Manana Arabuli
Grigol Robakidze University

"EXCESS" GENOME AND ITS BIOLOGICAL FUNCTION

The last decades of the past century was distinguished for a remarkable progress in the fields of science and technology. Major achievements were made in the field of biology and medicine. Development of modern research methodologies in the field of molecular biology enables scientists to revise traditional views about the field and rethink this field of science. One of the important topics that was carefully studied during this period was our understanding of living systems genomic composition.

Genome is an entirety of organism's hereditary information at a cell level. Genetic information is encoded in DNA as a nucleic acid sequence. DNA is one of the most important bio-molecule that is often referred as to hold "information about the secret of life" [1] as it defines all features of the living organism. Genome determines not only **what** proteins should be synthesized, but also **where** and **when** the synthesis should occur.

The term "Genome" was introduced by the German botanist Hans Winkler in 1920 to refer to the haploid chromosome set. However, as the field of molecular biology evolved, meaning of the term "genome" changed [2]. Genetic information was found not only in haploid chromosomes, but also in Extrachromosomal DNA, such as plasmids of bacteria or mitochondrial DNA in eukaryotes.

Genomes are mostly represented as DNA though in some viruses the function is served by the RNA [2]. DNA molecules are of great size and contain information about multiple types of protein that synthesize mRNA – the process called transcription. There are also genes that only define transcription of structural RNA that do not synthesize proteins. Collection of all these genes is called coding region/stand of DNA.

Coding strands of DNA (genes) are separated with **promotion** regions. These are nucleotide sequences placed closely, remotely or directly within the gene and control expression (transcription of mRNA) of genes, although, they do not contain coding information and are referred as non-coding region on DNA. So, structure of Genome contains **obligatory and facultative elements.** Obligatory elements are genes and promotion regions, while facultative regions contain multiple types of DNA and RNA number of topography of which would vary by cells and by organisms [2].

In 1990, for the purpose of studying human genomics, the international scientific-research project "Human Genome Project" headed by the James Watson was launched in the USA [3]. The project identified DNA sequence in all human chromosomes and mitochondrial DNA. The full structure of genome was published in 2003, although, the analysis of some segments is still ongoing. Findings of this study are widely used in biomedical research throughout the world.

As we know, nucleus of human genome contains 23 chromosomal pairs and mitochondrial DNA that represent 5% of the whole genome. Altogether, this is 3.1 billion pairs of nitric base. Human genome was 20 to 25 thousand genes, which is much less than it was expected in the beginning of the project. The expected number of genes in the beginning was over 100 thousand. One of the findings of the project was that coding elements in the human cell constituted only 1,5% of the whole genome and over 98% of DNA were non-coding, the so called "excess" nucleotide sequences that do not contain information about proteins and functional RNA. Except of promotion regions, facultative DNA contains segments the function of which is still unknown. 47% of these elements are mobile genetic elements, the so called "transposine" that often change their location [4]. It is considered that 1% of genome is represented by retroviruses.

Genome is not unchangeable. Volume of genetic information varies in somatic cells and gametes even in one organism. Furthermore, somatic cells might "lose" genetic information that they have inherited from gamete and segments of DNA might amplify (replicate/double), or change. Sex of the organism also plays an important role in definition of genome, as sex chromosomes also differ in size and volume of information.

Size of genome significantly varies among different species, although, there is no correlation observed between size of the genome and complexity of the organism. Prokaryotes have the smallest genome size. Prokaryotic DNA is circular and minuscule. It contains from 10^6 to 10^7 nucleotide pairs and has a capacity to code from 1000 to 4000 genes [1]. Non-coding segments are very small and represent from 5-20% of the genome.

DNA in eukaryotic cells is comparatively big and non-coding segments represent large part of the genome. These segments do not participate in protein synthesis, nor performs any other function. Compared to these simple organisms, human genome has 1000-times more nucleotide pairs than typical genome in bacteria, but only **20-times more genes**, although, non-**coding segments are 10 000-times more** [5].

In 1980' genetic scientists believed that genes pre-defined the development of the live organism. The main subject of the search was to find genes that were responsible for specific hereditary information and thus cure and treat genetic disorders. However, as genetic engineering developed, this belief was altered dramatically. New approach in science called **Epigenetics** changed number of core beliefs in biology and medicine [6].

The opinion that genome is unchangeable was proved to be wrong – biological structure is actively formed during the life-cycle and alters due to the influence of external factors. Genes "work" in coordination to help the organism to adapt to living conditions and some genes are more active, some more passive as required.

This indicates that cell regulates performance of the genome. As a result, the life-cycle of a human being can alter his/her genetic qualities. It was proved that as a result of the epigenetic influence the same genetic material can produce up to 30 thousand variations of protein. On the one hand, this means that a human might be born as healthy, but due to

number of epigenetic mechanisms can become ill, e.g. develop cancer, and on the other hand, cells that are "unhealthy" can produce normal protein [7].

To conclude, genes are not independent units and cannot regulate their functioning. **Genes do not direct biological systems, rather the bio-system utilizes the genes.** This epigenetic approach reveals that we, as humans are not slaves to our genetic inheritance, rather, we regulate their activity. This is a power that nature has – even from a mediocre genetic product a variety of organisms could be formed thus increasing a chance for their survival.

In eukaryotes a great part of non-coding segments of DNA, at the first glance, serves no function, but as suggested, they might be in charge of regulating which genes become active. They "oversee" active processing of information in a nucleus.

If we remember the differentiation of blastomeres during embryogenesis of mitosic multiplication, we can see that, although, all cells in our organisms are formed from one cell – **zygote** (thus have identical genome), the body will develop to have more than 200 different types of cells based on their morphology and function. All the cells have their "genetic instructions" from the same genome. This is a key to the formation of multi-cell, complex organisms.

"Excess" of genome has other functions as well. As it is known, zygote, whose size is 1000 times more in mammals than in other cells, replicates with higher speed. Definitely, this cell should have special mechanisms to support such a speedy multiplication despite its size.

One of the simplest strategies for rapid growth is to have addition genetic replicas in the cell. Therefore, **selective activation of RNM genes** takes place in the oocyte. For example, oocyte of ameba contains 1-2 million replicas of such genes. This excess accumulation of hereditary information is necessary for the **rapid speed of embryogenesis.**

Study of information encoded in genome is very important and informative [2]. Genome is like a universal machine that is very sensitive to the external signals and has a very complex mechanism to respond to these signals [2].

Therefore, genome provides us with the key to study phylogenesis of each species. Some genes are more prone to change due to environmental factors and thus they mutate. It was found out that segments of DNA that do not code protein are more prone to mutation, while segments that serve other functions are rather immutable.

In conclusion, genome is not a simple sum of genes. It is a complex and multi-functional system. Non-coding nucleotide sequences represent large part of the genome and they serve a regulatory function that defines targeted sorting of genetic information and selective expression of genes. This determines not only the qualities of each cell and its function, but also development of the whole organism, its coordination and finally aging and death.

References

1. Patrushev L. I. 2000. *Gene Expression*. Moscow.
2. Golubski M. D. 2000. *Century of Genetics: Evolution of Ideas and Concepts.* Saint Petersburg.
3. http://en.wikipedia.org/wiki/Human_Genome_Project
4. http://ru.wikipedia.org/wiki/Progect_«Human's Genome»
5. Aladashvili N. "Jumping" Genes and Modern Possibilities of Ancient Genes. http://mkurnali.ambebi.ge/skhvadaskhva/genetika/2918
6. http://en.wikipedia.org/wiki/Human_genome
7. http://mixednews.ru/archives/16985
8. http://www.medbiol.ru/medbiol/epi