International Journal of Agriculture, Environment and Biotechnology
Citation: IJAEB: 10(4): 469-479, August 2017
DOI: 10.5958/2230-732X.2017.00058.4

ENTOMOLOGY

Insect Molecular Markers and its Utility- A Review

Satyapriya Singh[1*], V.K. Mishra[2] and Tanmaya Kumar Bhoi[3]

[1,3]Division of Entomology, ICAR-Indian Agricultural Research Institute, New Delhi, India
[2]Department of Entomology and Agricultural Zoology, Institute of Agricultural Sciences**,** Banaras Hindu University, Varanasi, Uttar Pradesh, India

*Corresponding author: satyapriya_10797@iari.res.in

Paper No. 603 Received: 12-5-2017 Accepted: 3-7-2017

ABSTRACT

Insects represent a major life form on earth. Nearly one million insect species have been discovered which comprises 75% of all the recorded animal species. This biological success is accredited to the enormous diversity of their size and body structure, mating strategies, and the incredible feeding and adaptation behaviour. Significant progress has been made in understanding insect diversity and ecology by using classical genetic principles. Over the last past 15 years, DNA markers have made a significant contribution in molecular studies on genetic relatedness, phylogeny, population dynamics and genome mapping in insects. A brief account of DNA based marker systems, their utility in entomological research, with examples wherever available is presented to prompt further reading and applications. The techniques described include restriction fragment length polymorphisms (RFLPs), randomly amplified polymorphic DNAs (RAPDs), amplified fragment length polymorphisms (AFLPs), microsatellites/simple sequence repeats (SSRs), and expressed sequence tag (EST) based marker system, single nucleotide polymorphisms (SNPs) and other derived marker systems along with their genetic nature and relative comparison. Although a large number of samples can be analyzed quickly, a number of other factors such as cost, speed and requirement of technical skills are the major concern. In this review, we hope to inform the general reader about the importance and scope of the main molecular markers commonly in use, along with brief details of some other techniques which show great promise for entomological studies. Thereafter, we discuss their applications including suitability for particular studies, the reliability of particular techniques, the issues of safety involved, cost effectiveness and the statistical analyses utilized.

Highlights

- Genetic markers ditermine genetic differences between individual living organisms or species.
- Molecular marker were used to identify different alleles of a given gene.

Keywords: Molecular marker, insect individual, species, gene, phenotype, determination

DNA markers defined places along each chromosome and be used to delineate when one has reached or passed by a particular gene of interest. The utility of molecular markers as additional tools in basic and applied entomology has led to 'molecular entomology', also referred to as 'precision entomology' (Ananthakrishnan, 2005). The limitations in using morphological, physiological and cytological markers for assessing genetic diversity and population dynamics have been largely sidestepped by the developments in DNA-based markers (Fakrudin *et al*. 2006). Molecular marker is a piece of DNA with easily identified phenotype such that cells or individuals with different alleles. It can be a protein, isozyme, DNA sequence, such as a sequence surrounds a single base-pair or a long one, like minisatellites, whose inheritance can be determine. Genetic markers determine genetic differences between individual living organisms or species. They cannot represent the target genes themselves. Genetic markers that are located on genes, these referred to

as gene 'tags. All genetic markers occupy specific genomic positions within chromosomes (like genes) called 'loci' (singular 'locus'). Insects comprise the largest species nearly about 900,000 and composition possess a vast undiscovered 75% of all the recorded animal species, genetic diversity and gene pool that can be better explored using molecular marker techniques. Molecular marker were used to identify different alleles of a given gene such as insecticide resistance, pathogen identification, chromosome mapping, detection of prey in insect predators and level of polymorphism could be obtained by using (Molecular markers) thereafter. Because mutations in introns or even in the codons of a gene can potentially provide variation at the DNA level. Molecular marker are potential tools contributed towards genetic relation, phylogeny, population dynamics or gene and genome mapping in insects.

Important characteristics of molecular marker:

1. **Polymorphic Nature :** A polymorphism is a detectable and heritable variation at a locus Polymorphisms arise as a result of several classes of mutations like substitution of as little as a single nucleotide (point mutation), genome rearrangement due to insertion ordeletion or errors in replication of arrays of tandemly-repeated DNA.
2. **High Reproducin ability:** Should give similar results in different experiments irrespective of the time and the place. Preferably display co-dominant inheritance (both forms are detectable inheterozygotes).
3. **Easily detecble**: The detection of marker must be fast and inexpensive. It should detect difference in expression between trait types and/or alleles of interest, early in the development of the organism.
4. **Easily exchange of data between laboratory condition:** It is extremely difficult to find a molecular marker which would meet all the above criteria. Depending on the type of study to be undertaken, a marker system can be identified that would fulfill atleast a few of the above characteristics.

Types of Markers

Protein molecular markers

Isozyme markers: Multiple enzyme coded by the different genes- Isozyme: one enzyme, more than one locus (gene duplication; gene families)- Allozyme: one enzyme; one locus; two or more alleles in a population. Isozymes are proteins with same enzymatic function but different structural, chemical, or immunological characteristics (coded by the different genes). (Behura *et al.* 2006) To be usefulas marker, isoforms must be electrophoretically resolvable, and detectable. Isozyme markers are used as they are simple, inexpensive, and are used for detection of the gene introgression and recombination, for comparative mapping, and for determination of the genetic diversity and phylogenetic relationships. Because isozyme analysis does not require DNA extraction or the availability of sequence information, primers or probes, they are quick and easy to use, codominant markers that have high reproducibility. Restricted due to limited number of enzyme systems available. The main weakness of isozymes is their relatively low abundance and low level of polymorphism, affected by environmental conditions. For example, the banding profile obtained for a particular isozyme marker may change depending on the type of tissue used for the analysis. On the contrary, molecular markers, because they are based on differences in the DNA sequence, are not environmentally influenced, which means that the same banding profiles can be expected at all times for the same genotype.

Molecular markers are broadly classified into two categories:

1. Hybridization-based markers: DNA profiles are visualized by hybridizing the restriction enzyme-digested DNA, to a labelled probe, which is a DNA fragment of known origin or sequence.
2. Polymerase chain reaction (PCR)-based markers: PCR-based markers involve *in vitro* amplification of particular DNA sequences with the help of specifically chosen oligonucleotide sequence, thermostable DNA polymerase enzyme. The amplified fragments

are separated electrophoretically and banding patterns are detected by different methods such as staining and autoradiography.

Hybridization–based markers

Restriction Fragment Length Polymorphism: RFLP is a technique in which organisms may be differentiated by analysis of patterns derived from cleavage of their DNA. If two organisms differ in the distance between sites of cleavage of particular *Restriction Endonucleases*, the length of the fragments produced will differ when the DNA is digested with a restriction enzyme. They have their origin in the DNA rearrangements that occur due to evolutionary processes, point mutations within the restriction enzyme recognition site sequences, insertions or deletions within the fragments. Size fractionation is achieved by gel electrophoresis and, after transfer to a membrane by Southern blotting; fragments of interest are identified by hybridization with radioactive labeled probe. These probes are species-specific single locus probes of about 0.5–3.0 kb in size, obtained from a cDNA library or agenomic library. Normally genomic library probes exhibit greater variability than gene probes from cDNA libraries. The first step in the analysis is to derive a set of clones that can be used to identify RFLPs. Genomic clones are represent sequences at random are a poor choice as hybridization probes because plant genomes consist of a large percentage of repeated sequences. So, many of the clones will contain repeated sequences, and hybridizations with those clones containing repeated sequences generate many hybridization bands that are difficult to analyze genetically. RFLPs, being codominant markers and very reliable markers, are extensively used in linkage analysis and breeding and can easily determine if a linked trait is present in a homozygous or heterozygous state in an individual. The assay is time-consuming and labour-intensive and only one out of several markers may be polymorphic, which is highly inconvenient especially for crosses between closely-related species. RFLPs have been widely used in gene mapping. *Genetic linkage map study on* Colorado potato beetle *Leptinotarsa decemlineata* (Haw throne 2001). Phylogenetic study on Ticks and Mites (Cruickshank 2002). Population genetics studies on determination of male and female sexes in honey bee (Hall 1990).

Variable Number of Tandem Repeats (VNTR): Repetitive DNA is a major step forward in genetic identification is the discovery that about 30–90% of the genome of virtually all the species is constituted by regions of repetitive DNA, which are highly polymorphic in nature. These regions contain genetic loci comprising several hundred alleles, differing from each other with respect to length, sequence or both and they are interspersed in tandem arrays ubiquitously. The repetitive DNA regions play an important role in absorbing mutations in the genome. Of the mutations that occur in the genome, only inherited mutations play a vital role in evolution or polymorphism. Thus repetitive DNA and mutational forces functional in nature together form the basis of a number of marker systems that are useful for various applications in plant genome analysis. The markers belonging to this class are both hybridization-based and PCR based.

Minisatellites are tandem repeats with a monomer repeat length of about 11–60bp, while microsatellites or short tandem repeats/simple sequence repeats (STRs/SSRs) consist of 1 to 6 bp long monomer sequence that is repeated several times. Microsatellites and minisatellites thus form an ideal marker system creating complex banding patterns by simultaneously detecting multiple DNA loci.

VNTR: These loci contain tandem minisatellite repeats that vary in the number of repeat units between genotypes and are referred to as variable number of tandem repeats (VNTRs) i.e. a single locus that contains variable number of tandem repeats between genotypes. Minisatellites consist of chromosomal regions containing tandem repeat units of a 10–50 base motif, flanked by conserved DNA restriction sites. A minisatellite profile consisting of many bands, usually within a 4–20 kb size range, is generated by using common multilocus probes that are able to hybridize to minisatellite sequences in different species. Locus specific probes can be developed by molecular cloning of DNA restriction fragments, subsequent screening with a multilocus minisatellite probe and isolation of specific fragments. Variation in the number of repeat units, due to unequal crossing over or gene conversion, is considered to be the main cause of length polymorphisms. Due to the high mutation rate of minisatellites, the level of polymorphism is

substantial, generally resulting in unique multilocus profiles for different individuals within a population. The main advantages of minisatellites are their high level of polymorphism and high reproducibility. Disadvantages of minisatellites are similar to RFLPs due to the high similarity in methodological procedures.

PCR based-markers

Randomly-amplified polymorphic DNA (RAPD)

This procedure detects nucleotide sequence polymorphisms in DNA by using a single primer of arbitrary nucleotide sequence. In this reaction, a single species of primer anneals to the genomic DNA at two different sites on complementary strands of DNA template. If these priming sites are within an amplifiable range of each other, a discrete DNA product is formed through thermocyclic amplification. RAPDs are DNA fragments amplified by the PCR using short synthetic primers (generally 10 bp) of random sequence. These oligonucleotides serve as both forward and reverse primer, and are usually able to amplify fragments from 1–10 genomic sites simultaneously. Amplified products (usually within the 0.5–5 kb size range) are separated on agarose gels in the presence of ethidium bromide and view under ultraviolet light and presence and absence of band will be observed. These polymorphisms are considered to be primarily due to variation in the primer annealing sites, but they can also be generated by length differences in the amplified sequence between primer annealing sites.

On an average, each primer directs amplification of several discrete loci in the genome, making the assay useful for efficient screening of nucleotide sequence polymorphism between individuals. However, it is important to optimize and maintain consistent reaction conditions for reproducible DNA amplification. They are dominant markers and hence have limitations in their use as markers for mapping, which can be overcome to some extent by selecting those markers that are linked in coupling. Multiple Arbitrary Amplicon Profiling (MAAP) is the collective term for techniques using single arbitrary primers. Taxonomy and population genetics of aphids, moths and parasitoid detection (Black *et al*. 1992, Puterka *et al*. 1993, Stevens and Wall 1995). Social behavior in honey bees (Hunt and Page 1992). Genetic linkage and map construction in *Tribolium castaneum* (Herbst) (Beeman and Brown 1999).

Amplified fragment length polymorphism (AFLP)

Amplified fragment length polymorphism (AFLP) fingerprinting technique is an intermediate between RFLPs and PCR. This helps to identify DNA restriction fragments by means of PCR amplification. The process involves restriction of genomic DNA, followed by ligation of adaptors complementary to the restriction sites and selective PCR amplification of a subset of the adapted restriction fragments. These fragments are interpretad on denaturing polyacrylamide gels either through autoradiographic or fluorescence methodologies. Polymorphisms are recognised from differences in the length of the amplified fragments by polyacrylamide gelelectrophoresis (PAGE) or bycapillary electrophoresis. The technique consisted four steps: (1) Restriction of DNA using a rare cutting and a commonly cutting restriction enzyme simultaneously (such as *Mse*I and *Eco*RI*)* followed by ligation of oligonucleotide adapters, of defined sequences including the respective restriction enzyme sites. (2) Amplification of preselective sets of restriction fragments, using specifically designed primers. To access this, the 5′ region of the primer is made such that it would contain both the restriction enzyme sites on either sides of the fragment complementary to the respective adapters, while the 3′ ends extend for a few arbitrarily chosen nucleotides into the restriction fragments. (3) Amplification of selective sets with primers complementary to both the adaptors but having a few arbitrarily choosen nucleotides at the 3′ end. (4) Gel analysis of amplified fragments.

The potentials of AFLPs presnt in their high genomic abundance, considerable reproducibility, the generation of many informative bands per reaction, their enoromous applications, and the fact that no sequence data for primer construction are required. AFLPs can be applied in studies related to genetic identity, parentage, clone and cultivars identification, and evoulationary and phylogenetic studies of closely related species because of the highly informative fingerprinting profiles generally obtained. Their random distribution and high genomic abundance throughout the genome make

AFLPs a widely valued technology for gene mapping studies and related aspects. Genetic linkage map construction is well studied in mosquito (Severson *et al*. 1993), *Rhagoletis pomonella* (Roethele *et al*. 1997), *Heliothis virescens* (Fabricius), *Plutella xylostella* and *Helicoverpa armigera* (Heckel *et al*. 1998).

Simple Sequence Repeats (SSR)

Simple Sequence Repeats (SSRs) otherwise known as microsatellites are sections of DNA, made up of tandemly repeating mono-, di-, tri-, tetra- or penta-nucleotide units that are arranged throughout the genomes of most eukaryotic species. By PCR, amplification of the microsatellite is done through designing the specific primers (normally 20–25bp), if nucleotide sequences in the flanking regions of the microsatellites are known. The identification of microsatellites and their flanking sequences can be done by constructing a small-insert genomic library and screening the library with a synthetically labeled oligonucleotide repeat and sequencing the positive clones. Alternatively, microsatellite may be identified by screening sequence databases for microsatellite sequence motifs from which adjacent primers may then be designed. Microsatellite markers prepared from genomic libraries can belong to either the transcribed region or the non-transcribed region of the genome, and rarely is there information available regarding their functions. Moreover, primers may be used that have already been designed for genetically close related species. The main cause of variation in the number of repeat units of a microsatellite is due to polymerases lippage during DNA replication, or slipped strand mispairing, resulting in length polymorphisms and that can identified by gel electrophoresis.

The potentials of microsatellites include the allele co-dominance, their high genomic abundance in eukaryotes and their random distribution throughout the genome. Because the technique is PCR-based, only low quantities of template DNA (10–100ng per reaction) are required. Very low quantities of template DNA (10–100ng per reaction) are required because it is a PCR-based technique. Due to the use of long PCR primers, the reproducibility of microsatellites is high. The high development costs is one of the major limitations if adequate primer sequences for the species of interest are unavailable, making them difficult to apply to unstudied groups. It is well documented in paternity studies of Hymenoptera (Estoup *et al*. 1995) and genetic sexing of Lepidopteron insects (Ananthakrishnan 2005). Lehmann *et al*. (1997) utilised these markers in mosquitoes for genetic studies at population level. In Aphids, hymenopteran insects, mosquitoes, moths and butterflies these markers have provided useful information on genetics of populations (Black *et al. 2001*)

ISSRs are DNA fragments of about 100–3000bp located between adjacent, oppositely oriented microsatellite regions. In this technique, primers based on microsatellites are utilized to amplify inter-SSRDNA sequences. PCR is used to amplify ISSRs by taking microsatellite core sequences as primers with a few selective nucleotides as anchors into the non-repeat adjacent regions (16–18bp). Approximately 10–60 fragments from multiple loci are generated simultaneously, separated by gel electrophoresis and scored as the presence or absence of fragments of particular size. Now days, this markers are well studied in silkworm genomic fingerprinting (Nagaraja and Nagaraju, 1995), genome mapping & evolutionary studies (Nagaraju *et al.* 2001) and genetic diversity & phylogenetic studies.

Sequence Tagged Sites

It combines the advantages of both RAPDs (*i.e*. markers PCR based, no probe maintenance or distribution) and RFLPs (*i.e.* co-dominance mode of inheritance) markers, and that could efficintly be used. A STS is a unique, simple-copy segment of the genome whose DNA sequence is known and amplification is done by PCR. Three types of STS are mentioned below:

Sequence Characterized Amplified Region (SCAR)

In this technique, the RAPD marker termini are sequenced and longer primers are designed (22–24 nucleotide bases long) for specific amplification of a particular locus. These markers are DNA fragments amplified by the PCR using specific 15–30bp primers, designed from nucleotide sequences established from cloned RAPD fragments linked to a trait of interest. SCARs do not face the problem of low reproducibility as it uses longer PCR primers, which

is generally encountered with RAPDs. Obtaining a co-dominant marker is an additional advantage of converting RAPDs into SCARs, although SCARs may exhibit dominance when one or both primers partially overlap the site of sequence variation. Gel electrophoresis helps in detection of length polymorphisms. The potentials of SCARs is that they are quick and easy to use. In addition, SCARs have a high reproducibility, locus-specific and have been applied in gene mapping studies and marker assisted selection. Very low quantities of template DNA are required (10–100ng per reaction), as it uses PCR. Its limitation includes the need for sequence data to design the PCR primers.

Cleaved amplified polymorphic sequences (CAPS)

CAPS are DNA fragments amplified by PCR using specific 20–25bp primers, followed by digestion of the PCR products with a restriction enzyme. Based on the sequence information available in databank of genomic or cDNA sequences or cloned RAPD bands, PCR primer for this process can be synthesized. Subsequently, gel electrophoresis helps to identify length polymorphisms resulting from variation in the occurrence of restriction sites of the digested products. CAPS have also been referred to as PCR-Restriction Fragment Length Polymorphism (PCR-RFLP). The main advantages of CAPS are the involvement of PCR requiring only low quantities of template DNA (50–100ng per reaction), the co-dominance of alleles and high reproducibility. CAPS analysis is easy as compared to RFLPs as it is laborious and technically demanding steps of Southern blot hybridization and radioactive detection procedures. These markers are co-dominant in nature. Due to limited size of the amplified fragments, it is more difficult to find CAPS polymorphisms. The synthesis of the primers need Sequence data. It has been applied predominantly in gene mapping studies.

- ***Sequence-tagged microsatellite site markers (STMS):*** This technique includes DNA polymorphism using specific primers designed from the sequence data of a specific locus.
- ***Expressed sequence tag markers (EST):*** These markers are accessed by partial sequencing of random cDNA clones. They are very useful in cloning specific genes of interest and synteny mapping of functional genes in various related organisms. ESTs are popularly used in full genome sequencing and mapping programmes underway for a number of organisms and for identification of diagnostic markers and helping in identifying active genes. Moreover, an EST that appears to be unique in nature helps to isolate new genes.

Single Nucleotide Polymorphism (SNP)

An unique and novel class of DNA markers namely single nucleotide polymorphism in genome (SNPs) has recently become most popular in genomic studies. In many organisms most polymorphisms result from changes in a single nucleotide position (point mutations), has led to the development of techniques to study single nucleotide polymorphisms (SNPs). Single nucleotide polymorphisms, or SNPs are pronounced as "snips". When a single nucleotide (A,T,C, or G) in the genome sequence is altered then variations in DNA sequences occur and that leads to SNPs.

For example a SNP might change the DNA sequence A**A**GGCTAA to A**T**GGCTAA. For a variation to be entertained a SNP, it must occur in at least 1% of the population. Various types of SNPs are distinguished, according to their assignment to the structural element of genomic DNA or their functional effect. Oligonucleotide substitutions, involving regulatory regions that control gene expression, are also referred to as regulatory SNPs and designated rSNPs. They may include promoter SNPs and some intron SNPs. They are classified into anonymous SNPs (functional effect is unknown), candidate SNPs (presumably having a functional effect), and protein SNPs (single-nucleotide substitutions), depending on the presence of SNP in genic region and resulting in a change in the protein function or expression.

SNP detection is based on:

- overlapping genomic DNA sequences.
- overlapping EST sequences.
- unique (non-overlapping) genomic and EST sequences.
- "shotgun" sequencing.

Analytical procedures need sequence information for the design of alleles specific PCR primersor oligonucleotide probes. SNPs and flanking sequences can be determined by library construction and sequencing or through the screening of readily available sequence databases. Once the location of SNPs is identified and appropriate primers designed, it provides high throughput automation. To achieve high sample throughput, multiplex PCR and hybridization to oligonucleotide microarrays or analysis on automated sequencers are often used to interrogate the presence of SNPs. In the cultivated tomato, discrimination of cultivar is done by using SNP analysis as is difficult to find polymorphisms.

SNPs may also be helped to saturate linkage maps in order to locate relevant traits in the genome. For example, in *Arabidopsis thaliana* a high density linkage map for easy to score DNA markers was lacking until SNPs became available.

Two components have involved in SNP genotyping techniques—a method for determining the type of base present at a given SNP locus (allele discrimination), and a method for reporting the presence of the allele(s) (signal detection). There are normally three general allele discrimination methods: hybridization/annealing (with or without a subsequent enzymatic discrimination step), primer extension, and enzyme cleavage. In each case, the technology platform may be homogeneous (in solution) or heterogeneous(involving both a liquid and a solid phase, such as a microarray). Some of the analysis require prior amplification of the genomic target, whereas others are sensitive enough to work directly on genomic DNA or cDNA. There are many signal detection platforms, and most of these follow the fate of a label either in real time or at the assay end point. SNPs have been used in mites for discriminating species and subspecies (Navajas *et al*. 1998) phylogeography (Brumfield *et al.* 2003), biodiversity assessment linkage disequilibrium analysis (Akey *et al*. 2003). Population genetic parameters (Kuhner *et al*. 2000) and in ecology, evolution and conservation biology (Morin *et al*. 2004).

Major applications of molecular markers:

- ***Mating, parentage and kinship:*** DNA markers can unravel information to determine parentage and kinship relations in insects. In this regard, one of the innovative works done was to use RAPD markers to determine paternity in two odonate species of Anisopteran dragonflies, *Anax parthenope* (Julius Brauer) and *Orthetrum coerulescens* (Keeled Skimmer) (Hadrys *et al*. 1993). RAPD banding patterns were used to access paternity of 'synthetic offsprings' generated by quantitative mixing of genomic DNA from putative parents. Where no information on mating histories of both males and females are known, this approach has been helpful to establish the paternity of guarding males in species.
- ***Insect plant interaction:*** One of the most appealing applications of molecular markers in entomological studies is probably on those for insect–plant interaction. DNA markers provide utility in tagging and mapping genes in important crop plants that provide resistance to damaging insect pests, and are also useful in characterizing a virulence genes in insects

Table 1: Difference between various molecular markers

Characteristics	RFLPs	RAPDs	AFLPs	SSRs	SNPs
DNA required (?g)	10	0.02	0.5-1.0	0.05	0.05
DNA quality	High	High	Moderate	Moderate	High
PCR-based	No	Yes	Yes	Yes	Yes
Number of polymorph loci analyzed	1.0-3.0	1.5-50	20-100	1.0-3.0	1.0
Ease of use	Not easy	Easy	Easy	Easy	Easy
Amenable to automation	Low	Moderate	Moderate	High	High
Reproducibility	High	Unreliable	High	High	High
Development cost	Low	Low	Moderate	High	High
Cost per analysis	High	Low	Moderate	Low	Low

interacting with the host plants (Harris *et al*. 2003). The molecular genetic information generated by marker data is used to characterize phenotypic ability of insect to attack specific plant varieties. Specific examples of marker usage in these aspects can be evident from molecular studies on gall midges (*Cecidomyiidae* family of Diptera), major insect pests of rice and wheat.

- ***Insect pathogen interaction:*** Molecular markers are useful in understanding genetic interaction between disease-causing pathogens and the insect vectors that spread them (Crampton *et al*. 1997). *Triatoma infestans* (Klug) (*Reduviidae*) is a major insect vector of Chagas' disease in many South American countries. It transmits the *Trypanosoma cruzi* , the agent of Chagas' disease. Mixed and pure clones of *Triatoma cruzi* in the gut of *Triatoma infestans* have been studied by using RAPD profiles to provide information on the vectoral ability of the insects. Similarly, molecular markers were applied to determine the vectorial ability of mosquitoes (Severson *et al*. 2001) by means of mapping quantitative trait loci (QTL) that determined if a species could transmit the malaria parasite (Severson *et al*. 1995; Bosio *et al.* 2000). Expressed sequence tag (EST) markers have also been used to identify the potential target genes that are involved in immunity to specific pathogens causing malaria (Bartholomay *et al*. 2004). *Metarhizium anisopliae*is a causative agent of green muscardine disease, and is transmitted by a broad range of insects. To explore the molecular basis of this process, EST markers were used to study the pattern of gene expression in responses to diverse insect cuticles and were found that an array of secreted protein had potential activity in the hosts in these responsive actions.
- ***Insecticide resistance:*** Insecticide resistance is another important focus in entomological research and bears medical and agricultural importance. Molecular markers are used for identification and mapping of resistance genes in insects against insecticides. In malaria control programs, difficulties arise because of emerging resistance in the mosquito vectors to DDT. DDT resistance in the major malaria vector in Africa, *Anopheles gambiae*, is associated with an increased metabolism of the insecticide. Use of microsatellite markers in mapping experiments has identified QTL in *Anopheles gambiae* that determines the DDT resistance phenotypes (Ranson *et al*. 2000). Using RFLP markers, it was discovered that DDT resistance in houseflies, *Musca domestica* was associated with the 'knockdown' *kdr* trait (Knipple *et al*. 1994).
- ***Behavioural studies:*** Behavioural plasticity in social insects represents a complex biological phenomenon that is getting attention of molecular biologists. Honeybee (*Apis mellifera*) is an emerging model organism that is being studied for social behaviour at molecular level (Robinson *et al*. 2005). ESTs have been used as expression markers in microarray formats to predict the nursing and foraging behaviour in individual bees (Whitfield *et al*. 2003). In honeybee, these social behaviours are polygenic traits and are influenced by more than one gene referred to as QTL. The two major QTLs that determine the foraging behaviour in honeybee have been identified by employing RAPD markers in backcross populations between bees collecting nectar and those collecting pollen (Hunt and Page, 1995).

In insect ecology study

Ecological research on insect species unveil invaluable information on population structure, speciation, gene flow and genetic diversity, and provide explanation on insect diversity based on their interaction with environmental factors, either biotic (including other biological species) or abiotic. In entomology, DNA markers are used to provide raw information based on which an ecologist makes estimates of genetic diversity and gene flow between species, identifies haplotypes and lineages or predicts migration and colonization history (Hale and Singh, 1991; Yoon and Aquadro, 1994; Behura *et al*. 2001b; Salvato *et al*. 2002; Llewellyn *et al.* 2003; Bosio *et al*. 2005). Molecular data also provide the information to differentiate sympatric species from allopotaric species and parapatric species (Ballinger-Crabtree *et al*. 1992; Wilkerson *et al*. 1993; Favia *et al*. 1994; Apostol *et al*. 1996; Banuls *et al*. 1999; Ayres *et al*. 2003; Margonari *et al*. 2004). Gene flow and genetic variations within and between insect species, measured from marker data, are very much essential

to establish meaningful explanation for population structure and dynamics (Cervera *et al*. 2000; Wong *et al.* 2000; Takami *et al.* 2004; Mendelson and Shaw, 2005). Similarly, natural selection is a major factor that is regarded as the potential for population diversity. In this regard, molecular markers are utilized to infer phylogeny and biogeography of insect populations and to understand modes of evolution and evolutionary trajectories (Luque *et al.* 2002; Chatterjee and Mohandas, 2003; Chatterjee and Tanushree, 2004; Prasad *et al*. 2005). Based on linkage to certain traits or genes, diagnostic molecular markers are used for diagnostic purposes of individual insects (Hunt and Page, 1994; Behura *et al*. 1999; Kengne *et al*. 2001; Manguin *et al.* 2002; Kethidi *et al.* 2003; Ullmann *et al*. 2003).

Future Perspectives of insect molecular markers:

- *Drosophila* genomic sequence is a major milestone for genomics and it proved as new strategy for sequencing large eukaryotic genomes and as a model system to understand biological functions.
- At present, only three agricultural pest insect genomes, that of the red flour beetle, pea aphid and DBM have been fully sequenced, but within the next few years several pest sequences will be available with the help of molecular markers.
- Molecular markers information increases our knowledge for understanding the biology of insects and insecticide resistance.
- In future, Molecular markers will be an increasingly useful approach for pinpointing common and different genes across species.
- Molecular markers will be useful to explore gene functions.
- Molecular markers are being applied more and more in every aspect of life sciences research, including ecology and evolution.
- In future, there is a growing tendency for insect molecular scientists.

CONCLUSION

Because of the simple nature of DNA itself, we tend to assume that the technologies to detect DNA polymorphisms are not infinite. If this statement is correct, then we must be getting near the asymptote of the sigmoidal graph of DNA techniques vs. time (let's say since 1953), and that it is unlikely (although not impossible) that many more techniques will be developed in the next 20 years or so. However, at this point, the field of molecular marker technology is fast progressing by adopting new forms and innovative approaches of the existing genetic principles in detecting DNA polymorphism. One example of this scenario is the use of novel approaches for generating new class of genetic markers. Insect chromosomes are home to numerous DNA and RNA transposons, some of which are highly abundant in the genome. Hence, comparative sequence studies of the repeat elements in diverse insect species can provide useful information on how to make use of them for developing abundant markers that can be used in those species.

ACKNOWLEDGMENTS

Authors are extremely thankful to Dr. Amitava Rakshit Department of Soil Science and Agricultural Chemistry, IAS, BHU, Varanasi and Professor N.N. Singh, Department of Entomology and Agricultural Zoology, IAS, BHU, Varanasi, because give new thing and Idea.

REFERENCES

Akey, J.M., Zhang, K., Xiong, M. and Jin, L. 2003. The effect of single nucleotide polymorphism identification strategies on estimates of linkage disequilibrium. *Molecular Biololgyand Evolution*, **20**: 232–242.

Ananthakrishnan, T.N. 2005. Dimensions of molecular entomology. *Current Science*, **88**: 541–543.

Apostol, B.L., Black, I.V., William, C., Reiter, P. and Miller, B.R. 1996. Population genetics with RAPD–PCR markers: the breeding structure of *Aedes aegypti* in Puerto Rico. *Heredity*, **76**(4).

Ayres, C.F.J., Melo-santos, M.A.V., Solé-Cava, A.M. and Furtado, A.F. 2003. Genetic differentiation of *Aedes aegypti* (Diptera: Culicidae), the major dengue vector in Brazil. *Journal of Medical Entomology*, **40**(4): 430-435.

Ballinger-Crabtree, M.E., Black IV, W.C. and Miller, B.R. 1992. Use of genetic polymorphisms detected by the random-amplified polymorphic DNA polymerase chain reaction (RAPD-PCR) for differentiation and identification of *Aedes aegypti* subspecies and populations. *The American Journal of Tropical Medicine and Hygiene*, **47**(6): 893-901.

Banuls, A.L., Jonquieres, R., Guerrini, F., Le Pont, F., Barrera, C., Espinel, I. and Tibayrenc, M. 1999. Genetic analysis of Leishmania parasites in Ecuador: are *Leishmania* (Viannia) *panamensis* and Leishmania (V.) guyanensis

distinct taxa?. *The American Journal of Tropical Medicine and Hygiene*, **61**(5): 838-845.

Bartholomay, L.C., Cho W.L. and Rocheleau T.A. 2004. Description of the transcriptomes of immune response-activated hemocytes from the mosquito vectors *Aedes aegypti* and *Armigeres subalbatus*. *Infection Immunity*, **72**: 4114–4126.

Beeman, R.W. and Brown, S.J. 1999. RAPD based genetic linkage maps of *Tribolium castaneum*. *Genetics*, **153**: 333–338.

Behura, S.K. 2006. Molecular marker systems in insects: current trends and future avenues. *Molecular Ecology*, **15**(11): 3087-3113.

Behura, S.K., Sahu, S.C., Mohan, M. and Nair, S. 2001. Wolbachia in the Asian rice gall midge, *Orseolia oryzae* (Wood-Mason): correlation between host mitotypes and infection status. *Insect Molecular Biology*, **10**(2): 163-171.

Behura, S.K., Sahu, S.C., Rajamani, S., Devi, A., Mago, R., Nair, S. and Mohan, M. 1999. Differentiation of Asian rice gall midge, *Orseolia oryzae* (Wood-Mason), biotypes by sequence characterized amplified regions (SCARs). *Insect Molecular Biology*, **8**(3): 391-397.

Black, W.C., DuTeau, N.M., Puterka, G.J., Nechols, J.R.and Pettorini, J.M. 1992. Use of the random amplified polymorphic DNA polymerase chain reaction (RAPD-PCR) to detect DNA polymorphisms in aphids (Homoptera: Aphididae). *Bulletin of Entomology*, **82**: 151–159.

Bosio, C.F., Fulton, R.E. and Salasek, M.L. 2000. Quantitative trait loci that control vector competence for dengue-2 virus in the mosquito, *Aedes aegypti*. *Genetics*, **156**: 687–698.

Bosio, C.F., Harrington, L.C., Jones, J.W., Sithiprasasna, R., Norris, D.E. and Scott, T.W. 2005. Genetic structure of *Aedes aegypti* populations in Thailand using mitochondrial DNA. *The American Journal of Tropical Medicine and Hygiene*, **72**(4): 434-442.

Brumfield, R.T., Beerli, P., Nickerson, D.A. and Edwards, S.V. 2003. Single nucleotide polymorphisms (SNPs) as markers in phylogeography. *Trends in Ecology and Evolution*, **18**: 249–256.

Cervera, M.T., Cabezas, J.A., Simon, B., Martinez-Zapater, J.M., Beitia, F. and Cenis, J.L. 2000. Genetic relationships among biotypes of *Bemisia tabaci* (Hemiptera: Aleyrodidae) based on AFLP analysis. *Bulletin of Entomological Research*, **90**(5): 391-396.

Chatterjee, S.N. and Mohandas, T.P. 2003. Identification of ISSR markers associated with productivity traits in silkworm, *Bombyx mori* L. *Genome*, **46**(3): 438-447.

Chatterjee, S.N. and Tanushree, T. 2004. Molecular profiling of silkworm biodiversity in India. *Genetika*, **40**(12): 1618-1627.

Crampton, J.M., Beard C.B. and Louis C. 1997. *The Molecular Biology of Insect Disease Vectors: A Methods Manual*. Kluwer Academic Publishers, Boston, Massachusetts.

Cruickshank, R.H. 2002. Molecular markers for the phylogenetics of mites and ticks. System. *Appl. Acar*., **7**: 3–14.

Estoup, A., Solignac, M. and Cornuet, J.M. 1995. Precise assessment of the number of patrilines and of genetic relatedness in honeybee colonies. *Proceedings of the Royal Society b Impact Factor*, **258**: 1–7.

Fakrudin, B., Kotresh, H., Punnuri, S.M., Rajkumar, B.K., Thudi, M., Paramesh, H. and Kuruvinashetti, M.S. 2006. Identification of two RAPD markers genetically linked to a recessive allele of a Fusarium wilt resistance gene in pigeonpea (*Cajanus cajan* L. Millsp.). *Euphytica*, **149**(1-2): 113-120.

Favia, G., Dimopouios, G. and Louis, C. 1994. Analysis of the *Anopheles gambiae* genome using RAPD markers. *Insect Molecular Biology*, **3**(3): 149-157.

Hadrys, H., Schierwater B. and Dellaporta S.L. 1993. Determination of paternity in dragonflies by random amplified polymorphic DNA fingerprinting. *Molecular Ecology*, **2**: 79–87.

Hale, L.R. and Singh, R.S. 1991. A comprehensive study of genic variation in natural populations of Drosophila melanogaster. IV. Mitochondrial DNA variation and the role of history vs. selection in the genetic structure of geographic populations. *Genetics*, **129**(1): 103-117.

Hall, H.G. 1990: Parental analysis of introgressive hybridization between African and European honeybees using nuclear DNA RFLPs. *Genetics*, **125**: 611–621.

Harris, M.O., Stuart J.J. and Mohan M. 2003. Grasses and gall midges: plant defense and insect adaptation. *Annual Review of Entomology*, **48**: 549–577.

Hawthorne, D.J. 2001. AFLP-based genetic linkage map of the Colorado potato beetle *Leptinotarsa decemlineata* (Say): sex chromosomes and a pyrethroid resistance candidate gene. *Genetics*, **158**: 695–700

Heckel, D.G., Gahan, L.J., Daly, J.C. and Trowell, S. 1998. A genomic approach to understanding *Heliothis* and *Helicoverpa* resistance to chemical and biological insecticides. *Philosophical Transactions of the Royal Society London B Biological Sciences*, **353**: 1713–1722.

Hunt, G.J. and Page, R.E. Jr. 1992. Patterns of inheritance with RAPD molecular markers reveal novel types of polymorphism in the honeybee. *Theoretical and Applied Genetics Journal*, **85**: 15–20.

Hunt, G.J. and Page, R.E. 1994 Linkage analysis of sex determination in honeybee (*Apis mellifera*). *Molecular and General Genetics*, **244**: 512–518.

Kengne, P., Trung, H.D., Baimai, V., Coosemans, M. and Manguin, S. 2001. A multiplex PCR-based method derived from random amplified polymorphic DNA (RAPD) markers for the identification of species of the *Anopheles minimus* group in Southeast Asia. *Insect Molecular Biology*, **10**(5): 427-435.

Knipple, D.C., Doyle, K.E., Marsella-Herrick, P.A. and Soderlund, D.M. 1994. Tight genetic linkage between the kdr insecticide resistance trait and a voltage-sensitive sodium channel gene in the house fly. *Proceedings of the National Academy of Sciences*, **91**(7): 2483-2487.

Kuhner, M.K., Beerli P., Yamato, J. and Felsenstein, J. 2000. The usefulness of single nucleotide polymorphism data for estimating population parameters. *Genetics,* **156**: 439–447.

Luque, C., Legal, L., Staudter, H., Gers, C. and Wink, M. 2002. ISSR (inter simple sequence repeats) as genetic markers in Noctuids (Lepidoptera). *Hereditas*, **136**(3): 251-253.

Lehmann, T., Besansky, N.J., Hawley, W.A., Fahey, T.G., Kamau, L. and Collins, F.H. 1997: Microgeographic structure of *Anopheles gambiae* in western Kenya based on mtDNA and microsatellite loci. *Molecular Ecology*, **6**: 243–255.

Llewellyn, K.S., Loxdale, H.D., Harrington, R., Brookes, C.P., Clark, S.J. and Sunnucks, P. 2003. Migration and genetic structure of the grain aphid (*Sitobion avenae*) in Britain related to climate and clonal fluctuation as revealed using microsatellites. *Molecular Ecology*, **12**(1): 21-34.

Manguin, S., Kengne, P., Sonnier, L., Harbach, R.E., Baimai, V., Trung, H.D. and Coosemans, M. 2002. SCAR markers and multiplex PCR-based identification of isomorphic species in the *Anopheles dirus* complex in Southeast Asia. *Medical and Veterinary Entomology*, **16**(1): 46-54.

Margonari, C.S., Fortes-Dias, C.L. and Dias, E.S. 2004. Genetic variability in geographical populations of *Lutzomyia whitmani* elucidated by RAPD-PCR. *Journal of Medical Entomology*, **41**(2): 187-192.

Mendelson, T.C. and Shaw, K.L. 2005. Use of AFLP markers in surveys of arthropod diversity. *Methods in Enzymology*, **395**: 161-177.

Morin, A.P., Luikart, G., Wayne, R.K. and SNP workshop group 2004. SNPs in ecology, evolution and conservation. *Trends in Ecology and Evolution*, **19**: 208–216.

Nagaraja, G.M. and Nagaraju, J. 1995. Genome fingerprinting of the silkworm, *Bombyx mori*, using random arbitrary primers. *Electrophoresis*, **16**(1): 1633-1638.

Nagaraju, J., Reddy, K.D., Nagaraja, G.M. and Sethuraman, B.N. 2001. Comparison of multilocus RFLPs and PCR-based marker systems for genetic analysis of the silkworm, *Bombyx mori*. *Heredity*, **86**(5): 588-597.

Navajas, M., Lagnel, J., Gutierrez, J. and Boursot, P. 1998. Species-wide homogeneity of nuclear ribosomal ITS2 sequences in the spider mite *Tetranychus urticae* contrasts with extensive mitochondrial COI polymorphism. *Heredity*, **80**(6): 742-752.

Prasad, M.D., Muthulakshmi, M., Madhu, M., Archak, S., Mita, K. and Nagaraju, J. 2005. Survey and analysis of microsatellites in the silkworm, *Bombyx mori*. *Genetics*, **169**(1): 197-214.

Puterka, G.J., Black, W.C., Steiner, W.M. and Burton, R.L. 1993. Genetic variation and phylogenetic relationships among worldwide collections of the Russian wheat aphid, *Diuraphis noxia* (Mordvilko), inferred from allozyme and RAPD-PCR markers. *Heredity-London*, **70**: 604-604.

Ranson, H., Jensen B. and Wang X. 2000 Genetic mapping of two loci affecting DDT resistance in the malaria vector *Anopheles gambiae*. *Insect Molecular Biology*, **9**: 499–507.

Robinson, G.E., Grozinger C.M. and Whitfield C.W. 2005. Sociogenomics: social life in molecular terms. *Nature Reviews Genetics*, **6**: 257–270.

Roethele, J.B., Feder, J.L., Berlocher, S.H., Kreitman, M.E. and Lashkari, D.A. 1997. Toward amolecular genetic linkage map for the apple maggot fly (Diptera: Tephritidae): comparison of alternative strategies. *Annals of the Entomological Society of America*, **90**: 470–479.

Salvato, P., Battisti, A., Concato, S., Masutti, L., Patarnello, T. and Zane, L. 2002. Genetic differentiation in the winter pine processionary moth (*Thaumetopoea pityocampa*—wilkinsoni complex), inferred by AFLP and mitochondrial DNA markers. *Molecular Ecology*, **11**(11): 2435-2444.

Severson, D.W., Mori, A., Zhang, Y. and Christensen, B.M. 1993. Linkage map for *Aedesaegypti* using restriction fragment length polymorphisms. *Journal of Heredity*, **84**: 241–247.

Severson, T.F., Kamita, S.G., Hinton, A.C., Huang, Q., Inceoglu, A.B., Kang, K.D. and Hammock, B.D. 2001. Recombinant baculoviruses for insect control. *Pest Management Science*, **57**(10): 981-987.

Stevens, J., and Wall, R. 1995. The use of random amplified polymorphic DNA (RAPD) analysis for studies of genetic variation in populations of the blowfly *Lucilia sericata* (Diptera: Calliphoridae) in southern England. *Bulletin of Entomological Research*, **85**(4): 549-555.

Takami, Y., Koshio, C., Ishii, M., Fujii, H., Hidaka, T. and Shimizu, I. 2004. Genetic diversity and structure of urban populations of *Pieris* butterflies assessed using amplified fragment length polymorphism. *Molecular Ecology*, **13**(2): 245-258.

Ullmann, A.J., Piesman, J., Dolan, M.C. and WC IV, B. 2003. A preliminary linkage map of the hard tick, *Ixodes scapularis*. *Insect Molecular Biology*, **12**(2): 201-210.

Whitfield, C.W., Cziko A.M. and Robinson G.E. 2003. Gene expression profiles in the brain predict behavior in individual honeybees. *Science*, **302**: 296–299.

Wilkerson, R.C., Parsons, T.J., Albright, D.G., Klein, T.A. and Braun, M.J. 1993. Random amplified polymorphic DNA (RAPD) markers readily distinguish cryptic mosquito species (Diptera: Culicidae: Anopheles). *Insect Molecular Biology*, **1**(4): 205-211.

Wong, A., Forbes, M.R. and Smith, M.L. 2001. Characterization of AFLP markers in damselflies: prevalence of codominant markers and implications for population genetic applications. *Genome*, **44**(4): 677-684.

Yoon, C.K. and Aquadro, C.F. 1994. Mitochondrial DNA variation among the Drosophila athabasca semispecies and Drosophila affinis. *Journal of Heredity*, **85**(6): 421-426.