ABDUSA DANIELA
University of Academy of Science of Moldova, Functional Genetics Center,
Republic of Moldova, Chisinau, Academy str., 3/2, MD-2028, e-mail: abdusadaniela@yahoo.com

IDENTIFICATION OF GENES INVOLVED IN CARDIOVASCULAR DISEASES

Cardiovascular diseases (CVD) remain the leading cause of death all over the world. According to the World Health Organization (WHO), in 2012, more than 17.5 million people died from CVD [10]. The main forms of CVD are coronary heart disease and stroke, which caused 7.4 million deaths and, correspondingly, 6.7 million deaths [10].

Cardiovascular disease continues to cause a much greater mortality burden among Europeans than any other disease. Despite recent decreases in mortality rates in many countries, every year, CVD causes over 4 million deaths in Europe [6]. The proportion of all deaths that are attributable to CVD is substantially greater among women (51 %) than men (42 %), compared with 19 and 23 %, respectively, for all cancers (fig. 1).

The most up-to-date data on CVD in Europe show that the burden of mortality continues to show large geographic inequalities [6]. Updated data from Denmark and Norway show that they now have among the lowest rates of age-adjusted CVD mortality (180 per 100000 men at all ages, 120 per 100000 women), and Denmark in particular has joined countries, including France, Portugal, the Netherlands, and Spain, with the lowest rates of CVD. The highest rates of CVD mortality were found in the Russian Federation and Belarus for men (915 and 892 per 100000, respectively), and Uzbekistan and Kyrgyzstan for women (662 and 588 per 100000).

Moldova is Europe's leading country in terms of mortality due to vascular disease, registering a high rate of deaths, as cardiovascular and cerebrovascular disease [7]. According to the World Health Organization Mortality Database, available data show that in Republic of Moldova 790 deaths per 100000 men at all ages and 564 deaths per 100000 women were caused by CVDs in 2012. National Bureau of Statistics of Republic of Moldova reports that more 58 % of all deaths were CVDs as the cause in 2014 (http://www.statistica.md/, 2015). Thus, one of the priority research directions of the scientific community is cardiovascular diseases.

The identification of various genes linked with specific disorders has resulted in the development of sets for studying the expression of many related genes. The use of these sets could be helpful to identify genes with altered expression associated with other diseases. This method could attract attention to genes that have never been directly related to the disease studied. The aim of this study was to identify genes specifically linked with CVD conditions.

Materials and methods

Exploratory analysis of genes involved in cardiovascular pathologies was carried out according to the strategy of extraction and expression data analysis developed by our research group within Center of Functional Genetics, University of Academy of Science of Moldova [3–5].

The study material consisted of expression profiles of genes potentially involved in CVDs, from the microarray technology, stored in NCBI-GEO

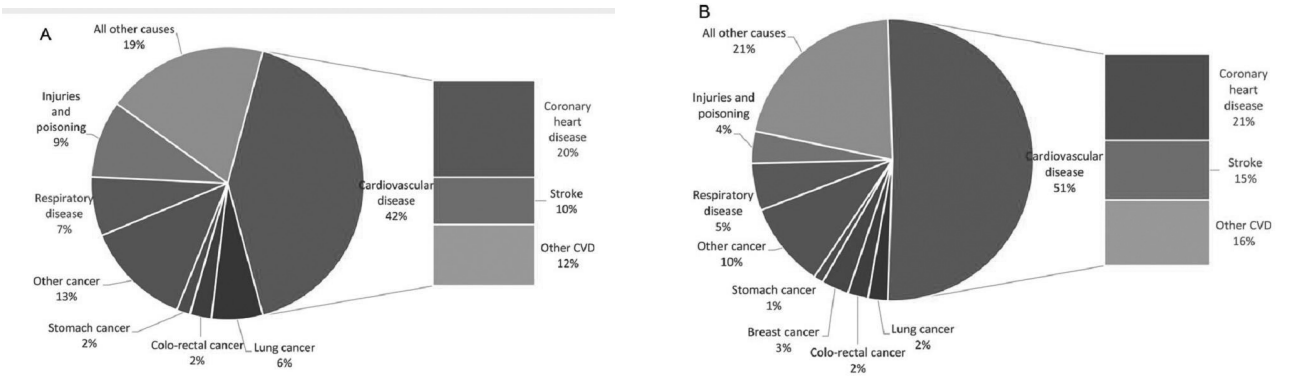

Fig. 1. Proportion of all deaths due to major causes in Europe, latest available year, among men (A) and women (B) [6]

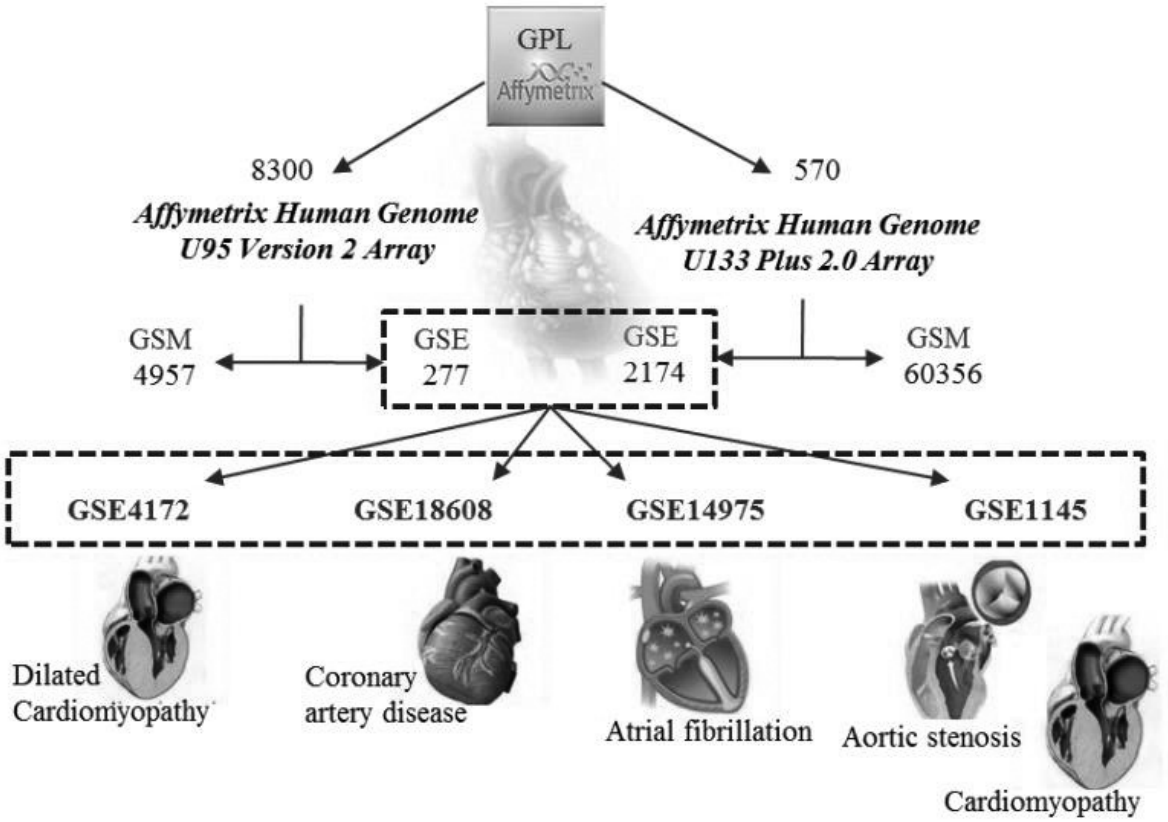

Fig. 2. Workflow for identification of genes involved in CVDs

database. Only datasets refereed to CVDs were used in the analysis.

Dataset extraction, statistical and exploratory analysis were done under R environment [1, 2, 8]. Gene annotations were extracted by NetAffx and GEOquery package [9] from Bioconductor (www.bioconductor.org). For genes without annotation, Probe IDs from the selected chip were used.

In order to identify genes involved in CVDs, two Affymetrix chips (fig. 2) were selected.

Their GPLs include most of all expression datasets from GEO: GPL570 includes 2173 GSEs (series), while GPL8300 – 277 GSEs. For this study GSE4172, GSE14975, GSE18608, GSE1145 were selected as referred to the experiments on the CVD investigation, including cardiovascular pathologies such as: coronary artery disease, atrial fibrillation, cardiomyopathy and aortic stenosis.

Results and discussion

Applying the IE threshold for selection, it was possible to determine a set of 6000 candidate genes hypothetically involved in CVDs. Following the validation of candidate genes set, there were established that 60 genes are primarily involved and 700 genes are potentially involved in CVDs. Following, it was possible to compare candidate gene lists between all pathologis studied. Thus, we obtained 50 genes differentially expressed in CVDs (fig. 3).

Using GeneCards http://www.genecards.org/ and DAVID Bioinformatics Resources https://david.ncifcrf.gov/ was selected 19 genes (table) that are related with important pathway associated with CVDs.

According to KEGG PATHWAY Database (http://www.genome.jp/kegg/pathway.html), the most significant metabolic pathways involving this genes are: Cytokine-cytokine receptor interaction, Chemokine signaling pathway, Phagosome, PI3K-Akt signaling pathway, AMPK signaling pathway, TGF-beta signaling pathway, Focal adhesion, ECM-receptor interaction, Signaling pathways regulating pluripotency of stem cells, etc.

In order to perform expression analysis of genes potentially involved in CVD, we enrolled 55 patients with cardiovascular pathology in the study. The cardiovascular patient group include

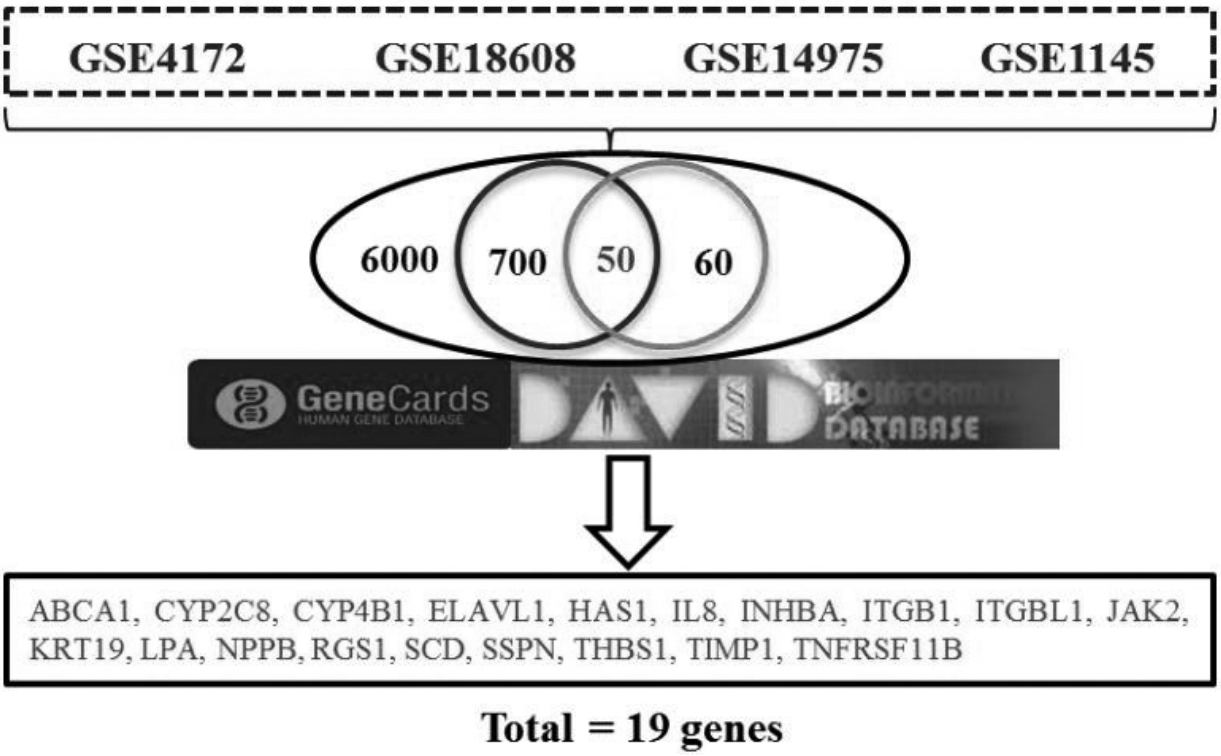

Fig. 3. Genes hypothetically involved in CVDs

Table

Genes involved in cardiovascular diseases

Gene symbol	Entrez_ID	Gene title
ELAVL1[1]	1994	ELAV like RNA binding protein 1
IL8[1]	3576	Chemokine (C-X-C motif) ligand 8
INHBA[1]	3624	Inhibin, beta A
JAK2[1]	3717	Janus kinase 2
HAS1[1]	3036	Hyaluronan synthase 1
SSPN[1]	8082	Sarcospan
KRT19[1]	3880	Keratin 19
TIMP1[1]	7076	TIMP metallopeptidase inhibitor 1
ITGBL1[1]	9358	Integrin, beta-like 1 (with EGF-like repeat domains)
CYP2C8[1]	1558	Cytochrome P450, family 2, subfamily C, polypeptide 8
CYP4B1[1]	1580	Cytochrome P450, family 4, subfamily B, polypeptide 1
TNFRSF11B[1]	4982	Tumor necrosis factor receptor superfamily, member 11b
RGS1[1]	5996	Regulator of G-protein signaling 1
ITGB1[1]	3688	Integrin, beta 1 (fibronectin receptor, beta polypeptide, antigen CD29 Includes MDF2, MSK12)
ABCA1[2]	19	ATP-binding cassette, sub-family A (ABC1), member 1
SCD[2]	6319	Stearoyl-CoA desaturase (delta-9-desaturase)
THBS1[2]	7057	Thrombospondin 1
NPPB[2]	4879	Natriuretic peptide B
LPA[2]	4018	Lipoprotein, Lp(a)

N o t e s: 1 – Genes potentially involved in CVDs; 2 – Genes primarily involved in CVDs.

14 patients with coronary artery disease, 13 patients with coronary artery disease associated with atrial fibrillation, 13 patients with cardiomyopathy and 15 patients with congenital aortic stenosis. The control group include 14 healthy individuals. The subjects enrolled were the in-patients or out-patients in Institute of Cardiology, Chisinau and Institute for Mother and Child Health Care, Chisinau, Republic of Moldova. Currently, it is working on expression data analysis of genes potential involved in CVD obtained by Real-Time PCR.

Conclusions

This study allowed to identify about 19 candidate genes potentially involved in manifestation of CVDs and can be recomanded for evaluation of the expression in patients with coronary artery disease, atrial fibrillation, aortic stenosis and cardiomiopathy. Knowledge about genes with different expression pattern in relation to cardiovascular diseases could potentially lead to new therapeutic targets and give researchers a better understanding of the pathogenesis of CVD.

REFERENCES

1. Arppe A., Milin R.P., Baayen H. Package Naive Discriminative Learning Version 0.1.1. – 2011 [Electronic resource]. – Mode of access: http://cran.r-project.org/web/packages/ndl/ndl.pdf.
2. Baayen H.R. Data sets and functions with „Analyzing Linguistic Data: A practical introduction to statistics". Package 'language R', version 1.4. – Cambridge University Press, 2011. – 134 p.
3. Duca M., Levițchi A., Martea R., Abduşa D., Dragomir L. Metodologia de utilizare a metadatelor experienţelor microarray în elaborarea ipotezelor ştiinţifice pentru cercetarea *Arabidopsis thaliana* // Buletinul Academiei de Ştiinţe a Moldovei. Ştiinţele Vieţii. – 2012. – Nr. 3 (318). – P. 105–112. ISSN 1857-064X.
4. Levitchi A., Abdusa D., Duca M. Exploratory analysis of gene microarray datasets for molecular diagnosis. Scientific annals of the „Nicolae Testemiюanu" State University of Medicine Pharmacy. – 2012, XIIIth edition. – 3. – P. 57–62.
5. Martea R., Abduşa D., Dragomir L., Levițchi A. Analiza datelor microarray pentru evidenţierea aspectelor moleculare legate de caracterele valoroase în ameliorarea plantelor // Culegere de teze. – Chişinău, 2011. – P. 43.
6. Nichols M., Townsend N., Scarborough P., Rayner M. Cardiovascular disease in Europe 2014: epidemiological update // European Heart Journal. – 2014. – 10 p. doi:10.1093/eurheartj/ehu299.
7. Protocol clinic naюional „Accident vascular cerebral ischemic". – Chişinău, 2008. – P. 8.
8. R Development Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Version 2.11.1 (2010-05-31), ISBN 3-900051-07-0.
9. Sean D., Meltzer P.S. GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. Bioinformatics. – 2007. – 23, N 14. – P. 1846–1847.
10. WHO. Global status report on noncommunicable diseases. – 2014. – 298 p. ISBN: 978 92 4 156485 4.

ABDUSA DANIELA

University of Academy of Science of Moldova, Functional Genetics Center,
Republic of Moldova, Chisinau, Academy str., 3/2, MD-2028, e-mail: abdusadaniela@yahoo.com

IDENTIFICATION OF GENES INVOLVED IN CARDIOVASCULAR DISEASES

Aim. The aim of the investigation was to identify genes hypothetically involved in the manifestation of cardiovascular diseases. ***Methods.*** In order to identify genes potential involved in CVDs we conducted exploratory analysis of microarray data sets according to the strategy of extraction and expression data analysis developed by our research group within Center of Functional Genetics, University of Academy of Science of Moldova. ***Results.*** Applying established methodology, a set of 6000 of candidate genes was found. These findings were validated through two sets of genes: cardiovascular priority genes and associated genes. The validation of the candidate genes set established that 90 genes are primarily involved and 450 genes are potentially involved in CVDs. ***Conclusions.*** Through analysis of expression profiles of genes potentially involved in CVDs, from the microarray technology, stored in NCBI-GEO database we obtained 50 genes differentially expressed in CVDs, of which we selected 19 genes for quantitative gene expression studies by Real-Time PCR.

Keywords: Cardiovascular diseases, exploratory analysis, microarray, qPCR.