AUCTORES
Globalize your Research

Biomedical Research and Clinical Reviews
Abyt Ibraimov,

Open Access

Review Article

Why are there so Many non-coding DNAs with Repeating Sequences of Nucleotides in the Genome of Higher Eukaryotes?

Abyt Ibraimov [1,2], Stalbek Akhunbaev [1], Orozali Uzakov [1]

[1]International Higher School of Medicine, Bishkek, Kyrgyzstan.

[2]National Center of Cardiology and Internal Medicine, Bishkek, Kyrgyzstan.

***Corresponding Author:** Abyt Ibraimov, International Higher School of Medicine, Bishkek, Kyrgyzstan

Received Date: April 24, 2022; **Accepted Date:** May 10, 2022; **Published Date:** May 31, 2022.

Citation: Abyt Ibraimov, Stalbek Akhunbaev and Orozali Uzakov, (2022) Why are there so Many non-coding DNAs with Repeating Sequences of Nucleotides in the Genome of Higher Eukaryotes?. *J. Biomedical Research and Clinical Reviews*, 7(2); Doi: 10.31579/2692-9406/119

Abstract

There are many questions regarding the biological roles of non-coding DNAs (ncDNAs) in the eukaryotic genome. It is noteworthy that ncDNAs, despite the overwhelming majority in the genome is still a mysterious object. From simple theoretical considerations, it follows that this vast class of ncDNAs should play some important role in the vital activity of higher eukaryotes. If by their nature they are not capable create specific biological products (proteins, enzymes or RNAs), then they must form some non-specific structures in the cell that are important in the vital activity of higher eukaryotes. Thus, from our point of view, excess ncDNAs were fixed in the genome of higher eukaryotes mainly due to the presence in their composition of sites with repetitive sequences of nucleotides and their wide variability in the population, which ultimately played a decisive role in the emergence of the eukaryotic nucleus and cells, biological sex and species, large multicellular and warm-blooded organisms, up to modern humans.

Key words: biological role of non-coding DNAs; non-coding DNAs; repeating sequences of nucleotides; condensed chromatin; chromocenter; nucleolus

Introduction

It has been established that the genes encoding proteins make up only about 2% of the DNA in the human genome; the remaining 98% do not encode any proteins. Almost the same situation is observed in other mammals. It is believed that most of this DNA has no real function — hence its name: "redundant", "silent", "garbage" or "selfish" DNA.

There are many questions concerning the biological role of non-coding DNAs (ncDNAs) in the eukaryotic genome, the most important are the following: a) why do most of these DNAs contain short sequences of nucleotides that are repeated hundreds and thousands of times? b) why is there so much ncDNAs in the genome of higher eukaryotes? and, c) why did they survive in the process of evolution if they are not able to encode specific products (phenotypes) useful for life? There are also many hypotheses trying to answer these questions. It is not our task to discuss these hypotheses in detail, since none of them has sufficient experimental confirmation yet. Here we will limit ourselves to presenting our answers to the questions listed above.

It is known to distinguish DNA with unique (encoding proteins, enzymes and RNAs) and non-coding repetitive sequences of nucleotides. The latter, in turn, are divided into highly repetitive DNA, such as satellite DNA and an extensive class of DNA with a moderately repetitive sequence of nucleotides. The role of DNA with uniquely repeating sequences of nucleotides – genes – is well known and is the main object of genetic and molecular biological research. However, despite its exceptional importance, the gene part of DNA in higher eukaryotes occupies a very small share of their genome. The rest of their genome is occupied by DNA with highly repetitive and middle repetitive sequences, unable to encode unique polypeptides or RNAs. It is noteworthy that this class of DNAs, despite the overwhelming majority in the genome and the simplicity of the structure (in the sense of the uniqueness of the alternation of nucleotide sequences), is still a mysterious object when it comes to their possible biological role in the life of higher eukaryotes.

From simple theoretical considerations, it follows that this vast class of ncDNAs should play some important role in the vital activity of higher eukaryotes. If by their nature they are not capable create specific biological products (proteins, enzymes or RNAs), then they must form some non-specific structures in the cell that are important in the vital activity of higher eukaryotes. For example, it is known that part of this DNA forms heterochromatin regions of chromosomes, making a dense layer of condensed chromatin around the interphase nucleus, chromocenters, nucleoli and B chromosomes. However, their biological role is still unclear.

Why do most ncDNAs contain short sequences of nucleotides that are repeated hundreds and thousands of times?

We believe that it is precisely in this – in the existence of sections of chromosomes consisting of short but repeatedly repeated sequences of nucleotides - is the exceptional value of ncDNAs in the vital activity of higher eukaryotes lies. The value of repeating ncDNAs primarily lies in the fact that, unlike DNA with unique nucleotide sequences, they are able to form complex forms of DNA organization, without which eukaryotic organisms do not exist: cell nucleus, mitotic chromosomes, condensed chromatin, chromocenters and nucleoli. After all, it is obvious that these specialized three-dimensional structures could not be the product of a directed favorable mutation of structural genes and/or their simple accumulation in the process of evolution. Prokaryotes have existed on Earth for about 3.5 billion years, while higher eukaryotes, having emerged much later (about 500 million years ago), became the dominant forms of life. However, the question remains, why have prokaryotes, like eukaryotes, with genes remained unchanged for 2-2.5 billion years? There are suggestions that they owe this to ncDNAs with repeating sequences of nucleotides, thanks to which the emergence of biological sex, species, cell thermoregulation, large multicellular and homoiothermic organisms, including humans, became possible [1-5].

It seems highly probable to us that one of the possible approaches to understanding the possible biological role of ncDNAs is to study their ability to non-specific fusion, well known by the example of the formation of a dense layer of condensed chromatin around the nucleus, chromocenters and nucleoli. However, the question remains unclear as to why DNA with repeated sequences of nucleotides is characterized by "stickiness" and their possible consequences (effects) for eukaryotic cells. In other words, what factors contribute to the molecules of the two helical DNA strands to approach each other, which ultimately lead to the formation of such dense structures as condensed chromatin, chromocenters and nucleoli?

The mechanisms of formation of two helical DNA using hydrogen bonds between adenine-thymine and guanine-cytosine pairs have long been established. However, little is known about the nature of convergence ("communication at a distance") between two helical strands of DNA and what underlies them. It is known that even in the absence of any "guiding" molecules or forces, two-helical DNA with identical nucleotide sequences are able to recognize each other at a distance and even converge. There is nothing surprising in the fact that the nitrogenous bases of different DNA chains form pairs: adenine corresponds to thymine, and cytosine corresponds to guanine, just as a glove from the right hand corresponds to a glove from the left. But when the pairs have already formed, i.e., when mutually complementary chains have combined into a double helix, nothing prompts the nitrogenous bases of different molecules to interact with each other. They are securely hidden inside the double helix and shielded by an electrically charged sugar-phosphate backbone.

Nevertheless, it was found that two-helical DNA with the same nucleotide sequences are about twice as likely to approach each other. It has been suggested that, depending on the nucleotide sequence, the double helix twists slightly differently, and although DNA molecules, being equally charged, repel each other (even weakly), the similar arrangement of grooves and ridges in identical molecules helps them find each other according to the similarity principle [6]. This circumstance, under certain temperature conditions, as we believe, contributes to the correct mutual alignment of DNA segments with highly repetitive sequences (hrDNAs), which is a necessary condition for the formation of condensed chromatin, chromocenters and nucleoli in interphase nuclei.

More specifically, the essence of the question is as follows; why and under what conditions do repeated sequences of nucleotides can form the aforementioned complex forms of DNA organization? In search of an answer to these questions, we turned to the long-standing experiments of molecular cytogeneticists on localization *in situ* of large blocks of hrDNAs in heterochromatin regions of mitotic chromosomes [7,8]. As is known, in such experiments they applied methods of denaturation and renaturation of nucleic acids (reannealing experiments).

Numerous studies have established that after denaturation, each component of the genome reanneals at a rate determined by the number of copies of the sequence per genome. The reannealing curve of human DNA, for example, reveals that about 10% of human sequences have a copy number greater than 100000, and 10–15% have a copy number between 100 and 100000. The rate of reannealing depends upon the probability of getting good base pairing. Highly repetitive sequences will reanneal quickly because there are more ways of getting alignment.

Why do we use reannealing experiments here to explain the possible biological role of ncDNAs? The fact is that it is still not clear under what conditions sections of chromosomes with hrDNAs are able to "communicate at a distance" in order to form complex three-dimensional structures in the cell nucleus?

As is known the reannealing of DNA is happening at temperatures lower than required for denaturation, but higher than normal cell temperatures. Such a temperature condition in the cell can occur in those areas of the interphase nucleus where active biochemical processes occur, such as in the nucleoli. Here the question may arise, what does the temperature have to do with it? The answer may be very simple; among other things, as the temperature increases, the movement of microparticles accelerates and the probability of collisions and convergence of neighboring molecules increases. In our case, this means "increased communication" of molecules, including DNA with sites with highly repetitive nucleotide sequences, with all the consequences that follow from this (for more details see [9-12]).

This can best be illustrated by the example of nucleoli, where, without a doubt, the temperature should be higher than in the rest of the interphase nucleus. For unknown reasons, the main body mass of the nucleoli is made up of heterochromatin regions of chromosomes. In humans, the short arms of the acrocentric chromosomes, mainly formed from heterochromatin, are frequently associated in the interphase nucleus with other chromosomes having a large heterochromatin block (1, 9, 16 and Y chromosome) [13]. In addition, it is known that the tandem repetition of a DNA sequence in a large number of copies is sufficient on its own to direct the formation of heterochromatin. Such repeated sequences could allow the chromatin to be compacted to a greater extent, by forming characteristic structures. These structures could be recognized by specific proteins, such as the HP1 proteins, which in turn direct the formation of a dense higher-order chromatin.

There are direct observations that hrDNAs from different chromosomes can contact each other. So, for example, using high-resolution microscopy, Jagannathan et al. also observed chromatin fibers that contain satellite DNA and the proteins D1 (in fruit flies) or HMGA1 (in mice). These fibers connected different chromosomes [22,23]. This suggests that during interphase, the chromocenters keep the genome within the nucleus by gathering pericentromeric DNA from different chromosomes.

In all this, we believe, lies the biological role of such a large number of ncDNAs, because they, due to the presence of repeating sequences of nucleotides in their composition, are able to form complex forms of DNA organization: bodies of mitotic chromosomes, condensed chromatin, chromocenters, nucleoli, as well as B chromosomes [4,6,12].

If the role of mitotic chromosomes as the main informational organelle in ensuring the heredity of eukaryotic organisms is indisputable, then the biological meaning of the existence of a dense layer of condensed

chromatin around the nucleus, chromocenters, nucleoli and B chromosomes needs further research. We have repeatedly discussed the possible role of these dense structures composed mainly of hrDNAs in maintaining intracellular temperature homeostasis, and therefore we will not dwell on them in detail here (for details see [9,10,14]). The question of where areas with a sufficiently high temperature can occur in interphase nuclei and how hrDNAs is collected there to form dense structures was considered in detail by the example of the formation of chromocenters and the appearance of chromosome bands [9-11].

Why are there so many non-coding DNAs in the genome of higher eukaryotes?

In short, they are necessary in large quantities for the formation of the body of mitotic chromosomes and, first of all, their vital components, such as centromeric, telomeric and nuclear organizing regions, without which eukaryotic nuclei, biological sex, cell thermoregulation, multicellular and homoiothermic organisms would not have arisen in the course of evolution (for more details, see [1-5, 14, 15]).

Why have non-coding DNAs been preserved in the process of evolution, if they are not able to encode specific biological products useful for the vital activity of eukaryotic cells?

From our point of view, non-coding DNA has been firmly entrenched and preserved in the process of evolution because they are able to create what genes cannot: nucleosomes and eukaryotic nucleus, mitotic chromosomes and condensed chromatin, without which there would be no eukaryotic organisms with biological sex and species, large multicellular and warm-blooded organisms with cell thermoregulation [1-2,12,15].

According to Modern Synthesis, genes are considered the material basis of evolution [19]. We believe that any scientific hypothesis trying to explain the evolution of eukaryotes must be able to explain the origin of eukaryotic-specific traits such as nucleosomes, mitotic chromosomes, chromosome bands, cell nucleus, eukaryotic cells, sex, species, multicellular and homeothermic organisms, including modern humans.

So, for example, despite the impressive successes of modern biological science the problem of sex origin of eukaryotes in the process of evolution still has not settled. Existing theories and hypothesizes mainly concern the maintenance and biological reasonability of sexual mode of replication. Their theoretic foundation is based on Darwin's and Mendel's ideas that sex was originated due to natural selection and genes [16-18]. From our point of view, sex of eukaryotes was originated as a result of long-term evolution of ncDNAs in a genome at one of the branches of prokaryotes. Non-coding DNAs accumulation and evolution in prokaryotes' ring chromosomes eventually led to emergence of mitotic chromosomes and mitotic way of cell division. Sex and sexual replication became possible since that time when modified variant of mitosis – meiosis – have originated. Separate stages of the proposed model may be exposed to experimental check [2-5,12**].**

Evolution is a historical process. It is unlikely that we will ever be able to fully understand it, but still, we are moving towards the truth. We have repeatedly discussed the assumptions that: a) evolution is not only an the origin of various species of animals and plants, but the major transitions in life history on the Earth, such as the emergence of mitotic chromosomes, mitosis, meiosis, eukaryotic cells, biological sex, multicellular organisms, homoeothermic animals, including humans; b) the material basis of the origin and further evolution of the eukaryotic organisms were, apparently, ncDNAs, while genes, though important, were of secondary importance; c) ncDNAs, due to their molecular structure (short, repeated thousands and millions times DNA sequences), behavior during the cell cycle (remain condensed through the cell), genetic organization (they do not encode polypeptide chains), cytological organization (an ability to dense packing in the cell), the localization features (centromere and telomeric sites, as well as the regions of nucleolar organizers of chromosomes) and wide variability, were the main genetic material in the evolution of eukaryotic organisms [1-5,9,12].

Evolution of early forms of life and prokaryotes were carried out based on genes, but the origin of eukaryotic organisms and their further development were made possible by emergence and further evolution of ncDNAs, which led to the formation of chromosomal C-, G-, Q-, R - and T-bands. Some of the chromosomal bands underwent further evolution and formed two types of constitutive heterochromatin: C-heterochromatin that exists in the genome of all higher eukaryotes and Q-heterochromatin, existing only in three species of higher primates (*Homo sapiens, Pan troglodytes* and *Gorilla gorilla*). Among these three higher primates, only human populations have a wide hereditary polymorphism of Q-heterochromatin, and this genetic material may have played an important role in the origin of modern man and his adaptation to different climatic and geographical conditions of the Earth (for more details see [12,19]).

Concluding remarks.

The main obstacle in elucidating the possible biological role of redundant DNA with repetitive nucleotide sequences is that: a) they do not produce specific biological products (proteins, enzymes or RNAs); and, b) some higher forms of ncDNAs organization, such as heterochromatin regions of chromosomes, are prone to wide hereditary variability in the population. Apparently, the time has come to apply a different, non-traditional for classical genetics and molecular biology conceptual approach. For example, in order to clarify the possible biological role of constitutive heterochromatin in the human genome, instead of searching for a visible specific phenotype, we searched for its non-visible physiological manifestation, which could manifest itself, for example, when a human adapts to different climatogeographic conditions [19-21]. Indeed, it turned out that human populations permanently residing in different environmental conditions differ in the heat conductivity of their bodies, the level of which is closely related to the number of heterochromatin regions of chromosomes in the genome [12,19]. We believe that it is precisely due to the presence of hrDNAs in the constitutive heterochromatin and their wide quantitative variability that the birth in a population of individuals with different levels of body heat conductivity that differ in their resistance to different environmental temperature conditions became possible [12,20,21].

Thus, from our point of view, excess ncDNAs were fixed in the genome of higher eukaryotes mainly due to the presence in their composition of sites with repetitive sequences of nucleotides and their wide variability in the population, which ultimately played a decisive role in the emergence of the eukaryotic nucleus and cells, biological sex and species, large multicellular and warm-blooded organisms, up to modern humans.

Acknowledgements.

I apologize to that authors whose work is not cited or is cited only through reviews. The reason for this is only the space limitations of the publication.

References

1. Ibraimov AI. (2004). The origin of condensed chromatin, cell thermoregulation and multicellularity. Complexus, 2: 23-34.
2. Ibraimov AI. (2012). On the origin of sex. GJSFR (C), 12(8) (Ver. 1.0), 1-6. [Global Journal of Science Frontier Research: C Biological Sciences].
3. Ibraimov AI. (2017). Cell Thermoregulation: Problems, Advances and Perspectives. J. Mol. Biol. Res., 7(1): 58-79. doi:10.5539/jmbr. v7n1p58.

4. Ibraimov AI. (2019). Evolution Without Genes: A Review. Current Research in Biochemistry and Molecular Biology, 1(1),45-68.
5. Ibraimov AI. (2021). Why should non-coding DNA be considered as a primary material in the evolution of eukaryotes? Current Research in Cytology and Histology, 2(1):10-16.
6. Baldwin GS, Brooks NJ, Robson RE, et al. (2008). DNA Double Helices Recognize Mutual Sequence Homology in a Protein Free Environment. J Phys Chem. B., 112 (4): 1060-1064.
7. Arrighi FE, Hsu TC. (1971). Localization of heterochromatin in human chromosomes. Cytogenetics, 10: 81-86.
8. Pardue ML, Gall JG. (1970). Chromosomal localization of mouse satellite DNA. Science, 170: 1356-1358.
9. Ibraimov AI. (2018). Chromocenters and Cell Thermoregulation. J Biol Med Res., Vol.2 No.3:19.
10. IbraimovAI. (2019). B-chromosomes and cell thermoregulation. Int J Biol Med., 1:99-106.
11. Ibraimov AI. (2020). Chromosome Bands and Cell Thermoregulation. Current Research in Biochemistry and Molecular Biology.
12. Ibraimov AI. (2020). Chromosomal Q-heterochromatin in the Human Genome. Cambridge Scholars Publishing.
13. Schmid M, Vogel W, Krone W. (1975). Attraction between centric heterochromatin of human chromosomes. Cytogenet Cell Genet., 15:66-80.
14. Ibraimov AI. (2003). Condensed chromatin and cell thermoregulation. Complexus, 1: 164-170.
15. Ibraimov AI, Akanov AA, Meimanaliev TS, Sharipov KO, Smailova RD, Dosymbekova R. (2014). Human Chromosomal Q-heterochromatin Polymorphism and Its Relation to Body Heat Conductivity. Int J Genet., 6(1): 142-148.
16. Maynard Smith J. (1978). The evolution of sex. Cambridge: Cambridge University Press.
17. Bell G. (1982). The masterpiece of nature: The evolution and genetics of sexuality. London: Croom Helm.
18. Michod RE, Levin BR, eds. (1988). The evolution of sex. Sunderland, Mass.: Sinauer Associates.
19. Rose MR, H Oakley H. (2007). The new biology: beyond the Modern Synthesis. Biol Direct., 2007; 2: 30. doi: 10.1186/1745-6150-2-30.
20. Ibraimov AI. (1993). The origin of modern humans: a cytogenetic model. Hum Evol., 8(2): 81-91.
21. Ibraimov AI. (2019). The origin of modern humans. What was primary: genes or heterochromatin? Hum Evol., 34(1-2): 39-58.
22. Jagannathan M, Cummings R, Yamashita YM. (2018). A conserved function for pericentromeric satellite DNA. eLife 7: e34122.
23. Jagannathan M. Yamashita YM. (2018). Function of junk: pericentromeric satellite DNA in chromosome maintenance. Cold Spring Harbor Symposia on Quantitative Biology 034504.

DOI: 10.31579/2692-9406/119