Know Your Chromosomes

6. But Why?

Vani Brahmachari

Vani Brahmachari is at the Developmental Biology and Genetics Laboratory at Indian Institute of Science. She is interested in understanding factors other than DNA sequence *per se*, that seem to influence genetic inheritance. She utilizes human genetic disorders and genetically weird insect systems to understand this phenomenon.

Previous articles of this series were:

1. Nature's way of packing genes, January 1996.
2. The strong holds of family trees, March 1996.
3. Hybrid cells and human genetics, June 1996.
4. The paths to disorder are many, October 1996.
5. The uniqueness of sex chromosomes, March 1997.

In this article, the last of the series, the implications of understanding our genome in its totality are examined. The concerns raised seem as important as the gains promised.

Apart from their relevance to health care, the quest to understand genes – their nature, function, organisation and location on chromosomes can be driven by various levels of curiosity. In the study of human biology one of the major motivations is disease management and cure. There have been several milestones in medicine like surgery by which the affected organ can be removed, the use of vaccines to protect against infectious diseases, and the use of antibiotics to cure infectious diseases. The new revolution that is in sight is the use of genes as therapeutic agents. Understanding the molecular basis of diseases is the first step in the forthcoming revolution. But it is important to remember that in terms of public health care, medical history records the establishment of proper sanitation system as a major milestone which protects people from infectious diseases. This perhaps continues to be of relevance in the Indian context.

From an overview of medical history, it is apparent that the easiest way to get to the basis of a disease is through an observable difference between a patient and a normal individual in extractable body fluids like urine, blood or saliva. However it is essential that the observed difference is not coincidental but is indeed a consequence of the disease and is seen in most patients affected by it. Fruitful examples of this approach are diabetes, hemophilia, and sickle cell anemia. But it is important to note that this understanding has often only helped clinicians to manage the disease but not to cure it. The story of sickle cell anemia is particularly fascinating.

The Trail Began with Altered Shape of RBCs.

In terms of public health care, medical history records the establishment of proper sanitation system as a major milestone which protects people from infectious diseases.

The first description of sickle cell anemia was made by a physician, James Herrick, at a hospital in Chicago in 1910. The patient reported persistent cough and fever and complained of fatigue and shortness of breath. Herrick failed to find any signs of infection when the urine and sputum of the patient were repeatedly examined, but he observed the peculiar shape of the red blood cells. The basic symptoms were anemia, fever and pain at joints. At that time it puzzled the medical community and the disease was simply called Herrick's anemia. The disease was common in Western Africa and was given names that were basically a description of the symptoms. People themselves had observed that the disease runs in families and that whenever one of the parents was suffering from it, the proportion of his/her children afflicted was higher. By 1949, studying several African families, the recessive mode of inheritance of sickle cell anemia was established. The infancy of genetics at that stage did not promise a fruitful path to the genetic basis of the disease.

At that time the protein involved in carrying oxygen within the red blood cells (RBC) namely haemoglobin had been fairly well studied. When normal haemoglobin and the one from sickle cell anemia patients were analysed, a difference showed up in the oxygen carrying capacity. In blood from patients, in the absence of oxygen, the protein molecules which are otherwise soluble came together, aggregated and resulted in a change in the shape of the RBC. This further resulted in blocking of the capillaries in addition to oxygen deficiency. Early studies had also shown that parents of the sickle cell anemia patients showed sickling of RBCs but did not suffer from the disease. They were *carriers* of the trait and turned out to be heterozygous for the mutated gene. Sickness manifests only when both copies of the gene are mutated, that is, in homozygotes. The mutation was traced to a change in the aminoacid sequence from glutamic acid to valine which could be traced to a single base change, CAG to CGG in the β-globin gene. Sickle cell anemia was the first single gene

Individuals carrying the sickle cell trait are also known to be resistant to malaria.

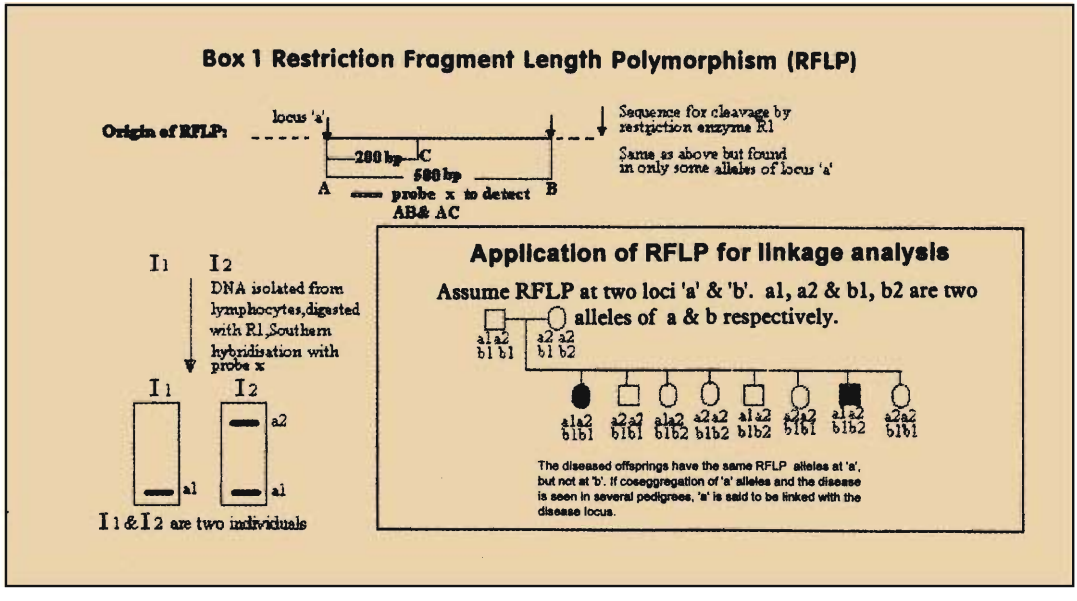
disorder described not involving a morphologically aberrant chromosome. Individuals carrying the sickle cell trait are also known to be resistant to malaria. It is interesting that several people in West Africa have the sickle cell trait and this region is infested with malarial parasites and mosquitoes. The resistance to malaria in carriers of the sickle cell trait is a consequence of the aggregation of proteins resulting in a decrease of pH, creating an acidic environment within the RBC unsuitable for the multiplication of the malarial parasite. Presently we understand several aspects of the β-globin gene and the protein. Hence, the sickle cell anemia is one of the diseases where a correction at the genetic level has been contemplated. Unfortunately the intricacies of globin gene expression have slowed down the progress in a cure for globin related diseases like sickle cell anemia and thalassemia.

From Gene to Protein

In diseases where there is no obvious clue to the defective protein or proteins, the search for the gene gets more challenging. This is well illustrated in the search for the cystic fibrosis (CF) gene. This is a recessive disease that affects 1/2500 Caucasians. The disorder results in accumulation of thick mucus in different ducts, the finer bronchi and trachea in the lungs. It is a fatal disease. Patients are prone to repeated infections, impairment of the pancreas and respiratory failure. The disease was described in 1938, but the clue for the accumulation of thick mucus in CF patients came only in 1983. Abnormalities relating to the secretory ducts of surface tissues and sweat glands were identified; sweat and saliva of CF patients had a higher chloride content than usual.

Unlike sickle cell anemia, in the case of CF the search started with an attempt to find abnormalities at the DNA level since the nature of the defective protein was not known. This endeavour is often compared with the task of finding an address when the only information available is the country. But around the 1980s

several molecular genetic tools were developed which provided a systematic trail leading to the needle in a haystack, so to speak. One such tool is based on single base pair differences in frequently occurring stretches of DNA containing di or tri nucleotide repeats, described as microsatellite DNA. Because of the single base pair change the stretch of DNA fails to serve as a substrate for DNA cutting enzymes that recognise specific sequence of bases (restriction endonucleases). This further results in a difference in the length of the DNA fragment generated after treatment with the given restriction enzyme and is described as *restriction fragment length polymorphism* (RFLP, *Box 1*). An approach similar to this showed that CF patients within a family have a particular pattern for a polymorphic locus. By *in situ* hybridization (Know Your Chromosomes, *Resonance*, Vol.1, No.6, 1996) this RFLP locus was assigned to Chromosome 7. However this was only associated with CF but was not the cause for it. It was like narrowing the search for an address from a country to a city. Using molecular genetic tools the gene responsible for CF was located and was shown to carry various kinds of mutations in CF patients like deletions of DNA sequences and change of basepair (point mutations) in CF The most frequent change recorded is a deletion of 3 base pairs.



Having located and isolated the gene, the amino acid sequence was derived from the DNA sequence and the protein was recognised as a transmembrane protein (a protein residing in the plasma membrane) that is involved in transport of chloride ions; it was called CFTR (cystic fibrosis transmembrane regulator). This is an example of a *reverse genetic approach* (positional cloning); the trail begins at the DNA level to pindown the defective protein. It is rather interesting that at least two CF patients themselves were part of the research groups which identified CFTR.

Where Does this Lead Us?

Of the innumerable insights that genetic research provides for understanding nature's design, let us consider the implications that are of immediate concern for human health.

The identification of a molecular basis for genetic disorders has two prominent implications. One concerns the ability to detect a defective gene as early in pregnancy as possible so that the birth of genetically impaired individuals is prevented. This is akin to a surgical approach in management of diseases like cancer or appendicitis where the affected organ is removed to cure the disease. Termination of pregnancy however involves many more ethical issues than say organ surgery. The other, a more challenging prospect is the ability to correct the defective gene known as *Gene therapy*. This is a conceptually new approach to treat human diseases.

Prenatal Diagnosis

This involves the detection of defects at the DNA level in the first trimester of pregnancy using foetal cells from amniotic fluid or chorionic villi. The technology demands little perturbation of the foetus itself and highly sensitive methods to assess changes at single base pair level in extremely small amounts of DNA. In order to eliminate any damage to the

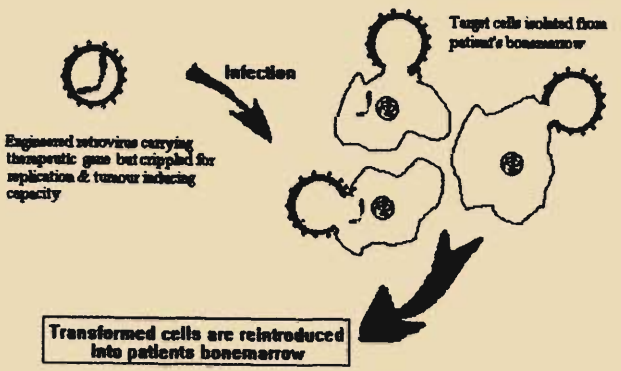
foetus, foetal cells in the maternal blood stream are used in prenatal diagnosis. The ability to carry out prenatal diagnosis is a large step in genetic counselling which is normally done based on family history and health of parents. The reliability of detection can be enhanced to a great extent by using DNA-based diagnostics. Presently it is possible to detect several genetic diseases by this approach. Early detection is vital not only to identify the congenital genetic disorders but also in the case of infectious diseases and cancer. There is a continuous effort to design extremely sensitive methods of detection based on antibodies and DNA.

Tackling the Disease at the Root

Gene therapy is a new concept in medicine where one attempts to correct a defect at the gene itself. The ability to carry out such corrections is intimately related to the nature of the tissue or the organ affected by the disease and the number of genes that cause the disease. The most suitable candidates are the diseases related to blood cells caused by defects in single genes. This is because the blood cells are continuously replenished by cells made in the bone marrow. The approach is to take out the bone marrow cells from the patient, introduce a healthy gene into it, and make sure that the gene persists within the cell as an integral part of the chromosome or by itself and is capable of making the functionally active protein. The corrected cells are reintroduced into the patient's bone marrow. In this treatment, care is taken to destroy the defect carrying cells within the patient's bone marrow before introducing the corrected cells (*Box 2*). This approach has been already used for SCID (severe combined immuno deficiency) syndrome, which is an immunological disorder caused by defects in the enzyme adenosine deaminase (ADA). A drawback in this approach is that the bone marrow cells after differentiation into red blood cells or lymphocytes survive only for a limited time period and therefore repeated treatment at regular intervals is required. The present aim is to manipulate and correct what are called the stem cells so that a

Gene therapy is a new concept in medicine where one attempts to correct a defect at the gene itself. The ability to carry out such corrections is intimately related to the nature of the tissue or the organ affected by the disease and the number of genes that cause the disease.

BOX 2
Gene Therapy



Retroviruses are most widely used vehicles for therapeutic genes. These are RNA viruses that normally cause cancer. They are preferred vehicles as they can infect a wide variety of cells. These recombinant viruses are not able to multiply without the assistance of an intact virus.

life long production of the therapeutic protein is ensured.

A further challenge has been to maintain a sustained production of reasonable amounts of protein from the introduced gene and to ensure that there is no activation of inappropriate genes in that tissue. There also needs to be an appropriate regulation of the genes that are manipulated. One can perceive how essential it is to understand the gene in totality in terms of its function, regulation, location, primary sequence as well as its interaction with neighbours and other gene products in the cell to make gene therapy effective.

It is essential to understand the gene in totality in terms of its function, regulation, location, primary sequence as well as its interaction with neighbours and other gene products in the cell to make gene therapy effective.

The approach outlined above is referred to as *ex vivo*, as the cells are manipulated outside the patient's body and there is an attempt to directly introduce the corrective gene into the affected tissue. This is under trial for cystic fibrosis. A list of diseases being treated in clinical trials of gene therapy is given in *Table 1*.

Table 1 Clinical trials of gene therapy.

Disease	Target cells/tissue
SCID	Bone marrow cells or T lymphocytes
Cystic fibrosis	lung cells
Hemophilia	liver cells or fibroblasts
Cancer	renal cells, brain, breast, colon (depends on nature of affected tissue)
Gaucher disease (lysosomal storage disorder)	brain cells
Familial hyper cholesterolemia	Systemic gene therapy
Chronic granulomatous disease	Bone marrow or leukocytes
Hunter syndrome	Bone marrow cells or lymphoblastoid cells
Fanconi anemia	Bone marrow cells

The list given is not exhaustive. Gene therapy is also being used to treat rheumatoid arthritis and peripheral vascular disease .

Challenges to Medical Genetics

The diseases that are difficult to understand are those that seem to have a genetic basis but occur only sporadically, not strictly conforming to the dictum of 'running in the family'. These are generally recessive disorders caused by mutation in several genes and are termed as polygenic disorders. The rate of mutations being as low as 10^{-6} per locus per generation, it is understandable that a polygenic disease, where mutations in several genes have to accumulate before a disease phenotype is seen, appears sporadic. This is true of coronary diseases and certain cancers. But genetic tools can help in assessing the risk factor or the genetic predisposition of an individual to certain diseases. This would simply mean that, of the several genes involved in the disease, if the person carries mutations in one or two of the genes, the chances of getting the disease increases. In addition, environmental factors can induce mutations. Polygenic disorders are particularly difficult to tackle in terms of gene therapy as well as in prenatal diagnosis.

A polygenic disease where mutations in several genes have to accumulate before a disease phenotype is seen, appears sporadic. Polygenic disorders are particularly difficult to tackle in terms of gene therapy as well as in prenatal diagnosis.

A well founded concern is the evolution of yet another criteria for discriminating humans on the basis of an acquired biological disposition over which the individuals themselves have very little control.

Yet another class of diseases that pose challenges to therapy are those that are caused by mutations in a single gene but that affect different organ systems and tissue types. Cystic fibrosis and the fragile X syndrome are examples. It is impossible to correct the gene in every affected tissue. Therefore novel approaches in gene therapy are required. An essential step in developing these approaches is to gain greater understanding of our genetic and biochemical endowment.

Future Direction

Over the last 7-8 years a world wide intiative has been taken up to map and decipher the complete DNA sequence of the human genome. This endeavour termed as the *human genome project* attracted considerable criticism regarding the cost and time involved. But presently it is well on its way and most scientific targets of the project are being met before the stipulated time. The development of technologies to achieve this massive task of sequencing three billion base pairs is stunning. However, it also throws up academic challenges not only to make sense of this massive data generated but also to the management of the data in a 'user-friendly' manner. In some ways it is similar to creating

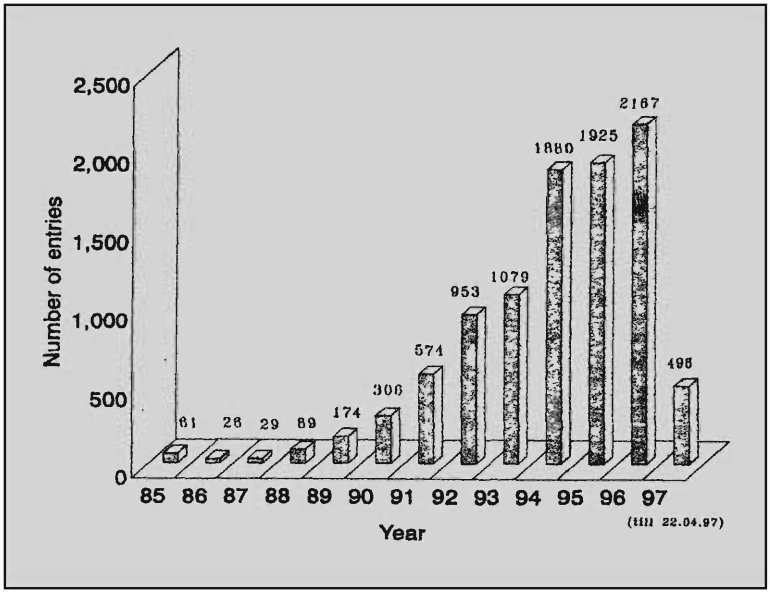


Figure 1 The total number of mutations entered in the human genetic mutation data base starting from 1985 are shown. Note the significant increase after 1989.

genomic DNA libraries of organisms so that an interested investigator can simply bypass the effort of sequencing genes of interest but utilize the available information for further research. One of the obvious changes seen is the rate at which mutations relating to genetic disorders are being discovered (*Figure 1*).

Increasing knowledge of gene sequences accompanied by the ability to predict predisposition to diseases and to manipulate genes in humans has raised numerous ethical issues. A well founded concern is the evolution of yet another criteria for discriminating humans on the basis of an acquired biological disposition over which the individuals themselves have very little control.

Suggested Reading

- Inder M Verma. Gene Therapy. *Scientific American*. November . pp.34-41, 1990.
- Lois Wingerson. *Mapping Our Genes: The genome project and the future of Medicine* . Penguin Group, 1991.
- Necia G Cooper Edt. *The Human Genome Project; deciphering the blueprint of heredity* . University Science Books. California, 1994.
- W French Anderson .Gene Therapy. *Scientific American*. September pp.96-99, 1995.

Address for Correspondence
Vani Brahmachari
Developmental Biology and Genetics Laboratory
Indian Institute of Science
Bangalore 560 012, India