BIOSCIENCES BIOTECHNOLOGY RESEARCH ASIA, December 2014. Vol. **11**(3), 1267-1272

The Molecular Genetic Study of Genetic Markers of Predisposition to Physical Loading

Elena Vladimirovna Vorobyeva, Anna Stanislavovna Yazeva, Ivan Vladislavovich Nikolaev, Ilnur Bikbaevich Gaizullin and Valentina Yurievna Gorbunova

M. Akhmullah Bashkir State Pedagogical University 3a, Oktyabrskoi Revolutsii street, 450000, Ufa, Russia

doi: http://dx.doi.org/10.13005/bbra/1515

(Received: 30 October 2014; accepted: 05 December 2014)

Polymorphic variants of 6 genes the products of which have an impact on the cardiovascular system have been studied: rs4646994 of angiotensin converting enzyme *ACE* gene, rs1801133 of methylenetetrahydrofolate reductase *MTHFR* gene, rs1800875 of chymase *SMA1* gene, rs6025 of coagulative factor V *F5* gene, rs1815739 of alpha-actinin *ACTN3* gene, rs121913641 of fibrous protein *MYH7* gene. Alleles combinations of studied genes that have a significant impact on the physical performance level of an individual with minimal adverse effects have been determined. Research paper also shows the potential of using haplotype analysis in the sports teams building for them to achieve high results without any disablement.

Keywords*:* athletes, cardiovascular system, haplotype, protein conformation, transcription factor.

One of the intensively developing areas of modern genetics is the development of molecular genetic approaches that make it possible to determine a person's predisposition to various activities, including the search for molecular genetic markers that determine a person's ability to perform high sport loads[1-3]. We have previously established the high heritability of characteristics: "the overall physical performance" -0.82 and "maximum oxygen consumption" - 0.72[4].

It is well known that professional sport makes high demands of the body, as it causes systems overstrain, which can lead to the serious diseases, especially in the presence of latent genetic predisposition. The most common complications are disturbances of the cardiovascular system (CVS), which may be the cause of oxymortia due to coronary vessels pathology and heart function disorders in case of the physical stress. On the basis above, studying of the individual characteristics of a person regarding the performance of maximum physical loads without any damage to health taking into account its genetic constitution is of special importance.

The purpose of this research work is the molecular genetic study of functioning of polymorphic genetic markers that control the cardiovascular system at high physical loading.

The following tasks have been set in accordance with the purpose:

1. Analyze the frequencies of genotypes and alleles of genes polymorphic variants Which regulate the performance of systems: the

* To whom all correspondence should be addressed.

renin-angiotensin rs4646994 of *ACE* gene, rs1800875 of *SMA1* gene, blood clotting: rs6025 of *F5* gene, rs1801133 of *MTHFR* gene, vascular tone: rs1815739 of *ACTN3* gene, rs121913641 of *MYH7* gene among athletes and in the control group.

2. Study the possibility of using haplotype analysis of alleles of *F5* and *MTHFR* genes, and also *CMA1* and *MYH7* in the selection of individuals for the professional sports.
3. Analyze the combination of alleles and evaluate intergenic interactions of genes that determine the performance of systems: the renin-angiotensin, blood clotting and vascular tone of athletes.
4. Identify the possible risk combinations of typed genes alleles that could lead to the development of cardiovascular diseases.
5. Conduct bioinformatic analysis of the protein products of the studied genes.

Methodology

We used DNA samples obtained from the whole venous blood of 570 people living in the Republic of Bashkortostan. The studied samples was divided into groups according to the level of sports achievements:

1) Candidate Masters of sports, Masters of sports, International Masters of sports;
2) second-class athletes that were engaged in sports activities for a long time, have been taken as the control group. All studies have been conducted at the Center for molecular genetic testing of M. Akmullah BSPU.

Cardiovascular system, which in many ways determines the achievement of high sports performance, is strictly controlled by the genes shown in Table 1.

Extraction of DNA was performed through the Mathew phenol-chloroform extraction[11], then Mullis method of polymerase chain reaction[12] was performed, Maniatis method of electrophoresis in 7% polyacrylamide gel[13]. Restriction fragment lengths polymorphism was determined through Langdahl method[14].

Data processing was carried out using the software package "Statistica for Windows 6.0" (StatSoft, 2007), MS Excel 2003 software (Microsoft, 2002). In assessing disbalance of linkage of each pair of polymorphic loci, coefficient D' calculated in the program "2LD" was used[15]. Haplotype frequencies of genes linked loci were calculated in the program "EH"[16]. In the analysis of genes interactions used for modeling of gene-gene and gene-environment interactions the program "MDR 2.0" was used[17]. The differences were considered significant at a significance level of $P < 0.05$. The bioinformatic analysis of the effect of studied polymorphic loci mutations on the functioning of CVS was carried out. The nucleotide sequences of typed polymorphic loci were analyzed for sites of alternative splicing (SAS) and the transcription factors binding sites (TFBS) and their changes under the influence of SNP-mutations. Nucleotide sequences analysis was carried out for both the normal and mutant allele locus. Nucleotide sequences were obtained from the database (DB) "GeneBank" (www.ncbi.nlm.nih.gov). The search of TFBS was carried out in the program "TFSCAN" (mobyle.pasteur.fr/cgi-bin/portal.py#forms::tfscan). For the search of SAS, "Alternative Splice Site Predictor" program was used (wangcomputing.com/assp/index.html).

Analysis of the impact of the studied mutations on genes protein products was carried out in two ways: on normal and mutant forms of the protein. Information about the amino acid (AA) sequence of protein as well as about constituent structural and functional domains was obtained from the "Universal Protein Resource" database (uniprot.org). Location of the intronic mutations in the *ACE* gene against the structure of the protein was determined by the translation of the nucleotide sequence of the gene encoding exons into the AK sequence till the border with the target mutation in "EMBOSS Transeq" program (www.ebi.ac.uk/Tools/st/emboss_transeq). Then the multiple alignment in "ClustalW" program of software complex «BioEdit 7.0.9.0» (BLOSUM62 alignment matrix) was performed. Position of the end of the translated sequence against the sequence of canonical protein isoform was considered to be the site of the likely impact of mutation on the protein structure. Assessment of the physical-chemical properties of protein structures was carried out in "ProtParam" program (web.expasy.org/protparam). The criteria for changing the physical-chemical properties were considered to be the changes in molecular mass, isoelectric point, aliphatic index and instability index of the protein. Spatial modeling of the protein

structures domains was held in "Schrödinger Suite 2013" (Schrödinger, 2013). Template files for the modelling were obtained from "Protein Data Bank" DB (pdb.org). Conformation change in the protein structures were analyzed in "Vadar 1.8" program (vadar.wishartlab.com/index.html). Criteria of information changing were considered to be the changes in total protein domain volume and change of its available area.

Main part

The basis of a genetic predisposition to the execution of high physical loading at a

Table 1. Genes controlling the work of the cardiovascular system

Gene/Polymorphism	Function of genes alleles	Ref.
(location)	(fragments length, restriction endonuclease)	
ACE (17q23)	**I* – hypoactivity of angiotensin converting enzyme in	[5]
InDel *Alu*	blood and tissues, wide vessels lumen (490 bps);	
(rs4646994)	**D* – increase in angiotensin II formation in vascular	
15 intron	endothelium, leads to luminal occlusion (192 bps).	
CMA1 (14q11.2)	**G* – normal strength of angiotensin II formation	[6]
G1903A	catalyzing enzyme (195+90 bps; BstXI);	
(rs1800875)	**A* – enzyme strength is increased, angiotensin II	
Promoter	accumulation, whereupon the ingress of oxygen through	
	heart vessels decreases and vessel constriction (285 bps) as a result.	
F5 (1q23)	**G* – normal strength of enzyme (130+30 bps; MnlI);	[7-8]
G1691A (rs6025)	**A* – high strength of enzyme, tendency to the	
10 exon	development of vascular thrombosis (160 bps).	
MTHFR (1p36.3)	**A* – normal strength of enzyme (220 bps);	[9]
C677T	**T* – deprivation of enzyme, associated with the	
(rs1801133)	predisposition to thromboses, development of	
4 exon	homocysteinemia (170+50 bps; HinfI).	
ACTN3 (11q13-q14)	**C* – presence of α-actinine-3 in skeletal muscles,	[2]
C1729T	improves speed-power physical qualities (205+85 bps);	
(rs1815739)	**T*– defective protein variant, its absence in skeletal	
16 exon	muscles (85+97+108 bps; BstDEI).	
MYH7, 14q12	**G* – synthesis of normally functioning protein of β-	[10]
G14692A	heavy chain of the heart myosin (156+51 bps; MspI);	
(rs121913641)	**A* – synthesis of protein with the changed conformation	
19 exon	β-heavy chain of the heart myosin, leads to the	
	compensatory enlargement of the cardiac muscle (207 bps).	

Table 2. Changing of proteins properties and structure in case of mutations

Protein structure (gene allele)	Molecular mass (Dalton)	Available area (A^2)	Total volume (A^3)	Theoretical pI (pH)	Instability index	Aliphatic index	Template file (homology %)
ACE (**D*)	149714.8	4727.2	6477.1	5.95	43.38	81.01	4APH(100%)
ACE (**I*)	152271.6	6582.0	9637.7	6.04	43.61	80.15	
F5 (**G*)	251703.4	5956.6	6560.4	5.68	48.94	73.09	1FV4(100%)
F5 (**A*)	251604.2	5820.0	6451.8	5.66	48.74	73.09	
MTHFR(**C*)	74596.5	4356.7	6676.4	5.22	49.02	80.72	1V93(100%)
MTHFR(**T*)	74624.6	4347.0	6692.9	5.22	49.05	81.01	
ACTN3*(*C)*	103241.2	4897.1	5813.0	5.37	44.10	84.84	1HCI(66.7)
ACTN3 *(*T)*	66474.7	2634.4	2891.4	5.55	44.14	86.79	
MYH7 *(*G)*	223097.3	4717.0	6282.8	5.63	47.05	82.19	1M8Q(78.4)
MYH7 *(*A)*	223069.2	4690.9	6223.1	5.61	47.05	82.19	

professional sports activities are the complex interrelationships of different genetic systems, analysis of which was carried out in "MDR 2.0" program (Fig. 1).

Analysis of the six loci model of genes interactions of athletes showed several risk combinations of genes alleles:

*ACE*D/*D//CMA1*A/*A//MTHFR*T/*T //F5*A/*A//MYH7*A/*A//ACTN*T/*T*; *ACE*I/*D//CMA1*A/*G//MTHFR*C/*T//F5*A/*A//MYH7*A/*A//ACTN*T/*T*; *ACE*I/*D//CMA1*A/*A//MTHFR*T/*T//F5*A/*A//MYH7*G/*A//ACTN*T/*T*. Interaction of the products of mutant genes alleles that determine blood clotting *MTHFR* (rs1801133) and *F5* (rs6025), increase the risk of thrombosis caused by the accumulation of homocysteine , violation of the elasticity of intravascular lining and intravascular activation of procoagulant factor V. Interaction of genes alleles of the renin-angiotensin system of ACE (rs4646994) and CMA1 (rs1800875) causes luminal occlusion due to the accumulation of angiotensin II in the blood, the contractile function of the heart is disturbed. Herewith the nonprotective alleles of *MYH7* and *ACTN3* genes have a negative redundant impact on the muscle tissue deformation, i.e. the absence of alpha-actin 3 breaks the connection of actin and myosin. Disturbance of hemostasis (*MTHFR* gene) and luminal occlusion (*CMA1* gene) increased by the compensatory exposure to hypertrophic changes in the heart muscle (*MYH7* gene) leads to the malfunction of the CVS. The significant increase in the frequency of *MYH7*G* allele (P=0.02; $[Chi]^2$ = 5.39) and *MYH7*G/*G* genotype frequencies has been detected in the group of athletes with high qualification (P=0.003; $[Chi]^2$=9.46), as well as significant decrease of *G/*A genotype frequency (P = 0.024; $[Chi]^2$=5.7). This is due to the fact that there is the allele *MYH7*A* in a genotype, which may lead to changes in conformation of protein of the bheavy chain of heart myosin. Probably there occurs the elimination of the high classification of genotype *MYH7*A/A** from the athletes environment due to the compensatory enlargement of the cardiac muscle during the execution of high physical loading in training. The *MYH7** A allele frequency in the control group is significantly higher than the frequency of the same allele (P=0.02; $[Chi]^2$=5.39).

Linkage analysis of 4 alleles was performed: *F5* (rs6025) and *MTHFR* (rs1801133) on the first chromosome and *SMA1* (rs1800875) and *MYH7* (rs121913641) on the 14th chromosome. Loci linkage analysis of these genes in the studied sample showed linkage disequilibrium: (D'=0.7). In case of the pair-wise comparison of *F5* and *MTHFR* gene haplotypes there was a significant increase of *F5*C/MTHFR*A* haplotype in the comparison group (P=0.02; $[Chi]^2$=5.8), and *F5*T/MTHFR*A* risk haplotype was often detected in the group of athletes with high athletic achievements with a frequency of 28,8% (P=0.001; $[Chi]^2$=11.78). According to the literature, *F5*T* allele is associated with the larger cross-sectional area of muscle fiber and their hypertrophy[18].

Linkage analysis between rs1800875 *CMA1* (14q11.2) and rs121913641 *MYH7* (14q12) gene polymorphic loci detected the high level of disequilibrium in linkage: D'= 0.9. Fairly significant differences were detected only on haplotype *CMA1*A/ MYH7*A* (P=0.04; $[Chi]^2$=4.02), carrying nonprotective alleles of polymorphic loci rs1800875 of *CMA1** A gene and rs121913641 of *MYH7** A gene. Co-inheritance of these alleles leads to the adverse functional disturbances in the CVS specifically during the execution of high physical loading, and the frequency of this haplotype in the group of athletes is significantly low.

Bioinformatic analysis revealed differences between normal and mutant alleles of studied genes loci. For mutant and normal forms of their protein products we determined physical-chemical properties for a mutant and normal forms of their protein products as the change of corresponding domains as a result of mutation.

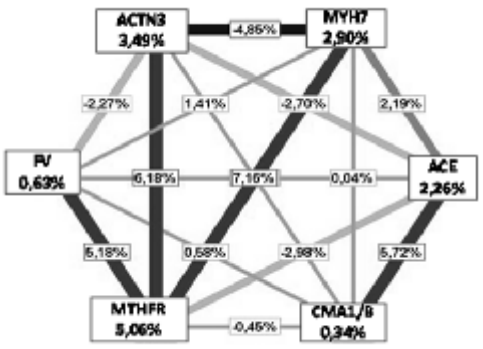

Fig. 1. Graph of loci nonprotective alleles interaction of the studied genes of athletes

These data as well as the template files used for modeling are shown in Table 2.

CONCLUSION

The problem of individual selection in sport is particularly acute, as in the last decade there is an increased incidence of cardiovascular diseases, and high physical activity is an additional factor that has a negative impact on the normal functioning of heart and blood vessels. Therefore the search for molecular markers of predisposition to cardio-pathology at athletes is a priority for research in the field of genetics sport. Works in this direction are of considerable practical importance since they can be applied to the design of algorithms of sports selection, building of sports teams and the personification of the training process, which will ensure the achievement of high sports results without athletes' health damage and to help to avoid their early incapacitation.

The conducted research allowed us to study the particularity of the effect of polymorphic loci mutations of *ACE*, *CMA1*, *F5*, *MTHFR, ACTN3, MYH7* genes on the CVS in determining the predisposition to high sports achievements. There were established the combinations of genotypes of studied genes loci that have a negative impact on CVS and the level of sports achievements. Using the bioinformatic methods we studied conformational changes of protein domains under the influence of typed loci mutations, as well as changes in the physical-chemical properties of these proteins caused by these mutations. Irregularities in the cardiovascular system defined by typed loci mutations are the result of the destabilization of regulation mechanism of this system, important for human, caused by an improper ratio of protein products involved in its work. CVS determines many aspects of human activity, including achievements in sports, so the study of molecular and genetic mechanisms of its functioning is a priority for a more precise and correct selection of candidates for professional sports.

Summary

1. There is a significant increase in the frequency of the homozygotic **G/*G* genotype on *F5* gene in professional athletes, which determines the proper functioning of the coagulation V factor.
2. It has been demonstrated that the protective **C*G* haplotype on *MTHFR* and *F5* genes, may be a marker of predisposition to the execution of high physical loading while doing the professional sports.
3. Complex analysis (molecular-genetic and clinical) revealed a combination of major *F5, MTHFR, ACE, CMA1* alleles, predisposing to execute physical loading: **C*G*I*G.*
4. It is found that at high physical loading, combination of *F5, MTHFR, ACE, CMA1*:**T*A*D*A* genes alleles is risky towards the development of cardiovascular diseases.
5. Analysis of gene interactions revealed the risk genotypes for the execution of high physical loading. These include the following combinations: *ACE*D/*D//CMA1*A/*A//MTHFR*T/*T// F5*A/*A; ACE*I/*D//CMA1*A/*G// MTHFR*C/*T//F5*A/*A; ACE*I/*D// CMA1*A/*A//MTHFR*T/*T//F5*A/*A.*
6. The modeling by means of bioinformatic methods was carried out and possible physical-chemical and conformational changes of gene products of typed genetic markers were identified.
7. There were revealed the significant associations of blood circulation durability ratio in case of aerobic activities with **I* (χ^2=6,29; P=0,01) allele and homozygous **I/*I* (χ^2=4,15; P=0,04) genotype on *ACE (rs4646994)* gene, which indicates the mobilization of blood systems of oxygen capture, transport and release to tissues in execution of high physical loading.

Credits

The project was performed within the framework of the Competition of fundamental and exploratory scientific researches in order to create the scientific reserve for the university teachers.

REFERENCES

1. Montgomery, H., P. Clarkson, M. Barnard *et al.*, Association of angiotensin-converting enzyme gene I/D polymorphism with change in left ventricular mass in response to physical training. *Circulation*. 1998; **96**: 741-747.
2. Rogozkin, V.A., I.B. Nazarov and V.I. Kazakov,

Genetic markers of a human's physical capability. *Theory and practice of physical education*. 2000; **12**: 33-36.

3. Lekontsev E.V., Kayumova L.R., Timkova A.V. and others, Methodology of children selection to sports groups: molecular-genetic aspects. *Agricultural Russia*. 2009; **1**:124-125.
4. Gumerova O.V., Stolbova O.V., Zaripova T.Yu., and others, Molecular-genetic analysis of associations of genes polymorphic markers of neurotransmitter systems with a human's development level. BSU reporter. 2007; **1**: 39-49.
5. Regat, B., C. Hubert, F. Allene-Gelas, *et al.,* An insertion/deletion polymorphism in the angiotensin I converting enzyme gene accounting for half the variance of serum enzyme levels. *The Journal of Clinical Investigation.* 1990; **86**: 1343-1346
6. Weidinger, S., L. Rümmler, N. Klopp, *et al.,* Association study of mast cell chymase polymorphisms with atopy. *Allergy.* 2005; **60**: 1256-1261.
7. van Dunne, F., C. Doggen, M. Heemskerk, *et al.,* Factor V Leiden mutation in relation to fecundity and miscarriage in women with venous thrombosis. Human Reproduction: *Oxford Journal.* 2005; **20**: 802-806.
8. Khan, S. and J. Dickerman, Hereditary thrombophilia. *Thrombosis Journal*. 2006; **4**: 15.
9. Böger, C., M. Stubanus, T. Haak, *et al.,* Effect of *MTHFR C677T* genotype on survival in type 2 diabetes patients with end stage diabetic nephropathy. Nephrology Dialysis Transplantation: *Oxford Journals.* 2006; **22**: 154-162.
10. Anan, R., G. Greve, L. Thierfelder, *et al.,* Prognostic implications of novel beta-cardiac myosin heavy chain gene mutations that cause familial hypertrophic cardiomyopathy. The *Journal of Clinical Investigation.* 1994; **93**: 280-285
11. Mathew, C.C., The isolation of high molecular weight eukaryotic DNA. Methods in molecular Biology. N.Y.: Humana press, 1984; **2**: 31-34.
12. Mullis, K.B., R.K. Saiki, S. Scharf, *et al.,* Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science*. 1985; **230**: 487-491.
13. Maniatis, T., E. Fritsch and J. Sambrook, Molecular cloning (Methods of genetic engineering). M.: «Mir». 1984; 220-228.
14. Langdahl, B.L., S.H. Ralston, S.F. Grant, *et al.,* An Spl binding site polymorphism in the *COLIA1* gene predicts osteoporotic fractures in both men and women. *Journal Bone and Mineral Research*. 1998; **13**: 1384-1389.
15. Zapata, C., C. Carollo, S. Rodriguez, Sampling variance and distribution of the D measure of overall gametic disequilibrium between multiallelic loci. *Annals of human genetics.* 2001; **65**: 395-340.
16. Xie, X. and J. Ott, Testing linkage disequilibrium between a disease gene and marker loci. *The American Journal of Human Genetics*. 1993; **53**: 1107.
17. Lou, X., G.B. Chen, L.Yan, *et al.,* A generalized combinatorial approach for detecting gene by gene and gene by environment interactions with application to nicotine dependence. *The American Journal of Human Genetics*. 2007; **80**: 1125-1137.
18. Ahmetov, I. I., N. V. Makarova, E. A. Khismatullina, *et al.,* The *MTHFR* gene *C677T* polymorphism is associated with athlete status and muscle fiber hypertrophy. *European Journal of Human Genetics.* 2012; **20**: 236.