Defence Science Journal, Vol 51, No 4, October 2001, pp. 341-351

REVIEW PAPER

Human Genome: A Gift of 20th Century and Challenge for 21st Century

M.W. Pandit and Lalji Singh
Centre for Cellular and Molecular Biology, Hyderabad - 500 007

ABSTRACT

Completion of the human genome sequencing has stirred the whole scientific community in many ways. Apart from deciphering the biological meaning of sequence-language written in three billion letters, which may take a century, scientists are essentially confronted with many challenges. It will not only revolutionise the field of genetics in terms of knowing ourselves better but also help us tremendously in identifying genetic diseases and preventing them by predicting and finding cure for them. Integration of medicine and molecular biology is expected to lead to better health care systems by preventing people from becoming sick by following a suggested way of life based on individual genetic makeup. Recently developed technologies are helping scientists to analyse the genome at an unprecedented scale and understand the function of various genes, influence of environmental factors on them and their correlation with human behaviour. Availability of human genome sequence has raised several ethical and moral questions regarding the confidentiality of the sharing of genomic information as well as whether society is ready for this. Real challenge, after knowing functions of all the genes, would be to make use of this knowledge for the welfare of human beings with the aim to enhance the quality of human life.

Keywords: Human genome, gene, medicine, DNA, genetic diseases, genomics, genetic engineering, pharmacogenomics, proteomics, DNA chip technology

1. INTRODUCTION

Scientifically, the year 2000 can be called the Year of Genome in its true sense. The Human Genome Project Consortium has already announced the draft of the human genome[1]. Assembling the data is on the verge of completion. The whole scientific community is amazed at this outstanding achievement. However, it will take some time for euphoria to die down. We will soon be confronted with the challenge of deciphering the information generated by the human genome project in terms of its meaning, its biological significance, and its scientific and social implications. Ethical issues[2] related to the outcome of human genome sequencing is another area of confrontation. The time has now come to look back and understand what this project is meant for and how it is going to shape the future of human beings.

The entire human genome is made up of 23 chromosomes; sequencing DNA distributed over all the 23 chromosomes was a gigantic task. Chromosome 22 was sequenced[3] in 1999. Chromosome 21 soon followed its footsteps and its sequencing was completed[4] by 2000. By this time, the technical snags had virtually disappeared and the project progressed with unprecedented speed. Scientists were simultaneously busy with the completion of other important genomes. *Drosophila's* genome consisting of four chromosomes with 120 million bases was completed[5] by March 2000. *Mycobacterium leprae* sequencing[6] also saw the light of the day in the same year (Fig.1).

Received 03 September 2001

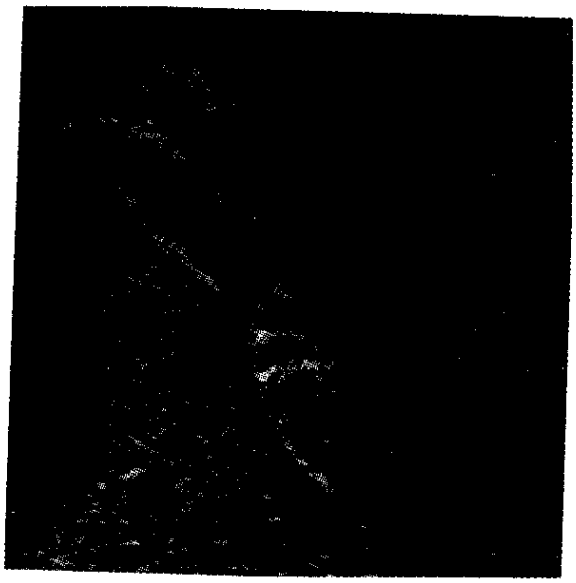


Figure 1. Schematic diagram of a single chromosome

2. HUMAN GENETIC CODE DRAFT

On 26 June 2000, scientists confirmed that the working draft of the human genetic code had been completed. This was a milestone that was supposed to transform the understanding, treatment and prevention of diseases. This was followed by two articles in leading scientific journals[1,7] (Fig. 2) which announced the completion of the draft sequence of the human genome. Participant countries (US, UK, France, Germany, and China) had every reason to be jubilant and reacted to the news with extreme satisfaction. The draft sequence provides the first comprehensive integration of diverse genomic resources. It was the story of three billion odd letters which tells us who we are. In one of the news conferences, Dr Michael Dextar, Director of the Welcome Trust–the organisation which partially funded the British part of the efforts–said that "mapping of human genome has been compared with putting the man on the moon; but I believe it is more than that". Sir Robert May, Chief Scientific Adviser, US Govt, confirming the completion of the project, compared it at par with Charles Darwin's 'Origin of species'. Dr John Toy, Medical Director of Imperial Research Fund in UK, while putting forth the challenges given by the project mentioned, "we have discovered the human alphabet–but now what we have to do is to put the letters in the right order and make sentences; only when all that is done shall we have the book of life to read". At the announcement of the completion of the working draft, Bill Clinton's reaction in his words at White House was "the most wondrous map ever produced by mankind". At the same time, British Prime Minister Tony Blair described the event as "the first technological triumph of the 21st century".

Human Genome Project is a publicly funded international effort involving several agencies, such as the US Dept of Energy, Welcome Trust in UK through its Sanger Centre in Cambridge, White House Institute for Biomedical Research at Mars Institute of Technology, The Washington University School of Medicine in St. Louis, Baylor College of Medicine in Houston, and other groups from France, Germany and Japan. Although formally proposed in 1993, large-scale operation of human genome sequencing began in 1997 with the final goal set for 2005 (it is now expected to be completed by 2003). Simultaneously, a private agency, Celera Genomics, in USA, participated in this venture. 'Speed matters, discovery cannot wait' being their motto, Craig Ventor, President of Celera Genomics, Rockville, Maryland said, "after mapping the genome, science would be confronted with an immense complex job of identifying, characterising and

Figure 2. Two scientific journals, viz., *Science* (16 February 2001) and *Nature* (15 February 2001) which carried the information about the human genome.

understanding thousands of proteins that are made by genes, and that may take a century". Celera's efforts were commendable as within seven months from the beginning, it shared the excitement of announcing the completion of the draft human genome sequencing.

2.1 Draft Sequence & its Accuracy

Human Genome Project was finally unveiled– a remarkable achievement. It is an outline of the information needed to create a human being. The sequence is available on internet. It is an achievement of the coordinated effort involving 20 laboratories from six countries and hundreds of researchers around the world. This draft sequence still has holes and imperfections. Nevertheless, the information provided by the draft sequence will be of interest to many investigators.

Human Genome Project was essentially an international consortium of laboratories–funded primarily by DOE[8] of the US Govt and Welcome Trust[9] –which endeavoured the sequencing of roughly 3.2 billion bases (gigabases) of the human genome, out of which 1.1 per cent to 1.4 per cent actually encodes proteins. At present, about 99 per cent of the sequencing job is complete. The remaining portion is going to be the hardest part of the game which involves the filling of thousands of gaps in the sequence. The questions being raised by other scientists are: How reliable are these going to be, and are they going to be useful for biomedical research? The criteria for complete sequence was set up to be 99.99 per cent, meaning that a sequence of 10,000 base pairs with one gap in between will be taken as completed. Chromosome 22 had only 11 gaps. Chromosome 21 had only 10 gaps. All that will soon be over and we will have a complete human genome sequence.

As the human genome sequencing data will be of utmost importance to the researchers as well as the health care agencies, it was originally proposed that the various regions in the genome will be sequenced 10 times (10 x coverage). However, the Human Genome Project Consortium has sequenced each base 7 times (7 x coverage) and has checked the reproducibility. In Celera sequencing data, each base was sequenced 5.1 times (5.1 x coverage), while plus 2.9 times (plus 2.9 x coverage) was taken from Human Genome Project Consortium data. Therefore, the human genome sequence reflects a fair amount of accuracy.

3. HISTORICAL PERSPECTIVE

It was in 1866 that an Augustinian monk named Gregor Mendel published results of his experiments on peas, which gave birth to the science of genetics. He not only named the genetic unit as gene but also formulated rules of segregation of genetic traits (Fig. 3). Through the experiments on *Drosophila*, Thomas Hunt Morgan in 1910 showed that genes are organised in a linear fashion in chromosomes. However, the real breakthrough came in 1944, when it was proved that the gene is nothing but DNA (and not protein). Subsequently in 1953, Watson and Crick[10] published a seminal paper on the structure of DNA showing that DNA has a helical structure made up of two strands, which they called double helix (Fig. 4).

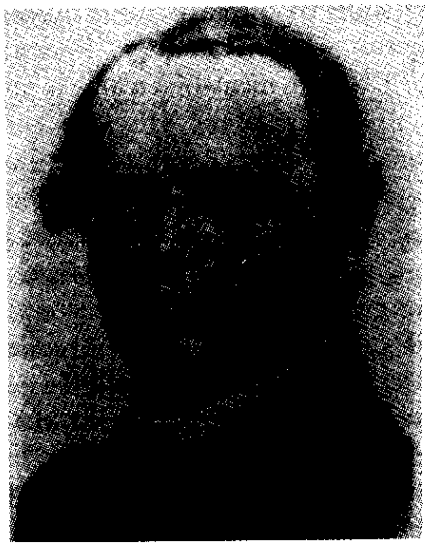


Figure 3. Johann Gregor Mendel (1822-84)

3.1 Importance of Human Genome Sequencing

Today, at the beginning of the century we have a complete sequence of the human genome which has revolutionised the field of genetics to the extent that it would help tremendously to identify inherited diseases, to prevent them and find a cure for them. There is striking correlation between the function of the gene product and

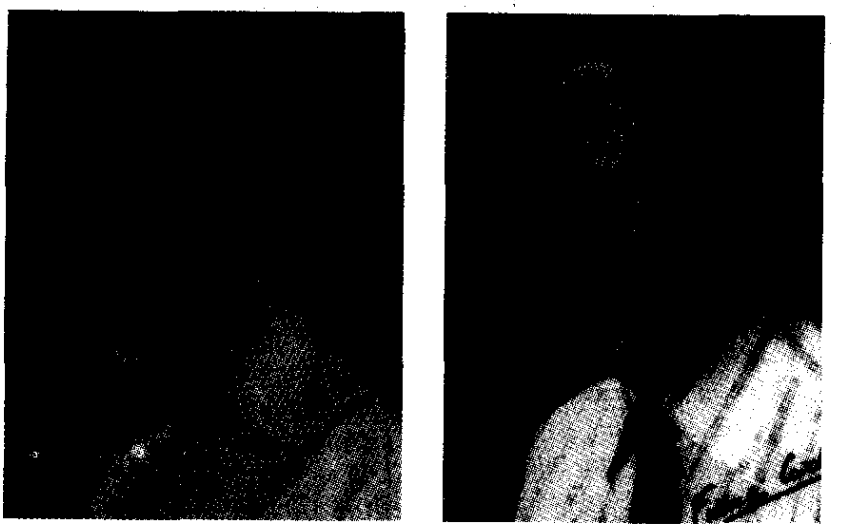


Figure 4. James Watson and Francis Crick won the Nobel Prize in 1962 for Physiology and Medicine along with Maurice Wilkins for solving the structure of DNA.

the features of disease, such as age of onset and mode of inheritance. Functional annotation of the human genome and a comprehensive list of human diseases genes should lead to much greater integration of medicine and biology. Most of the diseases are known to have genetic basis, including some of the infectious diseases, such as AIDS. But why AIDS only? Recently, some population has been identified which is not affected by HIV. Probably genes are also responsible for human behaviour. Practically every month new genes are discovered which have some connection with human welfare. Because of this, the human genome has taken a front stage in human activities so much so that criminals argue that what they have done is not due to their own fault, but it is because of the genes they carry. Would it therefore be essential for judges to look into the genetic makeup of a criminal before giving a judgement? Scientists are always in a dilemma as to what percentage of human characteristics are derived from genetic principles and what percentage is influenced by the environment in which they are brought up[11]. As more and more results of genome projects are coming, it is becoming increasingly apparent that it is a balance between genes and environmental influence which is responsible for making up the human characteristics.

3.2 Sharing of Human Genome

Based on earlier results, scientists had estimated that there are 100,000–150,000 genes in the human genome. However, after the completion of sequencing, they have realised that there are only 26,000–30,000 genes; the total human genome size is about 3.2 billion base pairs. Any two human beings differ in their genome only by 0.1 per cent. About 50 per cent of the human genome is derived from transposable elements. About 50 per cent of the total genome is made up of repetitive DNA. The regions on the human genome which code for proteins are less than 2 per cent. The *Drosophila,* which is used as a model organism by geneticists, has 13,600 genes, i.e. approximately half of the genes in human beings. The human genome is 200 times bigger than the yeast genome but it is 200 times as small as that of amoeba. About 0.5 per cent of human genes are copied into the genome from bacterial sources. Sharing of the human genome with genomes of other organisms in terms of proteomes are as follows:

Yeast	–	46 %
Worms	–	43 %
Drosophila	–	61 %
Chimpanzee	–	99 %

No doubt, the genomic view of our place in nature will be both a source of humility and a blow to the idea of human uniqueness.

3.3 About Human Genome

Our body is made up of living units called cells, their total number being about 10 trillion (10×10^{12}). Every cell contains a nucleus (with the exception of red blood cells and platelets). There are thread-like structures in the nucleus which are known as chromosomes (Fig. 5). There are 46 chromosomes in the nucleus of every cell of a human body, 23 coming from father through sperm and 23 from mother through egg. These chromosomes are made up of nucleic acids and proteins. The nucleic acid is known as DNA(deoxyribonucleic acid). Nucleic acids are made up of nucleotide units. The total length of DNA derived from all chromosomes would be around 2m. This 2m DNA is packaged inside the nucleus, the diameter around of which is barely 10 μm. This is one of the finest ingenuities of nature. Within this limited space, DNA is replicated to be passed on to daughter cell and transcribed into RNA which is translated into protein. It is not surprising that occasionally mistakes are committed during which sequence of bases in the DNA is altered and this results into a genetic disorder causing a disease. When a cell divides, the two strands of DNA get separated and making use of these strands as scaffold, two new DNA strands are made by a rule of complementary base pairing; that means in every new cell there will be one old strand and one new strand of DNA. DNA is essentially made up of a string of four nucleotides (nucleoside triphosphates). The backbone of the double helix of DNA is made of sugar and phosphate of the nucleotides which are held together by four bases of the nucleotides: Adenine (A), guanine (G), cytosine (C), and thymine (T). Adenine of one strand of DNA always pairs with thymine of the other strand and guanine pairs with cytosine. The chemical structure of these bases is such that A and T can never pair with G and C. Thus, when the base pairing takes place with a new strand, adenine complements with thymine and guanine with

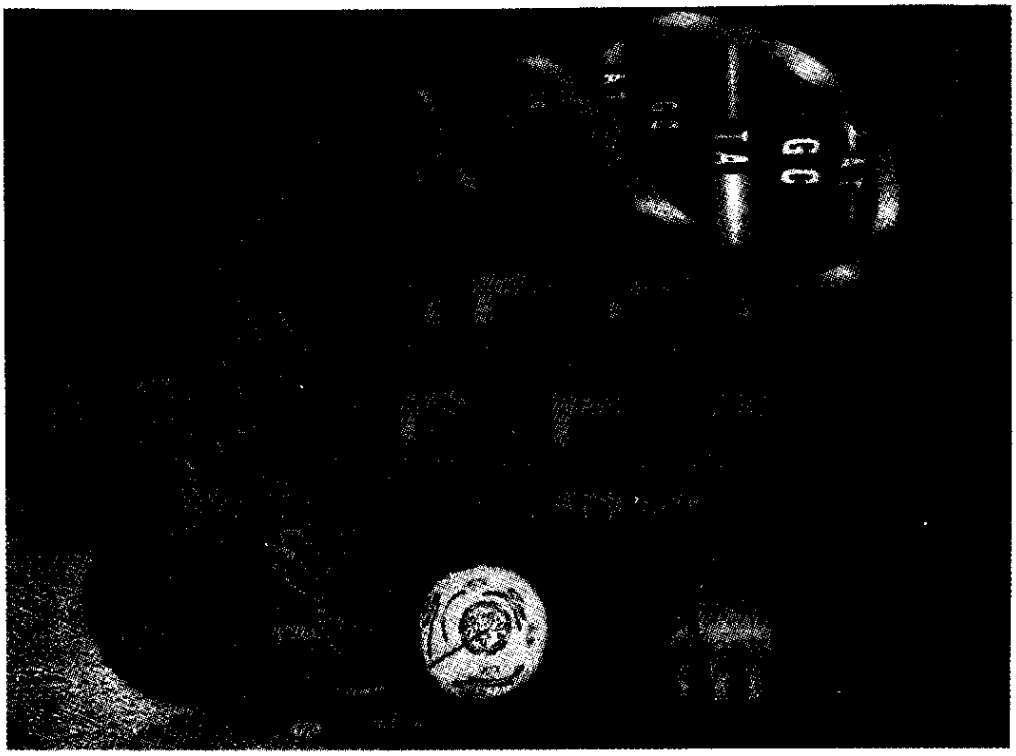


Figure 5. Schematic showing cell from the human body, its nucleus, chromosomes and DNA coding for protein through mRNA as an intermediary.

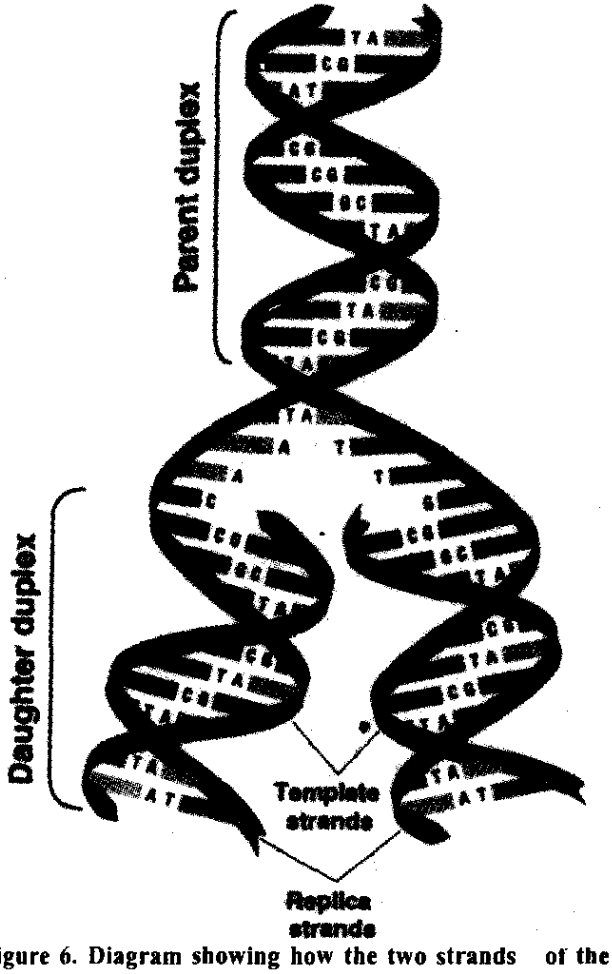


Figure 6. Diagram showing how the two strands of the parent DNA replicate to make two copies of complementary daughter strands.

with cytosine so that the new DNA strand is complementary to the old one. Therefore, if the sequence of bases of one strand of DNA is known, one can very easily work out the sequence of bases of the other (complementary) strand. DNA carries the genetic information from parents to the progeny in the form of the sequence of these nucleotides. The sequence in which these bases are arranged in a piece of DNA determines its function. A small part of this DNA which performs a particular function in the cell is called a gene. DNA first makes RNA (ribonucleic acid) and from that it makes protein which performs certain functions in the cell (in some cases RNA itself is useful to the cell and proteins are not made from such RNAs).

It is remarkable that all living organisms on this planet are made up of only four letters (A, G, C, & T) which are capable of generating so much biodiversity that it is mind-boggling. For example, there are six billion people in the world today, but one cannot find two individuals with identical physical appearance excepting identical twins. Think of the microbial diversity, plant diversity and animal diversity ; they all are the result of combination of only these four letters. This is the beauty of nature's careful selection of genetic material and marvel of its experimentation (simplicity in diversity).

Drosophila genome contains twice as many genes as in yeast. How many genes does the human genome have? Scientists are trying to answer these questions in various ways. Estimates based on the average occurrence of a number of genes in a given length of DNA sequence many a times mislead in arriving at the total number of genes. The analysis based on the expressed sequence tags gave an estimate of about 35,000 genes. Comparative genomics using other reference genomes lead to the estimate of 28,000–34,000 genes. At present, nobody knows the exact number of genes. However, the latest estimates indicate that the human genome probably contains an average of 30,000–35,000 genes[12]. The part of DNA in a cell which constitutes genes coding for proteins is less than 2 per cent of the total DNA. Function of the remaining DNA which is more than 98 per cent is not yet clear to the scientists. Therefore, they sometimes call it as junk DNA. The sequence of 99.9 per cent of DNA is the same in every human being; only 0.1 per cent which varies from one individual to the other is probably the basis of individual differences. Total human genome is made up of 3.2 billion base pairs. Scientists are therefore now in a position to read the complete genetic code. However, it is not as easy as it seems. If one decides to write this genetic code which is made up of various sequences of four letters, it may run into pages equivalent to about 200 telephone directories. It will take decades before one deciphers the biological meaning of the coded language made up of these four letters. Many investigators are engaged day and night in analysing this information. This analysis at the present level will allow them to answer many global questions fairly well but the details will still remain open for some time to come.

4. HUMAN GENOME & FUTURE OF MEDICINE

Human Genome Project Consortium estimated that there are 31,000 protein encoding genes in the human genome of which they can now provide a list of 22,000 genes. Celera Genomics, a company that simultaneously published the draft human genome sequence[7] finds 26,000 genes. There are about 740 identified genes that make the non-protein–coding RNAs in various housekeeping duties, with many more to be found. Yeast sequence has 5,800 genes, while *Drosophila* has about 13,000 genes and plant has 26,000 genes. This raises the question: What gives a human being its complexity, his enormous behavioural patterns, ability for physical coordination and ability to take conscious decisions, response to external stimuli, memory, etc. Understanding of these will be a challenge for the future.

Although genomes of several other organisms have been sequenced (Table 1), the completion of Human Genome Project has brought a new revolution in genetics. For the diagnosis of various diseases, it has become imperative that genetic information should be made available. It is not necessary that a particular medicine would be effective in case of every individual suffering from the same disease or that an effective dose for an individual will be the same in every case. The toxic effect of drugs could vary from individual to individual. Therefore, the nature of medicine in future will be personalised based on genotype. Unfortunately, the present professional training in medical colleges does not prepare the doctors for this new knowledge which is unravelled as a result of modern research in the medicinal field. For example, by examining certain genes it would be possible today, at least in some diseases, to predict that an individual would be affected in future by a particular disease or it would be possible to predict even at a prenatal stage that there is a potential danger of some disease to a particular child. After two to three decades perhaps it would be possible that the DNA of a newborn baby would be able to tell the story of its health in terms of heart disease, cancer, autoimmune, other metabolic disorders, etc. It would be possible to detect every defective gene and to work on these defects by available means through disciplines such as gene therapy. The nature of the medicine in future would be predictive and preventive rather than curative. Depending upon the analysis of the DNA, it would be possible to provide to the individual a list of factors which are dangerous for his health, the diet and lifestyle which he should adopt to avoid diseases in future. It appears that in future many of the things which appear impossible now will be a reality of life tomorrow.

Table 1. Genomic sequences (completed by May 2001)

Organism	Number
Viruses & Viroids	649
Plasmids	219
Organelles	192
Phages	89
Eubacteria	43
Archaea	9
Animals	3
Fungus	1
Plant	1

5. TOOLS/TECHNIQUES FOR GENOME STUDY

Today's challenge is to produce the information infrastructure needed to support the next generation of biomedical research. Some of the tools and techniques used for the study of human genome are described:

5.1 Comparative Genomics

Scientists have already studied the genes of insects and animals like *Drosophila,* rat and nematode in the laboratory. Comparing the information obtained from the genetic studies of these animals and the sequences of their genomes will certainly be useful to understand the genes in human beings as many of these are common.

5.2 Functional Genomics

Analysing the human genome sequence, it would be possible to understand which gene will make which protein, what would be its structure, and what role that protein can play in the human body. New computational tools are being developed and making use of bioinformatics–a new scientific discipline–which brings biology, mathematics,

computers, information science together, will be able to help to assign functions to individual genes using predictive methods with computerised tools.

5.3 Pharmacogenomics

Certain medicines work for a certain disease in one individual but not in another. Some medicines are toxic to one individual but effective in another individual. Why is it so? Some drugs work in a particular group of tribes but not in others. Pharmacogenomics will help to understand these phenomena. The drug testing regimes in future will be based on the genetic background of an individual or a group on which the tests are being carried out. Making use of pharmacogenomics, people are about to enter into a new age of individualised medicine.

5.4 Proteomics

Proteins are the ultimate functional units in the body. All the proteins present in a cell or a tissue are called proteome. It is a dream of scientists to study all the proteins of a cell or tissue at the time of a particular stage of development of the human body. Thus, a new discipline of proteomics is emerging by which thousands of proteins can be studied simultaneously within a short time. Since most functions of a cell in human body are performed by proteins, determining functions of all the proteins present in a cell will be the ultimate objective.

5.5 Micro-array, DNA Chip or Gene Chip

This technology was developed originally in USA by Afametrix, a private company. The technique allows one to study all the genes of an organism at a time by putting them on a few glass slides. Micro-arrays can be used for various purposes, such as to find a new gene, to find genetic basis of diseases or to find whether there is any genetic disorder in a newborn baby, to discover and test new drug targets, to understand differences in gene functions of the old and the young, of the normal and the cancer cells, etc.

There are a few other techniques/tools which are also used for studying human genome; important ones are gene knockout, transgenic animal models, etc. These models will be increasingly used to study the diseases and new drugs before they are recommended for human beings.

6. ROLE OF ENVIRONMENTAL FACTORS

As our knowledge about the development of human character and behaviour is increasing, it is becoming more and more obvious that factors, such as environment and the lifestyle, play a very important role in the development of human beings. Cancer and intelligence are two good examples of these. Scientists have shown that two clones of mouse who had the same genetic makeup when brought up in different environment behaved differently. This proves that the environment plays a very important role in determining characteristics. This means that whatever may be the extent to which one studies the genes of a particular disease, the story will never be complete. It would be essential to understand the balance between genetic factors and environmental factors and how they influence human nature, characteristics, behaviour, etc. To understand all these, the study of human genome will not be enough. The intricate relationship between genetics and environment has to be understood completely.

7. AFTER HUMAN GENOME SEQUENCE

The availability of complete human genome sequence is like entering into a huge library of scattered books where the information is written making use of only four letters. We have to know how the words are made, how to read them, how to make sentences out of them apart from commas, colons, fullstops, etc. To be able to read this information would be the first step towards understanding the human genome. The second step would be to understand the function of all the genes, while at present, the function of hardly 1000 genes is known. It would be a challenge to understand the working of the remaining 26,000–30,000 genes. The third and the most difficult part would be to understand the function of more than 98 per cent of the so called junk DNA. To get significantly more out of the human genome sequence, there needs to be a marked improvement in the ability to turn raw sequence data into biological knowledge. According to Craig Ventor[7], the ultimate use of the sequence will be to explain how our minds have come to organise thoughts sufficiently well to investigate the existence of our own.

The question to be answered would be that if all the human beings are made up of the same genes and all have the same proteins, then why are they so different from one another. The answer would probably lie in this so-called 'junk' DNA and scientists would have to take the help of chemistry, physics, mathematics, computer science and other scientific disciplines. It has become clear that genomes from two human beings most often differ in terms of single base changes, also known as single nucleotide polymorphism (SNP). Individual human beings differ from one another by about one base per thousand. About 1.42 million SNPs have been found[13] in human genome. This nucleotide diversity is a sensitive indicator of biological and historical factors that have affected the human genome. Variations in genome sequences underlie differences in susceptibility to, or protection from, all kinds of diseases, in the age of onset and severity to illness, and in the way our bodies respond to treatment. Today's excitement comes from where these sites of variation are in the genome.

8. SOCIAL & ETHICAL ISSUES

One cannot ignore the fact that availability of human genome sequence has raised more questions[14] than it has given answers. The results of scientific research have social consequences. Social impact of having decoded the human sequence is felt strongly in the area of biomedicine, particularly when physicians perform the DNA testing for asymptomatic individuals. Is it always advisable to tell the patient about the probability of his being affected in future by a specific disease and make his life miserable; or on the other hand, is it ethical to keep this information secret from him? Professional and public response to genetic testing is bound to be restrained because of the reluctance to give out one's genetic makeup and fears about discrimination;be restrained because of the reluctance to give out one's genetic makeup and fears about discrimination; the history of eugenics in the West is not very old and man's potential desire to interfere with procreation is well known. Biomedicine is faced with many challenges, such as ensuring that the patient understands the implications of genetic testing, facing the dilemma that results will be used to deny access to health assurance, using genetic differences as justification for treating people differently, etc.

Our society is based on the premise that people should be responsible for their own actions; if it is so, then, can youth and people with disorders such as insanity or mental retardation be excused from criminal responsibility for acts that harm others? Can a criminal get away by saying that his genes made him do that? In a number of recent cases, criminal defence attorneys have sought to introduce the evidences of family histories of violence asserting their relevance to establishing genetic causes of the defendants' action. How far would such claims be tenable? As a consequence of human genome sequencing, there is another important ethical issue emerging, viz., whether individuals whose DNA is used in the research should have, along with their families, some economic stake in the resulting products. In the post-genomic era, plethora of issues are bound to emerge and one has to learn to deal with these as and when understanding of how the genomics affects the society develops. People can only try, as best as one can, to ensure that this wealth of knowledge is used to benefit the people.

9. CAUTION TO HUMAN BEINGS

On one side, it appears that the study of genetics and the analysis of genome will help tremendously in reducing the agonies of human beings while on the other side one also sees some dangerous pitfalls of this knowledge. Unless the knowledge is used effectively for human welfare, it is of no consequence. Thus scientists, and everybody for that matter, should understand the power of knowledge. Our genome is a heritage of millions of years from our ancestors. We have to hand over this heritage to the next generation; and to do this effectively we should save it from pollution-pitfalls and without disturbing the harmonious balance with nature around us. Many species which existed on the earth at one time have become extinct. We should be careful not to do anything which will be harmful to human species so that it would face the danger of extinction. All this apart, the ultimate challenge to human

intelligence will be to bridge the gulf between a list of sequences and whole organism biology.

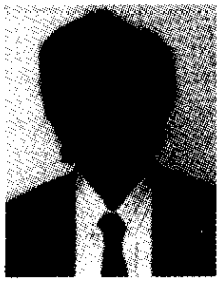
10. FUTURE CHALLENGE

There have been a lot of technological breakthroughs during the human genome sequencing programme which have led to the capability to analyse the genome at a scale which was unthinkable. These emerging technologies involving bioinformatics, comparative genomics, functional genomics, DNA chip, proteomics and structural biology also help in determining function of the genes. After knowing the function of all the genes, one has to make use of this knowledge for the discovery of new drugs by creating transgenic animal models for human diseases. Unfortunately, these technologies are very expensive. It is difficult for a poor country like India to provide these facilities in every university and research institution. Since future medicine is going to be based on individual genotype, in order to provide this facility to everyone, there is a necessity to create many resource centres in the country, network these and bring in peoples' participation (Team India) to make India a genome valley. Seamless collaborations among clinicians, epidemiologists, geneticists, mathematicians, and computer experts will be needed to solve the genetic underpinning of complex diseases that affect the lives of millions.

REFERENCES

1. International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature*, 2001, **409**, 860-921.
2. To know ourselves. The US Department of Energy and The Human Genome Project. July 1996. 32 p.
3. Dunham, I.; Shimizu, N.; Roe, B.A.; Chissoe, S.; Hunt, A.R.; Collins,J.E.; Bruskiewich, R.; Beare, D.M.; Clamp, M.; Smink, L.J. *et al.* The DNA sequence of human chromosome 22. *Nature*, 1999, **402**, 489-595.
4. The Chromosome 21 Mapping and Sequencing Consortium. The DNA sequence of human chromosome 21. *Nature*, 2000, **405**, 311–19.
5. Adams, M.D. *et al.* The genome sequence of *Drosophila melanogaster*. *Science*, 2000, **287**, 2162-195.
6. Hagmann, M. Intimate portraits of bacterial nemeses. *Science*, 2000, **288**, 800-01.
7. Ventor, J.G.; Adams, M.D.; Myers, E.W.; Li P.W.; Mural, R.J.; Sutton, G.G.; Smith H.O.; Yandell, M.; Evans, C.A.; Holt, R.A. *et al.* The sequence of the human genome. *Science*, 2001, **291**, 1304-351.
8. Website of the Dept of Energy, USA. http://www.er.doe.gov
9. Website of the Wellcome Trust, London, UK. http://www.wellcome.ac.uk./
10. Watson, J.D. & Crick, F.H.C. Molecular structure of nucleic acids. A structure of deoxyribose-nucleic acid. *Nature*, 1953, **171**, 737-38.
11. Baker, B.S.; Tylor, B.J. & Hall, J.C. Are complex behaviours specified by dedicated regulatory genes? Reasoning from *Drosophila*. *Cell*, 2001, **105**,13-24.
12. Semple, C.A.M.; Evans, K.L. & Porteous, D.J. Twin peaks : The draft human genome sequence. *Genome Biology*, 2001, **2**, 2003.1-2003.4.
13. The International SNP Map Working Group. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature*, 2001, **409**, 928-33.
14. Clayton, E.W. Through the lens of the sequence. *Genome Research*, 2001, **11**(5), 659-64.

Contributors



Dr MW Pandit was associated with the National Chemical Laboratory, Pune, Indian Institute of Chemical Technology, Hyderabad, and National Aeronautical Laboratory, Bangalore. At present, he is working as Consultant at the Centre for Cellular and Molecular Biology (CCMB), Hyderabad. He helped in establishing the Centre for DNA Fingerprinting and Diagnostics (CDFD), Hyderabad, and the Distributed Information Centre at the CCMB Hyderabad, as part of the Biotechnology Information System. His areas of specialisation include: Interaction of biological macromolecules, computational biology, and bioinformatics. His current interest is in the establishment of the Laboratory for the Conversation of Endangered Species (LaCONES) at Hyderabad. He has published a number of research papers in international journals.

Dr Lalji Singh obtained his PhD in Cytogenetics from the Banaras Hindu University, Varanasi. He is presently the Director of CCMB, Hyderabad. He was awarded the Honorary DSc of Purvanchal University. He is a fellow of the Indian Academy of Sciences, National Academy of Sciences, Indian National Science Academy, AP Academy of Sciences and National Academy of Agricultural Sciences. He has won several awards, including the *Outstanding Forensic Expert Award, Om Prakash Bhasin Award, Vasvik Research Award and CSIR Technology Award* . He is in the editorial board of several national and international journals. His research interests include: Molecular sex determination, DNA fingerprinting, wildlife conservation, ancient DNA studies, and human genome analysis. He isolated a highly conserved sex chromosome-specific satellite DNA, Bkm, from the Indian-branded krait which is considered to be a major breakthrough in the understanding of the molecular basis of sex determination. Dr Lalji Singh and his colleagues have developed a special probe for DNA fingerprinting, which is being extensively used for forensic investigations, paternity determination and seed stock verification. The work of Dr Lalji Singh and his group on genetic variability in Asian tiger and lion population by DNA fingerprinting has been of great use for the conservation breeding programmes. This work has led the Govt of India to decide on setting up LaCONES in Hyderabad.