

**ABSTRACT**  
**of the dissertation for the degree**  
**Doctor of Philosophy (Ph.D.)**  
**6D070100-Biotechnology**

**Yurikova Oxana Yuryevna**

**The interaction of miRNA with mRNA coding region of genes  
associated with some socially significant diseases**

**General characteristics of the work.** The work is devoted to the study of miRNA interaction with coding sequence of mRNAs of genes, involved in the development of some socially significant diseases and to the search for new effective associations of miRNAs and their target genes for diagnostics of the diseases.

**Relevance of the topic.** Socially significant diseases are among the most important problems of medicine. For the effective prevention of these diseases, it is necessary to involve authorities, employees of the medical, educational spheres, etc. Socially significant diseases lead to disability of various population groups and premature mortality. A large amount of funding is allocated to solve the problems of their prevention, treatment, and rehabilitation. Among the socially significant diseases, there are cancer, acute myocardial infarction, degenerative diseases of the nervous system, and demyelinating diseases of the central nervous system.

Cardiovascular diseases is one of the leading causes of death in Kazakhstan and around the world (17.9 million deaths in 2016). In 85% of cases, the causes are heart attack and stroke.

Cancer occupies the second place after cardiovascular diseases in the ranking of global deaths in Kazakhstan. Lung cancer is the most common cause of cancer death in the world. According to estimates from GLOBOCAN database, there were 1.76 million deaths due to lung cancer in 2018. There are two main histological types of lung cancer: small cell lung cancer and non-small cell lung cancer. Non-small cell lung cancer accounts for approximately 85% of all lung cancer cases.

Neurodegenerative diseases remain incurable, resulting in progressive degeneration and death of neuron cells. This creates an urgent need to improve an understanding of the pathogenesis of neurodegenerative diseases. Worldwide, there are about 50 million people with dementia. Medicines used in neurodegenerative diseases are aimed at correcting symptoms. One of the primary goals of care for dementia is an early diagnosis to ensure early and optimal management. Among neurodegenerative diseases, Alzheimer's disease is the most common form of dementia in the world - it accounts for 60-70% of all cases of dementia.

The search for biomarkers allowing minimally invasive early diagnosis of socially significant diseases, as well as predicting the effectiveness of therapy, remains one of the priorities of modern medicine. A lot of biomedical research is discussing the potential use of miRNA molecules as biomarkers. miRNA - non-coding RNA of 17-25 nucleotides in length found in many species.

It has been shown that miRNA molecules are expressed in various tissues and cell types. Changes in miRNAs expression profiles contribute significantly to many normal and pathological processes. Changes in miRNA levels have been found in the tissues and serum of patients with different types of diseases. miRNAs are highly stable in serum and plasma, resistant to endogenous ribonucleases and adverse physical conditions, which makes it possible to efficiently isolate these molecules from biological fluids; their level can be measured in real-time PCR with high sensitivity and specificity. To date, quantitative changes in circulating miRNAs levels have been identified for a number of diseases, and a further search for miRNA markers is an important area of medical biotechnology.

There are few publications devoted to the prediction and study of miRNA binding sites in the protein coding region of mRNA. Although it has been shown that the sequence of CDS (protein coding region of mRNA) is more conserved than 5'UTR (5'-untranslated region) and 3'UTR (3'-untranslated region) mRNA. Therefore, it is likely that miRNA binding sites in CDS, as well as the regulation of gene expression by established miRNAs, are an evolutionarily conservative. Consequently, the study of the structural and functional organization of miRNA binding sites in CDS of mRNA is relevant.

**The aim of the work:** to identify the structural and functional properties of miRNA binding sites in the coding region of mRNAs of genes associated with some socially significant diseases.

**Tasks of the work:**

1. To create databases of genes involved in the development of Alzheimer's disease, myocardial infarction, and non-small cell lung cancer.
2. To establish characteristics of the interaction of miRNA with mRNAs of genes associated with the development of Alzheimer's disease.
3. To establish characteristics of the interaction of miRNA with mRNAs of genes associated with the development of non-small cell lung cancer.
4. To establish characteristics of the interaction of miRNA with mRNAs of genes associated with the development of myocardial infarction.
5. To study the interaction of miR-1322 with genes involved in some socially significant diseases.
6. To study the features of the fully complementary interaction of miRNA with the coding region of mRNAs of genes.

**Objects of the research:** nucleotide sequences of 6272 miRNAs and genes associated with the development of some neurodegenerative, oncological and cardiovascular diseases.

**Subject of the study:** characteristics of interactions of miRNAs with nucleotide sequences of coding region of mRNAs of genes involved in the development of Alzheimer's disease, myocardial infarction, and non-small cell lung cancer.

**Methods of the study:** Methods of computer analysis and search for miRNA binding sites in mRNAs of genes based on modeling hydrogen bonds using the MirTarget program.

**The scientific novelty of the study.** The characteristics of the interaction of miRNA with CDS mRNAs of genes associated with the development of myocardial infarction, Alzheimer's disease, and non-small cell lung cancer were first established. Associations of miRNAs and their target genes are proposed for developing the basics of the diagnosis of some oncological, neurodegenerative and cardiovascular diseases. miRNA binding sites in mRNAs with free interaction energy equal to or more than -125 kJ/mol and binding sites with cluster organization were found. Fully complementary interactions of miRNAs with the coding region of mRNAs of genes involved in the development of the socially significant diseases were revealed. The binding sites of the miRNAs are conserved in mRNAs of orthologous genes.

**Theoretical significance of the work.** The nucleotide sequences of miRNAs and mRNAs of human genes and their orthologs associated with the development of myocardial infarction, Alzheimer's disease, and non-small cell lung cancer were selected. MiRNAs that bind to CDS mRNAs of genes involved in development of the diseases have been identified. The miRNA binding sites located in mRNA of genes with overlapping of nucleotide sequences (clusters) have been established. Based on the characteristics of the interaction of miRNA with target genes mRNAs, associations involved in the pathogenesis of myocardial infarction, Alzheimer's disease, non-small cell lung cancer are proposed. Evolutionary conservation of completely complementary miRNA binding sites was revealed, which indicates the early emergence of regulation of gene expression by miRNA molecules. The results of an analysis of miRNA binding sites in mRNAs of orthologous genes provide justification for selection of an experimental animal model to confirm the regulation of target genes by miRNAs in the experiment.

**Practical value of the study.** The revealed miRNA binding sites in mRNAs CDS are proposed for further experimental validation of associations of miRNAs and target genes of myocardial infarction, Alzheimer's disease and non-small cell lung cancer. The results were translated into Kazakh, introduced into the educational process and were used in teaching the elective course "Basics of genomics and proteomics" in 3-year undergraduate groups of specialty 5B060700 "Biology" (Appendix A).

**Basic statements for the defense:**

122 coding sequences of mRNAs of the studied genes involved in the development of myocardial infarction, Alzheimer's disease and non-small cell lung cancer are miRNAs targets.

In CDS mRNAs of genes involved in the development of some socially significant diseases, there are single, multiple and polysites for miRNA binding.

Cluster organization of miRNA binding sites in genes mRNAs of the studied diseases leads to compaction of the nucleotide sequence of mRNA, which is the target of several miRNAs, and to the emergence of competition between miRNA molecules for binding with the target gene.

Of the studied 17,508 human genes, 32 have fully complementary miRNA binding sites in CDS of mRNA. The miRNA and mRNA associations of the genes are phylogenetically conservative.

### **The main research results and conclusions:**

1. In the format available for the MirTarget program, a database of 6272 human miRNA nucleotide sequences was created containing 2565 sequences from the miRBase database and 3707 sequences identified by Londin et al. The databases of nucleotide sequences of 75 genes involved in Alzheimer's disease, 237 genes associated with myocardial infarction and 115 genes involved in the pathogenesis of non-small cell lung cancer has been created.

2. The interaction characteristics of 6272 miRNAs and 115 genes involved in the development of non-small cell lung cancer were determined; of which 64 mRNAs have miRNA binding sites in CDS. Of these, associations of 40 miRNAs and 15 target genes were established, the free interaction energy of which is 125 kJ/mol and higher.

3. 23 genes involved in the development of myocardial infarction and having miRNA binding sites in CDS of mRNAs were found. Of these, only one miRNA and one mRNA interact with a high value of free energy equal to 134 kJ/mol.

4. Of the 75 studied genes involved in the development of Alzheimer's disease, 35 genes were identified that have miRNA binding sites in CDS of mRNAs. Of these, associations of 11 miRNAs and 8 mRNAs have free interaction energy of 125 kJ/mol and higher.

5. A change in the number of miR-1322 binding sites in CDS mRNAs of orthologous animal genes was revealed. miR-1322 may be involved in the regulation of expression of candidate genes of some neurodegenerative, cardiovascular and oncological diseases.

6. Of the 17508 mRNAs studied human genes, 32 genes with completely complementary binding sites of the 41 miRNAs from 2565 studied miRNAs in CDS were found, for which phylogenetic conservation was revealed.

All the assigned tasks were completed.

One gene can be a target of many miRNAs, and one miRNA can bind to several alternative genes. Therefore, the identification of the expression level of one miRNA or one gene as markers for diagnosis is extremely difficult. In order to achieve greater adequacy of the molecular diagnosis of diseases, it is worth considering the effect of miRNA on the expression of alternative target genes. Based on the analysis, it is proposed to study the expression of associations of one miRNA with the target gene with a high value of the interaction energy; association of a single target gene with several miRNAs with binding sites in cluster organization; associations of one miRNA and several its target genes.

The idea of using specific miRNAs combinations with standard laboratory tests is aimed at developing early diagnostic methods, increasing the accuracy of diagnosis, as well as selecting and monitoring the effectiveness of the therapy. The use of identified miRNA-mRNA associations will expand understanding of the mechanisms of expression regulation of individual genes involved in the development of Alzheimer's disease, non-small cell lung cancer and myocardial infarction at the post-transcription level.

**Connection with the plan of basic scientific works.** The thesis is performed within projects "Development of a method for early diagnostics of cardiovascular

diseases based on microRNAs and their target genes» № 0115RK00286, «Development of test-systems for early diagnosis of cardiovascular, oncological and neurodegenerative diseases based on the associations of miRNA and their target genes» № 0118RK00034 of Ministry of education and science of the Republic of Kazakhstan.

**Approbation of the work.** Materials of the thesis are reported and discussed at:

- IX International Congress "Biotechnology: State and Development Prospects", Moscow, Russia, 2017;
- IV International Scientific Conference of Students and Young Scientists "Farabi Alemi", Al-Farabi KazNU, Almaty, Kazakhstan, 2017;
- International Conference “Clinical proteomics. postgenome medicine”, Russia, Moscow, 2017;
- International Forum “Biotechnology: State and Development Prospects”, Moscow, Russia, 2018;
- V International Scientific Conference of Students and Young Scientists "Farabi Alemi", Al-Farabi KazNU, Almaty, Kazakhstan, 2018;
- V International Conference “Molecular phylogenetics and biodiversity biobanking”, Moscow, Russia, 2018;
- International conference “Modern molecular-biochemical markers in clinical and experimental medicine”, Prague, Czech Republic, 2018;
- International Congress "Biotechnology: state of the art and perspectives", Moscow, Russia, 2019;
- International Symposium "Cell Symposium on Regulatory RNAs" - Berlin, Germany, 2019;
- VII International Scientific and Practical Conference "Biotechnology: Science and Practice", Sevastopol, Russia, 2019;
- International conference “New research: problems and prospects”, Prague, Czech Republic, 2019;
- International Conference "Modern Science: Problems, Ideas, Trends", Sofia, Bulgaria, 2019;
- International Symposium “EMBO | EMBL Symposium: The Non-Coding Genome”, Heidelberg, Germany, 2019.

**Publications and personal contributions of the author.** The main content of the thesis is reflected in 19 printed works, including 1 article in the international journal with the Impact Factor quoted in Scopus; 5 articles from the list of the Committee for Control in Education and Science; 14 abstracts in materials of international conferences, including 9 conferences in foreign countries.

The author has conducted an analysis of published data by investigated issues, the choice of objects of the study, the definition of aim and tasks, experimental studies and analysis of results of the research, writing, and design of the thesis manuscript by herself.

**The volume and structure of the dissertation.** The thesis is presented on 123 pages and consists of notations and abbreviations, introduction, review of

literature, materials and methods, results and discussion, conclusion and list of references including 178 titles, contains 44 tables, three figures and four appendixes.