Tools and techniques of Molecular Genomics and Proteomics: A brief review

Sravanthi Yeligeti*
Department of Bio-technology, Kakatiya University, Warangal, India

*Corresponding author: Sravanthi Yeligeti, Department of Bio-technology, Kakatiya University, Warangal, India, Tel: 9059884939; E-mail: sravanthiyeligeti1305@rediffmail.com

Received: December 20, 2016; Accepted: March 20, 2017; Published: April 02, 2017

Abstract
The breakthrough in Science and technology accelerated the understanding of many fields of Biology. The advancement in science and technology is also playing a major role in shaping our daily life, especially, for the present generation. Scientific in-depth analysis of cellular mechanism, interdisciplinary approaches and their impact on gaining specific knowledge for combating disease or exploring a microorganism revealed the infinite possibilities in front of us. One such scientific approach being implemented at present is molecular genomics. Molecular genomics received the much due attention of the scientific community after the human genome project (HGP) and the scientific community realized that there is much to know beyond the one-gene-one-protein hypothesis. Numerous scientific methodologies including both experimental and theoretical techniques aided in increasing the reliability and accuracy of the approaches and understanding. This review article provides an overview on the basic techniques and methodologies which became the synonymous with molecular genomics.

Keywords: Molecular genomics, molecular techniques, computational molecular genomics, polymerase chain reaction, genomics.

Introduction
The term “Molecular genomics” is the branch of genetics that deals with the expression of genes by studying the DNA sequences of chromosomes. It is the branch of genetics which mainly focuses on the intra cellular mechanisms such as chemical structure and the functions, replication, and mutations of the molecules involved in the transmission of genetic information [1], in the hereditary vehicles the DNA and RNA. Genomics is a much newer field than genetics and became possible only in the last couple of decades due to technical advances in DNA sequencing and computational biology. The branch Molecular genomics is concerned with the arrangement of genes on DNA, the process of replication, transcription and the translation process where the DNA is converted into RNA and RNA into proteins[2].

Genomics is a quickly evolving field of research that distributes the radical approach in understanding the relationship between the genetic code and diseases. The significant breakthroughs in the last decade have now positioned the animal health research community to capitalizing the infrastructure built by the human genome project [3]. Major milestones include the drafting of genome sequences for several animal species like the chicken, dog, cattle and the sequencing of the other animals like cat, pig, and sheep genomes has been initiated. Like animal genomics, plant genomics is also emerging its importance in the scientific community in production of hybrid varieties which are rich in energy supplements which are

Citation: Citation: Yeligeti S. Tools and techniques of Molecular Genomics and Proteomics: A brief review Biochem Mol Biol Lett. 2017 3(2):120.
© 2017 Trade Science Inc.
commonly required in human survival. Advancement in plant genomics had made a revolutionary impact in disease resistant plants and also increased the shelf life [4-6] of the outputs from the plant by inserting the desired gene in the genome of the plant cell.

Broadly applied molecular genomics may be of two major types, experimental methods applied for genomics and theoretical laboratory based analysis.

**Analysis of nucleic acids:**

**Polymerase chain reaction (PCR):**
The purpose of PCR is to create infinite copies of a particular target gene in order to use it for research, treatment and forensic investigations. Amplifying a small amount of DNA more easily and rapidly than cloning is possible by PCR [7, 8]. It is carried out in 3 steps namely- Denaturation, Annealing and Elongation.

**Annealing:**
The separated strands are renaturated with a primer by using heat stable enzyme called “Taq polymerase” which is extracted from thermophile bacteria called Thermus aquaticus [9].

**Elongation:**
The DNA with the primer is allowed to elongate the DNA strands into multiple strands by placing in thermo cycler.

**Gel electrophoresis:**
The use of agarose gel electrophoresis revolutionized the separation of DNA. Agarose gel electrophoresis is the highly effective technique used for the separation of DNA fragments varying sizes ranging from 100 bp to 25 kb. The Process in which amplified DNA or the desired DNA fragments of different lengths is runned on agarose gel and is separated by electrical current, negative charge of DNA and constant charge to-mass ratio [10, 11].

**Restriction Fragment Length Polymorphism (RFLP):**
The main purpose of RFLP is to differentiate between individuals. It is applicable in forensics and to certify relationships between individuals such as paternity tests, identification of criminals etc. [12, 13]. The desired DNA is cut by using the restriction endonucleases.

Restriction endonucleases (R.E’s) are the enzymes which are used to cleave the desired DNA at a specific recognition site. R.E’s are used to cleave genome at unique sites identical to the individual and the sample is sent through gel electrophoresis and the Banding pattern is observed through Southern Blotting and the DNA Patterns are compared.

**Blotting Techniques:**
Blotting is the commonly used technique in which nucleic acids or proteins are immobilized onto a solid support, usually on nylon or nitrocellulose filter membranes. Blotting is the central technique for hybridization studies of nucleic acid such as DNA and RNA [14, 15]. Labeling and hybridization of nucleic acids on membranes have formed the basis for a range of experimental techniques which are involved in understanding the organization, gene expression etc. [16, 17].
Three types of blotting techniques are commonly used for visualizing particular macromolecules namely: Southern Blotting, Northern Blotting and Western Blotting techniques.

**Southern Blotting:**
It is the technique which is used to detect the DNA from a complex mixture. In this method, the DNA is extracted from the cell and digested with the restriction enzymes. Then the DNA fragments are separated by gel electrophoresis. The agarose gel is washed with the washing buffer to remove the contaminants. Then the gel is removed carefully from the glass plate and placed on transfer buffer. Finally the nitrocellulose filter membrane is placed in the transfer buffer and a pile of blotting papers and a glass plate is placed on the nitrocellulose membrane to add the weight. Then the whole setup is placed in transfer buffer where the DNA is migrated on to the nitrocellulose membrane from the gel. This migration is due to the negative charge of DNA and the positive charge of the nitrocellulose membrane. After the immobilization is completed the nitrocellulose membrane is carefully removed from the transfer buffer and DNA bands on the nitrocellulose membrane is exposed under UV rays and the bands are detected by autoradiography.

**Northern Blotting:**
It is exactly similar to southern blotting but in this method the RNA is detected instead of DNA.

**Western Blotting:**
It is also similar to southern blotting but used in the detection of proteins. In this method the nitrocellulose membrane is not exposed to the UV rays but the antibodies are used to detect the proteins [18, 19]. The primary antibody binds to nitrocellulose paper while secondary antibody binds to the primary antibody. Non bound antibodies are usually washed off from the membrane. The presence of bound antibody is detected by incubating it with secondary antibody. Secondary antibody is usually attached with an enzyme so that its presence can be detected on a nitrocellulose membrane. Secondary antibody can also be attached with radioactive isotopes. The color emitted determines the presence of particular protein [20].

**Applications of blotting techniques:**
Southern blotting technique is widely used to find specific nucleic acid sequence present in different animals including man. Northern blotting technique is widely used to find gene expression and regulation of specific genes. Infectious agents present in the sample can be detected. Identification of inherited diseases is possible by this technique. It can be applied to mapping restriction sites in single copy gene.

**Gene expression analysis**
**Real-time PCR:**
It is based on PCR. In this process a fluorescent dye is used as a reporter molecule which is used to observe the PCR [21]. The fluorescent dye enhances the amplified desired gene with each cycle of amplification. Based on the enhancement of fluorescent dye, the amplified gene is easily detected. The real-time PCR is categorized under two type’s namely: Non-specific Detection using DNA Binding Dyes and Specific Detection Target Specific Probes. Real time PCR allows the detection of PCR product during the early phases of the reaction [22]. This ability of measuring the kinetics of reaction in the early phases of PCR provides a distinct advantage over traditional PCR detection [23].
Micro Arrays:
DNA microarrays consist of diverse individual gene sequences which bound closely to space regions on the surface of a glass microscope slide or synthesized sequences on a chip surface. A DNA microarray allows the analysis of the expression of thousands of genes simultaneously [24]. The combination of DNA microarray technology with genome sequencing projects enables scientists to study the complete transcription process carried out in an organism during a specific physiological response or developmental processes [25, 26].

Fluorescence in situ hybridization (FISH):
Fluorescence in situ hybridization (FISH) is a stalwart technique used in the detection of chromosomal abnormalities [27-29]. The FISH technique is dependent upon hybridizing a probe with a fluorescent tag, complementary in sequence, to a short section of DNA on a target gene. FISH can be applied to detect genetic abnormalities such as characteristic gene fusions, aneuploidy and loss of a chromosomal region or a whole chromosome or to monitor the progression of an aberration serving as a technique that can help in both the diagnosis of a genetic disease or suggesting prognostic outcomes. FISH can also be applied to such research applications as gene mapping or the identification of novel oncogenes or chromosomal aberrations that contribute towards various cancers [30-34]. FISH is based on DNA probes annealing to specific target sequence of sample DNA.

Recombinant DNA Technology:
Ectrophotometric methods [26-30] have been reported for their determination. These methods were laborious, time consuming, or/and require derivatization of the drug. Therefore, our laboratory involved to develop new simple spectrophotometric and fluorimetric methods that overcome these drawbacks. In our previous report [32], the development of a simple fluorimetric method for determination of acyclovir and amantadine HCL was described (TABLE 1). The present study was dedicated to the development of new simple spectrophotometric methods for determination of both drugs in their pharmaceutical dosage forms.

Plasmids:
Plasmids are the small single stranded circular DNA molecules which are capable of replicating themselves in the chromosome. Plasmids are mostly present in unicellular organisms like protozoans and also likely to be present in eukaryotic organisms also [35-37]. Some of the commonly used plasmids are pBR322, pUC19 etc.

Cosmids:
Cosmids are plasmid vectors that contain “cos” sites. The “cos” site is required for DNA packaging in a phage particle. Cosmids contain the cos sequences of lambda phage. pHC79 is a classic example for cosmid which can carry large copies of DNA compared to other cosmids and plasmids.

Phasmids:
A phasmid is a cloning vector which is similar to that of a plasmid which is circular and self- replicating. The replication is initiated at the origin of replication site and the genetic material present in a phasmid is double stranded. These can carry a large amount of DNA when compared to plasmids.
Some recombinant DNA products being used in human therapy: Numerous human genes have been cloned in E. coli or in yeast. This made it possible for the first time to produce infinite amounts of human proteins in-vitro. Cultured cells transformed with a human gene are being used to manufacture Insulin for diabetics [38-40], Factor VIII for males suffering from Hemophilia A, Factor IX for Hemophilia B, Human Growth Hormone (HGH), erythropoietin (EPO) for treating Anemia [41], several types of interferon’s, several interleukins [42,43].

**Theoretical Techniques used in molecular genetics:**

**Theoretical epigenetics:**

MicroRNA are small non coding RNA gene products which lengths about 22nt in length which are almost found in almost all varieties of organisms. They play a key role in the translation repression by targeting the mRNA. Many recent studies reported that there are many miRNAs are to be discovered, as probably they are expressed in lower levels and more specialized expression to accelerate the genomic sequence exploitation. By the use of computational methods, mirCoS- can detect the new miRNAs and predict the extensive set of hairpin structures that based on the homolog in human and mouse genomic comparisons. By the use of novel approach, the mirCoS performs a better prediction in detection of the miRNAs. The miRNA gene family may be more abundant than previously anticipated, and confer highly extensive regulatory networks on eukaryotic cells [44].

Flaviviruses cause complex viral diseases which results in heavy socioeconomic burden [45]. Understanding and identifying the viral genome structure of the conserved regions and predicting the putative miRNA targets on such genomic regions may prove to be an effective approach in controlling these viral diseases. Repression of candidate genes involved in pathogenesis or other associated activity by a single or a cluster of miRNA may aid in efficient control on several diseases caused by different strains of Flaviviruses. Shinde et al. proposed a complex and comprehensive miRNA-target genome interaction map for viral and human miRNAs for the conserved regions of the 15 species of Flaviviruses and Several physicochemical properties of top-ranking miRNAs were identified and analyzed specifically for this family and they concluded that the transient expression profiles of small miRNA markers can be used for easy diagnosis and can provide a platform for multi-dimensional treatment plan, which may aid in designing effective strategy against viral diseases. In spite of complex relationship between miRNAs and viral genomes, the emerging evidences suggested that the use of anti-miRs to target miRNA in vivo will provide a novel and potential therapeutic approach for these viral diseases [46].

**Comparative genomics:**

Comparative genomics is a field of biological research in which the genomic features of different organisms are compared which include the DNA sequence, genes, gene order, regulatory sequences, and other genomic structural landmarks. The major principle of comparative genomics is that common features of two organisms will often to be encoded within the DNA that is conserved between them in respective to the evolution. Comparative genomics has revealed high levels of similarity between closely related organisms, such as humans and chimpanzees, and, more surprisingly, similarity between seemingly distantly related organisms, such as humans and the yeast *Saccharomyces cerevisiae*.

The Major Histocompatibility Complexes is the region where the genes are found densely in all the gnathostomates. MHC contains numerous highly polymorphic immune genes where a particular gene is involved in antigen presentation which plays a key role in disease resistance.
Using computational and experimental methods, comparative genomics can enable the identification of a minimal set of genes that is necessary and sufficient for the sustainability of a functional cell. The concept of a minimal gene-set for cellular life is originated from the following straight forward ideas namely- the functional parts of a living cell are protein and RNA molecules, and the instructions for making these parts are encoded in genes [47].

**Comparative Genomics for Eukaryotes:**
The comparative analysis of the genomes of Drosophila, *C. elegans* and Yeast and the proteins they encode, was taken in the context of cellular development and the evolutionary process. This study of different categories explains that the genome size is similar for the arthropods and nematodes but differ in bacteria which are twice to the size of the flies and worms [48].

**Gene and Protein sequence analysis:**
Gene Sequence Analysis: Gene sequence analysis is the process of subjecting the nucleic acids or peptide sequence to a wide range of analytical methods to understand its features, function, structure, or evolution. This sequencing analysis includes a wide range of activities such as:
- The comparison of sequences to identify the homology of the genes.
- Identification of Intron sites and Exon Sites on the gene.
- Identification of evolutionary and molecular diversities of the organism

For example, when porcine Rspos2 (pRspos2) was amplified from cDNA derived from porcine brain tissue, the DNA sequenced indicated the open reading frame (ORF) of pRspos2 consists of 732 base pairs (bp), which is encoded with 243 amino acids protein with a putative molecular weight of 28.19 kDa and is highly identical to those of horse (98%), cattle (98%), human (96%) and rat (92%). pRspos2 is highly expressed in the brain tissue, whereas with comparatively lower level of expression in duodenum, thymus, liver, lung, lymph node, kidney and ileum tissues [49].

**Protein Sequence and structure analysis:**
Protein sequencing is a technique to determine the amino acid sequence of a protein, and also conformation of the protein adopts. Moreover, to the extent to which a protein forms complexes with any non-peptide molecules can be also assessed applying the latest techniques.
In parallel to the understanding of the sequences, it is important to know the structural information of a protein molecule and how it interacts with other molecules. Computational structural analyses for proteins successfully fulfilled the requirement of the analysis gap present [50-59]. Ample work has been done in this direction and excellent literature evidence is present where structural details and interactions with important ligands have been reported for significant protein molecules [60-67].

**Application of Proteomics in Human Diseases:**
Human proteomics points out on physiological conditions which show a high impact on medicine. This knowledge has the capacity to decode the pathogenic mechanism repressed diseases which interpret the potential risk factors and molecular targets for drug development and therapeutic interventions and identification of biomolecules that could be developed for diagnostic and prognostic purposes and for better disease management strategies. Thus proteomics can translate basic scientific discoveries into the clinical practice for precision medicine [68].
Proteomic analysis for inflammatory responses:
Prostaglandin is a well-known hormone like lipid compound which had proved its role in declining inflammation. Exact appurtenant proteins associated functionally are unknown in different organisms. One among the other important aspect where on spot decision from medical science is essential is acute inflammation. Sudden and general burn conditions due to accident are a cause of concern for emergency medicine discipline [69].

Molecular phylogenetic:
Molecular phylogenetic is the branch of phylogeny that analyses hereditary molecular differences, mainly in DNA sequences, to obtain the information regarding the evolutionary relationships of organisms. Molecular phylogenetic is one aspect of molecular systematics, a broader term which includes the use of molecular data in taxonomy and biogeography. Every living organism contains DNA, RNA, and proteins. Conserved sequences, such as mitochondrial DNA are expected to accumulate over time mutations and helps in assuming the constant rate of mutation. Sanger sequencing made it possible to isolate and identify these molecular structures. Another application of the techniques that make this possible can be seen in the very limited field of human genetics, such as the ever-more-popular use of genetic testing to determine a child's paternity, as well as criminal forensics focused on evidence known as genetic fingerprinting.

As described above, the genetic markers play a key role in the identification of species, paternity testing, Huang et al. [70] has described some effective strategies in the identification of novel genetic markers by the combination of DNA micro array and Single Nucleotide Polymorphism (SNP) techniques which has given a full-fledged information regarding the role of novel genetic biomarkers in genetic polymorphism [71].

In his study, he applied the micro array technique to carcinogen identification, toxicology and drug safety which helped him to identify the differentially expressed genes [72]. By the combination of micro array technique and genotyping he identified many novel genetic markers [73] in animal selection. In this technique different probes like oligonucleotides, cDNA or small fragments of PCR products that correspond to mRNA were identified.

With this study, he concluded that DNA polymorphism can be differentiated with insertions, deletions [74] of single base pairs in gene and that can be identified by the genetic markers like PCR, RFLP, RAPD, micro array techniques etc. [75,76].

Conclusion
The use of molecular genetic technologies potentially offers a wide spectrum to select the genetic component even at embryonic stage and to select for a wide range of traits and to enhance reliability in predicting the mature phenotype of the individual. Molecular genomics in the present generation is like a boon for the mankind with its numerous advantages. It is not possible to sum up all the techniques available along with the specific applications in a single article. In this article, an attempt was made to provide a brief overview of the most used and available popular techniques in the recent times for molecular genomics and proteomics.

REFERENCES


