**Биология. Текст 1.**

Abstract: The aim of this article is to present the current literature concerning the expression analysis

and methods of functional characteristics of genes. The progress in the analysis of gene expression

within cells or whole tissues is undisputed and leads to a constant improvement of our understanding

of the function of particular gene. The traditional methods of the functional characteristics of genes

such as homology, inactivation and overexpression are more frequently being replaced by microarray

and DNA chip analysis, which are extensively supported by bioinformatics tools. Knowledge of the

functions and changes in gene expression has applications in medical diagnostics, the pharmaceutical

industry and in plant and animal biotechnology.

Key words: gene expression, gene, DNA chips, microarray

INTRODUCTION

The spectacular reports of completely sequenced genomes of yet another

organism have been there since 2000, when the weekly "Nature" magazine

published the full sequence of the model plant organism - Arabidopsis thaliana [1].

Current literature shows that on average, once a month, global science is enriched

with full sequences of four genomes of new organisms [5]. However, the

recognition of the DNA sequence is merely the beginning on the way to

understanding how an organism functions at the molecular level [7]. The

identification of coding sequences within a genome and an analysis of gene

products are the next phases leading to the recognition of the functions fulfilled by

genes. The analysis of gene function, is a task for the newly emerging field

functional genomics. There are numerous methods of characterising gene functions

and their number is constantly increasing due to quicker and quicker progress of

technology and computerisation.

The process of determining the function of a studied gene is only possible

after obtaining the sequence and identifying where it is expressed [7]. It is

performed by means of computer analyses or experimental studies, associating a

gene and related phenotype. Various strategies and research techniques, such as

homology-based prediction, increased expression or gene inactivation have been

used for many years now. However, more and more modern microarrays and DNA

chips are used and they have become the basic technique of analysing gene

function [27].

**Биология. Текст 2.**

METHOD OF PREDICTING GENE FUNCTION THROUGH

HOMOLOGY

Communal access to genome sequences, which are stored in databases,

allows for a comparison with the studied sequences. An analysis showing similarity

with the recognised genes is helpful in the initial determination of a probable gene

product and its function [12].

The method of predicting through homology relies on the assumption that

if a newly sequenced gene is very similar to an already characterized and published

gene, the function of the new gene is probably similar.

In this method the first step is comparing the studied fragment of DNA

with the fragments of DNA of model organisms available in the computer

databases. DNA of most model organisms have been completely sequenced and at

least partially characterised with regard to gene function. The largest and most

popular databases, in which full sequences of genomes have been stored, are on the

NCBI (National Center for Biotechnology Information) server. The biggest

limitation in the use of homology-based methods is the presence of uncharacterised

sequences.

A comparison between an analysed sequence and described genes

contained in the databases also has certain risks. Although at the nucleotide level

the DNA sequences may have high similarity, they can encode different amino

acids, as well as proteins of various functions and structures. A protein is

comprised of many amino acids arranged in an specific order, thus a rule has been

adopted that the comparison uses not the sequences of nucleotides, but the amino

acids coded by them. It is known that the amino acids sequence is more conserved

then the nucleotide one. Use of the amino acids sequence allows detecting the shift

of reading frame between the non-homologous genes. Thus, the fragments for

comparison are selected in the databases on the basis of the highest possible

similarity of the content and order of amino acids.

**Биология. Текст 3.**

Using a technique based on homology we can also determine the

evolutionary relations between the studied and model organisms [12]. The method

of predicting functions by means of homology is the basis of phylogenetics, which

compares the sequences based on the common, evolutionary origin of genes.

Created phylogenetic trees simplify the search for gene functions, considering their

duplications, substitutions in the sequence of various species leading to divergence,

and resulting from that - speciation. The knowledge concerning the way of gene

flow and determining related species contributes to the initial identification of gene

function. On the basis of similarity of gene sequences we can conclude that

homologous genes, that is, ones having a common evolutionary ancestor and

occurring in the studied organisms, probably have the same or comparable

function. Unfortunately, a difficulty lies in the fact that with time, genes may

accumulate random mutations and though being highly similar in structure, they

may have distinct and separate functions [39]. Based on that we can distinguish

between:

Orthologs, are genes occurring in various species (which may also prove

that they have occurred in a common ancestor) and fulfilling the same or

comparable functions. The genes are created as a result of gene duplication

leading to speciation, but do not necessarily have the same function as the

gene of a common ancestor [17];

Paralogs, are genes occurring in various organisms or just in one, but due

to slight, significant changes in structure, they fulfill separate roles. In this

case the duplication leads to divergence, that is, a division of functions. An

example of two paralogs are human myoglobin and hemoglobin,

responsible for the storage of oxygen in skeletal muscles and transport of

oxygen between cells and pulmonary alveoli respectively [17];

Xenologs, are genes similar to one another due to the fact that they have

been acquired by organisms through horizontal gene transfer, which does

not, however, prove their common evolutionary origin [36].

**Биология. Текст 4.**

The requirement for the knowledge concerning gene functions and the

development of bioinformatic technologies has caused many scientific institutions

to engage into cataloguing knowledge of known homologues in databases

(Table 1) [12]. This has allowed for the grouping of data and common access to it.

Researchers comparing the genomes of alga Chlamydomonas with the

genome of humans and model plant species- Arabidopsis thaliana - used the

homology-based method of determining the gene function and BLAST algorithm.

The aim of the experiment was to determine the relationship between plant and

animal kingdoms and to assign the characteristic plant and animal gene functions.

As a result, 349 plant proteins engaged in the process of photosynthesis and 195

animal proteins responsible for the movement described [31].

The use of the BLAST algorithm while studying homology is limited when

the similarity between the studied sequences is low (reaching 20-30%) [19]. Such

proteins that differ significantly at the amino acid level can however, assume a

similar structure, fulfill similar functions and also be homologues. In common

classification, where proteins have similar sequence, structure and function, they

are combined into families, and then into super-families. These proteins that differ

in terms of sequences are described as so-called distant homologues. The studies

on evolution show that the structure of proteins is preserved better than their

sequence [16]. Due to this, studies on protein structure are important when

determining their functions. Accordingly, the requirement for the creation of

methods studying distant homology resulted in their quick development. Such

methods include searching for common motifs within a family, or identifying

conserved amino acid residues (e.g. Multiple Alignment – ClustalXhttp://

www.clustal.org/, [25]). An extraordinarily useful tool for the detection of

distant homologues and rating proteins as belonging to the same family, is the PSIBLAST algorithm and cascade PSI-BLAST [2,3].

**Биология. Текст 5.**

GENE INACTIVATION METHOD

The method of gene inactivation comprises of finding it and blocking

transcription, which allows for a comparison of the obtained phenotype of the

studied organism with the phenotype of the non-mutated organism. On this basis

we can determine what changes have occurred in an organism and attribute them to

the non-active gene. Currently, there are numerous methods of gene silencing used

[43]. A basic principle of this technique is to generate and introduce a gene

construction into an organism that will effectively block a specific gene. The effect

is the lack of synthesis of the protein encoded by the silenced gene, which often

result in phenotypic differences that can lead to a conclusion concerning the

function of a given gene.

One of the most common technique of gene inactivation is its

discontinuation by means of an artificially introduced DNA fragment through

insertional mutagenesis (knock-out) (Fig. 1), which is based on homological

recombination. This technique is generally performed in one-cell organisms to

avoid generating chimeras, whereby an organism is comprised of a mixture of

mutated and non-mutated cells. Insertional mutagenesis is based on an insertion of

the DNA fragment from a vector within the gene located on a chromosome. The

chromosomal DNA obtained in this way contains the discontinued gene, which

does not undergo expression and, in effect, there is no protein created. The

disorders caused by the lack of protein show its function in the organism. Gene

knock-out allows tracking phenotypical changes resulting from the exchange of

sequence fragments between chromosomal DNA and the vector. The vector usually

contains a gene discontinued by a selective marker allowing for an identification of

recombinants and at the same time causing an interruption of the gene and making

it inactive.