appeared in 2003

Tracing routes in the genetic maze

Decoding of the genome of the hepatitis virus the world over has been making news, says S.Ananthanarayanan.

Nearer home, a team in Hyderabad had decoded the genome of 'Indian Hepatitis C' some months ago. This is thought to be a step towards developing cures and preventives for the liver disease that this virus causes.

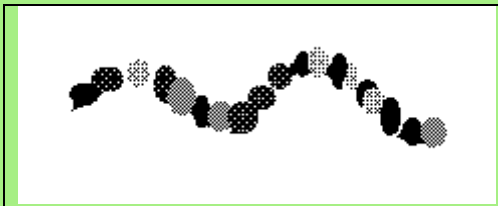
The Genetic Code

This is a compact and incredibly complex *mantra* that defines the nature and function of every cell in an organism. The code is packed inside a structure called the nucleus of the cell and consists, in fact, of a set of complex, chain molecules, some billions of atoms long. But all these billions of atoms are in a precise sequence and, in groups of a few tens, to a few thousand atoms, they help define the specific enzymes and proteins that the cells will produce.

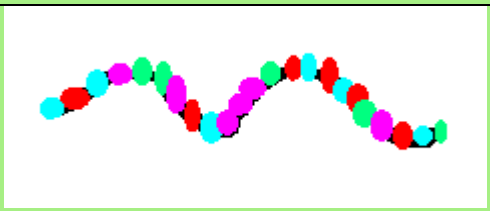
The specific enzymes, in turn, define the way various parts of the body will develop. What colour the hair or eyes of a person would be, for instance. Or whether the person's body would produce enough insulin or whether the person would be a diabetic. At a more basic level, it is the program locked inside the cells that decides whether the organism would be an amoeba, a silkworm or a beauty queen!

The structure

The structure of the molecules that pack this genetic information has been found to be a long chain of alternating sugar and phosphate molecules, with short branches sticking out from the sugar portion. A remarkable thing is that these branches all consist of one of just four clumps of atoms, Adenine, Guanine, Thymine or Cystine. The genetic molecule is thus like a long string, threaded with beads of four colours.

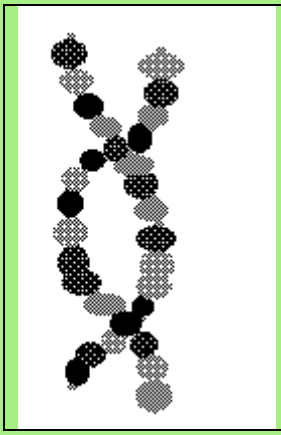


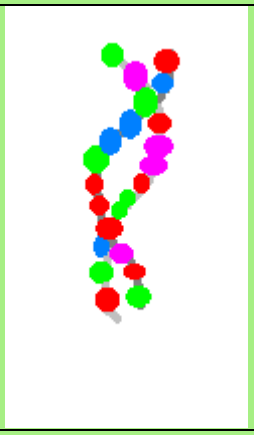
DNA molecule



DNA molecule

To get into a little more detail, the molecule is, in fact, two such chains, coiled around one another, like a spiral, with chemical bonds between the Adenine(A), Guanine(G), Thymine(T) or Cystine(C) groups. It is also found that the bonding is always between A and T or Between G and C. Thus, once the structure of one of the arms of the spiral is known, the structure of the other arm is clearly defined, as having a T or a C where the first has A or G, or vice versa.





Enables replication

This ‘complementary’ nature of the bonding in this giant molecule, which is called deoxyribonucleic acid, or DNA, makes it possible for the two arms of the spiral to split and then form again as a new pair of spirals. In cell division, the two arms separate and one part goes into each half of the divided cell. Each ‘half spiral’ now has unpaired stubs of A,T,G or C sticking out. In the ‘chemical soup’ of the cell, stray T, A C or G groups attach to the stubs and in a short time, the missing ‘better half’ gets generated. Here is the ‘self regulating’ mechanism that makes sure the genetic library of a cell is transferred faithfully from one cell generation to the next.

The code proper

It is found that shorter or longer sequences of A, T, G, C, along the length of the DNA molecule, specify amino acids, or the components of proteins. Cracking the genetic code consists of using a whole toolkit of chemical procedures to identify the actual sequence of a real DNA molecule. Knowing even bits and pieces is often useful in devising ‘genetic engineering’ interventions for diseases.