

# **Worms Have as Many Genes as We Do? But They Lack Alu Sequences**

***Israeli researchers identified the molecular mechanism which assists in the explanation of one of the great paradoxes of evolution: How is it possible that man performs such complex functions with a relatively low number of genes?***

By M. Sullivan  
Trans. Mical Raz

The DNA is constantly changing, and these changes are responsible for the immense diversity in nature, the creation of new species and the survival of existing ones. The principle cellular mechanism which contributes the nature's diversity is the creation of new genes. As far as is known today, this is done by means of the duplication of existing genes. When a gene is duplicated it forms two copies: the original copy which retains its primary function, whilst the new copy may undergo changes endowing it with new functions. The new gene may reside in the DNA for a prolonged period without having any effect whatsoever. If, however, at some stage environmental changes occur in which that gene is advantageous, the gene becomes essential for survival.

According to this model, the more genes one has, the greater the number of functions; as we ascend the evolutionary ladder, the number of genes increases. However, the elucidation of the human genome proved that this is not so. Prior to the mapping of the human genome, it was estimated that the human genome consists of approximately 100,000 genes; while today it is known that the number of human genes is less than 30,000. In contrast, parasitical intestinal worms, which have no head and do not perform any complex functions, have 18,000 genes.

The discovery that man has a relatively low number of genes raised the question of how the human complexity is formed. How were cognitive, communicative, memory, learning and other functions developed in a system with so few genes? What is the distinct mechanism, developed in primates, which enables this, and how less than 30,000 genes regulate the function of the 90,000 proteins which comprise and regulate the human body.

A research group from the School of Medicine in the Tel Aviv University, led by Dr. Gil Ast, discovered a new mechanism through which new proteins are formed on the basis of existing genes. The research, recently published in the journal "Science", provides answers to questions regarding human evolution, which previously had remained unanswered.

Biologists find it difficult to explain the fact that only one and a half percent of the human genome contains information to form proteins. In contrast, most of the genome is formed of insignificant sequences, which are repeated hundreds and thousands of times, without encoding proteins. Why does the body retain the excess baggage and expend so much energy in its maintenance? For years, no answer to this question was found, and the surplus genetic baggage was deemed “junk DNA”.

Amongst the reiterated sequences in the junk DNA, there are short sequences, comprising 11% of the entire human genome. These sequences are called Alu sequences. Approximately 1.4 million copies of them are found in the genome, all almost entirely identical, and existing only in humans and monkeys.

The discovery of the connection between Alu sequences and disease manifestation led to a change in researcher's perception of Alu sequences. They are no longer considered “junk DNA” and researchers are beginning to realize that they may effect organism development. Yet until now, no one knew how this function was performed. The Tel Aviv University research team succeeded in elucidating the mechanism responsible for this, and demonstrated that the Alu sequences in fact have the ability to create anew. They are responsible for creating new proteins in that they “shove” themselves into the genes and create new information without harming the existing genetic information. As these sequences are distinctive to primates, this new information probably contributed to some of our human properties.

The Alu sequences are used in a mechanism existing in the cell, which differentiates between the insignificant and significant genetic material. In each cell there are areas in the DNA which are informative, and areas which are meaningless. The genes contain the instructions for protein building, and RNA molecules facilitate the transition between the two. RNA molecules form an exact copy of the genetic code contained in DNA, and transfer it to the ribosomes, the protein-producing factories. Prior to the transfer to the ribosomes, the RNA undergoes modification. In a special process termed “splicing”, inconsequential areas are cut out and removed, and the cut edges are adhered. Thus, only the significant information in the cell is expressed.

Throughout the evolution of the human race, Alu sequences penetrated the insignificant areas in the DNA, and so became part of the “junk DNA”. Ast and his research team discovered that over time, the Alu sequences acquired mutations, and some of these mutations transformed these sequences from Alu sequences, to specific sequences, identified by the splicing system as significant. These significant sequences, as previously explained, are

not removed by the splicing mechanism, and the information they contain is translated into proteins.

Although Ast and his research team discovered that while Alu sequences are familiar to the splicing system, only in half of these incidences, the system deals with the sequences as “coding”, and adjoins them to existing genes, and so a new protein is formed. In the other half of the cases, the splicing system cuts out the Alu sequences, and removes them together with other non-coding areas. Thus, in the first case, an Alu sequence together with coding areas creates new information which leads to the production of a new protein; in the second case, the existing information remains unchanged in the genome.

“We have discovered a mechanism by which Alu sequences lead to the formation of new proteins, whilst ensuring not to harm the existing proteins”, says Ast. “If not for this mechanism, and the existing information was not retained, the original proteins would be changed”. This situation could lead to the development of disease”. Indeed, the researchers discovered three diseases caused in this fashion.

In the mechanism described by these researchers, there are two products from the same gene. The original copy retains its previous function, while the new copy, to which the Alu sequence was added, can obtain functions which previously did not exist. “According to our findings, some of the genes can instruct to form two proteins which enables the use of a relatively small number of genes to form many proteins”, says Ast. As this distinct mechanism is present only in monkeys and in humans, it explains how the human race can perform such complex functions with so few genes”.

The processes elucidated by Ast and his research team provide an explanation to the question why the cell does not rid itself of the immense surplus DNA in its genome. Amongst these surplus genes is the spare DNA, which makes the wheels of evolution go round.

In an adjunct article in the journal “Science”, Wojciech Makalowski, of the Foundation for Evolutional Molecular Genetics at the University of Pennsylvania, writes: “If we shall run the risk of personification of biological processes, we can say that evolution is too smart to waste valuable information. Thus, repeated DNA sequences should not be called ‘junk DNA’. Rather they should be called the genome’s ‘playground’, as they are a reservoir of pre-formed sequences, used for nature’s evolutionary experiences.”