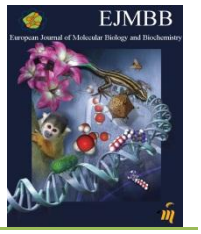
e - ISSN - 2348-2206
Print ISSN - 2348-2192

European Journal of Molecular Biology and Biochemistry

Journal homepage: www.mcmed.us/journal/ejmbb



THE PROMISE OF HUMAN GENOME PROJECT (HGP) IN MEDICINE

Matin Ahmad Khan[1], Binod Kumar[2], Uma Shanker Singh[3], A K Jha[4], Indrajeet Kumar[5], Fatin Raza Khan[6], Shahzeb Ahmad Khan[7], Shekh Arshad Ali[7]

[1]Associate Professor Deptt of Biochemistry, Patliputra Medical College, Dhanbad, Jharkhand, India.
[2]Asst Professor Dept of Pathology, MGM Medical College, Jharkhand, India.
[3]Asst. Professor, Dept of Pathology, MGM Medical College, Jamshedpur, Jharkhand, India.
[4]Associate Professor Dept of Pathology, MGM Medical College Jamshedpur Jharkhand, India.
[5]Lecturer Dept of PSM, MGM Medical College, Jamshedpur, Jharkhand, India.
[6]M.Tech (Bioinformatics) University of Hyderabad, Telangana, India.
[7]Yenepoya Medical College, Deralakatte Mangalore, Karnataka, India.

Article Info
Received 23/11/2016
Revised 16/12/2016
Accepted 29/12/2016

Key words:-
Human Genome Project

ABSTRACT
The worldwide scientific undertaking known as the Human Genome Project studies the genetic make-up of several organisms in addition to humans. The information derived from these studies has applications in the identification, characterization and understanding of genes that affect or cause human diseases. The year 2000 marked both the start of the new millennium and the announcement that the vast majority of the human genome had been sequenced. Much work remains to understand how this "instruction book for human biology" carries out its multitudes of functions. But the consequences for the practice of medicine are likely to be profound. Genetic prediction of individual risks of disease and responsiveness to drugs will reach the medical mainstream in the next decade or so. The development of designer drugs, based on a genomic approach to targeting molecular pathways that are disrupted in disease, will follow soon after. Potential misuses of genetic information, such as discrimination in obtaining health insurance and in the workplace, will need to be dealt with swiftly and effectively. Genomic medicine holds the ultimate promise of revolutionizing the diagnosis and treatment of many illnesses. The following statement which appeared in the January 15, 2001, issue of 'Time' magazine, as **:"***Doctors will treat diseases like diabetes and cancer before the symptoms ever begin, using medications that boost or counteract the effects of individual proteins and they will know right from the start how to select the best medicine to suit each patient".* aptly sums up the public perceptions about the implications of the Human Genome Project in the future practice of medicine.

INTRODUCTION

The history of the human race has been filled with curiosity and discovery about our abilities and limitations.

Corresponding Author

Matin Ahmad Khan
Email: - mak5962@hotmail.com

As an egotistical creature with a seemingly unstoppable desire for new accomplishments, we attempt feats with emotion and tenacity. People worldwide raced to be the first to discover the secrets and the ability of flight. Enormous amounts of monies were spent on sending people into space and the race to land on the moon. With the rapid growth of scientific knowledge and experimental methods, humans have begun to unravel and challenge

another mystery, the discovery of the entire genetic make-up of the human body [1].

This endeavor, the Human Genome Project (HGP), has created hopes and expectations about better health care. It has also brought forth serious social issues. To understand the potential positive and negative issues, we must first understand the history and technical aspects of the HGP [2].

The Human Genome Project (HGP**)** in an international scientific research project with the goal of determining the sequence of nucleotide base pairs that make up human DNA, and of identifying and mapping all of the genes of the human genome from both a physical and a functional standpoint. It remains the world's largest collaborative biological project. After the idea was picked up in 1984 by the US government when the planning started, the project formally launched in 1990 and was declared complete in 2003. Funding came from the US government through the National Institutes of Health (NIH) as well as numerous other groups from around the world. A parallel project was conducted outside of government by the Celera Corporation, or Celera Genomics, which was formally launched in 1998. Most of the government-sponsored sequencing was performed in twenty universities and research centers in the United States, the United Kingdom, Japan, France, Germany, Canada, and China [3,4].

The Human Genome Project originally aimed to map the nucleotides contained in a human haploid reference genome (more than three billion). The "genome" of any given individual is unique; mapping the "human genome" involved sequencing a small number of individuals and then assembling these together to get a complete sequence for each chromosome. The finished human genome is thus a mosaic, not representing any one individual [5].

Genome

A genome is an organism's complete set of deoxyribonucleic acid (DNA), a chemical compound that contains the genetic instructions needed to develop and direct the activities of every organism. DNA molecules are made of two twisting, paired strands. Each strand is made of four chemical units, called nucleotide bases. The bases are adenine (A), thymine (T), guanine (G) and cytosine (C). Bases on opposite strands pair specifically; an A always pairs with a T, and a C always with a G.

The human genome contains approximately 3 billion of these base pairs, which reside in the 23 pairs of chromosomes within the nucleus of all our cells. Each chromosome contains hundreds to thousands of genes, which carry the instructions for making proteins. Each of the estimated 30,000 genes in the human genome makes an average of three proteins [6,7].

History of the Human Genome Project: The HGP has an ultimate goal of identifying and locating the positions of all genes in the human body. A researcher named Renato Dulbecco first suggested the idea of such a project while the U.S. Department of Energy (DOE) was also considering the same project because issues related to radiation and chemical exposure were being raised. Military and civilian populations were being exposed to radiation and possible carcinogenic chemicals through atomic testing, the use of Agent Orange in Vietnam, and possible nuclear power facility accidents. Genetic knowledge was needed to determine the resiliency of the human genome [8-10].

Worldwide discussion about a HGP began in 1985. In 1986, the DOE announced its' Human Genome Initiative which emphasized the development of resources and technologies for genome mapping, sequencing, computation, and infrastructure support that would lead to the entire human genome map (3). United States involvement began in October 1990 and was coordinated by the DOE and the National Institute of Health (NIH). With an estimated cost of 3 billion dollars, sources of funding also include the National Science Foundation (NSF) and the Howard Hughes Medical Institute (HHMI). Because of the involvement of the NIH, DOE, and NSF who receive U.S. Congressional funding, the HGP is partly funded through federal tax dollars (1). Expected to last 15 years, technological advancements have accelerated the expected date of completion to the year 2003. This completion date would coincide with the 50th anniversary of Watson and Crick's description of the structure of DNA molecule [11-13].

The Human Genome Project (HGP) was the international, collaborative research program whose goal was the complete mapping and understanding of all the genes of human beings.

Genome

The HGP was the natural culmination of the history of genetics research. In 1911, Alfred Sturtevant, then an undergraduate researcher in the laboratory of Thomas Hunt Morgan, realized that he could - and had to, in order to manage his data - map the locations of the fruit fly (*Drosophila melanogaster*) genes whose mutations the Morgan laboratory was tracking over generations. Sturtevant's very first gene map can be likened to the Wright brothers' first flight at Kitty Hawk. In turn, the Human Genome Project can be compared to the Apollo program bringing humanity to the moon.

The hereditary material of all multi-cellular organisms is the famous double helix of deoxyribonucleic acid (DNA), which contains all of our genes. DNA, in turn, is made up of four chemical bases, pairs of which form the "rungs" of the twisted, ladder-shaped DNA molecules. All genes are made up of stretches of these four bases, arranged in different ways and in different lengths. HGP researchers have deciphered the human genome in three major ways: determining the order, or "sequence," of all the bases in our genome's DNA; making maps that show

the locations of genes for major sections of all our chromosomes; and producing what are called linkage maps, complex versions of the type originated in early *Drosophila* research, through which inherited traits (such as those for genetic disease) can be tracked over generations [14-17].

The HGP has revealed that there are probably about 20,500 human genes. The completed human sequence can now identify their locations. This ultimate product of the HGP has given the world a resource of detailed information about the structure, organization and function of the complete set of human genes. This information can be thought of as the basic set of inheritable "instructions" for the development and function of a human being.

The International Human Genome Sequencing Consortium published the first draft of the human genome in the journal *Nature* in February 2001 with the sequence of the entire genome's three billion base pairs some 90 percent complete. A startling finding of this first draft was that the number of human genes appeared to be significantly fewer than previous estimates, which ranged from 50,000 genes to as many as 140,000.The full sequence was completed and published in April 2003.

Upon publication of the majority of the genome in February 2001, Francis Collins, the director of NHGRI (National

Human Genome Research Institute) noted that the genome could be thought of in terms of a book with multiple uses: "It's a history book - a narrative of the journey of our species through time. It's a shop manual, with an incredibly detailed blueprint for building every human cell. And it's a transformative textbook of medicine, with insights that will give health care providers immense new powers to treat, prevent and cure disease."

The tools created through the HGP also continue to inform efforts to characterize the entire genomes of several other organisms used extensively in biological research, such as mice, fruit flies and flatworms. These efforts support each other, because most organisms have many similar, or "homologous," genes with similar functions. Therefore, the identification of the sequence or function of a gene in a model organism, for example, the roundworm *C. elegans,* has the potential to explain a homologous gene in human beings, or in one of the other model organisms. These ambitious goals required and will continue to demand a variety of new technologies that have made it possible to relatively rapidly construct a first draft of the human genome and to continue to refine that draft.

These techniques include:

- DNA Sequencing
- The Employment of Restriction Fragment-Length Polymorphisms (RFLP)
- Yeast Artificial Chromosomes (YAC)
- Bacterial Artificial Chromosomes (BAC)
- The Polymerase Chain Reaction (PCR)
- Electrophoresis

Of course, information is only as good as the ability to use it. Therefore, advanced methods for widely disseminating the information generated by the HGP to scientists, physicians and others, is necessary in order to ensure the most rapid application of research results for the benefit of humanity. Biomedical technology and research are particular beneficiaries of the HGP.

However, the momentous implications for individuals and society for possessing the detailed genetic information made possible by the HGP were recognized from the outset. Another major component of the HGP - and an ongoing component of NHGRI - is therefore devoted to the analysis of the ethical, legal and social implications (ELSI) of our newfound genetic knowledge, and the subsequent development of policy options for public consideration.

Human Genome Project Goals

The specific goals of the HGP are to:

- Identify all the estimated 80,000 genes in the human DNA,
- Determine the sequences of the 3 billion DNA chemical bases.
- Store this information in databases,
- Develop tools for data analysis, and
- Address the possible ethical, legal, and social issues (ELSI)

DISCUSSION

Technical Aspects of the HGP

The process of determining the human genome involves first mapping, or characterizing the chromosomes. This is called a physical map. The next step is sequencing, or determining the order of DNA bases on a chromosome. These are genetic maps.

a) Mapping Strategies

To sequence the human genome, maps are needed. Physical maps are a series of overlapping pieces of DNA isolated in bacteria (6). Physical maps are used to describe the DNA's chemical characteristics. Mapping involves dividing the chromosomes into fragments that can be propagated and characterized, and then ordering them to correspond to their respective chromosomal locations . Genetic markers are invaluable for genome mapping. Markers are any inherited physical or molecular characteristics that are different among individuals of a population. An example of a marker includes restriction fragment length polymorphisms (RFLP). RFLPs reflect sequence differences in DNA sites that can be cleaved by restriction enzymes. To be useful in mapping, markers must be polymorphic, or have more than one form among individuals so that they can be detectable in studies. Another marker is Variable Numbers of Tandem Repeats (VNTR), which are small sections of repeating DNA. VNTRs are prevalent in human DNA and can exist in wide

variance of numbers. This variability gives individuals unique VNTR regions. This is the application behind solving crime cases with blood samples. A genetic map shows the relative locations of these specific markers on chromosomes [18-22].

Used in RFLP markers are restriction enzymes. These enzymes recognize short sequences of DNA and cut them at specific sites. Since scientists have characterized hundreds of different restriction enzymes, DNA can be cut into many different fragments. These fragments are the DNA pieces used in physical maps.

Different types of physical maps exist. Low-resolution physical maps include chromosomal (or Cytogenetic) maps that are based on distinctive banding patterns of stained chromosomes. High-resolution physical maps represent sets of DNA fragments that were cut by restriction enzymes and placed in order as previously described.

b) Sequencing Strategies

To sequence DNA, it must be first be amplified, or increased in quantity. Two types of DNA amplifications are cloning and Polymerase Chain Reactions (PCR). Cloning involves the propagation of DNA fragments in a foreign host. Known as recombinant DNA technology, DNA fragments isolated from restriction enzymes are united with a vector and then reproduced along with the vector's cell DNA. Vectors normally used are viruses, bacteria, and yeast cells. Cloning provides an unlimited amount of DNA for experimental study With PCRs, DNA can be amplified hundreds of millions of times in a matter of hours, a task that would have taken days with recombinant DNA technology. PCR is valuable because the reaction is highly specific, easily automated, and capable of amplifying very small amounts of DNA. For these reasons, PCR has had major impacts on clinical medicine, genetic disease diagnosis, forensic science, and evolutionary biology [23-27].

PCR is a process through which a specialized polymerase enzyme synthesizes a complementary strand of DNA to a separate given strand of DNA in a mixture of DNA bases and DNA fragments. The mixture is heated, separating the two strands in a double-stranded DNA molecule. The mixture is then cooled and through the action of the polymerase enzyme, the DNA fragments in the mixture find and bind to their complementary sequences on the now separated strands. The result is two double helix strands from one double helix strand. Repeated heating and cooling cycles in PCR machines amplify the target DNA exponentially. In less than 90 minutes, PCR cycles can amplify DNA by a million fold.

Now that the DNA has been amplified, sequencing can begin. Two basic approaches are Maxam-Gilbert sequencing and Sanger sequencing. Both methods are successful because gel electrophoresis can produce high-resolution separations of DNA molecules. Electrophoresis is the process of using gels with stained DNA and then separating those DNA fragments according to size by the use of electric current through the gel. Even fragments that have only one single different nucleotide can be separated. Almost all of the steps in both of these sequences are now automated.

Maxam-Gilbert sequencing, also called chemical degradation method, cleaves DNA at specific bases using chemicals. The result is different length fragments. A refinement to this method known as multiplex sequencing enables scientists to analyze approximately 40 clones on a single DNA sequencing gel.

Sanger sequencing, also called the chain termination or dideoxy method, uses enzymes to synthesize DNA of varying length in four different reactions, stopping the replication at positions occupied by one of the four bases, and then determining the resulting fragment lengths.

A major goal of the HGP is to develop automated sequencing technology that can accurately sequence more than 100,000 bases per day. Specific focuses include developing sequencing and detection schemes that are faster, more sensitive, accurate, and economical .In 1991, computer technology entered the sequencing process at Oak Ridge National Laboratory where an artificial intelligence program called GRAIL was tested.

HGP in Clinical Medicine

Scientists estimate that chromosomes in the human population differ at about 0.1%. Understanding these differences could lead to discovery of heritable diseases, as well as diseases and other traits that are common to man Information gained from the HGP has already fueled many positive discoveries in health care. Well-publicized successes include the cloning of genes responsible for Duchenne muscular dystrophy, retinoblastoma, cystic fibrosis, and neurofibromatosis. Increasingly detailed genomic maps have also aided researchers seeking genes associated with fragile X syndrome, types of inherited colon cancer, Alzheimer's disease, and familial breast cancer [28-31].

If other disease-related genes are isolated, scientists can begin to understand the structure and pathology of other disorders such as heart disease, cancer, and diabetes. This knowledge would lead to better medical management of these diseases and pharmaceutical discovery. Current and potential applications of genome research will address national needs in molecular medicine, waste control and environmental cleanup, biotechnology, energy sources, and risk assessment.

Much of the benefit derived from genetic tests began long before the HGP got under a way. It was in the 1970s that newborn screening was established using Southern blot analysis; testing for phenylketonuria, for example, avoided detrimental effects to children's development simply with specialized diets [4]. Not long after, in the early 1980s, PCR techniques enabled the development of tests for Cystic Fibrosis, Huntington's disease, and Duchenne muscular dystrophy [5]. The further

characterization of genes through the HGP has aided the development of more genetic tests; however, some of these are for very rare conditions, limiting their widespread benefit to healthcare. The HGP has helped to pioneer new techniques, helping to expand genetic testing, for example with pre-implantation genetic testing [6]. More recently the Lancet has published results of the world's first bedside genetic test, which will be able to identify a particular allele in patients making them more susceptible to the adverse events of clopidogrel [32-35,7].

a) Identifying Disease : Genetic Testing

Much of the benefit derived from genetic tests began long before the HGP got under a way. It was in the 1970s that newborn screening was established using Southern blot analysis; testing for phenylketonuria, for example, avoided detrimental effects to children's development simply with specialized diets [4]. Not long after, in the early 1980s, PCR techniques enabled the development of tests for Cystic Fibrosis, Huntington's disease, and Duchenne muscular dystrophy.[5] The further characterization of genes through the HGP has aided the development of more genetic tests; however, some of these are for very rare conditions, limiting their widespread benefit to healthcare. The HGP has helped to pioneer new techniques, helping to expand genetic testing, for example with pre-implantation genetic testing.[6] More recently the Lancet has published results of the world's first bedside genetic test, which will be able to identify a particular allele in patients making them more susceptible to the adverse events of clopidogrel [36-38].

Genetic testing does give rise to ethical issues that perhaps undermine its overall benefit. Should we be telling patients about their predisposition to a disease with an unpredictable course? The possibility of both type I and type II errors can also be detrimental – for this reason, counselling and informed consent is imperative. There is little doubt that genetic screening in many circumstances has improved clinical care but how much of that is due to the HGP can be debated.

b) Better Classification of Diseases

Understanding of molecular mechanisms causing disease will lead to better classification of disease and better management. For instance the so-called "Essential Hypertension" is not a single entity but a heterogeneous mixed bag with several polygenic quantitative trait loci (QTLs). At the molecular level several mechanisms controlled by different genes have been identified which operate alone or in combination in different patients. Patterns of SNPs in candidate genes for blood pressure homeostasis have been studied. An interesting example is pregnancy hypertension caused by a mutation in the gene encoding mineralocorticoid receptors, which makes the receptors responsive to progesterone as well as to aldosterone. Patients with this mutation will respond to rising progesterone levels in the second trimester of pregnancy with rise in blood pressure, which comes down to normal after delivery with fall in progesterone levels. In future a drop of blood from the hypertensive patient will give the gene expression profile by cDNA micro-array. For instance it may reveal three particular gene polymorphisms relevant to hypertension. It may also reveal other polymorphisms that will predispose the patient to diabetes mellitus or myocardial infarction. The clinicians will then determine the drugs the patient will benefit from, and those which he should stay away from. He can also determine if the hypertension is salt-insensitive which means no need for salt-restriction. This scenario will be vastly different from the current "trial and error" method of matching a patient with a drug - thiazide diuretics, beta-blockers, calcium channel blockers, ACE inhibitors, angiotensin receptor blockers, centrally acting drugs like clonidine and reserpine etc. Most of the common and important diseases such as hypertension, diabetes mellitus, atherosclerotic coronary artery disease, cancer and neuropsychiatric disorders are polygenic with Quantitative Traits. Molecular markers will revolutionize the unravelling of quantitative traits into multiple Quantitative Trait Loci (QTLs), and an elaborate tool box for QTL mapping is now available All these diseases represent quantitative traits that are caused by interactions among genes and between genes and the environment. For instance QTL genes that contribute to elevated lipid levels in the blood may only be expressed if the individual eats a high fat diet- hence identification of such traits allows prevention by diet modification.

Understanding the molecular bases of QTLs remains a challenge. Developments in mapping across several model organisms have made it possible to go from QTL location to candidate genes.

c) Molecular Medicine

Through genetic research, medicine will look more into the fundamental causes of diseases rather than concentrating on treating symptoms. Genetic screening will enable rapid and specific diagnostic tests making it possible to treat countless maladies DNA-based tests clarify diagnosis quickly and enable geneticists to detect carriers within families. Genomic information can indicate the future likelihood of some diseases. As an example, if the gene responsible for Huntington's disease is present, it may be certain that symptoms will eventually occur, although predicting the exact time may not be possible. Other diseases where susceptibility may be determined include heart disease, cancer, and diabetes. Medical researchers will be able to create therapeutic products based on new classes of drugs, immunotherapy techniques, and possible augmentation or replacement of defective genes through gene therapy [39-41].

d) Search for New Drugs Through Proteomics

The availability of the entire DNA sequence of the human genome is only the first step. The genome only contains the recipe for making proteins. It is the 1278

families of proteins which do the cell's work. All cells have the same genome but they differ in which genes are expressed and which proteins are made. Likewise, diseased cells like cancer cells often produce proteins which healthy cells don't and vice versa. Transcriptomics is the study of variations in the expression level of different genes under different environmental conditions. Functional genomics or proteomics is the study of all the proteins made by a given cell, tissue or organism, determining how these diverse proteins join forces to form networks akin to electrical circuits. Determining the precise three dimensional structure of proteins by x-ray crystallography is structural Genomics.Two dimensional gel electrophoresis and mass spectometry have helped to identify which proteins are expressed in selected cells or tissues. In January 2002 Applied Biosystems unveiled its mass spectrometry-based 4700 Proteomics plants using robotic techniques borrowed from the assembly lines of the automobile industry. The effort is to decipher the entire human proteome in 3 years from now. Proteomics will be centre-stage for the drug industry in the coming decade. The Human Proteomic Organization (HUPO) aims to link public proteome projects in the same way as done for HGP human genome project. Like the companies that have automated the protein discovery program, x-ray crystallography has also been automated. Last December the Oxford Glyco Sciences in England filed patent applications for 4000 human proteins, a move that will shake up how intellectual property is defined in biotechnology.

e) Molecular Classification of Cancers

Cancer classification has been based primarily on morphological appearances, which have severe limitations. A new classification is based on global gene expression analysis using DNA microarrays. For instance a class discovery procedure automatically distinguished between acute myeloid leukemia (AML) and acute lymphatic leukaemia (ALL) without previous knowledge of these classes. Genes useful for cancer class prediction may also provide insights into cancer pathogenesis and response to a given chemotherapy. The technique of class prediction will help predict future clinical course and survival, provided that one has a collection of tumour samples for which eventual outcome is known [14].

f) New Drugs for Cancer

Functional genomics and proteomics reveal the differential metabolism related to relevant gene variations. Critical enzymes or proteins or receptors associated with the altered metabolism in cancer cells can be used as targets for new drug development.

The National Cancer Institute and Food & Drug Administration in USA have since July 2001 joined in a separate effort to focus on using proteomics to develop more targeted treatment and more reliable diagnostics for cancer. In this program, researchers will analyze tumour cells from individual cancer patients to come up with a roster of proteins expressed in cancer cells (altered gene expression) but not in normal cells. They will also search for protein markers that correlate with more aggressive cancers, perhaps leading to better and more specific diagnostic tests. A good example is the identification of HER2/ neu receptor expressed on only 25 -30% of breast cancer patients. They respond very well to a specific monoclonal antibody raised against the receptor. FDA has approved both the tests to detect the receptor, and the drug Herceptin (Trastuzumab). This is an example of "personalised medicine"; giving the drug only to those who test positive and not to others, thus avoiding wasteful use of drugs and their avoidable toxicity.

Genotyping or functional enzyme analysis has become standard practice in major cancer treatment centres, such as the Mayo Clinic in USA. Tissue transcription profiling is especially appropriate in cancers with inherent genetic instability, since mRNA can be extracted from biopsies or surgical samples. DNA micro-arrays can be used to measure the expression pattern of thousands of genes in parallel, generating clues to gene function that can help to identify appropriate targets for therapeutic interventions.

g) From Gene to Screen

Technologies such as differential gene expression, in situ hybridization, immune histochemistry and transgenic / knock out animal models are useful in proteomics to identify targets for new drug development. The general paradigm of new drug development is

1. Compare cancer tissue samples with normal tissue to identify over-produced protein in cancer cells not present in normal cells (by 2D gel electrophoresis and mass spectrometry- by charge and mass of the protein).
2. Isolate and crystallize the protein and study its diffraction pattern with x-ray crystallography to reveal the 3D structure of the protein.
3. Medicinal chemists, with the help of combinatorial chemistry libraries, look for candidate molecules for possible fit with the target molecule. High through put screening methods allow screening of thousands of candidate molecules in a very short time.

The challenge for drug discovery scientists is to identify these genes that play critical roles in normal physiology or in the causation of disease and to elucidate their function both biochemically and biologically. Sequence homology with proteins in other species such as fruit fly, yeast etc. and tissue distribution of a novel gene are two critical pieces of the large jigsaw puzzle of its function. A good illustrative example of a prototypic genomic-derived drug discovery target is Cathepsin K.

Surgical specimens of osteoclastoma tumour were used to extract osteoclasts (using osteoclast-specific antibodies attached to magnetic beads). Expression of human osteoclast genes was studied from its mRNA and making a complimentary DNA. It was found that about 4% of the EST (expressed sequence tags) generated encoded a

novel cysteine protease which was named Cathepsin-K. This enzyme is expressed exclusively in osteoclasts but not other bone cells. Immunohistochemistry revealed a polar distribution of the enzyme right at the site of contact between the osteoclast and bone resorption pit. This suggested that inhibition of Cathepsin K could be a strategy to treat osteoporosis. Supportive evidence came from a rare inherited human osteochon drodysplasia-pycnosynostosis. Mutants causative of the disorder were localised to the Cathepsin K gene. Osteoclasts from such individuals demineralized normally but do not degrade the bone protein matrix. Gene knockout mice for the Cathepsin K gene were constructed - these also exhibit an osteopetrosis phenotype and produce osteoclasts with impaired bone resorption activity. The SKB Laboratory determined the 3D crystal structure of Cathepsin K. Experienced medicinal chemists undertook the design and synthesis of an inhibitor. They demonstrated that inhibitors of Cathepsin K inhibit bone resorption both in vitro and in vivo. The gene-to function-to potential drug paradigm was thus demonstrated, in this case, a new treatment for osteoporosis [42].

h) Drug Design : Pharmacogenomics

It was predicted that the HGP would eventually reward clinical medicine with novel approaches to treatment. Pharmacogenomics is the concept that drugs can be tailored to an individual's genetic make-up to increase their efficacy and safety [16]; for example by identifying how cytochrome P450 variants metabolise drugs differently. Patients could also benefit from speedier recoveries without having to pursue different treatment regimens before the best is discovered. In the future, warfarin dosing could become more precise by analyzing a patient's genetic variation in drug metabolism prior to administration, rather than using the currently unpredictable loading schedules. As for current clinical practice, it is now becoming more common to screen patients for thiopurine methyltransferase deficiencies before instigating azathioprine in rheumatology patients and those with inflammatory bowel disease.[19] Patients with HIV can undergo assays to look for viral mutations that are causing resistance to their drugs, and subsequently be switched to a more effective regimen. However, a problem arises if there are no adequate alternatives to prescribe for their condition.

Although the idea of personalized medication sounds exciting, in reality there are many more hurdles to jump before we fully enter this era. Drug companies may not find it profitable to fund the development of drug variants for small populations of patients. In everyday clinical practice, it may also be confusing to have variants of the same drugs, potentially giving rise to further prescribing errors. Finally, the actual decoding of the genome in relation to drug metabolism is not straightforward; there are multiple genes involved with many different polymorphisms to be analysed, which is likely to be time-consuming and costly.

i) Making Effective Therapy Safer

The beneficial effects of many drugs are coupled with the inescapable risk of untoward or adverse drug reactions (ADRs). Most adverse drug reactions can be classified in two groups:

1. Exaggeration of a predicted pharmacological action.
2. Toxic effects unrelated to the intended pharmacological action - these effects are often unpredictable, are frequently severe and results from recognized as well as yet undiscovered mechanisms - these include direct cytotoxicity, initiation of abnormal immune response (e.g. drug-induced SLE) and perturbation of metabolic processes in individuals with genetic enzymatic or receptor defects.

Further, the effects of some drugs can be markedly altered by simultaneous administration of other drugs - such interactions can complicate therapy by adversely increasing or decreasing the desired action of a drug. During hospitalization patients receive as many as 10 different drugs. The sicker the patient, the more drugs he receives, with a corresponding increase in the likelihood of ADRs. The elderly patients as a group have a greater burden of disease, and receive a greater number of medications, hence the greater frequency of ADRs in them, which can be very subtle hence the clinician must be alert to the possibility that their symptoms and signs reflect an ADR [43].

The ADRs often present diagnostic problems because they can involve every organ and system of the body and are frequently mistaken for signs of the underlying disease. For instance, drugs can cause toxic effects which can mimic almost every naturally occurring liver disease in humans. About 2% of all cases of jaundice in hospitalized patients are drug-induced; about a quarter of fulminant hepatic failure are drug-related.

A report released by the Institute of Medicine (IOM) in USA in 1999, entitled "To err is human" estimated that 100000 deaths annually occur in that country due to adverse effects of therapy. Most ADRs are preventable and prior information about inter-individual differences in drug response and individualising the therapy dose to each individual based on his genetic and drug-metabolising profile will make effective drug therapy safer. Within a decade the use of DNA chips for SNP profile of the individual patient will become routine. Affymetrix in Santa Clara, California has developed a tiny chip that can analyze 10,000 to 20,000 genetic markers (SNPs) to probe for 6817 genes, in 15 minutes. Drug companies have an enlightened self-interest in accelerating this process because it will decrease their liability for damages through ADRs (e.g. troglitazone). The general public should also be interested in this development since the cost of ADRs related to morbidity and mortality in the ambulatory setting in USA alone ranges between $ 30- 130 billion.

j) Avoiding Drugs with Side Effects

Troglitazone, a PPARg agonist drug was approved by the FDA in USA but later on withdrawn from the market because of death due to hepatotoxicity in a number of patients. This negative aspect was not detected in the clinical trials before the drug received approval. In future, using proteomics techniques it should be possible to determine prospectively whether an investigational drug prompts production of possibly harmful metabolites or proteins. Comparison of a serum sample (with no drug) and a serum sample after investigational drug with two dimensional gel electrophoresis may show that the drug prompted the production of new proteins - some innocuous and some with potential to cause side effects. Mass spectrometry resolves individual proteins according to the masses of their constituent atoms.

In future medicines will be prescribed to only those patients in whom a high probability of efficacy without significant adverse events is predicted. This approach will change the design of future clinical trials. The future practice of medicine will be based on a different kind of evidence than that used at present [44].

k) Combating disease : Gene Therapy

Identification of genetic defect causing disease may logically suggest gene therapy either by augmentation of desirable but deficient genes, or blocking of harmful genes (through anti-sense oligoribonucleotides or transcription factor decoys, or specific aptamers). Currently the strategy of gene therapy is confined to somatic cells only (and excludes germ cells). Worldwide, more that 300 clinical trials involving over 2000 patients are currently underway, most of which are for the treatment of various cancers. A number of phase 1 and 2 clinical trials have provided the proof of principle of gene therapy. Apart from cancer, other active areas of gene therapy are : cardiovascular (promotion of angiogenesis in ischemic disorders and prevention of angiogenesis in proliferative retinopathy and cancer metastasis); rheumatoid arthritis (to suppress inflammation using IRAP-interleukin 1 receptor antagonist protein, or induce anti-inflammatory cytokines IL-4, IL-10, IL-13, metalloproteinase inhibitors); degenerative arthritis (regenerate catilage by TGFb gene, human bone morphogenetic protein-2 gene transfer to induce bone formation). With over 200 investigational new drug applications currently active, gene therapy research represents one of the fastest growing areas in therapeutic research.

Gene therapy involves replacing disease-causing genes with functional copies [8]. Gene therapy was first initiated in the early 1980s before the HGP was set up. However, techniques were cumbersome and yielded little success. The HGP, with its host of new DNA, certainly accelerated gene therapy and aided advancement in techniques for gene transfer. This still does not mean that technical difficulties have been avoided; the body has a tendency to mount an immune response against new DNA, and there is the risk of viral vectors reverting to their virulent form. In one case, gene therapy trials were halted when two subjects with X-linked severe combined immune deficiency developed leukemia due to the insertion of a transgene next to an oncogene.

As the HGP has aided discovery of genes for rarer conditions, some success stories have arisen, including a gene therapy trial for Leber's congenital amaurosis showing improvement in the sight of subjects.[11] There is also future promise with the recent development of nano-particles carrying tumourdestroying genes[12] and new viral vectors.

China has been the first to approve commercial gene therapy products. 'Gendicine' and 'Oncorine' target the p53 tumour suppressor gene to aid tumour lysis [14]. However, the tumour shrinkage seen has not necessarily translated to prolonged survival of cancer patients. Europe is also nearing the approval its first gene therapy drug 'Glybera' for the small population of patients with a lipoprotein lipase deficiency [15].

The HGP has driven gene therapy trials forward, but this has not yet greatly influenced clinical practice and any successes have only been for rare conditions. The polygenic inheritance of common conditions such as heart disease or diabetes makes the success of gene therapy in the near future seem doubtful due to the complexity of targeting multiple erroneous genes.

l) Ethical, Legal and Social Implications

Early planners of the HGP realized that human genomic mapping and sequencing would have profound implications for individuals, families and our society. Although this information can potentially and dramatically improve human health, it would raise a number of ethical, legal and social issues (ELSI) such as how this information would be interpreted and used, who would have access to it, and how can society prevent harm from improper use of genetic information. To address these issues, the ELSI Program was established as a part of the HGP. ELSI was created so that potential problem areas could be identified and solutions created before genetic information is integrated into modern health care practices (4). This is a unique aspect because the HGP is the first large scientific endeavor to address social issues that may arise from the project The DOE and NIH genome programs each set aside 3-5% of their annual budgets for the study of ELSI [40-42].

There are four major priorities being addressed by ELSI. The first is the issue of privacy and fairness in the use and interpretation of genetic information. As genetic information is being discovered, the risk of genetic discrimination increases as new disease genes are identified. The issue of privacy and confidentiality, including questions of ownership and control of genetic information becomes critical. Fair use of this information for insurance, employment, criminal justice, education,

adoption, and the military is necessary. Also, the impact of genetic information on psychological responses to family relationships and individual stigmatizations becomes an issue.

The second priority for ELSI is the clinical integration of new genetic technologies. It has been questioned if health professionals are adequately educated about genetics, genetic technologies and the implications of their use. Important issues include individual and family counseling and testing, informed consent for individual considering genetic testing, and the use of such genetic test for the use of reproductive risk assessment and making reproductive decisions.

The issues that surround genetic research are the third priority of ELSI. Such issues include the commercialization of the products from human genetic research. Examples are questions of the ownership of tissue and tissue derived products, patents, copyrights, and accessibility of data and materials.

The fourth priority is the education of the general public and health care providers. ELSI funded surveys have revealed that most of the public and health professionals are not knowledgeable about genetics, genetic technologies and the implications of having genetic information. It is essential that the public understands the meaning of genetic information and that the nation's health professionals have the knowledge, skills, and resources to integrate this new knowledge and technologies into diagnosis, prevention, and treatment of diseases [43-44].

CONCLUSIONS

The decoding of the human genome was undoubtedly a major intellectual advance for mankind, however, the practical applications have not been as clear cut. Genetic testing has been well established into clinical practice. With earlier diagnoses aiding disease management, it will continue to have a positive influence as long as we don't succumb to superfluous testing. Gene therapy is showing promise, however, is still in very experimental phases. It is likely that we will soon start to see more gene therapeutics on the market but they may not translate to actual clinical benefit. Pharmacogenomics has already started to be used for management of certain conditions. However, I believe individually tailored medication for all conditions will remain a sci-fi concept for now. As a result, twenty years later we are still waiting for the HGP to impress.

REFERENCES

1. Collins F, Green E, Guttmacher A. (2003). A vision for the future of genomics research. *Nature*, 422, 835–847.
2. McKusick V. (1997). Genomics, structural and functional studies of genomes. *Genomics*, 45, 244–249.
3. Boehm CD. (1989). Use of polymerase chain reaction for diagnosis of inherited disorders. *Clin Chem*, 35, 1843–1848.
4. Collins SL, Impey L. (2012). Prenatal diagnosis, types and techniques. *Early Hum Dev*, 88, 3–8.
5. Roberts JD, Wells GA, Le May MR. (2012). Point-of-care genetic testing for personalisation of anti-platelet treatment (RAPID GENE), a prospective, randomised, proof-of-concept trial. *Lancet*, 379, 1705–1711.
6. Cavazzana-Calvo M, Thrasher A, Mavilio F. (2004). The future of gene therapy. *Nature*, 26, 779–781.
7. Wilson J. (2009). Lessons learned from the gene therapy trial for ornithine transcarbamylase deficiency. *Mol Genet Metab*, 96, 151–157.
8. Hacein-Bey-Abina S, Fischer A, Cavazzana-Calvo M. (2002). Gene therapy of X-linked severe combined immunodeficiency. *Int J Hematol*, 76, 295–298.
9. Bainbridge J, Ali R. (2008). Success in sight, the eyes have it! Ocular gene therapy trials for LCA look promising. *Gene Therapy*, 15, 1191–1192.
10. Chisholm E, Vassaux G, Martin-Duque P. (2009). Cancer-specific transgene expression mediated by systemic injection of nanoparticles. *Cancer Res*, 69, 2655–2662.
11. Schirrmacher V, Fournier P. (2009). Newcastle disease virus, a promising vector for viral therapy, immune therapy, and gene therapy of cancer. *Methods Mol Biol*, 542, 565–605.
12. Shi J, Zheng D. (2009). An update on gene therapy in China. *Curr Opin Mol Ther*, 547–553.
13. Miller N. (2012). Glybera and the future of gene therapy in the European Union. *Nat Rev Drug Discov*, 11, 419.
14. Daudén Tello E. (2006). Pharmacogenetics I. Concept, history, objectives and areas of study. *Actas Dermosifiliogr*, 97, 623–629.
15. Hodgson J, Marshall A. (1998). Pharmacogenomics, will the regulators approve?. *Nature, Biotechnology*, 16, 243–246.
16. Eriksson N, Wadelius M. (2012). Prediction of warfarin dose, why, when and how?. *Pharmacogenomics*, 13, 429–440. [PubMed]
17. Pistoi S. (2002). Facing your genetic destiny, part II. Scientific American.
18. Haubrich R, Demeter L. (2001). International perspectives on antiretroviral resistance Clinical utility of resistance testing, retrospective and prospective data supporting use and current recommendations. *J Acquir Immune Defic Syndr*, S51–S59.
19. Collins FS. (1999). Shattuck Lecture, medical and societal consequences of the Human Genome Project. *N Engl J Med*, 341, 28-37.
20. Henig RM. (2000). *The Monk in the Garden*. New York, NY, Houghton Mifflin.
21. Watson JD, Crick FHC. (1953). Molecular structure of nucleic acids. *Nature*, 171, 737-738.

22. Sanger F, Nicklen S, Coulson AR. (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A*, 74, 5463-5467.
23. Maxam AM, Gilbert W. (1977). A new method for sequencing DNA. *Proc Natl Acad Sci U S A*, 74, 560-564.
24. Botstein D, White RL, Skolnick M, Davis RW. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet*, 32, 314-331.
25. Gusella JF, Wexler NS, Conneally PM, et al. (1983). A polymorphic DNA marker genetically linked to Huntington's disease. *Nature*, 306, 234-238.
26. National Research Council. (2000). Committee on Mapping and Sequencing the Human Genome. *Mapping and Sequencing the Human Genome*. Washington, DC, National Academy Press.
27. US Department of Health and Human Services and Department of Energy. (1990). *Understanding Our Genetic Inheritance. The U.S. Human Genome Project, The First Five Years.* Washington, DC, US Dept of Health and Human Services,.
28. Collins FS, Galas D. (1993). A new five-year plan for the U.S. Human Genome Project. *Science*, 262, 43-46
29. Collins FS, Patrinos A, Jordan E, Chakravarti A, Gesteland R, Walters LR. (1998). New goals for the U.S. Human Genome Project, 1998-2003. *Science*, 282, 682-689
30. Collins FS. (1995). Positional cloning moves from perditional to traditional. *Nat Genet*, 9, 347-350.
31. Hattori M, Fujiyama A, Taylor TD, et al. (2000). The DNA sequence of human chromosome *Nature*, 405, 311-319.
32. Dunham I, Shimizu N, Roe BA, et al. (1999). The DNA sequence of human chromosome. *Nature*, 402, 489-495.
33. Collins FS, Guyer MS, Chakravarti A. (1997). Variations on a theme, cataloging human DNA sequence variation. *Science*, 278, 1580-1581
34. Lockhart DJ, Winzeler EA. (2000). Genomics, gene expression and DNA arrays. *Nature*, 405, 827-836.
35. Bittner M, Meltzer P, Chen Y et al. (2000). Molecular classification of cutaneous malignant melanoma by gene expression profiling. *Nature*, 406, 536-540.
36. Pandey A, Mann M. (2000). Proteomics to study genes and genomes. *Nature*, 405, 837-846.
37. Druker BJ, Lydon NB. (2000). Lessons learned from the development of an abl tyrosine kinase inhibitor for chronic myelogenous leukemia. *J Clin Invest*, 105, 3-7
38. Roses AD. (2000). Pharmacogenetics and the practice of medicine. *Nature*, 405, 857-865
39. Verma IM. (2000). Gene therapy, beyond 2000. *Mol Ther*, 1, 493.
40. Kay MA, Manno CS, Ragni MV et al. (2000). Evidence for gene transfer and expression of factor IX in haemophilia B patients treated with an AAV vector. *Nat Genet*, 24, 257-261.
41. Cavazzana-Calvo M, Hacein-Bey S, de Saint Basile G, et al. (2000). Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease. *Science*, 288, 669-672.
42. Collins FS. (1997). Preparing health professionals for the genetic revolution. *JAMA*, 278, 1285-1286.
43. Hudson KL, Rothenberg KH, Andrews LB, Kahn MJE, Collins FS. (1995). Genetic discrimination and health insurance, an urgent need for reform. *Science*, 270, 391-393.
44. Rothenberg K, Fuller B, Rothstein M et al. (1997). Genetic information and the workplace, legislative approaches and policy challenges. *Science*, 275, 1755-1757.