Research Paper

Volume : 2 | Issue : 10 | October 2013 • ISSN No 2277 - 8179

Approches of Insect Genomics

Biotechnology

KEYWORDS : Genomics, Insects, Genome sequencing, Insect Genome Annotation.

Dipali B. Borkar	Ph.D. Scholars, Biotechnology Centre, Botany Department. Dr. PDKV Akola-444104
Vishal L. Bagde	Ph.D. Scholars, Biotechnology Centre, Botany Department. Dr. PDKV Akola-444104

ABSTRACT

Insects are the most successful group from all metazoans on earth, has important societal, as well as scientific benefits. They occupy a wide range of roles, which have an effect on human life either because the former pose serious threats to public health and commercial crops as well as in some cases represent the only way to propagate food resources. Despite their tremendous importance, insect genomics remained an uneven territory dominated by studies. The application of genomic technologies in insecticide resistance has greatly expanded on insects to evolve insecticide resistance. In so far as many detoxification genes occur in tightly linked clusters. These selective sweeps on the genetic variation available to these species to respond to future insecticide or other xenobiotic challenges. In sequencing of insect genome showcases the genetic breadth of insects and a trend towards sequencing organisms directly involved with human welfare. We describe traits in other insect species that make them important candidates for genomics projects, and review several recent workshops aimed at uniting researchers working with insect species to efficiently address problems in agriculture medicine and biotechnology. Recent advances in biotechnology have led to the development and evolution in the field of bioinformatics for the analysis and integration of information from genomic, transcriptomic, proteomic, metabolomic and phenomic data. Availability of whole genome sequences, expressed sequence tags, genetic linkage maps and insect transgenesis has opened up new vistas for fundamental research in entomology.

Introduction
The single or set of circular and/or linear chains of DNA (or RNA for some viruses), constitute the genome. Genomics is a discipline in genetics concerning the study of the genomes of an organisms simply as study of all the genes of a cell, or tissue, at the DNA (genotype), mRNA (transcriptome), or protein (proteome) levels." The actual term 'genomics' is thought to have been coined by Dr. Tom Roderick. Major branch of genomics is still concerned with sequencing the genomes of various organisms, but the knowledge of the full genome has been under the field of functional genomics. It mainly concerned with patterns of gene expression during various conditions. When people say that the genome of a sexually reproducing species has been "sequenced", typically they are referring to a determination of the sequences of one set of autosomes and one of each type of sex chromosome (Michelmore, 2000). The genomics field includes intensive efforts to determine the entire DNA sequence of organisms and fine-scale genetic mapping efforts. In the 1970-1980s techniques of sequencing, genome mapping, data storage, and bioinformatic analyses was carried out. The most important tools here are microarrays and bioinformatics (Zhou *et al*., 2010). In insecticide resistance three types of mechanism were revealed in early studies, two involving enhanced detoxification of the insecticide and one rendering the target site for the insecticide insensitive to its effects. In detoxification mechanism involving sequestration of the insecticide was seen in two cases of resistance to organophosphate and carbamate insecticides in aphids and culicine mosquitoes. The second detoxification mechanism, involving active degradation of the insecticide, In which structural mutations had arisen in specific detoxifying enzymes.

Genomic research is comparative by nature
Genomics is the most recent branch of biology to employ comparison-based strategies. Comparative genomics provides a powerful and general approach for indentifying functional elements without previous biological knowledge. It aids in the identification of genes, gene structure, regulatory elements, and evolutionary forces acting on an organism's biological processes its need for survival. The field of biology emerged as a discipline rooted in comparisons. Comparative physiology has assembled a detailed catalogue of the biological similarities and differences between species, revealing insights as to how life has adapted to fill a wide range of environmental niches (Batley *et al*., 2009). Attaining this information might finally bring us closer to understanding how species have managed to maintaining entire sets of genes which are conserved among many species, as well as developing sets of genes unique to each of them. Comparative genomics analyses have lead to changes in our understanding of some phylogenetic relationships along the branches of the evolutionary tree. Comparative genomics is the study of the relationship of genome structure and function across different biological species or strains. It exploits both similarities and differences in the proteins, RNA and regulatory regions of different organisms to infer how selection has acted upon these elements. For this reason comparative genomics studies of small model organisms, It has great importance to advance our understanding of general mechanisms of evolution (Michelmore,2000). Ranson *et al*. (2002) have compared the three major gene families that have so far been implicated in insecticide detoxification between the two insect genomes that have been fully sequenced. The three gene families are the cytochrome P_{450}, carboxylesterases and glutathione-S-transferases (GSTs), and the two species are the dipterans *Drosophila melanogaster* and the malarial mosquito *Anopheles gambiae*. Many of the esterases are also expected to have specialist non-detoxification functions. the P_{450} seem to be the main resource for the evolution of resistance by enhanced detoxification (Dary *et al*., 2005).

Computational genomics
Computational genomics refers to the use of computational analysis to decipher biology from genome sequences and related data, including both DNA and RNA sequence. As such, computational genomics may be regarded as a subset of bioinformatics, but with a focus on using whole genomes (rather than individual genes) to understand the principles of how the DNA of a species controls its biology at the molecular level and beyond. Their research developed a phylogenetic tree that determined the evolutionary changes that were required for a particular protein to change into another protein based on the underlying amino acid sequences. (Michelmore,2000).

Functional genomics
Functional genomics is a field of molecular biology that attempts to make use of the vast wealth of data produced by genomic projects (such as genome sequencing projects) to describe gene (and protein) functions and interactions. Unlike genomics and proteomics, functional genomics focuses on the dynamic aspects such as gene transcription, translation, and protein-protein interactions, as opposed to the static aspects of the genomic information such as DNA sequence or structures. Functional genomics attempts to answer questions about the function of DNA at the levels of genes, RNA transcripts, and protein products. (Bell'es 2009).

Genome sequencing
Almost any biological sample even a very small amount of DNA

or ancient DNA can provide the genetic material necessary for genome sequencing. Such samples may include saliva, epithelial cell, bone marrow, hair (as long as the hair contains hair follicle), seeds, plant leaves, or anything else that has DNA-containing cells. Whole Genome association studies can identify specific points of variation in eukaryotes in relation with DNA and genetic factors. Sequence a species without any known sequence information, by assembling the sequences with bioinformatics methods (Tomkins *et al*., 2002).

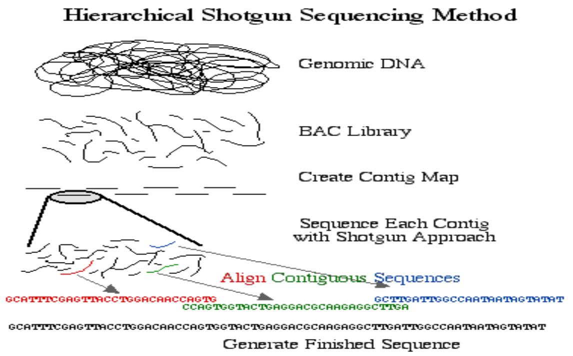
How do you sequence a Genome?

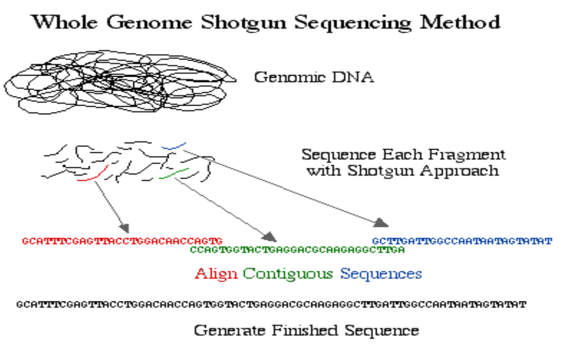
A) Hierarchical Shotgun Sequencing Method

In this approach, genomic DNA is cut into pieces of about 150 Mb and inserted into BAC vectors such as pbeloBAC 11, transformed into *E. coli*. The BAC inserts are isolated and mapped to determine the order of each cloned 150 Mb fragment. Each BAC fragment is fragmented randomly into smaller pieces and each piece is cloned into a plasmid and sequenced on both strands. These sequences are aligned so that identical sequences are overlapping. These contiguous pieces are then assembled into finished sequence. This approach was developed and perfected on prokaryotic genomes which are smaller in size and contain less repetitive DNA (Daborn *et al*., 2002).

B) Whole Genome Shotgun sequencing method

Shotgun sequencing randomly shears genomic DNA into small pieces cloned into plasmids Example, PSC 101 or PBR 322 on both strands, thus eliminating the BAC step from the Whole Genome Shotgun sequencing approach. Once the sequences are obtained, they are aligned and assembled into finished sequence. (Evans *et al*.,2003).





Why Insect genomics?

- **Evolution and comparative genomics:**

The complete sequence of an insect genome will be an important contribution to comparative genomics. Comparative studies on genome sequences between Insect, *Drosophila, C. elegans*, and vertebrates can point to insect-specific genes. *Drosophila* and *Anopheles* are distant evolutionarily from other insect, yet genome analysis of a species closely related to insect has not been performed (Smith, 2010). Several Bt-toxin resistance genes have been reported from *P. xylostella*, most of which probably encode the proteins that act as toxin-binding sites. Similarly, most of the laboratory-selected examples of resistance to Bt toxins involve multiple genes. Some of the individual genes underlying Bt-toxin resistance have now been mapped onto high-density linkage maps using QTL mapping.

- **Insects and Agriculture:**

One-third of the world's agricultural production is lost by insect pests, pathogens and weeds. The control of agricultural pest populations is achieved mainly by the application of chemical insecticides. The global insecticide market represents 12 billion dollar yearly expenditure. Biological control methods are also part of integrated crop protection strategies. Genetically modified crops with lepidopteran pests as the most common target. Despite environmental benefits of the reduced use of conventional pesticides, public acceptance of transgenic crops is low. Research is needed to develop new ways to enhance pest resistance of crops and allay concerns about genetically modified (George *et.al*., 1996).

- **Economically important pest**

The *Heliothines*, *Helicoverpa armigera* in Africa, Southern Europe, Asia, Australia and the Pacific, and *H. zea* and *Heliothis virescens* in the Americas, are the world's major crop pests. Orthologous genes of the these species are extremely similar, so that genomic information on one can readily be transfered to the other. These species cause damage amounting to billions of dollars yearly. Where they cannot be controlled, they affect the livelihood and cultural practices of entire regions (Smith, 2010).

- **Complex interactions**

Programs in plant, microbial or viral genomics can be productively interfaced with insect genomics. Insect are involved in an even broader range of interactions with other organisms. Tritrophic interactions between plants, insect and their parasitoids offer new paradigms for environmentally safe pest management (Zdobnov *et. al*., 2006)

- **Sericulture**

World silk production has approximately doubled during the last 30 years in spite of man-made fibers replacing silk for some uses. Applying methods of crop improvement yield and fiber quality can be tagged with molecular markers for rapid construction of genetically improved strains. Production of heterologous (e.g.spider) developmental and physiological processes that make the silk gland such an efficient bioreactor. (Grimmelikhuijzen, 2007).

- **Neurobiology and olfaction:**

Pheromones are chemicals emitted by living organisms to send messages to individuals of the same species. The discovery of pheromones and their binding proteins were achievements of insect biology. Genetic variance in the biosynthesis, reception or response to pheromones between closely and distantly related insect species is an area that genomics is uniquely suited to exploit. (Oakshott *et al*., 2003)

- **Insecticide resistance:**

Bt-producing transgenic cotton is currently effective in controlling pest damage while reducing chemical sprays; Bt-resistance threatens to force a return to the "pesticide treadmill". The Bt-resistance gene was cloned from an insect (*H. virescens*), and it will be a great challenge to apply this new knowledge soon enough to forestall or delay the appearance of resistance in the field (Gahan *et al*., 2001).

- **Evolution and Conservation:**

The discovery of mimicry in butterflies provided one of the most important early pieces of evidence for Darwin's theory of evolution by natural selection. More recently, butterfly wings have re-emerged as a model system in developmental genetics, Butterflies and moths figure as ecologically important and aesthetically appealing components of biodiversity and conservation projects worldwide (Sundaram *et al*., 2006).

- **Biotechnology:**

A complete knowledge of the insect genome enables further optimization of this system, for instance by providing a complete catalog of the enzymes used in post-translational processing of

Volume : 2 | Issue : 10 | October 2013 • ISSN No 2277 - 8179

the recombinant products (Wahlberg *et. al.*, 2008).

Insect Genome Annotation:

1. Cloning and sequencing

Whole insects were ground in liquid nitrogen and total RNA extracted using RNAqueous kit (Ambion, Huntingdon, UK) and treated with DNaseI (Sigma, St Louis, MO, USA). RT-PCRs were done with each primer pair using Hotstart Taq DNA polymerase (Qiagen, Valencia, CA, USA).

2. Data analysis

Insects constitute the vast majority of known species with their importance including biodiversity, agricultural, and human health concerns. However, proteome determination is an intensive undertaking. Computational method that uses genome analysis to characterize insect and eukaryote proteomes as an approximation complementary to experimental approaches, (Zhou *et. al.*, 2010).

3. Multiple sequence alignments

Two types of multiple sequence alignment (MSA) were generated, one with amino acid and one with nucleotide coding sequences (CDS). The MSA for the CDS orthologous regions was done by first aligning the amino acid sequences and then using this alignment to guide the nucleotide CDS alignment (Tittiger *et al.*, 2004)

4. Phylogenetic analysis

Phylogenetic relationships between homologous OBPs (both orthologs and paralogs) were obtained using the software MrBayes v3.1.2 (Ronquist & Huelsenbeck, 2003), under theWAG evolutionary model of amino-acid evolution (Torres, 2009).

Insect genome databases

Traditionally entomologist relies on textbooks and research articles as major resources for 'omics' information on insects. With the advancements in molecular biology and genomics technologies in the insect domain, lot of genomic data have been generated in the past decade.

This deluge of genomic information has led to an absolute requirement for computerized database to store, organize and index the data along with development of specialized tools to view and analyse the data. Genomic approaches are now becoming an important step for new developments in the biology and pathology of insects. The insect genomic databases contain information of all proteins, biochemical and physiological processes that occur in an insect (Heckel., 2008). Recently, efforts have begun to develop a comprehensive sequence database named Agricultural Pest Genomics Resources (Agripestbase) for storing genomic information from a broad range of pests, including insects, parasites and pathogens. Availability of genomic information on a broad range of agricultural pests will result in comparative genomics and further understanding of these species (http://agripestbase.org/). The information available from these databases will hopefully lead to better management strategies as well as new methods and targets for pest control.

Future aspects

Drosophila genomic sequence is a major milestone for genomics, as it indicates a new strategy for sequencing large eukaryotic genomes and as a model system to understand biological functions. In *Drosophila* post-genomic age, applications of genomics technology to entomology has added volume and quality to the data. In next few years several lepidopteran pest sequences will be available. This information will enormously increase our knowledge for understanding the biology of insects and insecticide resistance, which poses an increasing problem for pest control. In the future, where many genomes will be sequenced, a major application of bioinformatics will be the modeling of genetic and metabolic networks.

Conclusion

In conclusion, the application of genomic technologies in insecticide resistance has greatly expanded our views on the range of options available to insects to evolve insecticide resistance. On the other hand, however, we can also now see that the speed with which multiple resistance mutations are sweeping through some insect species will be substantially reducing the variation in linked genes. In so far as many detoxification genes occur in tightly linked clusters, these selective sweeps will impinge on the genetic variation available to these species to respond to future insecticide or other xenobiotic challenges Advances in sequencing technologies have provided opportunities in bioinformatics for managing, processing and analyzing the sequences. In this genomic era, bioinformatics is used as a bedrock of current and future biotechnology for finding new or better alternatives as designing potential target sites, safer insecticides and developing transgenic insects in applied insect science. The objective of this article is to provide comprehensive information on available insect genomic resources at one place to biotechnologists, molecular biologists, entomologists and physiologists for developing new methods in pest and disease management. The insect genomic databases are goldmines with information on all the proteins, biochemical and physiological processes of an insect. The newly sequenced insect genomes may harbor many surprises for biochemists, molecular biologists and insect physiologists. Insect pest control will soon enter the genomic era with all its surprises and discoveries, as pest and parasitoids genomes are now available thus, genomic advances during the last 10 years will revolutionize insect research.

REFERENCE

Batley J, Edwards D. Genome sequence data: management, storage, and visualization. Cell. 2009; 86, 521–529. | Bell´es X. Beyond. Drosophila: RNAi In Vivo and Functional Genomics in Insects ,Annu. Rev. Entomol. 2010; 55:111-128. | Daborn PJ, Yen JL, Bogwitz MR, Le Goff G. A single P450 allele associated with insecticide resistance in Drosophila. Science. 2002; 297:2253-2256. | Dary OG, Georghiou P, Parson's E, Pasture N. Microplate adoption of gomoris assay for quantitative determination of general esterase activity in single insect. J.Econ. Entomol. 2005; 83 : 2187-2192. | Evans JD, Rindal DG. Beenomes to Bombyx: future directions in applied insect genomics. Genome Biology. 2003;4:107. | Gahan LJ, Gould F, Heckel DG. Identification of a gene associated with Bt resistance in Heliothis virescens. Science. 2001; 293:857-860. | George L, Miklos G, Gerald M. The Role of the Genome Project Review in Determining Gene Function: Insights from Model Organisms. Current Biology.1996; 11:171–176. | Grimmelikhuijzen JP. The promise of insect genomics. Insect Biochemistry and Molecular Biology. 2007; 1–10. | Heckel DG, Gahan LJ, Daly JC, Trowell S. A genomic approach to understanding Heliothis and Helicoverpa resistance to chemical and biological insecticides. J. Biosci. 2008; 30(5). | Michelmore R, Genomic approches to plant disease resistance. Current Opinion in plant Biology. 2000; 3:125-131. | Oakeshott J G, Horne I, Sutherland T D, Russell RJ. The genomics of insecticide resistance. Genome Biology. 2003; 4:202. | Ranson H, Claudianos C, Ortelli F, Abgrall C, Hemingway J. Evolution of supergene families associated with insecticide resistance. Science. 2002; 298:179-181. | Smith H R, Genomics of Insect Pests to Agriculture Program. Biosci. 2010; 32:447–456. | Sundaram RM, Bentur JS. Biotechnology and insect pest management. Genome biology. 2006; 543-563. | Tomkins J, Luo M, Fang G, Main D, Goicoechea J, Atkins M. New genomic resources for the honey bee (Apis mellifera L.): development of a deep-coverage BAC library and a preliminary STC database. Genet Mol Res. 2002; 1:306-316. | Torres MC. Comparitive genetics and moleculer evolution, A Dissertation Presented to the Graduate School of Clemson University, 2009 . | Tittiger C. Functional Genomics and Insect Chemical Biology. Journal of Chemical Ecology. 2004; 30. | Wahlbere N. Genomic Outposts Serve the Phylogenomic Pioneers: Designing Novel Nuclear Markers for Genomic DNA Extractions of Lepidopter. Current Science. 2008; 89-7. | Zdobnov E M, Bork P. Quantification of insect genome divergence. Trends in Plant Science. 2006; 9-9. | Zhou JJ, Vieira FG, He XL. Genome annotation and comparative analyses of the odorant-binding proteins and chemosensory proteins in the pea aphid Acyrthosiphon pisum. IN Heong KL, Hardy B, editors. 2010; 429-446.