World Journal of Medical Sciences 1 (2): 72-81, 2006
ISSN 1817-3055

Biotechnology and Genomics in Medicine - A Review

Gulzar A. Niazi and S. Riaz-ud-Din

National Centre of Excellence in Molecular Biology, Punjab University, Lahore, Pakistan

Abstract: Biotechnology is associated with the manipulation of biological systems. The primitive use of biotechnology is documented as early as 6000 BC when people manipulated the innate properties of microorganisms, plants and animals to produce goods for their use. The modern biotechnology operates at the molecular level and within the short span of three decades, it has spread its wings into various fields of health and medicine especially in pharmacogenomics, development of cheaper and effective recombinant vaccines; screening, diagnosis and management of complex inherited diseases and product development of human needs. The human genome project that was completed in 2003 has led into next exciting phase of *Proteomics*, which is more challenging than the previous phase of *Genomics*. Consequently, new technologies and approaches are needed to handle the ever-increasing data generated by genomic projects. *Bioinformatics* is an example of this that combines the principles and concepts of computer science and biology. The public perceptions and fears about the abuse of biotechnology are there as the advanced knowledge in biotechnology can be used in biological warfare. Nevertheless future of the biotechnology is bright and it will certainly see significant strides that will be used for research and product development.

Key words: Molecular biology · biotechnology · genomics · proteomics · bioinformatics

INTRODUCTION

Biotechnology is scientific manipulations of living organisms, especially at the molecular level to produce useful products. The new tools of biotechnology, embodied in recombinant DNA technology (genetic engineering) have enabled the scientists to manipulate one gene at a time and create hosts that can harbor new genes. Genome is the complete DNA sequence, containing the entire genetic information of a gamete, an individual, a population or a species. Genomics is the study of genes and their role in organism's structure and function.

Biotechnology is an ancient art and a modern science [1]. Ancient biotechnology date back to several thousand years to when primitive humans used it for food production, medicinal purposes, for controlled breeding of plants and animals to select the ones with dominant trait and also to solve environmental problems.

The historical developments of modern biotechnology are as follows.

- The roots of modern biotechnology [2] date back about hundred years or so to the work of Louis Pasteur, Robert Koch and Gregor Mendel. While Pasteur and Koch laid the foundation of current microbiology. Mendel was first to describe the laws of heredity [3].
- The term biotechnology was coined in 1919 by Karl Erkey, a Hungarian agricultural engineer to refer all the lines of work by which products are produced from raw materials with the help of living organisms. Between the evolution of ancient biotechnology and modern biotechnology, there is piece of history that needs to be described. Although DNA was discovered in 1869 by Friedrich Miescher but it was isolated in pure state for the first time in 1935 by Andrei Belozersky. In 1953, Watson, Crick [4] described the double helical structure of DNA that marked the beginning of the modern era of genetics
- The techniques for insertion of foreign genes into bacteria were first developed in early 1970's when Hamilton Smith [5] discovered the restriction enzyme to cut DNA and then used ligase to paste two DNA strands together to form a hybrid circular DNA molecule (plasmid).
- The process of cutting DNA fragments was a mean to isolate genes or to alter their structure and

Corresponding Author: Dr. Gulazr A. Niazi, Medical Genetics Laboratory, National Center of Excellence in Molecular Biology, University of the Punjab, 78 West Canal Bank Road, Lahore, Pakistan

function. This was the first recombinant DNA molecule.

- Karry Mullis [6] developed polymerase chain reaction (PCR) that revolutionized the entire field of molecular biology.

The new biological techniques are permitting scientists to manipulate desired traits and the impact of modern biotechnology is now obvious on the society. In term of benefits to the society, many are yet to come and their applications are continuously discovered. In health and medicine, biotechnology is used for large-scale production of therapeutic proteins and other biopharmaceuticals; to diagnose and screen genetic disorders and to develop therapeutic modalities (gene therapy) to treat inherited conditions. Forensic applications of biotechnology have now entered in the judicial systems of many countries to solve crimes and legal problems. Microbes are being used to decompose and clean up contaminated sites by bioremediation technology [7]. Precisely both the biotechnology and genomics have potential of offering new therapeutic methods for the treatment of many diseases, as well, for developing new diagnostic techniques.

BIOTECHNOLOGY IN MEDICINE

Biotechnology is facilitating the development of new medicines, producing them faster, cheaper, safe and more efficient [8-10]. Instead of just studying the new drugs in clinical trials, scientists are identifying the generic cause of diseases. By stimulation, they are able to design and study the action of new products [11]. An estimated 400 pharmaceutical companies worldwide are conducting research and development into genetically engineered products and industry predicts that in few years several hundred genetically engineered products will be on the market. Many scientists believe that the impact of genetics in medicine will revolutionize the concept of human health and the scientific revolution in medical genetics is happening at a very fast pace. By year 2010 genetic tests will be available for 25 commonly encountered genetic disorders and by year 2020, drugs based on pharmogenomics will be routine part of common diseases such as diabetes and high blood pressure and by year 2040, individualized medicines will be produced [12].

Biotechnology revolution in health care and medicine started with the recombinant DNA technology (often referred as genetic engineering) revolution that began around 1970's. It allowed scientists to transfer genes from one organism to another, circumventing the sexual process. Several enzymes (such as restriction enzymes that cut DNA molecule at specific sites; DNA ligase that join DNA fragments end to end; DNA polymerase that synthesize DNA on a complementary template and DNA modification enzymes that are used to control ligation reaction and for DNA labeling) that bacteria use to manipulate DNA as a part of their normal cellular process enabled the recombinant DNA (rDNA) technology to flourish. In any exercise, the final objective of genetic engineering or rDNA technology is the stable and inheritable expression of a new trait in different organisms and certain basic steps are common to all rDNA experiments. To construct an rDNA molecule, DNA is isolated from a donor cell (animal or plant) and a plasmid is cut with the same restriction enzyme and mixed. "Sticky ends" of donor DNA form hydrogen bonds with the sticky ends of plasmid DNA fragment and recombinant molecule is sealed with another specific enzyme (ligases). Modified plasmid (rDNA) is introduced into a bacterium, which reproduces and clones the gene from donor cell that was spliced into the plasmid. The initial advances made in rDNA technology have increased our fundamental knowledge of the molecular basis of human diseases and this has resulted into the development of new field of medicine, known as molecular medicine. Stanley Cohen and Herbert Boyer [13] in 1973 and their work are the basis of much of the current work in biotechnology.

BIOPHARMACEUTICALS

There are four types of biopharmaceuticals: nucleic acid (e.g. DNA vaccines), proteins (therapeutic proteins including antibodies), viruses (e.g. bacteriophages and vectors) and cells (e.g. bacterial vaccines). With the exception of few nucleic acids that are synthesized commercially, all biopharmaceuticals are produced by a large-scale cultivation of microbial or animal cells (e.g. bacteria, yeast, animal cells). The major beneficial legacy of the early 20th century biotechnology was the discovery of penicillin by Alexander Fleming in 1928, an antibiotic derived from the mold of penicillium. Large scale production of penicillin was achieved in 1940s. The first application of rDNA technology was to mass-produce protein-based drugs and the first such drug was human insulin that was genetically produced in 1978 [14]. Insulin of rDNA origin (Humulin) is derived from plasmids in which the coding information for A or B chain is fused to

promoter and first few codons of the *E. coli* tryptophan (trp) gene. Separate cultures of *E. coli* are transformed with the A or B constructs and produces large amounts of the fusion peptides: either trp/A or trp/B. The tryptophan sequences are removed by treatment with cynogen bromide, which cleaves at methionine residue at the junction of the insulin gene. The A and B chains are mixed together at the junction and the chemical process forms disulphide bonds. Human growth hormone, a polypeptide of 191 amino acids, was also produced in *E. coli* by use of rDNA technique. The coding sequence for the first 24 amino acids of the expressed gene were synthesized chemically, where as amino acids from 25-191 were derived from cDNA copy of human growth hormone mRNA isolated from pituitary cells. The recombinant human growth hormone differs by one amino acid because *E. coli* is unable to remove the initiator methionine residue that is removed post-translationally in human cells.

Tissue plasminogen activator (t-PA) was produced in 1987 in transgenic mice, is another genetically engineered therapeutic protein marked under the name of Activase.

t-PA is a protease that attacks fibrin, a major protein involved in forming blood clot. Patients that demonstrate early signs of heart attack or stroke are administered t-PA, which acts by destroying small blood clots that can potentially form blockage of arteries. However several studies have showed little difference in recovery of patients treated with t-PA versus those treated with streptokinase. There are currently about two dozens genetically engineered therapeutic proteins that are marketed to treat the clinical conditions ranging from acute myocardial infarction to rheumatoid arthritis. Genetech was world's first genetic engineering company that was established in 1976 and there are now many companies who are cashing on the rDNA technology.

Biopharmaceuticals now-a-days are made by the production of large scale culture of microbial or animal cell and as such are expensive. A number of alternative systems have been developed experimentally. These include the production of therapeutic proteins in the milk of farm animals, in chicken eggs and in plants. This methodology is called pharming [15] in which certain farm animals are transformed with transgenes (foreign genes) from animals that encode inherited gene products that are used to treat diseases or improve quality of life. Some of the selected therapeutic proteins of animal pharming that are produced in the milk of animals are: ATT (in sheep), t-PA (in goat), factors VIII, IX (in sheep), hemoglobin (in pig), lactoferrin (in cow), CFTR (in sheep, mouse) and human protein C (in pig).

The study of the genetic basis of the differential response to therapeutics in different patients is allowing for the development of more efficient and safer medicines. Several pharmaceutical companies have conducted clinical trials for the treatment of different types of cancers e.g. IMC-C225 (ImClone System), GVAX (Cell Genesys), ABT 627 (Abbott) and SU 5416 (Pharmacia).

The use of antisense technology in drug discovery is a relatively new application of biotechnology. Antisense technology [16] selectively blocks the expression of a gene. It uses a piece of RNA that has complementary sequence to a sense RNA to stop expression of a particular gene. The rationale of this strategy is that all human diseases are the result of inappropriate protein production or improper protein performance. Host diseases e.g. cancer and infectious diseases e.g. HIV-AIDS, are protein based. These diseases can be treated at the root by preventing the undesirable production in the first place. Conventional drugs are designed to act on the disease causing proteins while antisense drug is designed to inhibit translation of specific mRNA targets into disease-causing proteins. For example, in case of cancer, antisense molecules comprising conventional single stranded anti-sense oligonucleotides (ASO) and small interfering RNA (siRNA) have shown to inhibit gene expression on the transcript level. This phenomenon is called post-transcriptional gene silencing [17].

MONOCLONAL ANTIBODIES

Monoclonal Antibodies (MAbs) are the antibodies that are identical because they are produced by one type of immune cell, all clones of a single cell. Given (almost) any substance, it is possible to create MABs that specifically binds to substance, they can serve to detect or purify that substance. They are widely used in diagnosis and research because of their specificity. In 1975, British scientists Cesar Milstein and George Kohler devised monoclonal antibodies (MAb) technology, which mass produced a single B cell, preserving its specificity and amplifying its antibody types. These monospecific antibodies are synthesized in hybridomas. In hybridoma technology, a B lymphocyte secreting antibodies against a specific antigen is fused with a Myeloma cell (a cancerous B-lymphocyte). The resulting cell if injected in a mouse's abdomen or if cultured in a bioreactor will grow and divide, indefinitely, producing large quantities of antibody, which can be harvested. The resulting proteins are called monoclonal (Mabs). The

most famous MAb containing diagnostic kit is the pregnancy test. Today MAbs are more like human antibodies as the original mouse versions caused allergic reactions. They are being used in different formats to treat cancer and various diseases [20, 21]. They can be coupled with radioisotopes to generate *in vivo* diagnostic imaging. Currently a large number of therapeutic antibodies have been introduced into human medicine e.g. to suppress immune system (*muromonab* inhibits the autoimmune destruction of beta cells in type 1 diabetes, *Infliximb* against inflammatory diseases, *Omalizuma*b against allergic asthma, *Duclizumab* against T-cell lymphoma); to kill or inhibit malignant cells (*Rituximab*, *Tositumomab*, to kill B dells, *Herceptin* binds to HER2,*Cetuximab* blocks HER1, *Mylotarg* binds CD33, *LymphoCide* binds CD22); as angiogenesis inhibitors (*Vitaxin, Bevacizumab*) and others (*Abiciximab*, inhibits the clumping of platelets by binding to receptors).

RECOMBINANT VACCINES

A vaccine is an inactive or partial form of pathogen that stimulates the immune system to alter B cell to produce antibodies. Two types of vaccines are commonly used: inactive vaccines and attenuated vaccines. They are administered prior to infection and provide varying degree of protection for varying duration of time. As such several major diseases have been eradicated (e.g. smallpox) while the spread of others have significantly been controlled (e.g. hepatitis, measles, typhus and tetanus). Vaccine technology is almost a thousand years old when Chinese used to collect the scabs from smallpox patients; crushed them into powder, which they either inhaled or rubbed into pricked skin. Modern vaccination dates to 1798, when Edward Jenner used non-lethal cowpox virus to induce immunity against structurally similar but deadly variola virus that causes smallpox in humans. Advances in genetic engineering in combination with new novel advanced technologies such as bioinformatics, microarrays and proteomics have revolutionized the approach to vaccine development. This has resulted in developing new vaccines and improving the quality of existing one [22, 23]. Subunit vaccines, recombinant vaccines, DNA vaccines and vector vaccines are rapidly gaining public acceptance and the new generation of vaccines are seriously considered alternatives to conventional vaccines. The difference between the conventional and genetically engineered vaccines is that conventional vaccines are prepared from killed or weakened versions or part of pathogens. Whereas DNA vaccines are prepared by inserting genes for one or two antigenic proteins into a plasmid that has been genetically constructed to lack the ability to reconstitute it and cause disease. The recombinant plasmids are delivered into the muscle cells of the host by a syringe or a gene gun. The recombinant plasmids enter the nucleus of the cell from where the host cell is compelled to express the enclosed antigen protein. The antigen-encoding genes are transcribed into mobile mRNA and subsequently translated into antigen protein. DNA vaccines have certain advantages, in addition to those of the conventional ones. They are capable of eliciting both humoral and cell mediated immune response. They can be engineered to carry genes both from different strains of a pathogen and thereby provide immunity against several strains at once.

Sequencing the genome of pathogens has allowed the development of more efficient vaccines and a new genome based approach "*Reverse Vaccinology*" has been successfully applied to produce more efficacious vaccines e.g. *N. meningitides* serotype B for which conventional strategies have failed [24]. Genomic sequences were obtained from many pathogens e.g. *P. falciparum* (malaria), *C. diphterie* (diphtheria), *N. meningitidis* (meningitis), *E. faecium*, which, has resulted in acceleration in development of new vaccines. The production of vaccines in recombinant microbes, animals and plants is now drawing great interest from WHO which recommends the use of vaccine as one of the efficient strategies for disease prevention. Many DNA vaccines have been developed and have undergone successful clinical trials. Currently genetically engineered vaccines are available for anthrax, whooping cough, tetanus, diphtheria, meningitis, tuberculosis, paratyphoid A and B, typhoid, pneumonia, cholera and hepatitis. Other have been developed as therapy for HIV, B-cell lymphoma, adenocarcinomas of breast and colon, prostrate cancer and cutaneous T-cell lymphoma and are in early stages of clinical trials. Research on edibles vaccines has been focused on GI diseases caused by *E. coli*, *V. cholerae* and retroviruses, Hepatitis B, type I diabetes and autoimmune diseases are also subject of investigation. But one main aspect still need to be understood with edible vaccines is how to guage the correct dosage of plant tissue that should be ingested to provide significant immunity.

DIAGNOSIS AND SCREENING OF GENETIC DISORDERS

Genetic disorders are broadly classified into five groups (a) single gene defects [2] chromosomal disorders

(c) polygenic disorders (d) disorders of mitochondrial DNA and (e) disorders due to somatic cell mutations. There are around 6000 diseases that have been designated as genetics and out of these single gene defects are most common. The recent advances in DNA recombinant technology have greatly improved our knowledge of gene organization, function and regulation, identifying several molecular lesions for many inherited disorders. Both prenatal and postnatal diagnostic methods are currently available for more than two dozens disorders that are highly prevalent in most world populations (e.g. sickle cell disease, thalassemias). For prenatal diagnosis, fetal cells are obtained by aminocentesis or by CVS and for carrier analysis; a blood sample is collected for DNA analysis. Allele specific oligonucleotides (ASO) probes are identifying alleles that differ by as little as one nucleotide [25]. A method using ASOs and PCR analysis is now available to screen for many genetic disorders. RFLP linkage analysis has opened a new era to demonstrate the inheritance of a particular disorder in the family and it is now being applied to several single gene disorders [26].

High throughput technologies such as DNA microarrays (also called DNA chips) have been developed for the profiling of gene expression pattern in whole organism or tissues [27]. Numerous applications of protein microarray-based assays are desired in basic biological research and in medicine to identify diagnostic biomarkers of inflammatory and cancerous pathologies and to find out new drug and new therapeutic targets [28]. DNA Microarrays are made of glass divided into small squares (fields) containing a set of target genes. In DNA microarray, DNA extracted from blood sample is amplified with PCR. The PCR product is tagged with a fluorescent dye and injected into microarray. The fragments with nucleotide sequence that exactly match the probe will hybridize. The pattern of and colors of the spot on the microarray reveal the resulting hybridization and are analyzed by the software linked to DNA microarray system. The US FDA has recently approved DNA chip based diagnostic test (AmpliChip CYP 450-Roche Pharmaceuticals) that would allow doctors to consider unique genetic information for patients in selecting and measure doses of medication for variety of common conditions.

BIOTECHNOLOGY IN TREATMENT OF GENETIC DISORDERS

Gene therapy is a molecular biotechnology technique for correcting genetic disorders by replacing defective genes with functional or normal genes. Gene therapy has some requirements, which should be met. First of all, genes of interest must be cloned; treatment should deliver sufficient copies of normal genes to target cells; transferred genes should have stable expression; modified cells must have survival advantage over unmodified cells and finally gene expression must correct or reverse the disease. Of the viral or non-viral methods of gene transfer, retrovirus and adenovirus based vectors had produced the best clinical results, but there remained always concern about the safety of these viruses that were used as vectors in the delivery of genes. Results also indicated that the somatic cell gene therapy would be practical and safer approach over germ line therapy [29]. The 1990s dawned with first successful trial of gene therapy on a patient with SCID and this brought the hope for many patients with inherited disorders. Successful trials were performed on the patients with different form of SCID, Hemophilia, Canavan disease and Cystic fibrosis. The present data shows that 632 gene trials were underway worldwide (mostly for cancer treatment) before the unfortunate mishap that resulted in the death of a patient with inborn errors of metabolism. The patient received large number of adenovirus vectors carrying orthinine transcarbamylase (OTC) genes that were inserted in the hepatic artery. Within hours, a massive immune response surged and the patient died of multiple organ failure. In another patient with SCID diseases who underwent gene therapy, leukemia-like condition developed [30]. As a result of that all the gene therapy trials were either suspended or shut down on USA till all the programs were thoroughly evaluated [31].

BIOTECHNOLOGY AND DNA PROFILING

A DNA profile of an individual reveals the detailed genetic pattern that is unique to the individual, enabling it to be used for a variety of diagnostic and identification purposes. It has rapidly become a standard and powerful tool used in forensic investigations. Alec Jefferys and his colleagues at the university of Leicester in England pioneered the technique of RFLP analysis and its application in forensic. His technique was based on the hypervariable regions, composed of short repeat sequences of DNA, minisatelite or cluster of 10-100 nucleotides that are widely spread in human genome. Modern DNA profiling is PCR based and uses short tendem repeats (STRs), which are very similar to variable number tendem repeats (VNTRs) but the repeated motif is shorter between two and four base pairs. The technology

was introduced in 1994 and the advantage of PCR is that it is quick to perform and allows a trace sample to be typed. STR loci have fewer alleles than VNTRs. PCR data is less compelling in criminal cases, except for excluding a suspect from the pool. However, a variation introduced in 1999 in Britain uses 10 STR loci. This guarantees that the odds of someone sharing the same results are less than one in a billion. In USA, the FBI developed 13 tetrameric (four base pair repeat) STR loci, called *CODIS Panel* in which the probability of that any two individuals having same 13 markers by chance is 1 in a trillion. If the STR pattern of a sample is from toothbrush of the victim, then the identification is fairly certain [32, 33].

GENOMICS IN MEDICINE

Genetics is a rapidly evolving field where new state of art research in biotechniques keeps on emerging with the passage of time [34]. Thomas Roderick was the first who in 1986 coined the term Genomics to describe the mapping, sequencing and characterization of genomes. Genome is the totality of genes in an organism. The scientists believe that the aim of genome is to produce biological periodic table that will represent an inventory of all genes that are involved in assembling of an organism. The first genome to be sequenced in its entirety was that of bacteriophage OX174 (5,368 kb) in 1980. The first free living organism to be sequenced was that of Haemophilus Influenza (1.8 Mb) in 1985 and since then the genomes are being sequenced at a rapid pace. Genomics has helped us to identify and map all genes in an organism and application of the sequence information can be highly valuable in basic research, biology and health. For example, comparison of a human DNA sequence to those of other organisms can lead to identification of disease causing gene. Genomics has some major challenges, which probably will be met in due course of time. In research, for example, genomics may provide some insight into structural and functional components of the human genome and it may help us to identify the pathways and networks through which genes and their protein products interact in different cell types, under different conditions. In health, genomics may provide information about genes that cause or contribute to disease. It may help us to understand which gene variants can provide resistance to particular diseases; how genome information can be used to predict disease susceptibility and drug response and how identification of genes can be used to develop new drugs.

HUMAN GENOME PROJECT

The human genome project was initially started by the United States Department of Energy to study the genetic and health effects of radiation and chemical products of energy production and it was determined that the best way was to study the DNA directly. International HGP to map and sequence human DNA that was launched in late 1990 was largely publicly funded. Two preliminary drafts of the physical map of human DNA were published in 2001 in the issue of *Science* and *Nature* [35, 36]. In 2003, the complete human DNA sequence that coincided with the discovery golden jubilee of the discovery of the DNA was produced [37]. Some of the salient information of the human genomes are: the human genome is largest and it contains 3,164.7 million nucleotide bases; of the total number of bases, 99.9% are exactly identical for all humans; the human genome is now estimated to contain 30-40 K protein coding genes (Fig. 1); average gene size including introns and exons is 27 kb; the largest human gene (1.25 Mb) is dystrophin gene associated with muscular dystrophy; only about 2% of the human genome is made up of protein coding and at least 50% 0f the genomes does not code for proteins (the so called "junk DNA"). The impact of that HGP has now and will be, that it will help the scientists in better understanding of the molecular mechanisms of diseases; markers for disease can be discovered; will help in understanding the biology of genome organization and gene regulation; will be useful in developing new technologies and finally HGP can increase the public awareness of development and application of biotechnology [38].

Genomics can be classified into:

- Structural genomics
- Functional genomics
- Comparative genomics

- Structural genomics [39] is concerned with the activities at the initial phase of genome analysis mapping (the construction of high resolution genetic, physical and transcript maps of an organism). Human genome project that has been discussed above is the example of structural genomics. Other examples of model organisms used in biotechnology research for which genome sequences have either been completed or in progress are: Plants (*Arabidopsis thaliana*, *Oryza sativa*, *Zee mays*); Mammalian

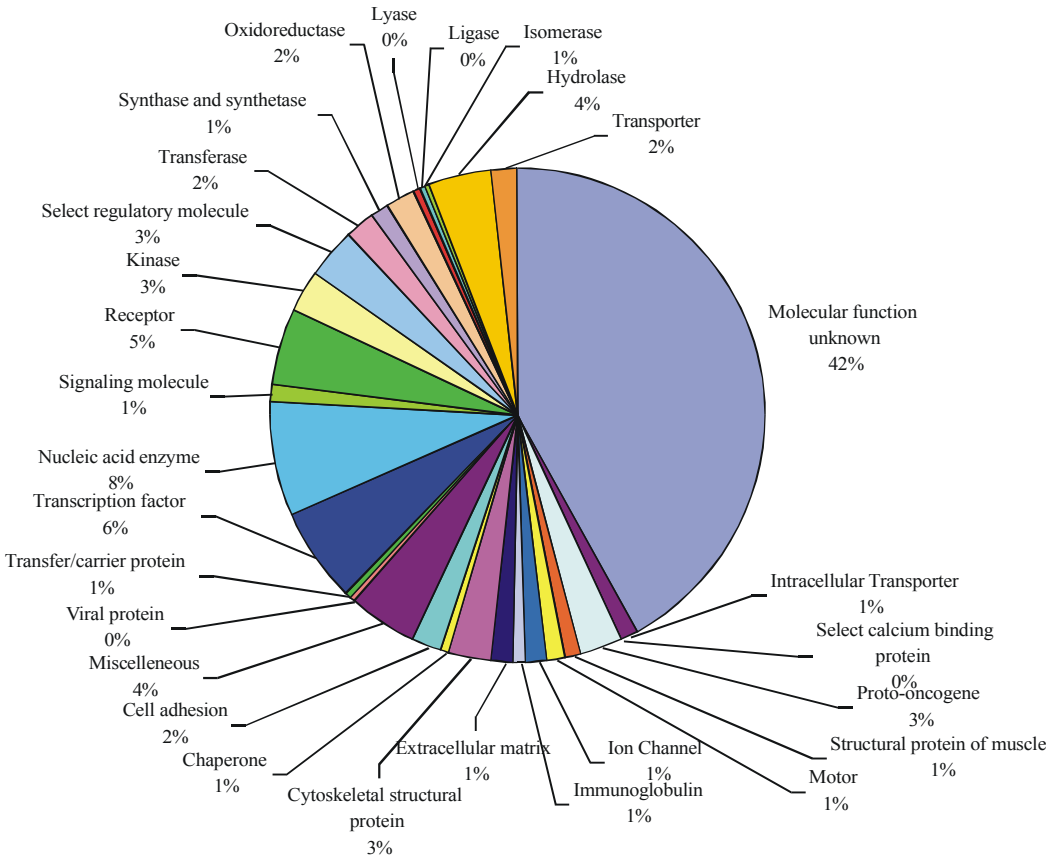

Figure 1: Human Genome, showing assigned functions of the genes

(*Homo sapiens*, *Mus musculus*, *Rattus*); Non-mammalian (*Drosophila melongaster*, *Echerichia coli*, *Caenorhabditis elegans*, *Xenopus laevis*, *Danio rerio*, *Saccharomyces cerevisiae*, *Anopheles* spp).

- Functional genomics [40] is the effort at understanding the functions of gene sequences using new, high-throughput technologies; these technologies therefore must represent the new tools of medical discovery. We all know that genes provide instructions for making specific proteins; however in the process of carrying these instructions, additional proteins can be produced. These uncertainties make linking genes to function a complex undertaking. Nonetheless, understanding the genome structure alone is insufficient; it is critical to identify the proteins the genes encode. For example in case of Alzheimer disease, several genes that increase the risk of having the disease have been identified. However the conclusive diagnosis of the disease comes from the presence of beta amyloid protein fragments but there is no beta amyloid gene and consequently Alzheimer's disease can not be identified through DNA microarray technology.

Many techniques have been developed and continue to be developed, for deciphering gene function. Bioinformatics is knowledge based theoretical discipline that attempts to make predictions about biological function using data from DNA sequence analysis. It uses supercomputers and sophisticated software to search and analyze databases accumulated from genome sequence projects and other sources. Two general categories of information are used in bioinformatics research: primary and secondary databases. The primary database consists of original data such as raw DNA sequences and protein structure information from crystallography while secondary databases use the information in primary database to compile profiles of highly conserved families [41]. In both types, a good database should have the original sequence and an annotation of description of the biological context of

the data. Proteome is the complete set of proteins encoded by a genome. Proteomics [42, 43] is the study of proteins encoded by a genome, including a list of which genes are expressed, their time of expression, the type and extent of any post-translational modification of the gene product, the function of the encoded proteins and its location in various cellular compartments. Two generalized categories of analysis are identified in proteomic: (i) expression proteomic that is involved in studying global changes in protein expression (ii) cell-map proteomics which entails a systematic study of protein-protein interactions through the isolation of protein complexes. Proteomics track protein expression by cells, beginning with the functional protein and backtracking to the gene that encoded it. Unlike genome, which entails a process in which there is a well defined ultimate end-point (i.e. the complete sequence of an organism's DNA) such a finality is impossible to define for proteomics. This is because, unlike DNA that is fixed, the proteome depends on when the sample is taken from the organism and kind of tools used for the investigation. The basic techniques in proteomics involve separating and identifying proteins isolated from cells. The most commonly used combination of techniques involves two-dimensional gel electrophoresis (2DGE) and mass spectrometry. With the draft sequence of human genome, it is possible to perform high throughput mapping of protein-protein interactions in humans, which is known as functional proteomics. As we emerge into the post-genome era, proteomics finds itself as the driving force field as we translate the nucleic acid information archive into understanding how the cells actually works and how disease processes operate. Clinical application of proteomics emphasizes the use of proteomic technologies at bed side with the ultimate goal to characterize information flow through the intra- and extracellular molecular network that interconnect organ and circulatory systems together. Proteomics-based analysis of traditional sources of biomarkers, such as serum or tissue lysates, has resulted into tremendous amount of information and serum proteomics pattern diagnosis is an emerging trend that employs low resolution mass spectrometry to generate a set of biomarker classified for early detection of disease. Studies have shown [44] that the low molecular weight (LMW) range of the circulatory proteome are the rich source of information that may detect the disease at an early stage and stratify risk. As mentioned above serum proteomic pattern diagnostics is a new technique in which proteomic signatures are used as a diagnostic classifier and the surface enhanced laser desorption/ ionization time of flight (SELDI-TOF) mass spectrometry holds a promise as modality for biomarker discovery of early stages of cancers.

- Comparative genomics [40, 45] is a field that searches for similarity and differences among genomes. It is often stated that a particular organism shares X percent of its DNA with humans. This number indicates the percentage of base pairs that are identical between two species: for example, there is 98.4% genetic similarity between the chimpanzee and human. The principle of comparative genomics is that similarities between species extend far beyond the level of individual gene and include whole genomes. This can be useful in number of ways. In closely related species e.g. man Vs baboon, comparative genomes can help to identify genes and their regulatory elements, since only sequences with an evolutionary conserved function would be found in both genomes while others would have diverged significantly. Although sequences and/or structural comparison can be useful first recourse for functional annotation, but for several reasons, the information must be treated with caution. For example low complexity regions that are found in many proteins with extremely diverse functions; that sequence similarity does not guarantee functional similarity and some proteins can acquire additional functions during evolution.

BIOTECHNOLOGY: THE FUTURE

Biotechnology is like an emerging star on the horizon that has potential ability to shine in the future. The human genome that has now been sequenced has not only the challenges of understanding how the newly discovered genes function but also it will have large impact on various drug-related fields including drug discovery and clinical medication. Advances in "Omics technologies (genomics, transcriptomics, proteomics and metabonomics) will revolutionize our approach in development of novel therapeutics. The applications that are currently experimental that will become commercialized or at least further advanced toward practical application. Gene therapy therefore will be revisited. Bioinformatics and proteomics are newly emerging fields and it is hoped

that future research in proteomics will shed more information that how bacterial proteomes change with alterations in the environments. Pharmagenomics will make significant strides in advancing the concept of personalized therapy. Novel classes of drugs will widen the scope of therapeutic action beyond merely modifying transmitter function and stem cell therapies could offer an even more selective mode of targeting. High throughput technologies such as DNA microarrays will have potential applications in diagnostics, prognosis and drug therapy. Proteomic application in clinical care medicine will expand and recent advances in molecular biology and nanoscience will be applied to diseases of CNS, hematological malignancies and others. Rapid process is being made to modify the aging process and with the current knowledge, it will be possible to delay the onset of the life threating diseases. Xenotransplantation research, whose key goal is to overcome post-transplantation organ rejections, will continue in both the public and private sectors. Molecular nanotechnology will be used to analyze, understand the molecular machinery of the human body. Advances in biotechnology will be helpful in the development of more affective vaccines and as such will provide a new impulse for microbial research. The new generation vaccines, especially rDNA ones however are reportedly to be less reactogenic than conventional vaccines, but also less immunogenic. Therefore, there is an urgent need for development of new and improved class of adjuvants.

Biotechnology will shape the future research as this field is being driven forward both by private biotechnology companies and also by academicians who are introducing new technologies required for the parallel identification of individual proteins. Economic impacts of biotechnology will surpass the information technology and there will be a greater demand for qualified biotechnologists. As such the developing countries like Pakistan need to focus on the production of highly trained manpower in the field of biotechnology not only for local market but also for export purposes.

REFERENCES

1. Stubee, H., 1972. History of Genetics, From Prehistoric Time to the Rediscovery of Mendel. (Translated by T. R. Waters), MIT press, Cambridge, MA.
2. Glick, B.R. and J.J. Pasternak, 1994. Molecular Biotechnology, Principles and Applications of Recombinant DNA. 1994; ASM press, Washington DC.
3. Bateson, W., 1909. Mendel's Principles of Heredity. Cambridge University Press. Cambridge, MA.
4. Watson, J.D. and F.H.C. Crick, 1953. Molecular structure of the nucleic acids. A structure for deoxyribose nucleic acid. Nature, 171: 737-738.
5. Smith, H.O. and K.W. Wilcox, 1970. A restriction enzyme from Hemophilus Influenza 1, purification and general properties. J. Mol. Biol., 51: 379-391.
6. Mullis, K.B., 1990. The unusual origin of polymerase chain reaction. Sci. Amer., 262: 56-65.
7. Deutschbauer, A.M., D. Chivian and A.P. Arkin, 2006. Genomics for environmental biology. Curr. Opin. Biotechnol., 17: 229-235.
8. Drew, J., 2003. Strategic trends in the drug discovery. Drug Discov. Today, 8: 411-420.
9. Knowles, J. and G. Gromo, 2003. A guide to drug discovery: target selection in drug discovery. Nature Rev. Drug Discov., 2: 1-10.
10. Castle, A.L., M.P. Carver and D.L. Mendrick, 2002. Toxicogenomic: a new revolution in drug safety. Drug Discov. Today, 7: 728-736.
11. Ohlstein, E.H., A.G. Johnson, J.D. Elliot and A.M. Romanic, 2006. New strategies in drug discovery. Methods Mol. Biol., 316: 1-11.
12. Shastry, B.S., 2006. Pharmacogenetics and concept of individualized medicine. Pharmacogenomics J., 6: 16-21.
13. Cohen, S.N., A.C.Y. Chang and H.W. Boyer, 1973. Construction of biologically functional bacterial plasmids invitro. Proc. Nat. Acad. Sci., USA., 70: 3240-3244.
14. Keen, H., A. Glynne and J.C. Pickup, 1980. Human insulin produced by recombinant DNA technology: Safety and hypoglycemic potential in healthy men. Lancet, ii: 398-401.
15. Crommelin, D.J.A. and R.D. Sindelar, 2002. Pharmaceutical Biotechnology, 2nd edn.
16. Carmichael, G., 2003. Antisense starts making more sense. Nature Biotechnol., 21: 371-372.
17. Wacheck, V. and U. Zange Meister-Witte, 2006. Antisense molecules for targeted cancer therapy. Crit. Rev. Oncol. Hematol., 59: 65-73.
18. Borrebaeck, C.A.K. and R. Carlson, 2001. Human therapeutic antibodies. Curr. Opin. Pharmacol., 1: 404-408.
19. Fischer, R., R.M. Twyman and S. Schillberg, 2003. Production of antibodies in plants and their use for global health. Vaccine, 21: 820-825.
20. Funaro, A., A.L. Horensyein and P. Santoro, 2000. Monoclonal antibodies and therapy of human cancers. Biotechnol. Adv., 18: 385-301.

21. Lewis, R., 2001. Herceptin earns recognition in breast cancer arsenal. The Scientist, 15: 10.
22. Capechi, B., D. Serruto, J. Adu-Bobie, R. Rappouli and M. Pizza, 2004. The genome revolution in vaccine research. Curr. Issues Mol. Biol., 6: 17-27.
23. Grandi, G., 2001. Antibacterial vaccine design using genomics and proteomics. Trends Biotechnol., 19: 181-188.
24. Scatselli, M., M.M. Giuliani, J. Adu-bobbie, M. Pizza and R. Rappuoli, 2005. The impact of genomics on vaccine design. Trends Biotechnol., 23: 84-91.
25. Antonarakis, S., 1989. Diagnosis of genetic disorders at the DNA level. N. Eng. J. Med., 320: 153-163.
26. The Huntington's Disease Collaborative Research Group, 1993.. A novel gene containing a trinuclotide repeats that is expanded and unstable on Huntington's disease chromosome. Cell, 72: 971-983.
27. Zhu, H. and M. Snyder, 2003. Protein chip technology. Curr. Opin. Chem. Biol., 7: 55-63.
28. Lueking, A., D.J. Cahill and S. Mullner, 2005. Protein biochips: A new and versatile platform for molecular medicine. Drug Discov. Today, 10: 789-94.
29. Niazi, G.A., 1997. Gene therapy: recent advances, future directions and concerns. Saudi Med. J., 18: 1-8.
30. Check, E., 2002. Shining hopes dented-but not dashed. Nature, 420: 735.
31. Juengst, E.T., 2003. What next for human gene therapy. Brit. Med. J., 326: 1410-1411.
32. Findlay, I., A. Taylor, P. Quircke, R. Frazier and A. Urquhart, 1997. DNA fingerprinting from single cells. Nature, 389: 555-556.
33. Lee, H. and F. Tirnady, 2003. Blood evidence. Cambridge, Mass.
34. Spalvieri, M.P., R.G. Rotenberg, 2004. Genomic medicine. Polymorphisms and microarray applications. Medicina (B Aires), 64: 533-542.
35. International Human Genome Consortium, 2001. A physical map of the human genome. Nature, 409: 934-941.
36. Venter, J.C., M.D. Adams and E.W. Myers, 2001. The sequence of the human genome. Science, 291: 1303-1351.
37. Collin, F.S., 2003. A vision for the future of genome research. Nature, 422: 835-847.
38. Van Ommen, G., 2002. The human genome project and future of diagnostics, treatment and prevention. J. Inherit. Metab. Dis., 25: 183-188.
39. Baker, D. and A. Sail, 2001. Protein structure prediction and structural genomics. Science, 29: 93-96.
40. Schoulink, K.G., 2002. Functional and comparative genomics of pathogenic bacteria. Curr. Opin. Microbiol., 5: 20-26.
41. Blundell, T.L. and J.H. Abell, 2002. High throughput crystallography for lead discovery in drug design. Nature Rev. Drug Discov., 1: 45-54.
42. Hanash, S., 2003. Disease proteomics. Nature, 422: 226-232.
43. Abersold, R. and B.F. Cravatt, 2002. Proteomics-advances, applications and challenges that remain. Trends Biotechnol., 20: s1-2.
44. Petricoin, E.F. and L.A. Liotta, 2004. SELDI-TOF based serum proteomic pattern diagnostics for detection of cancer. Curr. Opin. Biotechnol., 15: 24-30.
45. Clark, M.S., 1999. Comparative genomics: Key to understanding the Human Genome Project. Bioassay, 21: 21-30.