**BUSINESS GAME HUMAN GENOME FOR EDUCATION IN BUSINESS, LAW AND COMMERCE.**

PONOMARENKO M.P.+, FROLOV A.S., PONOMARENKO J.V., MININA A.V.1, IVASHIN S.A.1, MIKHAILOV Yu.I.1

Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences, 10 Lavrentiev Ave., Novosibirsk, 630090, Russia;

1Siberian Trade University Novosibirsk, Russia;

+Corresponding author

Keywords: human genome, information analysis, education, business game, economics, commerce, law

**Abstract**

The business game "HUMAN GENOME" is a new approach for training in economics, commerce and law via demonstrating the common features of the self-reproducing systems in nature. Its basic idea is in consideration of the genetic information descent law as the natural manner of the regulation cycle of the birth, maturation, self-reproduction, aging and, finally, altruistic death of any living organisms. The information-analytic methods widely used in bioinformatics becomes now dominant in economics, commerce and law, too. This training is, hence, carried out in a game form by solving practical problems of decoding human genetic information using INTERNET-available databases created within the International Human Genome Initiative. The game allows the participant to simulate the work of a researcher in the field of genetics who compares a previously unknown human genomic fragment with thoroughly studied fragments of the human genome to extract the new information on human genome. The business game "HUMAN GENOME" is designed to give the acquaintance with the informational-analytic methods of genetic information decoding. At the very beginning of the professional education in economics, commerce and law, the game will allow the participants to gain the skills in understanding of textual messages in an unknown natural language by analogy with the known messages in this language. These skills can be useful in their future professional practice, for example, for adequate consideration of a diverse information in making substantiated decisions. The INTERNET-based version of this business game employs the databases and tools integrated within the computer system GeneExpress (<http://wwwmgs.bionet.nsc.ru/Educate/HumGen/>).

**1. Introduction**

Each cell of the human organism contains 23 pairs of the giant DNA molecules. They comprise all the genetic information of the human organism being. This information is used by other molecules to provide the coordination and regulation of all the molecular processes within the human organism. Among such processes are protection of genetic information; its translation, transcription, recombination, repair; and transmission from the "parental" cells to their "progeny" cells; catalysis of biochemical reactions; maintenance of the "normal" values of physiological parameters of the organism (for example, body temperature, oxygen concentration in blood, etc.); maturation and differentiation of cells, tissues, organs; and forming the standard orgasmic reactions in response to typical environmental changes. All the processes regulated by the genetic information are focused on the self-reproducing of the organism generations from one another that is the basic difference between the living and non-living nature.

DNA molecules are linear polymers constituted by four monomers, named "nucleotides" and designated by four Latin letters {A, T, G, C}. Such a simple chemical DNA structure has allowed the complete genomic DNA sequences of a number of microorganisms (bacteria, viruses, phages; the genomes of 106-107 nucleotides long) to be already identified. It is expected that the yeast genomes (107-108 nucleotides long) will have been identified by the end of this XX century; the genomes of human and the agriculturally and industrially valuable animals and plants (1010-109 nucleotides) will have been identified by the very beginning of the XXI century.

Occurrence of a great number of common regular patterns in organization and function of naturally arisen self-reproducing systems are now being more and more widely accepted. When the human genome is one of the eldest self-reproducing natural systems, their genetic information decoded can show the greatest experience in solving the self-reproduction problems that is helpful for understanding human languages, social communities, legal systems, economics, industries and commerce markets.

That is why we have developed the business game "HUMAN GENOME", in which we suggest a new approach for training in the economics, commerce and law to deal with informational-analytical practice by the example of decoding the human genetic information. Participants of the game imitate the work of a researcher in genetics who compares the previously unknown fragment of the human genome with the fragments thoroughly studied to extract the novel genetic information on human. Thus, a specialist in economics, commerce and law acquires in a game form the skills to understand messages in an unknown natural language by analogy with the known messages in this language.

**2. Materials and methods**(game training)

To carry out the business game "HUMAN GENOME", a INTERNET-connected computer is required. When it is not available, the special cards having a fragment of the human DNA sequence with and without their genetic information cam imitate the INTERNET-available database on human genome. The classic training scheme, INTRODUCTION -> TRAINING-> CONTROL, is used for this game, and, hence, it is intended for three training sessions with their preliminary self-preparation of the participants.

***2.1 Introduction session****( 2 school hours)*

This starting session is to acquaint the participants with the variety of human genetic information elements (genes, regulatory regions, signals, etc.) and their interrelationships with molecular processes realizing this information to control all the vital processes within cells, tissues, organs, and organisms.

For preliminary self-preparation of the participants of this session, they should read and learn by themselves the topic **«Theoretical Principles of the Human Genetic Information (simplified for the economics, commerce and law education)»**. During this session, instructor through the quiz reveals the items appeared most difficult and requiring a detained explanation. Then, each participant receives a card with **"previously unknown"** fragments of the human genome and a hypothesis on the human genetic information encoded in these fragments to be tested.

After that, the participant is provided with the access to the databases on **"thoroughly studied"** genomic fragments of various organisms to select those fragments that he considers "useful" for decoding the **"previously unknown "** fragment he received (databases of the computer system GeneExpress are used in the Internet version of the game).

***2.2. Training session****(2 school hours)*

This main session is to use the informational-analytic methods for decoding textual messages in an unknown language by analogy with the known texts in this language including practical application of these methods for decoding human genetic information.

The preliminary self-preparation for this session is that the participants read by themselves the topic «Informational-Analytic Methods Revealing the Similarity in DNA Fragments (simplified for the economics, commerce and law education)». At this session, Instructor through the quiz finds out the difficult items concerning the principles of these methods and gives the necessary explanations. Then each participant reveals the similarity elements in the "thoroughly studied" genomic DNA fragments selected during INTRODUCTION SESSION using these informational-analytic methods.

To increase the interest to the game, the instructor may propose the participants to use their own intuitive methods in addition to conventional methods of heuristic search and informational-analytic techniques based on formal mathematical calculations. In this case, it is reasonable to recommend **individual** participants of the game to use **different** informational-analytic methods to compare the similarity elements found relative, for example, to the clearness of their perception. In the Internet-based version, the participants use the computer programs of the system GeneExpress.

***2.3. Control session****(2 school hours).*

This final session is, in the conference game form, to evaluate and repeat the material studied, when all the milestones and phases of the game are simulated considering their interrelationships.

The preliminary self-preparation of this session is that each participant applies by his own the revealed elements of the similarity between the **"thoroughly studied"** genomic fragments to decode his **"previously unknown"** human genomic fragment and prepares the results obtained for the control session. This session is carried out in a form of the conference, where each participant gives brief oral presentation on the new human genetic information that he managed to decode demonstrating the similarity elements between the **"thoroughly studied"** genomic fragments he found and explaining the operation of the informational-analytic method he applied to reveal this similarity. Participation in a game scientific conference can be useful for specialists in economics, commerce and law to acquire the experience of public oral presentation, clear formulation of the goals and objectives of their practical activity, preparation of illustrative demonstrations (tables and figures), and convincing argumentation on the conclusion made, and, also, discussing audience to explain the most fuzzy points of his decision.

***2.4. Marking the education results***

The instructor gives the mark for the public presentation of the participants examined considering the accuracy and completeness of decoding displayed by a participant of the game, the level he managed to master the informational-analytic methods for decoding textual messages in an unknown language by analogy with the known texts in this language, and the skill to present his own substantiated results.

**3. Educational topics**

For preliminary self-preparation for this business game session, the of the participants should read by themselves the educational topic simplified for the economics, commerce and law education)».



Figure 1. The scheme of human genetic information employment in regulation of the vital activities of cells, tissues, organs, and the entire organism (simplified for training in business, law, and negotiations).

|  |  |
| --- | --- |
| *HG-12345* | **Thoroughly studied fragment of the human genome** |
| *ORGANISM* | *GENE* | **FUNCTION** | *FRAGMENT* |
| Human | alpha1-globin | Oxygen transport from the lungs to all organs via the blood | Region where the copying starts |
| Agcg**ccgccc**ggcc**gggcgtgccc**ccgcgccccaagc**ataaacc**ctggcgcgctcgcggcccggc**a**ctcttctggtccccacagactc**.** |
| **SP1-signal** | **aIRP-signal** | **ATAAA-signal** | **Transcription START** |

|  |  |
| --- | --- |
| *HG-54321* | ***Previously unknown fragment of the human genome*** |
| **ORGANISM: human** | **ASSUMPTION:**the fragment can contain the transcription start of a human gene |
| ...TCAAAAACTAAATCCCACATTGAAAATATTTCTTATAGAGTCATGCAAAATAGACTACAAATATAAGATTTGTCACCCT... |

Figure 2.An example of **"thoroughly studied"** and **"previously unknown"** fragments of the human genome.

***3.1. Theoretical principles of human genomic information****(simplified for the economics, commerce and law education)*

Genetic information of the organisms is "written" in their DNA in a form of a definite order of "letters" A, T, G, and C following one another similar to "words", "phrases", and "sentences" (similar to the description using Latin letters and punctuation marks of the business game "HUMAN GENOME" in this topic). The human genetic information is realized to provide the regulation of the vital activities of all the cells, tissues, organs, and the entire organism in three stages, shown schematically in Fig. 1. **First,** the genetic information is "transcribed" from the DNA molecule into a molecule of pre-mRNA. **Then,** the pre-mRNA molecules are "processed" by excision of their noncoding regions to produce the mRNA molecule. And **finally,** the genetic information is "translated" from the mRNA molecule into protein molecule, which performs directly the corresponding vital function. For example, the protein gamma-crystalline fills the lens cells of the eye and focuses the image of the external world on its photosensitive cells. As is evident from Fig. 1, each of these three stages of genetic information realization (transcription, processing, and translation) has its own signals of switching on (**START)** and switching off (**STOP**). The region of the DNA molecule between the transcription **START** and **STOP** is named "**GENE**". Each human gene has essential differences from all the other genes in the regions encompassing its transcription, translation, and processing signals. It is these differences that allows this gene to be used in the organism for performance of only those biological functions that are encoded in this particular gene. For example, the gene coding for gamma-crystalline is used only is the human lens cells but are not in use in all the other cells. However, all the human genes display also an essential similarity of their DNA regions in the proximity of their transcription, translation, and processing signals. It is this similarity that allows the genes to be used by a universal molecular machinery, consisting of a great number of RNA, protein molecules, and low-molecular weight substances.

Let's consider the example of genetic information coding in the neighborhood of transcription **START**s. The human genome comprises over 50 000 genes. The genetic information of each gene serves to the performance of a stringently determined biological function of the corresponding cells, tissues, organs, and the entire organism. Selection of one or another gene depending on the state of the organism (for example, adult organism) and environmental conditions is first provided for by transcription initiation signals. Transcription **START** has frequently the following form: a nucleotide "letter" **C** or **T** is located two positions before the **START** followed by **C** or **T** again; the first transcribed "letter" (**START**) is usually **A,** the third, **A** or **T,** the fourth, again **C** or **T;** the fifth, again **C** or **T,** and further no preferences occurs. A very important signal **TATA-box,** designated according to its most frequent letters (first, **T**; second, **A**; third, **T**; and fourth, **A**), is usually located approximately 30 "letters" before the transcription **START.** The next four "letters" of this box are usually **A;** however, they were not included into the name of this box for the sake of convenient pronunciation. The **TATA-box** is usually surrounded by the region with an increased frequency of the "letters" **G** and **C**, which is likely to "aid" its recognition. These numerous **G** and **C** themselves form the signals for transcribing genetic information, for example:«**CCCCAGGC**» is the signal of «AP2»; «**CCGCCC**», of «SP1»; «**GGGATTCCC**», of «NF-kB». The signals very similar to **TATA-box** can be found near the box; they are likely to hinder its recognition. For example,«AGATAAG» is the signal of «GATA»; «TGTAAAG», of «C/EBP»; «TATxCAT», of «Pit-1» («x» indicates the lack of preference for the "letters"), «TCATA», of «YY1», named "yin-yang", as in certain genes it is the signal for switching on the transcription, whereas in other genes, for switching off.

On the whole, the fragment of the human genome encompassing the transcription**START** is organized according to the stick and the carrot principle: it contains a huge number of transcription signals, and some of them switch on the transcription of genetic information from the DNA molecule into the pre-mRNA molecule; others, switch the transcription off. Since these signals differ in the degree their efficiency depends on the human organism state and environmental conditions, each gene is poising between "**SWITCH-ON**" and "**SWITCH-OFF**" of its transcription after each instance the external and internal environments are changed. In the course of the evolution, only those of our progenitors managed to survive who "**SWITCHED ON**" the transcription of each gene only in case the biological function of this gene became vitally important and "**SWITCHED OFF**" this gene immediately when it became useless. This was the way to achieve the evolutionary balance between the general principles of genetic information coding in the human genome and the specific manifestation of these principles for each individual gene that corresponds strictly to the vitally important biological function of this gene. It is this evolutionary compromise of the general principles of genetic information coding and specific biological functions that allows the previously unknown human genome fragments to be decoded by analogy with thoroughly studied fragments of the human genome.

***3.2. Informational-Analytic Methods Revealing Similarity in DNA Fragments****(simplified for the economics, commerce and law education)*

Hundreds of various informational-analytic methods are currently used. The simplest and most widely used is the consensus method. "Consensus" of a similarity element of different genomic fragments is a sequence of the most frequently occurring nucleotide "letters" in each position of this similarity element. Consensuses of the commonly known similarity elements, designated as "transcription signals" were shown above: **«(C/T)(C/T)Ax(A/T)(C/T)(C/T)»** is the consensus of transcription START (here, the two most frequent nucleotide "letters" occurring in the same position are indicated in parentheses); **«TATAAAAA»,** consensus of **TATA-box**; **«CCCCAGGC», «CCGCCC», «GGGATTCCC», «AGATAAG», «TGTAAAG», «TATxCAT»**, and **«TCATA»**, consensuses of several other transcription signals.

To test the hypothesis "Does a given human genome fragment codes for the transcription initiation **START** of a gene?", one needs to count the number of regions within the fragment in question that display considerable number of coincidences with the consensuses of transcription signals. Similar processing of "thoroughly studied" genomic fragments containing the transcription initiation **START** and of random sequences consisted of the letters A, T, G, and C should be carried out. If the "previously unknown" genomic fragment is more similar to the **"thoroughly studied"**genomic fragment containing the transcription initiation START than to random sequences, this indicates a possibility that it also encodes the transcription initiation start of a gene.

**4. Discussion**

Description of the training process demonstrates that a participant of the business game "HUMAN GENOME" gains the concept of the structure of DNA molecules, their evolution, structure-function organization, general principles of the encoding of human genetic information in these molecules, molecular mechanisms through which this information is realized to regulate all the vital activities of the human organism, and informational-analytic methods used to extract genetic information from DNA.

The main goals of the instructor in the game are: (1) TO GIVE the possibility for a participant to search for the necessary information in the INTERNET-accessible databases and to make informational-analytic conclusions by his own basing on this information; (2) TO HELP the participant to comprehend and generalize the informational-analytic methods for decision making he used (by the example of decoding of genetic information); and (3) TO DEMONSTRATE how the informational-analytic skills acquired can be helpful in his future professional activities in the field of economics, commerce and law.

The participant has the following possibilities: (1) ***TO BECOME ACQUAINTED WITH*** the interrelationships of the problem of coding and realization of genetic information including various ways to settle these two interrelated problems found by "nature" in a variety of living organisms; (2) ***TO SEARCH BY HIS OWN FOR*** similar textual fragments in a "natural" language while solving informational-analytic problems;
(3) ***TO LEARN***to dovetail concrete practical problems with the necessity to provide the informational-analytic adequacy of their solutions; (4) ***TO LEARN*** the operation with the INTERNET and have the notion of the relevant difficulties in searching for adequate databases and tools for information search in these databases, clear and unambiguous formulation of the queries, making of prompt and clear informational-analytic conclusions, and performance of informational-analytic tasks in a specified terms; and (5) through the business game at early stages of the professional education, ***TO LEARN*** to consider adequately a diverse information to increase the efficiency of the professional activities in the field of economics, commerce and law.

For conclusion, the participant of the business game "HUMAN GENOME" gains the practical skills of searching for the information of interest in the databases of the INTERNET, masters the informational-analytic methods for extraction of the useful information from textual messages in an unknown natural language by analogy with the known texts in this language and substantiation of his conclusions. All these gains can be useful for the students in their future professional activities in the field of economics, commerce and law.

**Acknowledgements**

We are grateful to Ms. Galina Chirikova for help in translation. This work was supported by Russian National Human Genome Project, Russian Ministry of Science and Technical Politics, Siberian Branch of Russian Academy of Sciences (grants IGSBRAS-97 N13), and Russian Foundation for Basic
Research (grants 96-04-50006, 97-07-90309, 97-04-49740, and 98-07-90126).