycle denaturation at 94°C for 1min., annealing at 36°C for 1min., extension at 72°C for 2 min. and final extension at 72°C for 5 mins. The amplification products were separated using 1.2% agarose gel in 0.5 x TBE buffer with ethidium bromide staining to visualize the product separation using a Bio-Rad's Fluor S imager. The molecular weight of each band was estimated by comparing with a co-migrating 100-bp ladder (Amersham). RAGEP fingerprints of each sample from different regions were then interpreted using Fingerprint type module of Bionumerics software (Applied Maths Kortrijk Belgium, ver.2.0).

Table 1. Standardized PCR reaction mixture

Components	**Concentration**	**Company**	**Working concentration**	**Volume**
Taq Polymerase (BG)	3units/1µl	Banglore Genei	1.2 units	0.4µl
Taq Reaction Buffer	10x with 1.5mM $Mgcl_2$	Banglore Genei	1x	3.0µl
DNTPs (Amersham)	2.5mM	Amersham Pharm.	0.25mM	3.0µl
Primer		UBC	12 p moles	3.0µl
Sterile distilled water				18.60µl
Total Volume				28µl

A preliminary screening with 50 nuclear RAGEPs and 37 mitochondrial RAGEP primers were evaluated for polymorphism and repeatability. Only 11 nuclear (Table 2) and 11 mitochondrial RAGEPs (Table 2 & 3) from each group were selected for the study, as they showed constant repeatability of highly polymorphic patterns. Species specificity was evaluated by comparing the banding patterns in *H.puera* with those from the Teak skeletonizer, *Eutectona. machaeralis*, Leaf roller, *Sylepta derogata* F., Leaf folder, *Cnaphalocrocis medinalis* Guenée, and the Silk worm, *Bombyx mori* L.

Deleted: ¶

Table 2. Insect nuclear gene specific Primers sequences used in RAGEP PCR

Deleted: ¶

Sl.No	Primer	Sequence
1	EFS599	ATC TCC ggA Tgg CAC gg(CT) gAC AA
2	22.5drc	gAA CCA (Ag)TT (Ag)AC (Ag)Tg (Ag)AA gAT C
3	LEPWG1	gA(Ag)Tg(CT)AA(Ag)Tg(CT)CA(CT)gg(CT)ATgTCT gg
4	4 CK6–5’	gAC CAC CTC CgA gTC ATC TC(Cg) ATg
5	CK7–3’	CAg gTg CTC gTT CCA CAT gAA
6	CytC-B–3’	CAT CTT ggT gCC ggg gAT gTA TTT CTT
7	EF1–5’	gAC AAC gTT ggC TTC AAC gTg AAg AAC g
8	8 Tub3–5’	gAT TTg gAg CC(AgCT) gg(AgCT) ACC ATg gA
9	18S-A1984	TCC CTg gTT gAT CCT gCC AgT A
10	S1124	AgC gTA Tgg C(AC)T C(Ag)A AgAACT g
11	rcM4	ACA gC(CgA) AC(gT) gT(TC) Tg(CT) CTC AT(Ag) TC

Table 3. Insect mitochondrial gene specific primers sequence used in RAGEP PCR

Sl.No	Primer	Sequence
1	C1-J-2183	CAA CAT TTA TTT TgA TTT TTT gg
2	TL2-J-3034	AAT ATG gCA gAT TAg TgC A
3	C2-N-3661	CCA CAA ATT TCT gAA CAT TgA CCA
4	TK-N-3785	gTT TAA gAg ACC AgT ACT Tg
5	N4-N-8924	AAA gCT CAT gTT gAA gCT CC
6	CB-J-10612	CCA TCC AAC ATC TCA gCA TgA TgA AA
7	LR-J-12887	CCg gTC TgA ACT CAg ATC ACg T
8	LR-J-13417	ATg TTT TTg TTA AAC Agg Cg
9	LR-N-13398	CgC CTg TTT AAC AAA AAC AT
10	SR-J-14233	AAg AgC gAC ggg CgA TgT gT
11	CB-N-10920	CCC TCA gAA TgA TAT TTg TCC TCA

2.2.3. Genetic variation studies using *Hyblaea* population samples

Five endemic populations, 25 epicentre populations and 7 outbreak populations of the year 2002 were included in the study. Fig. 3 shows the location of some of the samples in the study area. The relationship between endemic and epicentre populations, epicentre and first outbreak populations was estimated using the nuclear and mitochondrial primers.

2.2.4. Data Analysis

The polymorphic content for nuclear and mitochondrial primers was analyzed using Bionumerics software (http://www.biosystematica.com/bionumer.htm, Applied Maths Kortrijk Belgium, ver.2.0). Band search parameters were kept constant as 5% minimum profiling for all the gels. The position tolerance for selection of bands in constructing a dendrogram was kept constant at 1% through out the interpretations. Only bands showing clear and reproducible patterns were included in the final analysis and these were scored. Real-time normalization of gel electrophoresis patterns and band position for all the gels was based on the reference system for the species specific bands. Normalization helped to control the brightness and streakiness of bands without altering the lighter bands and also the inter-gel mobility shifts. Subsequently a data matrix of similarity values was produced for each individual for each marker. The Dice coefficient was used to analyze the similarities of the banding patterns. Consensus similarity matrix and dendrogram based upon individual matrices from different markers were used for pair wise clustering based on Unweighted pair group method (UPGMA) with average linkages (11). The UPGMA dendrogram prevails on the assumption that nucleotide substitution rates are same across all branches. It employs a sequential clustering algorithm, in which local topological relationships are identified in order of similarity, and the phylogenetic tree is built in a stepwise manner. All analyses were done using Bionumerics software V-2.

Deleted: And t

3. RESULTS AND DISCUSSION

3.1. Ecological studies

The teak defoliator populations, which occurred during the years 2001, 2002 and 2003, were classified as endemic, epicentre and epidemic (outbreak) populations (Table 4). The sequences of infestation during the year 2001, 2002 are presented in Fig. 2a &b.

Deleted: Details of the data generated under the population aspects are presented in Table 1.:¶

Table 4. Teak defoliator populations during 2001-2003 in Nilambur teak plantations

Nature of Population	2001		2002		2003	
	No. of patches	Area infested (ha)	No. of patches	Area infested (ha)	No. of patches	Area infested (ha)
Endemic*	20	-	14	-	1	-
Epicentre**	29	42.5	27	84.25	5	17.5
Epidemic***	54	548	138	7675	64	2869

*Stray occurrence of insect in various life stages
**Aggregated insects of same life stage occurring in small patches – First sighting of patches- preoutbreak period
***Aggregated insects of same life stages occurring in patches of varying size (generally larger than the epicenters) during outbreak season

In the year 2001, twenty endemic samples were collected from the study area during the months of February and March. The epicentre populations occurred in the month of April in 29 discrete ocations. Subsequently, from April onwards, 54 large-scale outbreaks occurred in an area of 548 ha. In 2002. Fourteen insect samples were collected in the endemic phase during January, February and early March. A total of 27 discrete epicentre patches were observed during late March, infesting an area of 84.25 ha. The series of outbreaks, which occurred subsequently, infested an area of 7675 ha. In 2003, one endemic sample could be collected. Epicentre populations occurred during the month of March in 5 discrete patches infesting 17.5 ha. Large scale outbreaks during the year was spread out in 2869 ha. The temporal relationship between the endemic population and the epicentre populations and

that of the epicentre populations with the large-scale outbreaks were first worked out, the result of which is given in Table 5. Accordingly the larval samples from the populations which were suspected to be related were listed out (Table 6). Five endemic populations and 26 epicentre populations and 7 outbreak populations of the year 2002 were included in the study. Fig. 3 shows the location of the epicenters from where larval samples were collected for the study. The relationship between endemic and epicentre populations, epicentre and first outbreak populations was estimated using the nuclear and mitochondrial primers.

Table 5. List of teak defoliator populations sampled in 2002 with computed starting date/s and probable F_1date/s (next generation)

Sl.No	Code	Area	Class	Start date/s (Date of eggs noticed)	Probable F_1 dates (Date of appearance of next generation)
1	C	-	endemic	4-5/3	25/3-1/4
2	C1	-	endemic	4-5/3	25/3-1/4
3	B	-	endemic	5-12/3	25/3-8/4
4	D	-	endemic	5/3	26/3-1/4
5	E	-	endemic	6/3	27/3-2/4
6	K	0.5	epicentre	18-20/3	8/4-16/4
7	L	6.0	epicentre	19-21/3	9-18/4
8	N	5.0	epicentre	17-21/3	7-17/4
9	O	2.0	epicentre	17-20/3	7-16/4
10	P	0.5	epicentre	17-20/3	7-16/4
11	Q	5.0	epicentre	17-20/3	7-16/4
12	R	2.5	epicentre	17-20/3	7-19/4
13	S	0.5	epicentre	18-19/3	8-15/4
14	T	0.5	epicentre	18-19/3	8-15/4
15	U	5.0	epicentre	18-20/3	8-16/4
16	V	.25	epicentre	17-18/3	7-14/4
17	W	2.5	epicentre	17-20/3	7-16/4
18	Y	0.25	epicentre	17-21/3	7-17/4
19	Z	0.25	epicentre	17-21/3	7-17/4
20	AF	20.0	epicentre	17-19/3	7-15/4

Continued…….

Table 5 (continued). List of teak defoliator populations sampled in 2002 with computed starting date/s and probable F_1 date/s (next generation)

Sl.No	Code	Area	Class	Start date/s (Date of eggs noticed)	Probable F_1 dates (Date of appearance of next generation
21	AE	2.0	epicentre	17-19/3	7-15/4
22	AD	0.25	epicentre	17-19/3	7-15/4
23	AG	0.5	epicentre	17-19/3	7-15/4
24	AI	0.25	epicentre	17-19/3	7-15/4
25	AH	7.0	epicentre	17-19/3	7-15/4
26	AC	20.0	epicentre	17-19/3	7-15/4
27	AK	0.25	epicentre	20/3	10-16/4
28	M	0.25	epicentre	19-21/3	9-17/4
29	AB	2.0	epicentre	22-24/3	12-20/4
30	AJ	0.25	epicentre	20-24/3	10-16/4
31	AN	0.25	epicentre	19/3	9-15/4
32	AP	9.0	epidemic	7/4	28/4-4/5
33	BX	0.25	epidemic	18/4	29/4-5/5
34	BM	3.0	epidemic	6-8/4	27/4-5/5
35	AO	10.0	epidemic	7/4	28/4-5/5
36	BN	8.0	epidemic	6-8/4	27/4-5/5
37	BR	0.5	epidemic	6-8/4	27/4-5/5
38	BY	0.25	epidemic	10/4	1-7/5

Table 6. List of teak defoliator populations that could cause epidemic populations as per on temporal analysis.

Epidemic (outbreaks)	Suspected epicenter parent population/s
AJ	C, C_1, B
AP	N,O,P,Q,R,V,W,Y,Z,AF,AE,AD,AG,AI,AH,AC,B
AO	N,O,P,Q,R,V,S,T,U,W,Y,Z,AF,AE,AD,AG,AI,AH,AC,B
BX	L,R,AB
BM	K,N,O,P,Q,R,V,S,T,U,W,Y,Z,B,AF,AE,AD,AG,AI,AH,AC
BN	K,N,O,P,Q,R,V,S,T,U,W,Y,Z,B,AF,AE,AD,AG,AI,AH,AC
BR	K,N,O,P,Q,R,V,S,T,U,W,Y,Z,B,AF,AE,AD,AG,AI,AH,AC
BY	K.I.N.O.P.Q.R.S.T.U.V.W.Y.Z AF,AE,AD,AG,AI,AH,AC,AK,M,AJ,AN

3.2, Molecular studies

In this study, a variant of the RAPD approach involving various nuclear and mitochondrial gene specific primers was used to trace the origin of teak defoliator outbreaks.

The nuclear and mitochondrial gene specific primers chosen did not produce any amplification product when used in combination with the corresponding primers as described in the UBC primer set kit (University of British Columbia http://www.michaelsmith.ubc.ca/services/ NAPS/Primer_ Sets/). This resulted in our devising a novel PCR, which we have named RAGEP-PCR. In RAGEP-PCR, we used single nuclear and mitochondrial gene encoding primers at low stringency annealing temperatures. Unlike RAPDs, in RAGEP longer nuclear (21-26 nucleotide) and mitochondrial (19-26 nucleotide) gene encoding primers were utilised, and which we have here extensively employed to evaluate the species taxonomic specificity/reproducibility and to discriminate the endemic, epicenter and epidemic populations of teak defoliator from one another.

RAGEP markers were first tested for polymorphisms, species-specificity and constant reproducibility (Fig. 4). Similar fingerprinting patterns were observed in subsequent PCR for the same individual using the same primers, which display overall robustness and repeatability with RAGEP-PCR. It was also possible to discriminate various moth species based on their species-specific DNA fingerprint pattern (Fig.5). The bands scored for each nuclear RAGEP used in the present study were of a size range 200bp to 1500bp (Fig.6 A-D). With nuclear RAGEP markers, an average of 2-3 monomorphic bands were observed, except in the case of primer CK6–5'. In each marker, the average number of bands scored varied from 7-16. The maximum number of bands were detected using primer cytC-B-3', while the maximum number of monomorphic bands were detected using primer EFS599.

Each individual RAGEP marker gel was screened and a similarity matrix generated. Subsequently similarity matrixes of all experimental patterns were combined to generate a UPGMA (Unweighted pair-group mathematical average) tree. While evaluating the similarity

matrix based on the Dice coefficient for all nuclear specific RAGEP markers and whilst constructing a UPGMA tree, it was observed that the various population groups of *H.puera* fall in two clusters, which are further divided into two major sub clusters. Average similarity between the two major clusters was 20%, while that between the two sub clusters was 34%. In one of the major clusters, we observed all the endemic insects clustering together with some of the populations from the epicentre insects; however, both populations fall in two distinct sub-clusters (Fig 7). Similarly in the second major cluster, the remaining populations from the epicenter and entire epidemic insect populations, were likewise seen to fall into two distinct sub- clusters.

Using the mitochondrial RAGEP markers, the average numbers of bands scored for each primer ranged from 6-15 (Fig.8 A-H). All bands scored were of size range 300bp to 1600 kb. The maximum numbers of bands detected were found using primer SR-J-14233, the minimum numbers using marker N4-N-8924. Among mitochondrial markers, an average of 1-2 monomorphic bands were observed. The maximum number of monomorphic bands was observed using marker CB-N-10920. Two distinct clusters were observed in the UPGMA dendrogram for mitochondrial markers. Similarity between the two clusters was only 20%.

One of these clusters comprised the majority of the endemic samples with a few samples from epicenter insects, whilst the other cluster was comparatively larger and had two major sub clusters. Both these sub-clusters have insects from epicenter and epidemic populations (Fig- 9). From this dendrogram, it may be deduced that all the seven epidemic population samples, tested in the study shared the same gene pool with sets of epicenter populations. In contrast, the endemic populations are genetically distant from the epicentre populations.

Using the data from the molecular studies described above, the outbreak populations (epidemic) which could have been caused by epicentre populations (Table 3), were verified and the outcome is given in Table 7. It can be seen that all the outbreak populations tested, share the same gene pool as that of one or the other epicentre population suggesting their relationship.

Table 7. Genetic relationship between epicenter and outbreak *Hyblaea* populations

Outbreak population	**Suspected epicenter parent population**	**Genetically related epicenter parent populations**
AP	N,O,P,Q,R,V,W,Y,Z,AF,AE,AD,AG, AI,AH,AC,B	AF,AE,AD,AG,AI,AH,AC
AO	N,O,P,Q,R,V,S,T,U,W,Y,Z,AF,AE,A D,AG,AI,AH,AC,B	AF, AE,AD,AG,AI,AH,AC
BX	L,R, AB	R, AB
BM	K,N,O,P,Q,R,V,S,T,U,W,Y,Z,B,AF,A E,AD,AG,AI,AH,AC	AF, AE,AD,AG,AI,AH,AC
BN	K,N,O,P,Q,R,V,S,T,U,W,Y,Z,B,AF,A E,AD,AG,AI,AH,AC	AF, AE,AD,AG,AI,AH,AC
BR	K,N,O,P,Q,R,V,S,T,U,W,Y,Z,B,AF,A E,AD,AG,AI,AH,AC	N,O,P,Q,R,S,T,U,V,W,Y,Z
BY	K.I.N.O.P.Q.R.S.T.U.V.W.Y.Z AF,AE,AD,AG,AI,AH,AC,AK,M,AJ,A N	N,O,P,Q,R,S,T,U,V,W,Y,Z

4. GENERAL DISCUSSION

The origin and spread of the teak defoliator have always been a mystery. In large teak plantations vast area would get infested suddenly. Attempts have been made in the past to untie the knot in the dynamics of this pest. Studies carried out by Nair *et al* (1998) revealed that prior to the large- scale outbreaks, *H. puera* infestation begin in small epicenters and later spread to larger areas. It was then suspected that population build-up in the early outbreak epicenters might account for the subsequent widespread epidemic. In the study using the time lapse (developmental time) between two epidemics to determine whether an earlier epidemic was responsible for causing the subsequent outbreak showed that though not all, some of the subsequent outbreaks could be related to the epicenter populations. This study not so conclusive had to be repeated using advanced techniques. The present study addressed the above issue using molecular techniques.

Several technical advancements on the DNA fingerprinting methodologies have been established to resolve the taxonomic uncertainties and address the issue on species variability and migration.The RAGEP-PCR method described here uses gene-specific primers and randomly amplifies the nuclear and mitochondrial-like gene products. Longer mitochondrial (19–26 nucleotide) gene encoding primers are likely to increase the reproducibility and specificity when compared to RAPD technique. This method was found to be efficient, simple and highly reproducible. Here it has been effectively used to discriminate the various population groups of *H. puera* infesting teak plantations in south India. It can also be used to discriminate taxonomically various closely – related moths to the species level. Mitochondrial RAGEP fingerprints are derived from the randomness of RAGEP-PCR. It is difficult to predict with certainty that the bands are diagnostic features of the mitochondrial genome, but since RAGEP PCR uses gene specific primers, the PCR products could therefore be a result of amplification of homologous genes or pseudogenes which could represent nuclear mitochondrial –like sequences (NUMTs). Mitochondrial DNA sequences are frequently transferred to the nucleus-giving rise to NUMTs, which are considered to be common in eukaryotes (Richly and Leister, 2004). Very high rate of horizontal transfer

between organellar and nuclear genomes has been reported in the brown mountain grasshopper, *Podisma pedestris* (L.) (Bensasson *et al*., 2000). Age groups, sexes, life history variants, etc. and the processes including birth, death, immigration and emigration as different phenotypic classes have been very well defined (Roderick, 1996). While studying the differentiation process of grain aphid, *Sitobion avenae* (F.) populations across agricultural ecosystems using DNA fingerprinting [(GATA)4] and RAPDs, it was possible to discriminate the micro- and macro geographical heterogeneity (De Barro *et al*., 1994). Highly diagnostic banding patterns in individuals of *S. avenae* on wheat and cocksfoot grass, *Dactylis glomerata* (L.) were observed during the early months of infestation, which declined as the season progressed, largely as a result of genetic drift and local movement between adjacent host species (De Barro *et al*., 1995). Monophyly and a strong biogeographic pattern of each biotype have been reported in whitefly, *Bemisia tabaci* (Gennadius) populations studied throughout the world (De Barro *et al*., 2000).

While evaluating the genetic structure in introduced population of the fire ant, *Solenopsis invicta* (Buren) using different classes of markers, it was confirmed that both mitochondrial and nuclear markers display the same hierarchical structure (Ross *et al*., 1999). Distinct mitochondrial and nuclear DNA sequence divergence patterns for phylogenetic inference has been established among nymphalid butterflies (Brower and DeSalle, 1998). The present study using RAGEP-PCR provides a tool for a logical continuation of the earlier works (Nair *et al*., 1998) to trace the relationship of endemic, epicenter and epidemic populations of the teak defoliator. The dendrogram produced from nuclear RAGEP clearly indicates that the endemic insects are not involved in causing the epidemic; however, they are apparently involved in the localized spread by building up small epicenter populations. Similarly, while evaluating the observation based on mitochondrial RAGEP's,it is further apparent that endemic populations were not involved in causing the epidemic. This suggests that all the epidemic insects, which are spatially distinct, but temporally co-occurring, share the same gene pool. Randomness of genome amplification methods have been efficiently used in constructing the phylogenetic history in the weevil, *Aubeonymus mariafranciscae* (Roudier), which had diverged recently (Toberner *et al*., 1997), whilst the origin of the Argentine stem

weevil, *Listronotus bonariensis* (Kuschel) in New Zealand, was traced to the eastern coast of South America (William *et al*., 1994). Use of RAPDs to examine, for example, population subdivision of the saw toothed grain beetle, *Oryzaephilus surinamensis* (L.) (Brown *et al*., 1997), characterization and identification of Asian and North American gypsy moth, *Lymantria dispar* (L.) (Garner and Slavicek, 1996), host based genotype variation in *S. avenae* (Lushai et al., 2002), and genotypic variation among different phenotypes of asexual adult winged and wingless of some clones of cereal aphid species (Lushai *et al*., 1997), has been well documented. Earlier reports involving molecular DNA markers mention the use of these markers in the detection of sibling species of black flies, *Simulium* spp. (Brockhouse *et al*., 1993), whilst the dynamics of colonization of *Drosophila subobscura* (Collin) in the west coast of North America and its impact in the sibling species *D. athabasca* Sturtevant and Dobzhansky, and *D*. azteca Sturtevant and Dobzhansky has been extensively studied by allozymes,mitochondrial DNA (mtDNA) and RAPD markers (Pascual *et al*., 1997).

With the Teak defoliator, earlier studies based on temporal and spatial distribution of the larvae indicated that the epicenters were not constant over the years and did not represent highly favourable local environments (Nair and Sudheendrakumar, 1986). The present study found little evidence to show that the endemic populations cause the epicenter populations. On the other hand, it reconfirmed the relationship between epicenter populations and epidemic outbreaks. However some of the epicenter populations predicted to be related to some of the outbreaks based on ecological data were found to be not truly related. This suggests that the inflow of moths from far away places to cause major outbreaks could not be ruled out Thus it appears that under a single demographic structure, two phenotypic classes of *H. puera* coexist during the outbreak season. The degree of variability observed for RAGEPs also argues that this technique could be useful for a variety of questions, including individual identification, strain identification and phylogenetic analysis.

5. CONCLUSIONS

The present results appear to validate the hypothesis, that control of *H. puera* epicenter populations would help prevent large-scale outbreaks of the teak defoliator in teak plantations. Therefore, appropriate strategies should be adopted to control the epicenter populations, which occurs in a smaller area. This appears to be a more practical and economical approach for teak defoliator management when compared with management of the pest in the total plantation area covering thousands of hectares. Thus the molecular markers detected using RAGEP-PCR can enhance the understanding of insect population dynamics and aid in tracing the spread and cause of epidemics.

6. ACKNOWLEDGEMENTS

We are thankful to the Department of Biotechnology, Govt. of India for financial support extended for carrying out this project. Our sincere thanks are also due to the DBT Task Force Members for their valuable suggestions and encouragements during the course of this work Particularly we acknowledge the support and encouragement given by Dr. Seema Wahab, Director, DBT. Dr. J.K. Sharma, Director, KFRI and Dr. R.V.Thampan, Director, RGCB extended us all Institutional support during the course of the work.

The officers and field staff of the Kerala Forest Department. Nilambur were highly helpful to us. We thank Drs. George Mathew, M. Balasundaran and K. Swarupanandan for editorial scrutiny of the report and for useful suggestions to improve the text. Last but not the least, we place on record our sincere thanks to the staff of Library, administrative and LAN wings of KFRI for their cooperation and help to make this project a success.

7. REFERENCES

Andrew, F.C., Gary, A. F. (1995). Isolation and purification of Insect DNA. In: Methods in Molecular Biology, (ed) Clapp JP. Humana Press Inc. University of York, UK. pp.15-23.

Avise, J.C. (1994). Molecular markers, natural history and evolution. Chapman & Hall. New York.

Brockhouse, C.L., Vajime, C.G., Marin, R., Tanguay, R.M. (1993). Molecular identification of onchocerciasis vector sibling species in black flies (Diptera: Simuliidae). *Biochem. Biophys. Res. Commun*. 628-634.

Bensasson, D., Zhang, D.X., Hewitt, G.M. (2000). Frequent assimilation of mitochondrial DNA by grasshopper nuclear genomes. *Mol. Biol. Evol*. 17: 406-415.

Brower, A.V., DeSalle, R. (1998). Patterns of mitochondrial versus nuclear DNA sequence divergence among nymphalid butterflies: the utility of wingless as a source of characters for phylogenetic inference. *Insect Mol. Biol*. 7: 73-82.

Brown, R.J., Malcolm, C.A., Mason, P.L., Nichols, R.A. (1997). Genetic differentiation between and within strains of the saw-toothed grain beetle, Oryzaephilus surinamensis (Coleoptera: Silvanidae) at RAPD loci. *Insect Mol. Biol.* 6: 285-9.

Caetano-Anolles, G., Bassam, B.J., Gresshoff, P.M. (1991). DNA amplification fingerprinting using very short arbitrary oligonucleotide primers. *Biotechnology* 9: 553-557.

Daly, J.C., Gregg, P. (1985). Genetic variation in *Heliothis* in Australia: species identification and gene flow in two pest species *H.armigera* (Hubner) and *H. punctigera Wallengren* (Lepidoptera: Noctuidae). *Bull. Entomol. Res*. 75: 169-184.

De Barro, P.J., Sherratt, T.N., Wratten, S., Maclean, N. (1994). DNA fingerprinting of cereal aphids using (GATA)4. *Eur. J. Entomology* 91: 109-114.

De Barro, P.J., Sherratt, T.N., Brookes, C.P., David, O., Maclean, N. (1995). Spatial and temporal variation in British field populations of the grain aphid Sitobion avenae (F) (Hemiptera: Aphidae) studied using RAPD -PCR. *Proc. of R. Soc. Lond*. (B), 262:321-327.

De Barro, P.J., Driver, F., Trueman, J.W., Curran, J. (2000). Phylogenetic relationships of world populations of *Bemisia tabaci* (Gennadius) using ribosomal ITS1. *Mol. Phylogenet. Evol*. 16: 29-36.

Garner, K.J., Slavicek, J.M. (1996). Identification and characterization of a RAPDPCR marker for distinguishing Asian and North American gypsy moths*. Insect. Mol. Biol*. 5:81-91.

Liebher, J.K. (1988). Gene flow in ground beetles (Coleoptera:Carabidae) of differing habitat preference and flight-wing development. *Evolution*, 42: 129–137.

Loxdale HD, denHollander, J (1989) Electrophoretic studies on agricultural pests. Systematics. Association Special volume No 39, Clarendon Press, Oxford.

Loxdale, H.D., Lushai G, (1998). Molecular markers in entomology. *Bull. Entomol. Res*. 88: 577-600.

Loxdale, H.D., Lushai, G. (1999). Slaves of the environment: the movement of insects in relation to their ecology and genotype. Phil. *Trans. R. Soc. Ser.(B).* 354: 1479–1495.

Loxdale, H.D., Lushai, G. (2001). Use of genetic diversity in movement studies of flying insects. In: Insect Movement: Mechanisms and Consequences (eds) I.P.Woiwod, D.R.Reynolds & C.D. Thomas),CAB International, Wallingford,UK: 361–386.

Lushai, G., Markovitch, O., Loxdale, H.D. (2002). Host-based genotype variation in insects revisited. *Bull. Entomol. Res*. 92:159-64.

Lusha, G., Loxdale, H.D., Brookes, C.P., von Mende, N., Harrington, R., Hardie, J. (1997). Genotypic variation among different phenotypes within aphid clones. *Proc R. Soc. Lond. (B) Biol. Sci*. 264: 725-30.

Nair, K.S.S., Mohanadas, K. (1996). Early events in the outbreak of teak caterpillar *Hyblaea puera*. *Internat. J. Ecol and Environ. Sci.* 22: 271-279.

Nair, K.S.S., Sudheendrakumar, V.V. (1986). The teak defoliator, *Hyblaea puera* : Defoliation dynamics and evidence for short-range migration of moths. *Proc Indian Natl Sci Acad Part (B) Biol Sci*. 95: 7-21.

Nair, K.S.S., Sudheendrakumar, V. V., Varma, R.V., Chacko, K.C. (1985). Studies on the seasonal incidence of defoliators and the effect of defoliation on volume increment of teak. KFRI Research Report No. 30, Kerala Forest Research Institute, Peechi, Kerala : 78 pp.

Nair, K..S..S., Sudheendrakumar, V.V., Varma, R.V., Mohanadas, K. (1998). Tracing the epicentres of teak defoliator outbreaks in Kerala. KFRI Research Report No. 147, Kerala Forest Research Institute, Peechi, Kerala : 27 pp.

Osborne, J.L., Loxdale, H.D., Woiwood, I.P. (2001). Monitoring insect dispersal: methods and approaches. In: Dispersal ecology, (Eds. J.M. Bullock, R.E.Kenward & R.S.Hails) pp 24-49, British ecological Symposium, 3-5, 2001, Reading University, U.K. Blackwell publishing, Oxford, U.K. pp 458.

Pascual, M., Balanya, J., Latorre, A.. Serra, L. (1997). Diagnosis of sibling species of Drosophila involved in the colonization of North America by *D. subobscura*. *Mol Ecol*. 6:293-296.

Richly, E., Leister D. (2004). NUMTs in sequenced eukaryotic genomes. *Mol. Biol. Evol*., 21: 1081-1084.

Roderick G.K. (1996).Geographic structure of insect populations: Gene flow, phylogeography and their uses. *Ann. Rev. Entomol*. 41: 325-352.

Ross, K.G., Shoemaker, D.D., Krieger, M.J., DeHeer, C.J,, Keller, L. (1999). Assessing genetic structure with multiple classes of molecular markers: a case study involving the introduced fire ant *Solenopsis invicta*. *Mol. Biol. Evol*. 16: 525-543.

Stone, G.N. and Sunnucks, P. (1993). Genetic consequences of an invasion through a patchy environment—the cynipid gall wasp *Andricus quercuscalicis* (Hymenoptera: Cynipidae). *Mol. Ecol*. 2: 251–268.

Taberner, A., Dopazo, J., Castanera, P. (1997). Genetic characterization of populations of a de novo arisen sugar beet pest, *Aubeonymus mariaefranciscae* (Coleoptera, Curculionidae), by RAPD analysis. *J. Mol. Evol*. 45: 24-31.

Welsh, J., McClelland, M. (1990). Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Res*. 18: 7213-7218.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., .Hornes, M., Frijters, A, Pot, J, Peleman, J., Kuiper, M., *et al* (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res*. 23**:** 4407-4414.

Williams, C.L., Goldson, S.L., Baird, D.B., Bullock, D.W. (1994). Geographical origin of an introduced insect pest, *Listronotus bonariensis* (Kuschel), determined by RAPD analysis. *Heredity* 72: 412-9.

Williams, J,G,K,, Kubelik, A.R., Livak, K.J., Rafalsk, J.A. and Tingey, S.V. (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res.* 18: 6531–6535.

Deleted: ,K.S.S