THE FUTURE OF MEDICINE AND BIOMEDICAL RESEARCH IN PAKISTAN: Ist Draft

Dr. S. Qasim Mehdi with contributions by Dr. Qasim Ayub and Aiysha

Abid

INTRODUCTION AND BACKGROUND.

The revolution of modern genetics started in 1865 with the discovery and formulation of Mendel's 'laws of inheritance'. However, his work went unnoticed until Garrod (1902) found that alkaptonuria, an inherited disorder, followed Mendelian laws. Further support came from the work of Karl Landsteiner (1901) whose work on the inheritance of human blood groups was also interpreted to obey Mendel's laws.

To formulate the laws of inheritance, Mendel painstakingly spent many years to create pure lines of the pea plant that carried one or two clearly observable properties- plants that produced only round or wrinkled seeds or those with either long or short stems. From his meticulously planned experiments, Mendel concluded that a pair of "factors", one of which was inherited from each of the parents, controlled the characters that were being examined.

The chemical nature of the "factor" remained unknown until Frederick Griffith made a landmark discovery in 1928. He observed that when mice were injected with a mixture of heat killed virulent *pneumococci* (S type) and a non-pathogenic strain (R type), most of the animals died and surprisingly, the blood cells of the mice had live S type *pneumococci*. He concluded that a substance from the dead strain transformed the R type to become virulent. Using "purified" deoxyribonucleic acid (DNA) that contained no proteins or ribonucleic acids (RNA), Avery et al. (1944) showed that the transforming agent was in fact DNA. Additional support for DNA came from Hershey and Chase (1952) who infected *E.coli* with bacteriophage T2 in which the coat protein was labelled with ^{35}S and the DNA with ^{32}P. They showed that DNA, not the protein, was required for the production of the phage progeny. Finally, the elucidation of the double helical structure of DNA (Watson & Crick, 1953) laid the foundation of the 'Golden Era' of modern genetics (biotechnology and genetic engineering) and gene hunting in the 1970's and 1980's.

The discovery of enzymes that cut DNA (Arber 1974; Roberts 1981) founded the recombinant DNA technology (Cohen, 1973; Morrow, 1974; Berg et al. 1974) and led to the production of recombinant human insulin, which is of immense clinical benefit (Johnson 1983). Sanger's success in sequencing the genome of a bacteriophage (Sanger, 1977) and the ability to make millions of copies of a specific piece of DNA in a test tube without using the machinery of any living cell by Kary Mullis (Saiki, 1985) are a few hallmarks of this era.

When nucleic acids duplicating enzymes were discovered (Kornberg 1960; Okazaki et al. 1968; also see Kornberg and Baker, 1992), Watson and Crick's model of the structure of DNA provided an explanation for a number of vitally important questions. The two antiparallel strands of deoxyribonucleotides

that strictly obey the hydrogen bonding of A-T and G-C, explained the very basis of DNA replication and hence reproduction. The transcription of DNA into a messenger RNA (mRNA) molecule (Ochoa 1963; Holley et al. 1965) unravelled its intermediary role in protein synthesis. Soon after, it was shown that the sequence of a triplet of nucleotides (codon) code for an amino acid in the biosynthesis of proteins (Nirenberg 1963; Khorana et al. 1967) and this established the co-linearity of the gene and the protein that it codes for.

During the second half of the twentieth century, biology and therefore medicine found a universal language- the language of chemistry. This attracted a large number of workers to move into biology from other scientific disciplines. In addition, modern biologists became increasingly at ease to borrow and use techniques from the physical and chemical sciences.

The awesome power of modern technologies for cutting, patching, sequencing, synthesizing, making millions of copies of a specific piece of the genome and the availability of dedicated computer programmes lead to the completion of the sequencing of the human genome and that of other organisms (Venter et al. 2001; The Human Genome Project: International Human Genome Sequencing Consortium, 2001, 2004 and Dhand, 2006).

During the last two decades, rapidly emerging technologies used in the identification, characterization and analysis of genes have been made possible by advances in bioinformatics. Easy and immediate access to genetic data through the Internet (www) gave a real pace to these advancements. We do not have to wait long to find out whether a gene has been mapped and cloned, which genes are present on a particular region of a chromosome, which new regulatory factors are found in gene expression and what the new prospects are in therapies for genetic disorders. Primary genetic information such as DNA sequence of various organisms, genetic maps and information about genes and their protein products are stored in dedicated and up to date databases, which are accessible to anyone. Through these, researchers can match any new sequence that they find with that already present on these databases. They can find out which genes are responsible for a certain phenotype and even where they stand in the evolutionary tree. For human geneticists working on inherited disorders, OMIM (Online Mendelian Inheritance in Man) developed by Victor McKusick is a worthwhile database where human Mendelian characters are elaborately catalogued. Another important website for the researchers is Medline/Pub-Med, where online contents and abstracts of millions of research papers published in thousands of biomedical journals are easily available.

With the publication of the sequence of the last human chromosome (chromosome 1, Gregory et al., 2006), the primary goals of the HGP have been achieved. Briefly, the aim was the determination of the sequence of all the 3 billion base pairs of the human genome. This was accomplished by the development of fast and efficient methods for high throughput DNA sequencing and powerful computer programs to collect organize and interpret the enormous amount of data that was being generated by many dedicated laboratories worldwide. More than 99% of the human genome sequence with an error rate of 1 in 100,000 bases is freely available in various public databases. The human genome contains 20,000-25,000 genes but the functional annotation of many

genes is yet to be discovered. This is the main objective of molecular biologists in the post genome era.

Beside the HGP many other related projects were also initiated to aid the human genome project. These projects include the International SNP map working group (2001), The International HapMap Consortium (2005), Human Genome Diversity Project (HGDP, Cavalli-Sforza, 1990; Cavalli-Sforza et al., 1991), all three aim to catalogue the variation in DNA sequences between world populations (Cavalli-Sforza 2005). The Biomedical and Genetic Engineering Division (B & GED- KRL, Islamabad) has made pivotal contributions towards the HGDP (Cann et al. 2001). Many other projects sequenced the genome of model organisms as part of the comparative genomics efforts. These include some of the prokaryotes like *Escherichia coli* and *Haemophilus influenzae* as well as some eukaryotic organisms such as *Saccharomyces cerevisiae*, *Caenorhabditis elegans, Drosophila melanogaster* and *Mus musculus* etc. The study of the comparative genomics has provided vital information about the expression and function of homologous genes. Mapping of the human homolog of genes from other organisms provide candidate genes for the study of their functions in human development and their dysfunction in causing inherited diseases.

As a direct consequence of the finished human genome sequence, molecular biology is in a transition phase from the study of the structure of genes and DNA sequence information to the functions of their protein products. This area of study is called **proteomics**. This is the identification of the end product of genes, the proteins, their involvement in different biochemical pathways and interactions in development, differentiation, growth and disease pathobiology. Proteomics and transcriptomics are now the main fields of investigation for the purpose of functional genomics.

The '**proteome**' encompasses the entire complement of proteins present in a cell. Different proteins are translated in different cell types at different times as needed in the life cycle of a cell. The study of proteomics involves the characterization of proteins, their location, post-translational modifications and finally their specific function. Though the field of proteomics is in its infancy, tremendous developments have been made in the recent years in this field.

The '**transcriptome**' constitutes the entire mRNA transcript that is transcribed from the genome. Although all the cells of any organism harbour the same genome, the genes that are expressed in a particular organ or tissue vary depending on its function. A gene can produce many different transcripts either by the use of alternative promoter sites, alternative splicing or polyadenylation sites. Messenger RNA (mRNA), ribosomal RNA (rRNA) and transfer RNA (tRNA) have been known to have a key role in transferring genetic information through generations. Recently a new class of small units of RNA molecules were discovered that are thought to have a role in gene expression and gene regulation. These are micro (miRNA), small interfering (siRNA) and short hairpin (shRNA) RNA molecules.

Proteomics is much more complicated than genomics because multiple forms of proteins can be encoded by a single gene using alternative splice sites, by exploiting different reading frames and by virtue of having more than one translation initiation and termination sites. The post-translational modifications in the proteins also complicate the processing of the end product. An individual has

an identical genome in almost all the cells in the body, but the total contents of the transcriptome and proteome vary significantly in different cell types and in different cellular environments according to the specific requirements, location or function of the cells. These make the field of functional genomics a very dynamic and rapidly growing area in biology. Functional genomics has benefited from the advancement of novel experimental techniques to investigate gene structure and function and the use of bioinformatics.

Modern biology and genetic medicine have grown exponentially in the last few decades and details would be beyond the scope of this presentation. Therefore, only topics relevant to "Medicine in the 21st Century" will be covered briefly.

The human genome is composed of three billion base pairs that are packaged in 23 pairs of the human chromosomes. Twenty-two pairs of chromosomes are called 'autosomes' as they are the same in both males and females. One pair, the sex chromosome, is different. Human females have two X chromosomes while males have one X and one Y chromosome.

The genetic information carried by DNA does not get incorporated directly into proteins, which are the major structural and functional end points of this process. DNA always remains in the nucleus of the cell and is transcribed into mRNA. The mRNA migrates into the cytoplasm and is translated into polypeptide chain and proteins. A triplet of nucleotides code for an amino acid and therefore there are 64 possible combinations of the four bases that code for 20 different amino acids. It was also established that there are "start" and "stop" codons and because many different codons can code for the same amino acid, the code is redundant (Khorana, 1967).

In humans only 2-3% of the DNA codes for the proteins, while the remainder, collectively known as "junk" DNA contains non-coding regions, regulatory sequences, repetitive sequences and pseudogenes etc. Surprisingly, intermediate to large scale DNA copy number and sequence variants are also prevalent in the human genome and these segmental distributions are an additional source of individual human variations (She et el. 2004).

A gene is a stable entity and transfers its information from generation to generation in a specified manner. Due to environmental mutagens or errors in DNA replication and repair, the base composition changes and as a result of the mutation a new allele of the gene appears. The new allele is as stable as its older version and occupies the same position on the chromosome. The frequency of mutations generally lies between one in a hundred thousand and one in a million bases. Some mutations result in structural changes in the protein product that cause an inherited disease.

In many instances, clinicians observed a disease phenotype, but the discovery of its genetic defect took half a century or more. An example is sickle cell anaemia. In 1904, James Herrick came across a case of severe anaemia and observed that the patient's blood- smear contained red cells that were "thin, elongated, sickle-shaped and crescent shaped". Several years later he reported that some unrecognized change in the composition of the corpuscle itself might be the determining factor (Herrick, 1910). After years of efforts by researchers, Linus Pauling concluded that the "determining factor" was haemoglobin. Using electrophoresis to examine the physical-chemical properties of haemoglobin from

normal individuals (HbA) and sickle-cell patients (HbS), Pauling and co-workers (1949) revealed "*a clear case of a change produced in a protein molecule by an allelic change in a single gene".* Ingram (1956 &1957) finally sequenced the ?-chain of haemoglobin and showed that the glutamate at position 6 of normal HbA was replaced by valine in the HbS molecule. This results from a change in a single base, G**A**G to G**T**G in the gene. Thus, sickle cell anaemia became the first genetically transmitted disease to be characterized at the molecular level. Prenatal diagnosis of this disease using restriction enzymes is currently a common laboratory diagnostic procedure (Kan and Dozy 1978).

Since the physiological and biochemical cause of sickle cell anaemia was known before the gene was identified, it is an example of ***forward genetics***. However, in instances when the biological defect is not know then, as described later, procedures such as ***linkage analysis*** are used to find the position of the disease-causing gene on a particular chromosome and the function of its protein product. This procedure is called ***reverse genetics*** and was used, e.g., for mapping and identifying the genetic defect that is responsible for cystic fibrosis (Riordan et al., 1989). Using these approaches, more than 1700 human disease genes have now been identified and there has been an exponential increase in the rate at which single-gene disorders are being discovered.

One of the major concerns of modern molecular biologists is to be able to find the cause of incurable diseases and their treatment. Genetically inherited diseases that are caused by mutations in genes require detailed information about the gene, its exact location in the genome, the nature of the mutation(s) as well as the function of the final product of the gene. Applications of some approaches like linkage mapping, positional cloning and mutation screening of candidate genes have been very successful in identifying and characterizing disease-causing genes.

The completion of the ambitious Human Genome Project and its associated spin off projects have opened up several vistas for biomedical research in Pakistan. Being a developing country Pakistan's major disease burden accrues from communicable, maternal, childhood and nutritional disorders. The major challenge is to make the benefits arising from these projects safe, affordable and accessible. A brief overview of what benefits Pakistan can derive from the revolution in biomedicine follows:

Genetic Testing: An immediate benefit of the completion of the human genome sequencing and its associated projects is the provision of a vast array of genetic tests for patients with inherited disorders. Most genetic tests employ the polymerase chain reaction (PCR) to amplify a short, well-defined part of a DNA strand up to approximately 10,000 bases (Mullis 1990). This strand can be part of a gene, or from intergenic regions. These tests identify whether an individual has a disease causing mutation (or change in DNA sequence) and they have been developed for pre and postnatal analysis of genetic predisposition of affected family members and carriers. Couples can also choose to be tested for this risk before they marry, especially in communities where cousin marriages are rampant and feared diseases such as thalassemia or haemophilia are particularly common.

As described above for sickle cell anaemia, in order to apply genetic testing clinically the change in DNA sequence responsible for any disease needs to be known. The Biomedical and Genetic Engineering Division, KRL, Islamabad, was the first laboratory in Pakistan to identify disease-causing mutations in Pakistani families with deafness (Veske et al. 1996) and blindness (Leutelt et al. 1995). In addition, it was the first laboratory to identify novel mutations in rhodopsin (Bessant et al. 1999) and breast cancer (Khaliq et al. 2000) in Pakistan. Such studies provide the framework on which to base diagnostic tests tailored for our population.

It may be pointed out that not all disease causing mutations in Pakistani patients suffering from inherited disorders have been characterized and much needs to be done in this respect. High throughput technologies for genetic screening such as automated DNA sequencing, microarrays, mass spectrometry and denaturing high performance liquid chromatography (DHPLC; Underhill et al. 1997) should be promoted in a few centres of excellence catering to their specific needs. DHPLC has already been established at the Biomedical and Genetic Engineering Division, Islamabad (Khaliq et al. 2003; Mohyuddin et al. 2006). Centres for pre and post natal diagnosis of inherited disorders should be set up at all major public hospitals and genetic counselling should be offered to patients and families.

As mentioned previously, infectious diseases place a higher burden than inherited disorders on the health care and economy and are the leading cause for infant mortality in Pakistan. For example, according to the World Health Organization (WHO) Pakistan has the sixth highest global burden of tuberculosis with approximately 250,000 new cases being detected annually (WHO Report 2004). The remarkable progress in the sequencing of the genomes of bacteria, viruses, parasites and other infective agents as well as their hosts have opened up rapid, cost effective methods for molecular diagnosis of these diseases such as tuberculosis, viral hepatitis etc. which are now being routinely used in Pakistan, especially in major urban centres. However, Pakistan severely lacks expertise and laboratories that can diagnose hazardous infectious agents such as the avian influenza virus and bunyaviruses that cause the recurring tick borne Crimean-Congo hemorrhagic fever. For aerosol transmitted and other life threatening infectious diseases the WHO recommends that Biosafety Levels 3 or 4 be available (WHO 2000). Such facilities require specialized construction and limited personnel access along with laboratory staff highly trained in the handling of infectious agents.

It is known that certain genes and their products play an important role in the susceptibility or immunity to a given infection. Unravelling the genetic basis for susceptibility to infectious diseases would identify immunological or molecular markers for early disease detection. Diagnosing ailments more precisely will lead to more reliable predictions about the course of a disease and such genetic information will help patients and doctors weigh the risks and benefits of different treatments. The aim should be elucidation of disease mechanisms and development of diagnostic tests and therapeutic targets for evaluation. The emphasis should be on infectious diseases such as tuberculosis, hepatitis,

typhoid fever, rotaviruses and emerging infectious diseases such as avian influenza H5N1 strain virus and hemorrhagic fevers. The Biomedical and Genetic Engineering Division, Islamabad, has started investigations into the genetic basis of mycobacterial disorders (Barreiro et al. 2006; Shaw et al. 1993) in order to determine the role of various genes of the human immune system in susceptibility to tuberculosis in Pakistan.

In addition to the infectious diseases Pakistan is also facing a rising burden of diseases that normally afflict the developed world such as coronary heart disease, diabetes and cancer. Susceptibility to these diseases is attributable to the interaction between several different genes and multiple environmental factors. Research in Pakistan should aim at identifying persons who have a genetic risk using DNA variants specifically associated with the disease. Identification of such molecular markers in susceptible individuals can then help in reducing the incidence of these diseases by preventive measures such as changes in diet and lifestyle. For example, a genetic test can inform a patient with hypertension how damaging that condition is likely to be and doctors treating breast cancer will be able to predict by genetic tests how aggressive a tumour might be.

Another benefit involves the study of genetic basis of patient responses to drugs. Such genetic tests would help in identifying patients who would benefit from drug therapy and those that would suffer adverse drug reactions. This would lead to improvement in drug safety and efficacy and circumvent therapeutic failures (Marsh et al. 2006).

DNA evidence is also proving to be a very useful tool in the arsenal of crime fighting organizations (Jobling and Gill 2004). DNA fingerprinting, as it is commonly referred to, requires matching suspect DNA with biological evidence found at the crime scene. Such physical evidence includes bloodstains, semen, hair, or in cases of mass disasters or terrorist attacks bone and/or tissue samples. DNA matching has also helped in cases of paternity testing, missing persons investigations and in identification of relatives of individuals killed in mass disasters like terrorist bombings or natural disasters like the October 8, 2005 earthquake in Kashmir. The Biomedical and Genetic Engineering Division with its extensive experience in the field of population genetics (Mohyuddin et al. 2001; Underhill et al. 2000) is ideally suited to become a national reference laboratory for DNA matching. Since 2001 this Division has been assisting the law enforcing agencies in cases requiring DNA matching. Of the several novel genetic markers identified by researchers at this Division (Ayub et al. 2000; Kayser et al. 2004), two are recommended for routine forensic analysis of male individuals by the Department of Justice, USA, and other institutions worldwide.

Bioinformatics: The marriage between computers, mathematics and biology gave birth to bioinformatics in which experts try to make sense of the enormous data being generated in the biological sciences such as the complete human genome sequence. Most of this data is stored in public databases that are freely

accessible on the World Wide Web for use by the research community worldwide. In addition, software for analyzing such data is also freely available.

Researchers in Pakistan are already benefiting from this information by using on-line search engines to search, access and download research articles and abstracts from on-line international journals. The complete human genome sequence, and that of other organisms, can be downloaded from websites maintained by the National Centre for Biotechnology Information, National Institutes of Health, Bethesda, USA (GenBank), European Bioinformatics Institute (ENSEMBL) and the DNA Data Bank of Japan (Wheeler et al. 2005). However, in order to benefit from these databases and data retrieval and computational resources Pakistan also needs to develop databases that are specific for its population.

In particular there is an urgent requirement for voluntary national donor registries with the goal of providing centralized information on transplantation antigen (Human Leukocyte Antigen; HLA) phenotypes and making this information easily accessible to the physicians and surgeons who carry out tissue (kidney, bone marrow etc.) transplants. Pioneering research has been carried out on the HLA allele frequency in various Pakistani populations at the Biomedical and Genetic Engineering Division, Islamabad (Mohyuddin et al. 2002) and DNA based HLA typing is routinely carried out for patients referred by the Armed Forces Bone Marrow Transplant Centre, Armed Forces Institute of Urology, Armed Forces Institute of Pathology, Rawalpindi, and Mayo Hospital, Lahore.

In addition, there is a need to establish a national forensic criminal database similar to that of the USA, Federal Bureau of Investigation's Combined DNA Index System (CODIS; http://www.fbi.gov/hq/lab/ codis/program). This database blends forensic science and computer technology into an effective crime-solving tool by using DNA technologies for the identification and evaluation of biological evidence in criminal matters. Characterization of common polymorphisms and CODIS genetic markers in our population would assist the law enforcement agencies in identifying violent criminals and terrorists and bring the nation's crime fighting branches into the 21st century to combat the increasing sophistication of such criminals. Such a database will also complement studies on genetic diversity and allow us to study these individuals from an anthropological point of view.

Another useful database would be an online catalogue of disease causing mutations in Pakistani families on the lines of the Human Gene Mutation Database (HGMD; http://www.hgmd.cf.ac.uk). Such a database would aid local researchers, physicians and genetic counsellors in the development of specific genetic tests for our populations and allow molecular diagnosis of inherited diseases.

An electronic database similar to the India's Traditional Knowledge Digital Library (Jayaraman 2006), that lists traditional herbal medicines and other natural products for human ailments would be of assistance in the development of suitable drug candidates for our local healthcare needs.

Proteomics: While continuing with identification of disease causing mutations and genetic variation among its population more attention should be paid towards the protein products that are responsible for the disease so that they can be identified and targeted for therapeutic intervention. In collaboration with researchers from Europe and USA, researchers at the Biomedical and Genetic Engineering Division have identified protein products that are responsible for blindness (Sohoki et al. 2000) and deafness (Scott et al. 2001) and such biological molecules can be potential targets for interventional therapy.

The complete sequence of human genome revealed one surprising feature, i.e. the low number of genes (20,000-25,000). Only now researchers are unravelling the intricate mechanisms by which such a low number of genes produce the 400,000 or so biologically active protein molecules and pinpointing their functions. An international collaboration in understanding what proteins are present in a cell and how they function is being co-ordinated by the Human Proteome Organisation (HUPO).

Such a large number of protein products are thought to arise by use of alternate splicing during transcription and due to post-translational modification of proteins. Proteomics will identify a surfeit of targets for development of drugs molecules. Mass spectroscopy and microarrays are central to proteomic strategies that help in large-scale identification and characterization of proteins (Steen and Mann. 2004).

Vaccines and Biologicals: To reduce the burden of infectious diseases, such as tuberculosis, basic and applied research should be encouraged. In particular research into development of a vaccine against tuberculosis should be a priority. Tuberculosis is a major cause of morbidity and mortality worldwide and the WHO estimated that 1.9 million deaths resulted from the disease in 2004. The problem that Pakistan faces is becoming more acute with the appearance of multi-drug resistant cases of this disease.

Besides, development of indigenous vaccines for prevention of typhoid, malaria, rheumatic heart disease, hepatitis and rotavirus, which are the leading cause of morbidity and mortality in Pakistan, should be a priority. Local production of vaccine against rabies and anti-snake venom can help in reducing the incident of deaths due to dog and snakebites, respectively.

In addition, Pakistan also needs to develop generic biological drugs in particular recombinant insulin, interferon alpha and erythropoietin, which are widely used in clinical practice. Production of these biological reagents can help in building an infrastructure for large-scale production of other recombinant reagents.

Novel Therapies: Gene therapy involves replacing a defective, non-functional gene with its functional counterpart. Some patients have undergone gene therapy in clinical trials but this experimental treatment has suffered some recent setbacks in human trials Concerns regarding safety of the vectors used to deliver the normal gene to the target tissue have to be addressed before such therapy

can be routinely used (Scharschmidt and Lo 2006). However, its potential in treating single gene disorders cannot be discounted. Pakistan should only invest in this technology with specific targets. Emphasis should be on single gene disorders that are common in Pakistan because of cousin marriages and for which the mutation, leading to the genetic defect in our population, is already known (such as thalassemia, some forms of congenital blindness and non-syndromic deafness).

Another promising area for medical research is microRNA as potential therapeutic targets. Studies on gene expression have revealed the family of small regulatory ribonucleic acid (RNA) molecules. These molecules function to inactivate specific mRNA transcripts and regulate several cellular and developmental processes (Czech 2006). Intriguingly, introduction of exogenous microRNA abolishes molecular pathology in animal models. Their promise as relatively economical therapeutic targets in cancer, diabetes, hepatitis, blindness and other single gene disorders can be investigated in our population (Cheng et al. 2003).

Immunomodulatory therapy using monoclonal antibodies, animal lectins (Ilarregui et al. 2005) etc., for chronic inflammatory disorders, autoimmune diseases and cancer have shown promise in clinical trials. Antagonists to a proinflammatory cytokine, tumour necrosis factor-alpha (TNF-?), have been shown to be effective in the treatment of rheumatoid arthritis (Khanna et al. 2004). The use of such potential treatments for leukaemia, rheumatoid arthritis and other debilitating diseases should be explored in our population. However, clinical trials should be carefully monitored by a regulatory authority in order to prevent mishaps, like the one that recently occurred in the United Kingdom, where the clinical trial of a monoclonal antibody went horribly wrong (Wadman 2006).

Stem Cell: Stem cells have the potential of providing treatment, perhaps even a cure, for debilitating human diseases such as spinal cord injuries, strokes, diabetes, Parkinson's disease. These cells can be obtained from human embryos or adults. Embryonic stem cells are pluripotent, meaning that they have the potential of developing into any cell, tissue or organ in the human body and in that sense offer an unlimited source of replacement parts for humans.

To alleviate the moral concerns of working with embryonic stem cells scientists are also trying to develop stem cells from embryos that have been manipulated to prevent its implantation or development into a foetus. Such basic research on stem cells will also aid in solving the enigma of how a single zygote differentiates into an adult.

Researchers in Pakistan need to explore the possibility of developing stem cell therapy to cure paralysis due to spinal cord injuries, replace damaged heart cells in myocardial infarction or pancreatic islet cells in diabetes. These therapies offer much promise for the countless paraplegic and quadriplegic patients resulting from the recent devastating earthquake in Kashmir. However, there is a

lack of trained specialists in this field and human resources must be developed in order to benefit from these developments.

A major obstacle to stem cell therapy is that, unless they come from a compatible donor, the recipient's immune system may consider them as "foreign" and reject them. Somatic cell nuclear transfer in which cell nuclei from adult sources is transferred into an egg, from which the nucleus has been removed, circumvents this possibility and opens avenues for therapeutic cloning. However, similar techniques can be used for human reproductive cloning, which is forbidden by Islam. Recent incidents of scientific fraud carried out by stem cell researchers in Korea (Wohn and Normile 2006) and the possibility of cancer resulting from such therapies have raised alarm and caused major set backs for this area of research (Burkert et al. 2006).

Nanobiotechnology and Biomedical Engineering: Bioengineering and other diagnostic and therapeutic innovations such as implantable drug dosing devices, telepathology, advances in biomedical imaging and robosurgery offer a variety of therapeutic and diagnostic options. Telemedicine and telepathology would help link experts in industrial and developing countries and aid in dissemination of good medical practices. Robosurgery would be useful in the performance of certain tasks requiring great precision.

Nanobiotechnology involves the development of novel nanodevices and technologies for application in medicine and biology. These devices have at least one dimension of 100 nm or less. Materials such as metals, carbon, DNA and peptides have been used to develop microscopic machines that monitor the human body and correct any problem that may arise. The technology has attracted scientists from such diverse fields as bioengineers, metallurgists, biochemists and molecular biologists. Lipid-based nanoparticles capable of delivering drugs or genes have been developed for use in *in vivo* diagnostics and other devices capable of homing on to an appropriate target, such as a tumour, respond to its recognition by sending a signal, and effect a treatment are in development. Such technology is likely to usher in a new era in medicine, provided their safety in humans is ensured (Nel et al. 2006).

Bioethics: Scientists cannot work in isolation. Discussions between scientists, legislators and theologians on ethics and morality of novel biotechnological methods such as stem cell and gene therapy and the use of genetic tests should be encouraged and the results publicized in order to circumvent the public's fear of the unknown. It is necessary and crucial for all subjects to be fully informed about how their DNA data may be distributed, and to decide with whom they want their data shared. At the same time the researchers and clinicians should make full use of developing biotechnology without losing their empathy.

Pakistan still lacks the basic structure for the application of ethical practices in biomedical research and clinical care. Institutes in a few urban areas have ethics committees but governmental regulatory bodies, and independent bioethical research bodies are lacking. Our approaches to developing a more adequate

ethical framework for much of medical decision-making, whether it involves preventive medicine, clinical practice, or research, constitute another neglected area that requires research input from many different disciplines. The Fogarty International Centre at the National Institutes of Health, USA, funds the Pakistan Bioethics Programme that aims to build expertise in this field.

A completely independent bioethics task force with representative scientists, religious scholars lawyers and bioethicists who can debate the issues, uninhibited by pressures from government or commerce, should assist the legislative branch of government. They should be entrusted the task of forming a national consensus on the use of modern biotechnology and address issues such as privacy concerns of patients, protection against genetic discrimination, use of embryonic stem cells, therapeutic cloning etc. The legislature should frame laws based up on their recommendations.

FUTURE PROSPECTS:

Pakistan can benefit from the revolution in molecular medicine by designing tangible projects with attainable goals and collaborating on some projects with renowned world laboratories to enable exchange of personnel, ideas and transfer of novel technology, especially high throughput screening methodologies and development of novel therapies and nano-biotechnology.

The control of infectious diseases should remain the top priority. This goal can be partly achieved by preventing such diseases with the provision of safe drinking water and improving nutrition and sanitation. However, the development of vaccines or better chemotherapeutic agents must also remain a high priority. Exploring the genetic basis of these diseases may reveal potential drug targets.

The major issue regarding medical research in the future would be concerns on how to finance it and make it cost-effective. In Pakistan the bulk of the research is carried out in the public sectors by universities or government research institutes, and centres. The private sector pharmaceutical industries are not investing in basic medical research in the country. Charitable foundations are generous in donating funds for tertiary care facilities but are not investing in preventive measures, which are invariably more effective. In addition, centralization of decision making and bureaucratic red-tapism cause numerous headaches and delays. Head of research institutes and Departmental Heads should be given the academic and economic freedom to conduct research and held accountable for the public money they use.

A major bottleneck in implementation of this vision would be the lack of qualified researchers and clinicians with proven ability and productivity. This can be partly addressed by initiating linkages and collaborative research projects with renowned foreign laboratories and/or scientists on diseases affecting our population. Collaborative research projects “circumvent any possibility of insensitivity to the religious and cultural norms and minimise the possibility of any exploitation of the gene pool of a developing country by foreign commercial interests”.

The exemplary success of the Biomedical and Genetic Engineering Division, Islamabad, should be used as a model. This laboratory was established in 1992 with the primary aim of locating and sequencing human genes in order to identify the mutations responsible for different inherited diseases in Pakistan and to study the genetic diversity in the different ethnic groups of Pakistan from a clinical and anthropological point of view. The laboratory has been extremely successful in identifying research problems related to the local environment and in obtaining funding for these through international competitive grants and collaborations (Hagen and Carlstedt-Duke 2004). In order to maintain its cutting edge the emphasis has to shift on high throughput screening, gene expression analysis proteomics and eventually drug and vaccine development.

References:

Arber W. (1974). DNA modification and restriction. Progress in Nucleic Acid Res and Mol Biol 14:1-37.

Avery OT, CM McLeod and M McCarty. (1944). Studies on the chemical nature of the substance inducing transformation of pneumocolccal types. J Exp Med 79:137-138.

Ayub Q, A Mohyuddin, R Qamar, K Mazhar, T Zerjal, SQ Mehdi and C Tyler-Smith. (2000). Identification and characterisation of novel human Y-chromosomal microsatellites from sequence database information Nucleic Acids Res 28(2):e8-e14.

Barreiro LB, H Quach, J Krahenbuhl, S Khaliq, A Mohyuddin, SQ Mehdi, B Gicquel, O Neyrolles and L Quintana-Murci. (2006). DC-SIGN interacts with *Mycobacterium leprae* but sequence variation in this lectin is not associated with leprosy in the Pakistani population. Human Immunol 2:102-107.

Berg P, D Baltimore, HW Boyer, SN Cohen, RW Davis, DS Hogness, D Nathans, R Roblin, JD Watson, S Weissman and ND Zinder (1974). Potential biohazards of recombinant DNA molecules. Proc Nat Acad Sci USA 71:2593-2594.

Bessant DAR, S Khaliq, A Hameed, K Anwar, SQ Mehdi, AM Payne and SS Bhattacharya. (1998). A locus for autosomal recessive congenital microphthalmia maps to chromosome 14q32. Am. J. Human Genetics 62:1113-1116.

Bessant DAR, S Khaliq, A Hameed, K Anwar, AM Payne SQ Mehdi and SS Bhattacharya. (1999). Severe autosomal dominant retinitis pigmentosa caused by a novel rhodopsin mutation (Ter349 Glu). Human Mutation 13: 83-87.

Burkert J, NA Wright and MR Alison. (2006). Stem cells and cancer: an intimate relationship. J Pathol. 209:287-297.

Cann HM, C de Toma, L Cazes, M-F Legrand, V Morel, L Piouffre, J Bodmer, WF Bodmer, B Bonne-Tamir, A Cambon-Thomsen, Z Chen, J Chu, C Carcassi, L Contu, R Du, L Excoffier, GB Ferrara, JS Friedlaender, H Groot, D Gurwitz, T

Jenkins, RJ Herrera, X Huang, J Kidd, KK Kidd, A Langaney, AA Lin, SQ Mehdi, P Parham, A Piazza, MP Pistillo, Y Qian, Q Shu, J Xu, S Zhu,JL Weber, HT Greely, MW Feldman, G Thomas, J Dausset and LL Cavalli-Sforza. (2002). A human genome diversity cell line panel. Science 296:261-262.

Cavalli-Sforza LL**.** (1990) How can one study individual variation for three billion nucleotides of the human genome? Am. J. Human Genet. 46: 640-451.

Cavalli-Sforza LL, Wilson AC, Cantor CR, Cook-Deegan RM and King MC. (1991). Call for a worldwide survey of human genetic diversity: a vanishing opportunity for the Human Genome Project. Genomics. 11: 490-491.

Cavalli-Sforza LL. (2005). The Human Genome Diversity Project: past, present and future. Nature Reviews Genetics 6:333-340.

Cheng JC, TB Moore, KM Sakamoto. (2003). RNA interference and human disease. Mol Genet Metab. 80:121-128.

Cohen SN**,** Chang AC, Boyer HW and Helling RB. (1973). Construction of biologically functional bacterial plasmids in vitro. Proc Natl Acad Sci USA. 70(11): 3240-3244.

Czech MP. (2006). MicroRNAs as therapeutic targets. New England Journal of Medicine 354:1194-1195.

Dhand R. (2006). The 'finished' landscape. Nature Supplement S1: 7.

Garrod AE. (1902). The incidence of alcaptonuria: a study in chemical individuality. Lancet. 11: 1616-1620.

Gregory et al**.** (147 authors, 2006). The DNA sequence and biological annotation of human chromosome 1. Nature. 441: 315-321.

Griffith F. (1928). The significance of pneumococcal types. J Hygiene 27:113-159.

Hagen H-E and J Carlstedt-Duke. (2004). Building global networks for human diseases: genes and populations. Nature Medicine 7:665-666.

Herrick JB**. (**1910). Peculiarly elongated and sickle-shaped red blood corpuscles in a case of severe anemia. Arch Intern Med. 6: 517-521.

Hershey AD and M Chase. (1952). hdependent functions of viral proteins and nucleic acids in growth of bacteriophage. J Gen Physiol 36:39-56.

Holley RW, J Apgar, GA Everett, JT Madison, M Marquisee, SH Merrill, JR PEnswick, A Zamir. (1965). Structure of a ribonucleic acid. Science 147:1462-1465.

Ilarregui JM, GA Bianco, MA Toscano and GA Rabinovich. (2005). The coming of age of galectins as immunomodulatory agents: impact of these carbohydrate binding proteins in T cell physiology and chronic inflammatory disorders. Ann Rheum Dis 64 Suppl 4:96-103.

Ingram VM (1956). A specific chemical difference between the globins of normal human and sickle-cell anaemia haemoglobin. Nature 178:792-794.

Ingram VM (1957). Gene mutations in human haemoglobin: the chemical difference between normal and sickle-cell haemoglobin. Nature 180:326-328.

International HapMap Consortium. (2005). A haplotype map of the human genome. Nature 437:1299-1320.

International Human Genome Sequencing Consortium. (2001). Initial sequencing and analysis of the human genome. Nature 409:860-921.

International Human Genome Sequencing Consortium. (2004). Finishing the euchromatic sequence of the human genome. Nature 431:931-945.

International SNP map working Group. (2001). Nature 409:928-933.

Jayaraman KS. (2006). Break with tradition. Nature 442:342-343.

Jobling MA and P Gill. (2004). Encoded evidence: DNA in forensic analysis. Nature Reviews Genetics 5:739-751.

Johnson IS. (1983). Human insulin from recombinant DNA technology. Science 219:632-637.

Kan YW and AM Dozy. (1978). Antenatal diagnosis of sickle-cell anaemia by D.N.A. analysis of amniotic-fluid cells. Lancet. 8096:910-912.

Kayser M, R Kittler, A Erler, M Hedman, AC Lee, A Mohyuddin, SQ Mehdi, Z Rosser, M Stoneking, MA Jobling, A Sajantila and C Tyler-Smith. (2004). A comprehensive survey of human Y-chromosomal microsatellites. Am. J. Human Genetics. 74: 1183-1197.

Khaliq S, A Hameed, T Khaliq, Q Ayub, R Qamar, A Mohyuddin, K Mazhar and SQ Mehdi. (2000). p53 mutations, polymorphisms and haplotypes in Pakistani ethnic groups and breast cancer patients. Genetic Testing 4:23-29.

Khaliq S, A Abid, A Hameed, K Anwar, A Mohyuddin, Z Azmat, SA Shami, M Ismail and SQ Mehdi. (2003). Mutation screening of Pakistani families with congenital eye disorders. Experimental Eye Research 76:343-348.

Khanna D, M McMahon and DE Furst. (2004). Safety of tumour necrosis factor-alpha antagonists. Drug Saf 27:307-324.

Khorana HG, H Buchi, H Ghosh, N Gupta, TM Jacob, H Kossel, R Morgan, SA Narang, E Ohtsuka and RD Wells. (1967). Polynucleotide synthesis and the genetic code. Cold Srpring Harbor Symposium 31:39-49.

Kornberg A and Baker TA. (1992). DNA replication. 2nd edition. W. H. Freeman and Company.

Landsteiner K. (1901). Ueber Agglutinationserscheinungen normalen menschlichen Blutes. Wien Klin Wochenschr. 14: 1132-1134.

Leutelt J, R Oehlmann, F Younus, LI van den Born, JL Weber, MJ Denton, SQ Mehdi and A Gal. (1995). Autosomal recessive retinitis pigmentosa locus maps on chromosome 1q in a large consanguineous family from Pakistan. Clinical Genetics 47:122-124.

Marsh S, DJ Van Booven, HL McLeod. (2006). Global pharmacogenetics: giving the genome to the masses. Pharmacogenomics 7:625-631.

Mohyuddin A, Q Ayub, R Qamar, T Zerjal, A Helgason, SQ Mehdi and C Tyler-Smith. (2001). Y-chromosomal STR haplotypes in Pakistani populations. Forensic Science International 118:141-146.

Mohyuddin A, Q Ayub, S Khaliq, A Mansoor, K Mazhar, S Rehman, and SQ Mehdi. (2002). HLA polymorphism in six ethnic groups from Pakistan. Tissue Antigens 59:492-501.

Mohyuddin A, Q Ayub, PA Underhill, C Tyler-Smith and SQ Mehdi. (2006). Detection of novel Y SNPs provides further insights into Y chromosomal variation in Pakistan. J. Human Genetics 51:375-378.

Morrow JF, Cohen SN, Chang ACY, Boyer HW, Goodman HM and Helling RB. (1974). Replication and transcription of eukaryotic DNA in Escherichia coli. Proc Natl Acad Sci. USA. 71(5): 1743-1747.

Mullis KB. (1990). The unsual origin of the polymerase chain reaction. Sci American 262 (4):56-61.

Nel A, T Xia, L Madler and N Li. (2006). Toxic potential of materials at the nanolevel. Science 311:622-627.

Nirenberg MW. (1963). The genetic code:II. Sci American 190(3):80-94.

Ochoa S. (1963). Synthetic polynucleotides and the genetic code. Fed Proc 22:62-74.

Okazaki R, T Okazaki, K Sakabe, K Sugimoto and A Sugino. (1968). Mechanism of DNA chain-growth.I.Possible discontinuity and unusual secondary structure of newly synthesized chains. Proc Nat Acad Sci USA 59:598-605.

Pauling L, HA Itano, SJ Singer, and IC Wells. (1949). Sickle cell anaemia, a molecular disease. Science 110:543-548.

Riordan JR, Rommens JM, Kerem B, Alon N, Rozmahel R, Grzelczak Z, KielenskiJ, Lok S, Plavsic N, Chou JL, et al. (1989) Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. Science. 245: 1066-1073.

Roberts RJ. (1981). Restriction and modification enzymes and their recognition sequences. Nuc Acids Res 9:r75-r96.

Saiki RK, Scharf S, Faloona F, Mullis KB, Horn GT, Erlich HA and Arnheim N. (1985). Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science. 230(4732): 1350-1354.

Sanger F, GM Air, BG Barrell, NL Brown, AR Coulson, CA Fiddes, CA Hutchison, PM Slocombe and M Smith M. (1977). Nucleotide sequence of bacteriophage phiX174 DNA. Nature 265:687-695.

Scharschmidt T and B Lo. (2006). Clinical trial design issues raised during recombinant DNA advisory committee review of gene transfer protocols. Hum Gene Ther 17:448-454.

Scott HS, J Kudoh, M Wattenhofer, K Shibuya, A Berry, R Chrast, M Guipponi, J Wang, K Kawasaki, S Asakawa, S Minoshima, F Younus, SQ Mehdi, U Radhakrishna, M-P Papasavvas, C Gehrig, C Rossier, M Korostishevsky, A Gal, N Shimizu, B Bonne-Tamir and SE Antonarakis. (2001). Insertion of ?-satellite repeats identifies a transmembrane protease causing both congenital and childhood onset autosomal recessive deafness. Nature Genetics 27:59-63.

Shaw M-A, S Atkinson, H Dockrell, R Hussain, Z Lins-Lainson, J Shaw, F Ramos, F Silveira, SQ Mehdi, F Kaukab, S Khaliq, T Chiang and J Blackwell. (1993). An RFLP map for 2q33-q37 from multicase mycobacterial and leishmanial disease families: no evidence for an *Lsh/Ity/Bcg* gene homologue influencing susceptibility to leprosy. Annals of Human Genetics 57:251-271.

She X, Z Jiang, RA Clark, G Liu, Z Cheng, E Tuzun, DM Church, G Sutton, AL Halpern, and EE Eichler. (2004). Shotgun sequence assembly and recent segmental duplications within the human genome. Nature 431:927-930.

Sohoki MM, SJ Bowne, LS Sullivan, S Blackshaw, CL Cepko, AM Payne, SS Bhattacharya, S. Khaliq, SQ Mehdi, DG Birch, WR Harrison, FFB Elder, JR

Heckenlively and SP Daiger. (2000). Mutations in a new photoreceptor-pineal gene on 17p cause Leber congenital amaurosis. Nature Genetics 24:79-83.

Steen H and M Mann. (2004). The ABC's (and XYZ's) of peptide sequencing. Nature Reviews Molecular Cell Biology 5:699-711.

Underhill PA, L Jin, AA Lin, SQ Mehdi, T Jenkins, D Vollrath, RW Davis, LL Cavalli-Sforza and PJ Oefner. (1997). Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography. Genome Research 7:996-1005.

Underhill PA, P Shen, AA Lin, L Jin, G Passarino, WH Yang, E Kauffman, B Bonne-Tamir, J Bertranpetit, P Francalacci, M Ibrahim, T Jenkins, JR Kidd, SQ Mehdi, MT Seielstad, , RS Wells, A Piazza, RW Davis, MW Feldman, LL Cavalli-Sforza and PJ Oefner. (2000). Y chromosome sequence variation and the history of human populations. Nature Genetics 26:358-361.

Venter, JC et al. The sequence of the human genome. Science 291:1304-1351 (2001).

Veske A, R Oehlmann, F Younus, A Mohyuddin, B Muller-Myhsok, SQ Mehdi and A Gal. (1996). Autosomal recessive non-syndromic deafness locus (DFNB8) maps on chromosome 21q22 in a large consanguineous kindred from Pakistan. Human Molecular Genetics 5:165-168.

Wadman, M. (2006). London's disastrous drug trial has serious side effects for research. Nature 440:388-389.

Watson JD and FHC Crick. (1953). A structure of DNA. Nature 171:737-738.

Wheeler et al. 2005. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res 33:D39-D45.

WHO. (2000). Laboratory Safety Manual, 3rd Edition, WHO, Geneva.

WHO. (2004). WHO Report 2004; Global tuberculosis control. Country Profile:Pakistan. WHO, Geneva.

Wohn DY and D Normile. (2006). Korean cloning scandal. Prosecutors allege elaborate deception and missing funds. Science. 312:980-981.