Indian J Physiol Pharmacol 2000; 44(2) : 121–124

Guest Editorial

Impact of Human Genome Project on Biological Sciences

Genes were once hypothetical entities which transmitted characters from the parents to the progeny. Today, a gene is a segment of deoxyribonucleic acid (DNA) on a chromosome. But how specific can we get on that point ? Happily, amazingly so for several genes. For example, the 'address' of the gene for insulin-like growth factor II (IGF-II) is 11p11. What does this 'address' mean ? It means that the gene for IGF-II is located on chromosome number 11, on its *p* arm (that is the short arm), region 1 and band 1. Regions and bands on the chromosome are detected by using different staining methods. Further, the part of chromosome number 11 coding for IGF-II has 2,066-bp nucleotides; we now know the complete sequence of the four bases - adenine (A), thymine (T), guanine (G) and cytosine (C)-for the nucleotide sequence of IGF-II at 11p11. We hope to be that specific about every single gene by the year 2005, thanks to the Human Genome Project (HGP).

HGP is a multinational three billion-dollar venture launched in 1990 by USA. It will require the participation of several countries around the world, including India. HGP is said to be the most ambitious scientific project ever undertaken in the human history (1). The goals of the HGP are 'to complete a detailed human genetic map, acquire the genome on clones, determine the complete sequence and find all genes' (2). The entire sequence of human genome has been estimated to be three thousand megabases long.

Although experts in this field believe that HGP has been progressing rather slowly, it appears that this project will be completed by year 2005 (3). With this achievement within reach in the near future and with the recent completion of two sequencing projects-one of the *Saccharomyces cerevisiae* and the other of the *Caenorhabditis elegans*, one of the most frequently asked questions is what the sequence data can tell us about human biology and disease.

It appears that DNA sequence data from a single organism by itself reveals very little, unless we also understand the meaning and nuances of genomic informatics. This point is best illustrated by 'junk' DNA. A substantial (90% or more) amount of DNA in human cells has no clearly known purpose. This DNA is termed 'junk' DNA. If more than 90% of human genomic sequence is 'junk', what is the significance of deciphering it! It may seem more reasonable to identify and sequence only the functional genes. True, except that the notion of 'junk' DNA is only provisional. The truth is that we are not absolutely sure of all sense and nonsense held in the genetic code. Some sequences may appear 'nonsense' for the simple reason that we are ignorant about their functions. We have reasons to believe that genes cannot and do not function without interacting with noncoding DNA and adjacent genes (4). It is anticipated that we shall be able to examine relatively easily what different sections and segments of the human genome actually do when the HGP is complete. This is termed 'functional genomics'. It is quite likely that our ability to read the entire human genome will prompt us to understand the functions of all segments of DNA. This, in turn, can influence our approach to human disease. For example, targeted gene manipulation can correct some of the genetic disorders (5). Additionally, the knowledge of genomic sequences of different peptide hormones and other proteins can be utilized for developing pharmacological products for the treatment of some genetic disorders. Also, it may help to find new genes of high clinical relevance; discoveries of the fragile-X gene, the breast cancer gene and Alzheimer's genes support this possibility (3).

Although the potentiality of HGP appears enormous, the prediction that through HGP 'half of the total knowledge of the human organism will be available in five to seven years, and all of it by the end of the decade' (6) is seductively naive and dangerously reductionistic in nature. Let us amplify this point with some examples.

When the sex-determining gene on the Y-chromosome (*Sry*-gene) was injected into fertilized oocytes, only five out of ninety-three zygotes that grew to birth were transgenic (7). Of the five, two were normal males with XY-chromosomes, with no extra masculine features attributable to the extra-dose of *Sry*-gene. The other two were normal XX-females, with however, many copies of *Sry*-genes. Only one was an XX-female with male anatomy and behaviour; it had smaller testicles and was sterile, but with typical male mating behaviour. It is doubtful if we will be ever able to understand the biological basis of such epigenetic and genetic influence and resistance in cells, based only on gene mapping and sequencing.

Will it be possible to identify genetic predisposition based on gene mapping and sequencing? Identification of such predisposition will rest upon identification of 'normal', which in many cases may not be straightforward and simple. Let us take an example. More than four hundred variants of hemoglobin have been reported, and about half of them are 'normal'. Which one of these should be taken as the reference normal in order to predict genetic predisposition towards abnormality (8)?

Again, gene behaviour in a whole organism is often very complex, non-linear

and individualistic. Take sickle cell anaemia for example. Individuals having the sickle cell gene mutation can display an astonishing range of clinical profiles ranging from early childhood mortality to remaining unrecognised till very advanced age (9). Obviously, only sequence data cannot explain the remarkable variations in clinical picture. It is only a modern myth that DNA contains all the information necessary to shape an organism (10, 11).

If we were to extend the study of human genome sequences to large populations, will it be possible to identify genes that affect penetrance and expression of particular hereditary conditions ? Can the genome tell us why certain disease alleles are expressed in higher frequency in certain populations, or which are the really meaningful genetic differences between different human populations. Clearly, mere gene sequence data will not tell anything about these questions ? We need to study each phenotypic variation and to correlate it with genotypic variation.

Thus the inherent complexities of biology cannot be simplified in one stroke by mapping and sequencing of genes. Studies in other biological sciences, such as cell biology, developmental biology, anatomy, biochemistry and physiology will continue to contribute to improved understanding of gene function and other aspects of life processes. In other words, gene sequences may provide us hints about a gene's function based on similarities with genes of known functions; nevertheless, experimental confirmation will be required, which may involve genetic or biochemical studies. Some believe that comparative genomics is the key to understanding the functional aspects of the human genome (12). In this, we need to look at similar genes in different organisms and to determine how function and position changed over the course of evolution. Furthermore, it is true that, of the estimated 10,000 genes in human cells, only a small fraction is involved in the causation of human disease. Furthermore, the major human health problems in today's world are not fundamentally genetic in nature.

These examples and issues, however, are not meant to underestimate the tremendous potential of the HGP in any way. But along with the genomic sequence, it is important to understand the functional aspects of genetic information in the context of the whole organism. Otherwise, 'having sequenced the genome we may be in the position of a non-musician faced with the score of Wagner's Ring cycle: information making no sense at all, but in fact containing an amazing tale – if we only knew what it meant' (13).

The most important aspects of the HGP in modern biology are several-fold. First, the difference between the capability of reading gene sequence – nucleotide by nucleotide, versus not being able to do so is a fundamental one in biological sciences. To this effect, HGP is a heroic expedition. Secondly, the technological improvements that have been occurring in the process of working on the HGP have unquestionably enhanced progress in related fields. This has resulted in a general boost to modern biological research. Thirdly, our ability to read human genome will now compel us to confront the serious issues of functional genetics, which might otherwise have been

delayed or ignored. Finally, and perhaps most importantly, with the progress of HGP, there has been a new way of viewing any biological problem and issue from the angle of genomics. Such paradigm shift in biological sciences will positively influence its growth and advancement for quite some time to come. Moreover, it is possible that alternative explanations for previous observations may emerge under the influence of new groundwork of genomic conceptions and information (14). Thus, the fallout of HGP appears to be significant in the progression of modern biological sciences. There has been some resentment among a group of biological scientists and philosophers about the HGP; it generally occurs during the phase of any paradigm shift (14). It does however appear, as stated by Dr. Francis Collins, the director of the National Centre for Human Genome Research, that the HGP is 'the most important and the most significant project that human kind has ever mounted (15).

ARUNDHATI
Department of Zoology
Hans Raj College
Delhi University, Delhi 110 007

REFERENCES

1. Bijlani RL, Manchanda SK. *The Human Machine*. New Delhi: National Book Trust, 1999.
2. British Medical Association. *Our Genetic Future*. Oxford: Oxford University Press, 1992.
3. Interviews on Genomics. *BioEssays* 1999; 21: 175–182.
4. McCarthy MM. Molecular approaches to control of reproductive behaviour and sexual differentiation of brain in rodents. In: Ferraris JD, Palumbi SR (Editors). *Molecular Zoology - Advances, Strategies and Protocols*. New York: Wiley-Liss, 1996: 387–400.
5. Noble SA, Babiss LE. The potential of antisense oligonucleotides in therapy. In: Rapley R, Walker MR (Editors). *Molecular Diagnostics*. Oxford: Blackwell Scientific Pub, 1993: 431–449.
6. Gilbert W. Towards a paradigm shift in biology. *Nature* 1991; 349: 99.
7. Koopman P et al. Male development of chromosomally female mice transgenic for *Sry*. *Nature* 1991; 351: 117–121.
8. Sarker S and Tauber A. Fallacious claims for HGP. *Nature* 1991; 353: 691.
9. Lander ES, Schork NJ. Genetic dissection of complex traits. *Science* 1994; 265: 2037–2038.
10. Pritchard D. Modern myth. *Nature* 1991; 354: 170.
11. Hubbard R, Wald E. *Exploding the Gene Myth*. Boston: Beacon Press, 1993.
12. Clark MS. Comparative genomics: the key to understanding the Human Genome Project. *BioEssays* 1999; 21: 121–130.
13. Jones JS. Songs in the key of life. *Nature* 1991; 354: 323.
14. Kuhn TS. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press, 1970.
15. Kotala G. Unlocking the secrets of the genome. *New York Times* 1993; November 30: C1.