ISSN 1022-7954, Russian Journal of Genetics, 2012, Vol. 48, No. 10, pp. 1035–1047. © Pleiades Publishing, Inc., 2012.
Original Russian Text © E.A. Trifonova, M.G. Spiridonova, T.V. Gabidulina, F.D. Urnov, V.P. Puzyrev, V.A. Stepanov, 2012, published in Genetika, 2012, Vol. 48, No. 10, pp. 1207–1220.

HUMAN GENETICS

Analysis of the *MTHFR* Gene Linkage Disequilibrium Structure and Association of Polymorphic Gene Variants with Coronary Atherosclerosis

E. A. Trifonova[a], M. G. Spiridonova[a], T. V. Gabidulina[b], F. D. Urnov[c], V. P. Puzyrev[a], and V. A. Stepanov[a, d]

[a] *Research Institute of Medical Genetics, Russian Academy of Medical Sciences, Tomsk, 634050 Russia; e-mail: ekaterina.trifonova@medgenetics.ru*

[b] *Siberian State Medical University, Department of Obstetrics and Gynecology, Tomsk, 634050 Russia*

[c] *University of California, Department of Molecular and Cell Biology, Berkeley, CA, 94720 USA*

[d] *Tomsk State University, Department of Cytology and Genetics, Tomsk, 634050 Russia*

Received January 13, 2012; in final form, April 16, 2012

Abstract—Analysis of the genome-specific linkage disequilibrium patterns in certain populations is a highly promising approach to the identification of functional variants that underlie susceptibility to complex diseases. In the present study, the linkage disequilibrium patterns of the methylenetetrahydrofolate reductase gene (*MTHFR*) were examined in a group of patients with coronary atherosclerosis (coronary artery disease, CAD) and in a control sample from the Russian population. It was demonstrated that in the samples from one population, which were differentiated by the presence or absence of CAD, the *MTHFR* linkage disequilibrium patterns had similar features. Association of the *MTHFR* rs7533315 and rs2066462 polymorphisms with CAD was demonstrated. In addition, the evolution of the haplotypes and their role in the formation of CAD in the Russian population was reconstructed. The data on the association between genetic variability in the *MTHFR* locus and pathogenetically important indices of lipid metabolism were obtained. The high informativeness of the haplotype approach in case–control tests for associations with CAD was demonstrated.

DOI: 10.1134/S1022795412100122

INTRODUCTION

The search for cardiovascular disease candidate genes started about 20 years ago. At that time it was considered that genetic polymorphisms, along with the already known factors that predispose a person to cardiovascular disease, would be very promising for the evaluation of individual risk and the elaboration of new means of prevention and treatment. However, despite years of intensive studies, at present, none of the genetic risk factors is widely used in clinical practice. In time, the focus of genetic investigations of common cardiovascular diseases gradually shifted from primary risk stratification to a deeper understanding of the etiology and pathophysiology of the disease and their diagnostic and therapeutic sense [1].

Unraveling the genetic determinants of common diseases is one of the key frontiers in modern human genetics. The classical approach to this problem, which is based on case–control disease–marker association studies, still remains in force. However, in recent decade, the methods of genetic analysis of multifactorial diseases were supplemented with highly effective approaches, such as whole-genome mapping and meta-analysis, as well as cohort and multiple replicative studies [2–5]. Analysis of linkage disequilibrium patterns in the candidate gene regions and identification of the disease-associated haplotypes and their tagSNPs is one of the most promising strategies for the identification of the genetic variants that predispose a person to complex diseases [6–9].

The effects of the individual markers that are identified by classical association analysis are usually low. Furthermore, these effects can be associated with a linked functionally important variant (mutation or polymorphism) rather than with the marker of interest. For these reasons, association analysis at the haplotype level can be a more powerful and informative tool than the analysis of individual markers. If the candidate genes play a definite role in disease susceptibility, some suggestions can be made. First, certain haplotypes may demonstrate more significant associations and higher relative risk compared to marker alleles. Second, linkage disequilibrium patterns in patients may be more distinct compared to either control or population samples. Third, different ethnic populations can be different with respect to their linkage disequilibrium patterns and haplotype–disease associations.

To test this hypothesis, in the present study, haplotype structures at the methylenetetrahydrofolate

Table 1. Characteristics of the *MTHFR* SNPs we examined

No	SNP ID (generally accepted name)	Chromosomal position (according to the NCBI database)	dbSNP alleles	Ancestral	Mutation type	Location in the *MTHFR* gene (according to the NCBI database)
1	rs3753588	11863904	*A/G*	*G*		Intron 1
2	rs2066470	11863057	*C/T*	*C*	Synonymous (39 Pro/Pro)	Exon 2
3	rs17037397	11862163	*A/C*	*C*		Intron 2
4	rs7533315	11860683	*C/T*	*C*		Intron 3
5	rs4846052	11857951	*C/T*	*T*		Intron 4
6	rs1801133 (C677T)	11856378	*C/T*	*C*	Nonsynonymous (222 Val/Ala)	Exon 5
7	rs6541003	11855867	*A/G*	*G*		Intron 5
8	rs2066462	11854896	*C/T*	*C*	Synonymous (352 Ser/Ser)	Exon 7
9	rs1801131 (A1298C)	11854476	*A/C*	*A*	Nonsynonymous (429 Ala/Glu)	Exon 8
10	rs17375901	11852516	*C/T*	*C*		Intron 9
11	rs2274976 (G1793A)	11850927	*A/G*	*G*	Nonsynonymous (594 Gln/Arg)	Exon 12
12	rs1537516	11847861	*C/T*	*C*		(3'UTR)

reductase gene (*MTHFR*) locus were examined in populations of different ethnic origins and in patients with coronary atherosclerosis (coronary artery disease).

MATERIALS AND METHODS

In this study, a group of CAD patients and a control sample were examined. The patient group was comprised of unrelated Caucasoid individuals (128 males and 13 females) with coronary artery disease, who were observed at the Research Institute of Cardiology of the Russian Academy of Medical Sciences (Tomsk). The mean patient age to the moment of observation was 49 ± 8 years. The presence of coronary artery stenosis was confirmed by angiography. In most of the patients, the minimum degree of arterial stenosis corresponded to grade II lesions, according to the generally accepted classification of stenosis (stenosis of more than 50%). In 87% of the patients that were observed, the first clinical traits of ischemic heart disease appeared at an age below 55 years. In these cases, the pathology was considered as early-onset disease. The control group was comprised of 126 unrelated individuals (123 males and 3 females) of Caucasoid ethnicity (the mean age at the moment of observation was 41 ± 8 years) with no clinical symptoms of cardiovascular abnormalities. The paraclinical characteristics that were examined in the individuals tested included total serum cholesterol and triglyceride values, systolic (SBP) and diastolic (DBP) blood pressure, and anthropometric indices. Genealogic data were obtained based on family history and the clinical and anthropometric data were based on clinical records.

The molecular markers that were used consisted of 12 *MTHFR* SNPs (single nucleotide polymorphisms), including rs3753588, rs2066470, rs17037397, rs7533315, rs4846052, rs1801133 (C677T), rs6541003, rs2066462, rs1801131 (A1298C), rs17375901, rs2274976 (G1793A), and rs1537516. These markers are relatively randomly distributed over all exons, introns, and the 3' noncoding region of the *MTHFR* gene. For most of the loci tested, the frequency of the minor allele was about 5% (according to the NSBI database). The locations of the polymorphic loci that were examined in the *MTHFR* gene are presented in Table 1. Genotyping was performed using the polymerase chain reaction (PCR) and restriction fragment length polymorphism (RFLP) was performed according to previously described protocols [10–12].

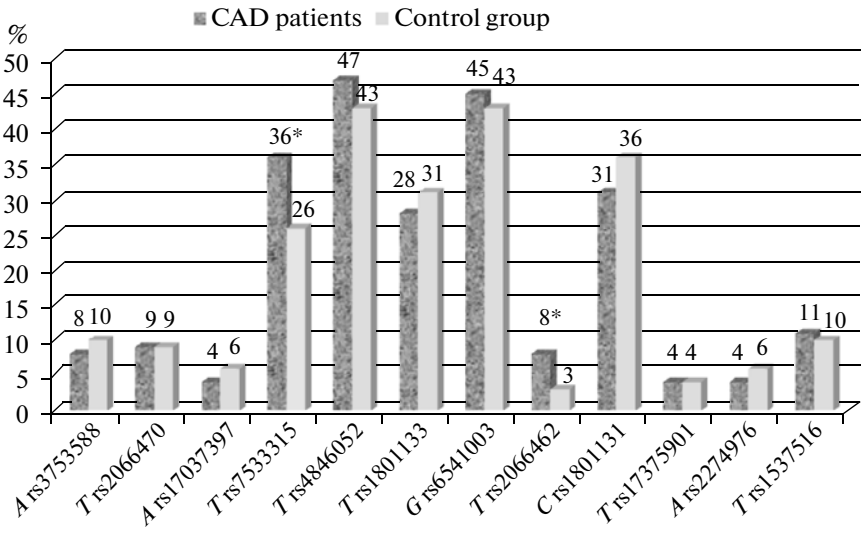

Fig. 1. Minor allele frequency distributions in CAD patients and the control group. The statistically significant differences ($P < 0.05$), which were obtained upon comparison of allele frequencies in CAD patients and the control group, are designated by *.

Statistical treatment of the data was performed using the Statistica 7.0, ARLEQUIN, and Haploview 4.0 software programs. The data distribution patterns were evaluated using Kolmogorov–Smirnov statistics. The role of polymorphic haplotypes in quantitative trait variability was determined using the Kruskal–Wallis test [13]. For the polymorphic variants in the groups that were examined the odds ratio (OR) and the confidence interval (CI) for the odds ratio (at the 95% CI) were calculated. Haplotype frequencies were determined using the EM algorithm. Linkage disequilibrium between the SNP pairs was evaluated using the Lewontin's coefficient D' and Pearson's correlation coefficient r^2. Haplotype block structure was determined using the solid spine option [14], as implemented in the Haploview 4.1 software program with the preset threshold level of $D' \geq 0.75$. The level of genetic diversity and interpopulation differentiation were calculated using the analysis of molecular variance (AMOVA). The selective neutrality of polymorphism was examined using the Ewens–Watterson test [15].

RESULTS AND DISCUSSION

The allele and genotype frequencies for all polymorphisms examined in the CAD patients and control group are shown in Table 2 and Fig. 1. In both groups, at nearly all loci no deviation from the Hardy–Weinberg equilibrium was observed (the exceptions were the rs4846052 and rs6541003 polymorphisms in the group of CAD patients). The possible cause of the deviation could be either the small sample size or the linkage of SNPs that were analyzed with selectively important markers.

The Selective Neutrality of the MTHFR Polymorphisms

The selective neutrality of the *MTHFR* polymorphisms was evaluated using the Ewens–Watterson test [15]. Among all the SNPs that were examined, deviation from neutrality in both the control sample and in the CAD patients was observed only for two polymorphic variants, viz., rs4846052 and rs6541003. All three functionally important SNPs that determine an elevated homocysteine level in human plasma were found to be selectively neutral. Analogous data were obtained by Spirovski (2008) for a Macedonian population. Specifically, the C677T and A1298C polymorphisms were selectively neutral in both the venous thrombosis patients and the control group [16]. These findings can be explained in terms of the suggestion that some phenotypic changes can be selectively neutral if they do not affect the effectiveness of reproduction. At the same time, data on the selection of the 677*T* allele and accumulation of the 677*TT* homozygotes in Spanish population exist. These data are based on an analysis of the changes in the allele and genotype frequency distribution patterns of this SNP during the 20th century. The authors of the study explained the data were obtained in terms of the increase in the intake of folic acid by women during the periconceptional period in the last quarter of the 20th century, which resulted in the increased viability of 677*TT* fetuses during early stages of embryonic development [17]. The selective importance of the 677*T* allele was confirmed in a study that analyzed the allele and genotype frequency distributions of five *MTHFR* polymorphisms in the populations of Israel, Japan, and Africa. In this study, it was demonstrated that the 677*T* allele was found in haplotypes that had selective advantages [18].

Table 2. The genotype distribution of *MTHFR* polymorphisms in CAD patients and the control group

No	SNPs examined	Genotype	Frequency, %		χ^2 value with Yates' correction (*P* significance level)*	Odds ratio—OR, (95% confidence interval)**
			Control group (*N* = 126)	CAD patients (*N* = 141)		
1	rs3753588	*AA*	1	–	0.34 (*P* = 0.561)	0.80 (0.43–1.50)
		AG	18	16		
		GG	81	84		
2	rs2066470	*CC*	83	84	0.26 (*P* = 0.607)	0.97 (0.51–1.82)
		CT	16	15		
		TT	1	1		
3	rs17037397	*AA*	–	–	0.51 (*P* = 0.474)	0.69 (0.31–1.55)
		AC	11	8		
		CC	89	92		
4	rs7533315	*CC*	53	38	**4.38 (*P* =0.036)**	1.60 (1.08–2.36)
		CT	42	52		
		TT	5	10		
5	rs4846052	*CC*	30	21	2.68 (*P* = 0.101)	1.19 (0.83–1.70)
		CT	53	63		
		TT	17	16		
6	rs1801133 (C677T)	*CC*	55	52	0.54 (*P* = 0.763)	1.14 (0.78–1.67)
		CT	39	40		
		TT	6	8		
7	rs6541003	*AA*	29	26	0.08 (*P* = 0.774)	1.09 (0.76–1.56)
		AG	56	59		
		GG	15	15		
8	rs2066462	*CC*	94	84	**4.82 (*P* = 0.028)**	2.71 (1.13–6.72)
		CT	6	16		
		TT	–	–		
9	rs1801131 (A1298C)	*AA*	40	48	1.12 (*P* = 0.290)	0.80 (0.55–1.17)
		AC	48	42		
		CC	12	10		
10	rs17375901	*CC*	91	92	0.01 (*P* = 0.958)	0.98 (0.38–2.55)
		CT	9	8		
		TT	–	–		
11	rs2274976 (G1793A)	*AA*	–	–	1.78 (*P* = 0.182)	0.35 (0.13–0.90)
		AG	13	7		
		GG	87	93		
12	rs1537516	*CC*	82	80	0.06 (*P* = 0.810)	1.12 (0.62–2.03)
		CT	17	18		
		TT	1	2		

Note: *N*, number of individuals in the group.

* χ^2 values with the Yates correction and *P* significance levels were obtained through comparison of genotype frequencies in CAD patients and the control group. Statistically significant differences ($P < 0.05$) are in bold type.

** odds ratios and confidence intervals are shown for minor alleles, shown in Fig. 1.

The Association of MTHFR Alleles and Genotypes with Coronary Atherosclerosis

Comparison of the genotype and allele frequency distributions in the control group and CAD patients revealed statistically significant differences at two *MTHFR* polymorphisms (rs7533315 and rs2066462). Specifically, the frequencies of the *T* allele and *CT* and *TT* genotypes were statistically significantly higher in the group of CAD patients (Table 3). For all markers examined, the odds ratio for the minor allele varied from 0.35 to 2.71 (Table 2). Analysis of the odds ratio confirmed a possible association between the genetic variation of rs7533315 and rs2066462 and hereditary susceptibility to CAD (OR = 1.60; 95% CI: 1.08–2.36; and OR = 2.71; 95% CI: 1.13–6.72, respectively).

In the present study, no statistically significant association of CAD with any of the three most thoroughly studied nonsynonymous substitutions in the *MTHFR* gene (C677T, A1298C, and G1793A) was observed. However, CAD was reported to be associated with the rs7533315 polymorphism, which is located in the third intron of the *MTHFR* gene, as well as with synonymous substitution in exon seven, rs2066462. It can be suggested that the effect of the *MTHFR* synonymous substitution rs2066462 on the synthesis of the encoded protein consists either in a splicing abnormality or in the need for the cell to recruit a rarer tRNA, which slows protein synthesis. Concerning the rs7533315 marker, it can be suggested that this association is caused by the fact that this SNP is closely linked ($D' = 1$) to rs2066462 and the nonsynonymous substitution rs2274976 (G1793A). Due to low frequency of the rs2274976 allele *A*, association of this missense mutation with CAD was not identified in a small sample. Literature data on rs2274976 are very scanty. It is known that the frequency of the G1793A polymorphism varies from 1.3% in Ashkenazi Jews to 26.6% in Indonesians from Java Island. It was demonstrated that heterozygosity for G1793A results in the elevation of plasma homocysteine level by 40% [19]. Associations of this SNP with endometrial cancer, squash cell carcinoma, colorectal cancer, and female infertility have also been reported. Some researchers recognize rs2274976 as a clinically important polymorphism, especially upon an analysis of genetic susceptibilities to complex diseases, whose pathogenesis involves folate metabolism defects [20].

It is also possible that rs7533315 is linked to another functionally important *MTHFR* polymorphism, which was not examined in the present study. The *MTHFR* resequencing that was performed in 2006 in four populations (Caucasoids, Africans, Chinese, and Mexicans) identified 11 nonsynonymous substitutions at this locus [21]. Of these, only three SNPs were the most widespread (C677T, A1298C, and G1793A). The remaining polymorphisms were found in individual populations at frequencies from 0.8 to 2.5%.

The data on the association of the *MTHFR* allelic variants with cardiovascular disease are rather controversial. Many authors consider the *MTHFR* C677T polymorphism as an independent risk factor of cardiovascular and cerebrovascular diseases, including coronary atherosclerosis. However, several other authors found no association between C677T and A1298C and vascular pathologies [22–28]. For example, a meta-analysis that was performed by Klerk et al. (2002) showed an association between the *MTHFR* C677T polymorphism and IHD in the European population and the absence of such an association in the North American population [29]. Another meta-analysis of 15 association studies also pointed to a considerable association between the 677*TT* genotype and CAD (OR, 1.67; 95% CI, 1.21–2.31) [30]. A large-scale analysis that explored the probability of significant effects in 23 case–control studies, including 5869 patients and 6644 controls, showed that the C677T marker was not a risk factor of cardiovascular disease, while the 677*TT* genotype was associated with a 25% elevated homocysteine concentration, compared to the *CC* genotype [31]. At the same time, a meta-analysis of studies, which included a total of 5644 patients, that was performed by Wu and Tsongalis in 2001 [32], showed that the homozygous variant 677*TT* was associated with an increased risk of CAD (OR, 1.30; 95% CI, 1.11–1.52). Opposite results were reported by Nakai et al. (2001) [33]. These contradictions can be explained by different approaches to research design, as well as by the involvement of genetic markers into CAD pathogenesis in some populations, while in other populations this mechanism does not work. Population-genetic processes, which can also lead to controversial results, deserve special interest. In different populations and ethnic geographic groups, associations are formed on different genetic and environmental backgrounds. The mutations that are expressed, for example, as an increased susceptibility to CAD, can occur more than once and in different allelic surroundings. It is known that any new allele is initially associated with other alleles, on whose background it appeared. Although each SNP can be examined independently, it seems more informative to explore them within haplotypes that are population-specific due to their unique demographic histories.

Haplotype Structure and Linkage Disequilibrium of the MTHFR Gene in the Groups We Examined

In the present study, genotypes were analyzed at 11 SNPs of the *MTHFR* gene (the rs3753588 marker was excluded from the analysis, because it was absent from the genotyping panel of the HapMap project). In CAD patients, a total of 14 haplotypes were discovered out of 4096 theoretically possible ones, while in the control sample 15 haplotypes were identified (Fig. 2). Nine of the haplotypes were identical for both groups. The number of major haplotypes (with a frequency

Table 3. Comparison of the groups we examined with respect to quantitative traits

No	SNPs examined	Genotypes	Mean quantitative trait values in the groups examined					
			Total cholesterol, mol/L		Triglycerides, mol/L		HDL, mol/L	
			Control group	CAD patients	Control group	CAD patients	Control group	CAD patients
1	2	3	4	5	6	7	8	9
1	rs3753588	*AA*	6.75 ± 0.00	–	1.53 ± 0.11	–	1.42 ± 0.07	–
		AG	6.11 ± 0.14	5.84 ± 0.28	1.76 ± 0.37	1.72 ± 0.22	1.35 ± 0.04	0.89 ± 0.08
		GG	5.78 ± 0.35	5.43 ± 0.12	1.22 ± 0.08	1.50 ± 0.80	1.32 ± 0.05	0.97 ± 0.04
		H	2.06	1.46	3.21	0.65	0.34	1.84
		p	0.361	0.235	0.201	0.423	0.843	0.175
2	rs2066470	*CC*	5.74 ± 0.15	5.44 ± 0.21	1.23 ± 0.06	1.64 ± 0.08	1.33 ± 0.03	0.98 ± 0.04
		CT	6.34 ± 0.33	5.89 ± 0.14	1.51 ± 0.15	1.97 ± 0.15	1.34 ± 0.07	0.94 ± 0.06
		TT	6.74 ± 0.48	6.39 ± 0.82	0.62 ± 0.11	2.36 ± 0.08	1.42 ± 0.32	0.68 ± 0.02
		H	3.86	5.17	2.81	4.47	0.16	2.90
		p	0.145	0.076	0.245	0.107	0.922	0.241
3	rs17037397	*AC*	5.80 ± 0.18	5.71 ± 0.45	1.47 ± 0.35	1.92 ± 0.37	1.32 ± 0.08	0.83 ± 0.11
		CC	5.86 ± 0.15	5.47 ± 0.11	1.29 ± 0.10	1.50 ± 0.07	1.33 ± 0.03	0.97 ± 0.03
		H	0.06	0.14	0.09	1.32	0.04	1.96
		p	0.80	0.71	0.77	0.25	0.834	0.162
4	rs7533315	*CC*	5.83 ± 0.18	**5.24 ± 0.17**	1.29 ± 0.08	1.53 ± 0.11	1.37 ± 0.04	**1.12 ± 0.06**
		CT	5.85 ± 0.21	**5.72 ± 0.14**	1.21 ± 0.09	1.81 ± 0.09	1.32 ± 0.04	**1.04 ± 0.04**
		TT	5.85 ± 0.61	**5.87 ± 0.33**	1.55 ± 0.27	1.85 ± 0.21	1.22 ± 0.13	**0.84 ± 0.08**
		H	0.05	**5.97**	2.88	4.21	0.50	**10.67**
		p	0.974	**0.048**	0.237	0.121	0.780	**0.005**
5	rs4846052	*CC*	5.74 ± 0.24	5.26 ± 0.21	1.26 ± 0.16	1.41 ± 0.13	1.34 ± 0.05	0.89 ± 0.07
		CT	5.94 ± 0.21	5.52 ± 0.14	1.27 ± 0.11	1.55 ± 0.10	1.36 ± 0.09	1.02 ± 0.04
		TT	5.77 ± 0.18	5.67 ± 0.29	1.53 ± 0.35	1.66 ± 0.21	1.24 ± 0.07	0.88 ± 0.07
		H	0.40	1.82	0.14	0.99	2.40	3.16
		p	0.82	0.40	0.93	0.61	0.300	0.206
6	rs1801133	*CC*	6.03 ± 0.20	5.66 ± 0.16	1.42 ± 0.17	1.61 ± 0.10	*1.33 ± 0.04*	0.97 ± 0.04
		CT	5.63 ± 0.23	5.30 ± 0.17	1.19 ± 0.12	1.49 ± 0.12	*1.30 ± 0.05*	0.92 ± 0.05
		TT	5.88 ± 0.24	5.37 ± 0.27	1.24 ± 0.12	1.33 ± 0.27	*1.19 ± 0.08*	1.11 ± 0.16
		H	3.36	1.15	1.77	1.94	*5.07*	1.95
		p	0.19	0.56	0.41	0.38	*0.079*	0.377

Table 3. (Contd.)

No	SNPs examined	Genotypes	Mean quantitative trait values in the groups examined					
			Total cholesterol, mol/L		Triglycerides, mol/L		HDL, mol/L	
			Control group	CAD patients	Control group	CAD patients	Control group	CAD patients
1	2	3	4	5	6	7	8	9
7	rs6541003	*AA*	5.77 ± 0.25	**5.08 ± 0.20**	1.26 ± 0.11	**1.41 ± 0.13**	1.30 ± 0.05	0.91 ± 0.09
		AG	5.91 ± 0.18	**5.67 ± 0.13**	1.24 ± 0.08	**1.79 ± 0.09**	1.37 ± 0.04	0.98 ± 0.04
		GG	5.73 ± 0.23	**5.87 ± 0.26**	1.38 ± 0.15	**1.88 ± 0.17**	1.24 ± 0.07	0.94 ± 0.07
		H	0.13	**7.24**	1.42	**9.57**	2.57	0.12
		p	0.936	**0.027**	0.491	**0.008**	0.276	0.943
8	rs2066462	*CC*	5.83 ± 0.14	5.48 ± 0.11	1.27 ± 0.06	**1.64 ± 0.07**	1.34 ± 0.03	**1.07 ± 0.04**
		CT	6.01 ± 0.05	5.94 ± 0.26	1.24 ± 0.24	**2.04 ± 0.17**	1.21 ± 0.11	**0.83 ± 0.07**
		H	0.65	2.02	0.13	**5.77**	1.22	**6.26**
		p	0.421	0.155	0.723	**0.029**	0.269	**0.012**
9	rs1801131	*AA*	5.76 ± 0.20	**5.17 ± 0.16**	1.22 ± 0.13	**1.42 ± 0.12**	1.23 ± 0.09	0.93 ± 0.05
		AC	6.04 ± 0.22	**5.70 ± 0.15**	1.43 ± 0.16	**1.57 ± 0.11**	1.33 ± 0.08	1.00 ± 0.05
		CC	5.50 ± 0.20	**6.12 ± 0.41**	1.15 ± 0.20	**1.96 ± 0.30**	1.11 ± 0.17	0.91 ± 0.08
		H	1.57	**8.17**	1.01	**6.30**	0.800	1.54
		p	0.46	**0.017**	0.60	**0.043**	0.670	0.462
10	rs17375901	*CC*	**5.74 ± 0.13**	**5.40 ± 0.15**	1.29 ± 0.09	**1.46 ± 0.08**	1.27 ± 0.09	0.98 ± 0.04
		CT	**7.09 ± 0.57**	**6.67 ± 0.32**	1.55 ± 0.63	**2.40 ± 0.23**	1.23 ± 0.20	0.89 ± 0.07
		H	**8.31**	**10.20**	0.56	**12.53**	0.67	2.60
		p	**0.0039**	**0.0014**	0.45	**0.0004**	0.979	0.107
11	rs2274976	*AG*	5.53 ± 0.27	5.61 ± 0.15	1.66 ± 0.37	1.52 ± 0.08	1.52 ± 0.16	0.88 ± 0.17
		GG	5.90 ± 0.15	5.48 ± 0.50	1.26 ± 0.10	1.80 ± 0.37	1.23 ± 0.06	0.97 ± 0.09
		H	0.17	0.001	0.27	0.66	0.10	0.705
		p	0.68	0.97	0.60	0.42	0.933	0.401
12	rs1537516	*CC*	5.75 ± 0.15	**5.37 ± 0.16**	1.22 ± 0.08	**1.45 ± 0.09**	1.22 ± 0.06	0.97 ± 0.04
		CT	6.37 ± 0.35	**5.85 ± 0.24**	1.83 ± 0.41	**1.79 ± 0.19**	1.52 ± 0.14	0.97 ± 0.06
		TT	5.84 ± 0.91	**6.69 ± 0.86**	0.74 ± 0.20	**2.67 ± 0.46**	0.85 ± 0.46	0.81 ± 0.13
		H	3.03	**5.93**	**2.48**	**8.64**	0.11	1.94
		p	0.22	**0.049**	**0.29**	**0.013**	0.995	0.378

Note: *H*, Kruskal–Wallis test value; statistically significant differences ($P < 0.05$) are in bold type.

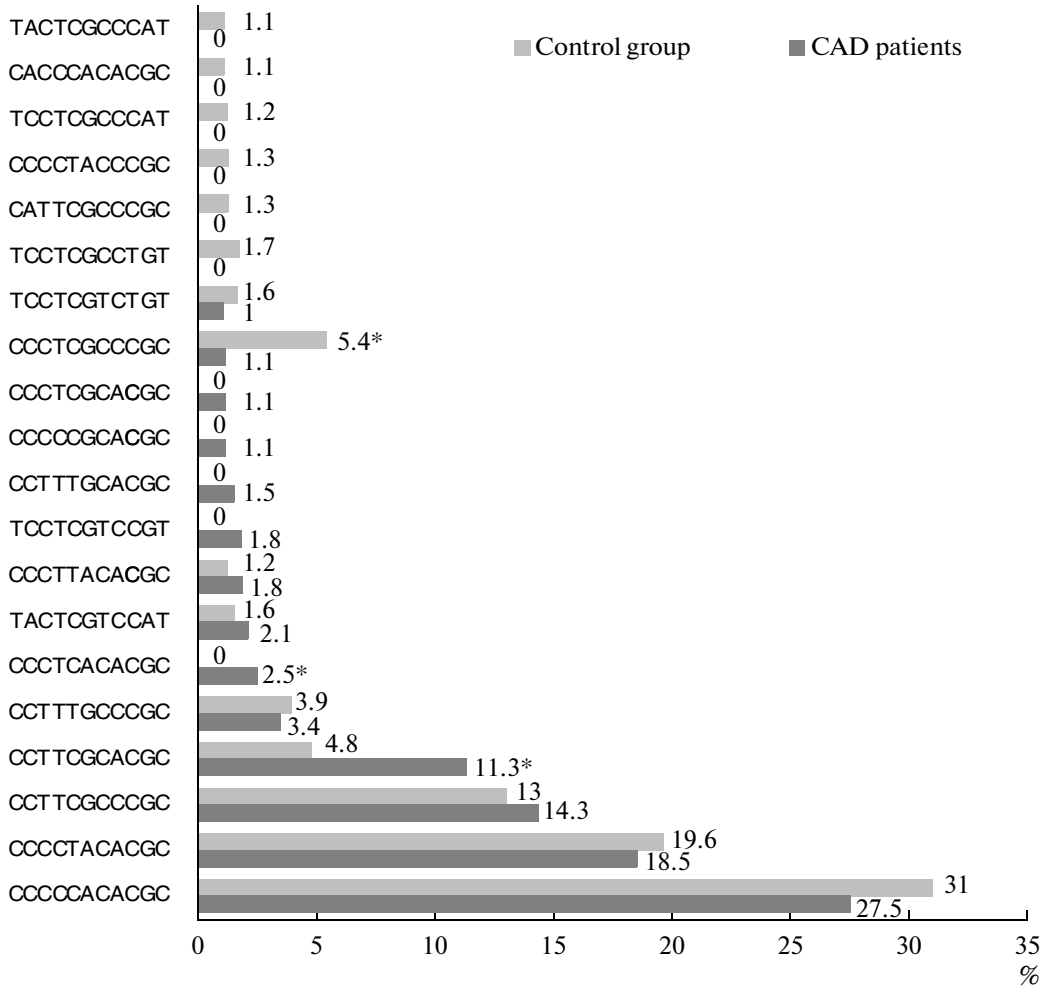

Fig. 2. Haplotype frequency distributions in CAD patients and the control group.

above 10%) equaled four in CAD patients and three in the control group. A comparative analysis of the haplotype frequencies revealed a statistically significant association of the CCTTCGCACGC (χ^2 with a Yates' correction, 6.91; $p = 0.009$; OR, 2.57; 95% CI, 1.24–5.39) and CCCTCACACGC (upon comparison using Fisher's test, $p = 0.017$; RR, 1.92; 95% CI, 1.77–2.08) haplotypes with CAD. In addition, a protective haplotype, CCCTCGC**C**CGC, (upon comparison using Fisher's test, $p = 0.003$; OR, 0.18; 95% CI, 0.04–0.69) was identified. For other haplotypes, no statistically significant differences were detected. The risk-associated haplotype CC**T**TCGCACGC contains the mutant allele *T* (in bold type) of the rs7533315 marker, for which an association with CAD was recorded in the present study. However, all other alleles of this haplotype are ancestral, suggesting close linkage between rs7533315 and some functionally important SNPs of the *MTHFR* gene, which was not explored in the present study. The haplotype CCCTC**A**CACGC, which includes a mutant allele of the rs6541003 locus, was also associated with CAD. The data of the present study support the idea that this locus is under the effect of stabilizing selection in both the control and CAD patient groups.

The haplotype CCCTCGC**C**CGC includes the single mutant allele *C*, which belongs to the functionally important *MTHFR* marker rs1801133 (A1298C). Nevertheless, in our study, an expressed protective effect of this haplotype relative to CAD was demonstrated. However, it should be noted that this haplotype is not detected in some Mongoloid populations (Buryats, Chinese, and Japanese). Moreover, it is more frequent among Russians at a statistically significant level, compared to the other populations that were examined [34]. At present, the influence of the *MTHFR* A1298C polymorphism on the plasma homocysteine level remains an open issue. Studies that concern this issue have produced controversial results [35, 36]. It was suggested that the effect of the 677*TT* genotype on the activity of MTHFR is much higher compared to that of the 1298*CC* variant. This is because 677*TT* is located in the catalytic domain of the enzyme, while the A1298C polymorphism is

located in the region that encodes the C terminal regulatory domain. This polymorphism provides only small changes in MTHFR activity through abnormal binding of the enzyme with its inhibitor, *S*-adenosylmethionine [37].

In our earlier study, a phylogenetic analysis of *MTHFR* haplotype relationships was performed [34]. This analysis demonstrated that the three haplotypes that were discussed were located one mutational step from the presumptive founder haplotype. They conditionally split the phylogenetic tree into three main clusters. This observation is quite interesting in the context of the evolutionary history of the genetic architecture of atherosclerosis and deserves further investigation.

It was demonstrated that the use of haplotypes instead of SNPs in association studies could substantially increase the statistical power of the test, especially if the predisposing polymorphism was not examined directly, or in case of a high level of multilocus linkage disequilibrium [38, 39]. Simulation modeling performed by Akey et al. (2001) showed that the power of haplotype tests was influenced by the genetic distance between the observed markers and the causative mutation, as well as by the allele frequencies and the age of the causative mutation [40]. One of the interesting haplotype features is the nonrandomized association among the SNPs that comprise its so-called linkage disequilibrium.

Earlier, we described population-specific *MTHFR* linkage disequilibrium patterns in different populations of Eurasia [34, 41]. The data of the present study indicate that *MTHFR* linkage disequilibrium patterns can be also characterized by intergroup differences (Fig. 3). In the control sample, whose characteristics are similar to those of the Russian population of Tomsk [42], three linkage disequilibrium blocks were identified. The first block consisted of two closely located SNPs (rs3753588 and rs206647), the second block included five polymorphisms (rs7533315, rs4846052, rs1801133, rs6541003, and rs2066462), and the third block consisted of four SNPs (rs1801131, rs17375901, rs2274976, and rs1537516). It is noteworthy that the C677T (rs1801133) variant, which is considered to be the most functionally important *MTHFR* polymorphism, showed no close linkage with any of the markers that were examined. It should be noted that the linkage disequilibrium patterns in CAD patients and controls had many features in common. In both cases a recombination hotspot occurs between the rs2066462 and rs1801131 markers. In addition, in both groups a haplotype block in the 3' gene region, including identical SNPs, was observed. At the same time, in the 5' gene region the linkage was more expressed in CAD patients.

In this study, an association between the *MTHFR* markers we examined and pathogenetically important lipid metabolism indices was observed. This finding serves as an additional confirmation of the association between the haplotype structure of the *MTHFR* gene we observed and coronary atherosclerosis.

Analysis of Associations with Pathogenetically Important Quantitative Traits

Using multiple comparison, in the present study an association of the genotypes of six SNPs, including rs7533315, rs6541003, rs2066462, rs1801131, rs17375901, and rs1537516 with variable plasma total cholesterol (TC) and (or) triglyceride (TG) levels in the group of CAD patients was identified (Table 3). Interestingly, an association with the lipid metabolism indices was found for the genotypes of three out of four SNPs that are contained in the second linkage block of the CAD patients. In the control group, a statistically significant elevation of the TC values was observed in the rs17375901 heterozygotes. It should be noted that among the CAD patients, maximum TC and TG levels corresponded to the homozygote for the mutant allele (rs7533315, rs1801131, and rs1537516), or to the heterozygote that carries this allele (rs2066462 and rs17375901). For rs6541003, the opposite tendency was observed. Specifically, in the carriers of the ancestral genotype of this locus a statistically significant elevation of the TC and TG levels compared to those in mutant homozygotes was observed. In the carriers of different *MTHFR* genotypes from both groups, the values of SBP and DBP indices, plasma LDL, and the body mass index (BMI) were not different at a statistically significant level.

In CAD patients, only two polymorphisms (rs7533315 and rs2066462) correlated with the HDL level. In this study, it was demonstrated that these polymorphisms were associated with CAD. In the control group, only a tendency towards association of the C677T genotypes with the HDL level was observed. According to the data we obtained, 677*TT* homozygotes were characterized by lower HDL levels compared to heterozygotes and homozygotes for the ancestral allele. Analogous data were obtained by Real et al. (2009), who examined the relationships of the C677T genotypes and plasma homocystein and lipoprotein levels in familial hypercholesterolemia. In this study, the 677*TT* genotype and an increased plasma Hcy level were associated with decreased HDL values [43]. In a study by Japanese researchers, a statistically significant association between the 677*TT* genotype and HDL upon the development of ischemic stroke was demonstrated [44].

The results of a large-scale seroepidemiological study that was performed with a Spanish population showed that combination of such cardiovascular disease risk factors as folate (<5.3 nmol/L) and HDL (<35 mg/dL) deficiency with the 677*TT* genotype increased the risk of hyperhomocysteinemia (HHcy) by 87 times [45]. At present, the mechanisms that underlie the association between HHcy, *MTHFR* polymorphisms, and plasma lipoprotein levels, are still

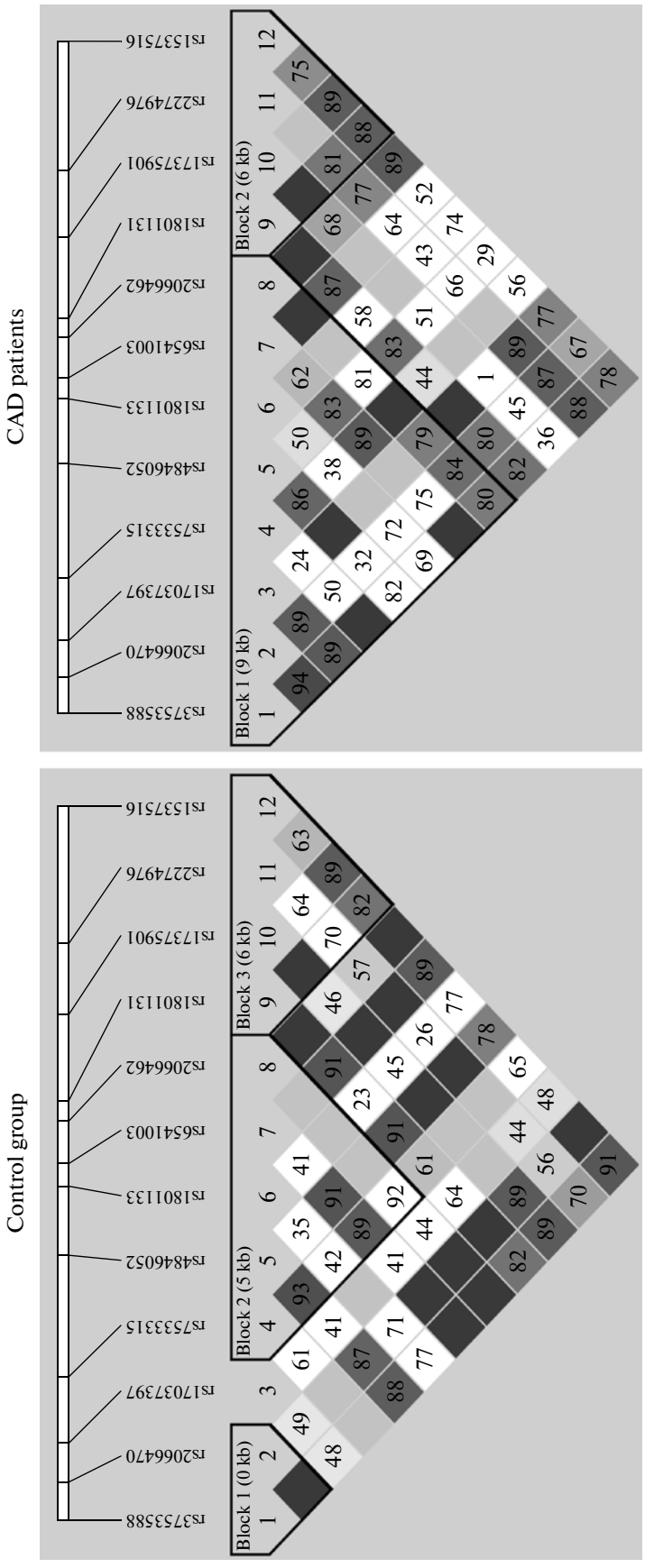

Fig. 3. The *MTHFR* linkage disequilibrium patterns in the populations we examined. The values of linkage coefficient $D' \times 0.01$ are in the cells (the empty cell designates $D' = 1$); color gamma reflects the linkage value between the SNPs: dark gray, close linkage ($D' = 1$; LOD > 2); light gray, considerable linkage ($D' < 1$; LOD > 2); white, weak linkage ($D' < 1$; LOD < 2).

unknown. Numerous experiments that were carried out using endothelial cell cultures showed that HHcy was accompanied by the generation of oxidants. These oxidants, in addition to their cytotoxic effect, were able to induce the oxidation of low-density lipoproteins (LDL). Moreover, in the case of HHcy, in the membranes and intercellular space of endotheliocytes elevated concentrations of LDL and very low density lipoprotein (VLDL) were observed due to a methyl deficit upon synthesis of the protein component of lipoproteins. Recent in vitro experiments have demonstrated Hcy-induced elevation of gene expression of the genes that are responsible for biosynthesis and absorption of Hcy, TG, and the intercellular accumulation of cholesterol [46]. It was also demonstrated that one of the metabolic pathways of Hcy in the human body was its transformation into the thioester Hcy–thiolactone. In 1997, Jakubowski et al. suggested a hypothesis on the involvement of Hcy-thiolactone in the development of atherothrombosis and lipoprotein modification. More recently, elevated tissue Hcy and its metabolite, Hcy–thiolactone, levels were observed in IHD patients [47].

Thus, many modern biochemical and genetic data indicate that elevation of the Hcy level, which is often determined by mutations in the folate metabolism genes, leads to the formation of proatherothrombotic phenotype. The data of the present study point to the involvement of the *MTHFR* polymorphic variants in the determination of CAD. However, from a clinical point of view, associations of genetic markers with CAD, while they are evidence in favor of increased risk of the disease, are not sufficiently informative to provide additional diagnostic information. In addition, the data indicate that the *MTHFR* linkage disequilibrium patterns in populations of different ethnic origins are substantially different. At the same time, the structure of the *MTHFR* haplotype patterns in the samples from one population that were differentiated by either the presence or the absence of coronary atherosclerosis, was similar. Furthermore, a highly statistically significant association of certain *MTHFR* haplotypes and SNPs with coronary atherosclerosis was demonstrated. A relationship between the *MTHFR* genetic variability and pathogenetically important indices of lipid metabolism was observed. Finally, the present study demonstrates that the haplotype approach is highly informative in case–control tests for associations with complex diseases. In our opinion, the data presented here are of special interest for the understanding of a number of genetic phenomena, including interpopulation differences of linkage disequilibrium patterns, the structure of the genetic component of complex diseases in terms of the comparative informativeness of the association of individual markers and haplotypes with the disease and the functional importance and pleiotropic effect of *MTHFR* activity.

Surprisingly, no association of the C677T polymorphism with atherosclerosis and its endophenotypes was revealed in either in the analysis of individual SNPs and during the assessment of *MTHFR* haplotype variability. Moreover, in the patient group, no statistically significant linkage of the C677T polymorphism to all the other polymorphisms examined was detected. At the same time, most of the collected data on the association of *MTHFR* with multifactorial states, including cardiovascular diseases, concern precisely this missense mutation, which results in the synthesis of a thermolabile variant of the enzyme. It can be suggested that the effect of this SNP with respect to cardiovascular pathology, if it is an attribute of the substitution itself, is not so large that it can be fixed in relatively small samples, like that examined in the present study. It seems likely that due to its population-specific linkage disequilibrium pattern, the C677T polymorphism may be a member of disease-associated haplotypes in some populations, while it is absent from these haplotypes in other populations, as was demonstrated in the present study.

ACKNOWLEDGMENTS

This work was supported by the Federal Target–Oriented Programs on Scientific and Scientific–Pedagogical Personnel of the Innovative Russia (State Contract no. P321) and Research and Development in Priority Fields of Science and Technology Complex of Russia for 2007–2013 (State Contract no. 11.519.112036).

REFERENCES

1. Cambien, F. and Tiret, L., Genetics of Cardiovascular Diseases from Single Mutations to the Whole Genome, *Circulation*, 2007, vol. 116, pp. 1714–1724.
2. Carlson, C.S., Eberle, M.A., Rieder, M.J., et al., Additional SNPs and Linkage-Disequilibrium Analyses Are Necessary for Whole-Genome Association Studies in Humans, *Nat. Genet.*, 2003, vol. 33, pp. 518–521.
3. Collins, A., Lau, W., and De La Vega, F.M., Mapping Genes for Common Diseases: The Case for Genetic (LD) Maps, *Hum. Hered.*, 2004, vol. 58, pp. 2–9.
4. Altshuler, D., Daly, M.J., and Lander, E.S., Genetic Mapping in Human Disease, *Science*, 2008, vol. 322, pp. 881–888.
5. Petterson, F.H., Anderson, C.A., Clarke, G.M., et al., Marker Selection for Genetic Case-Control Association Studies, *Nat. Protoc.*, 2009, vol. 4, no. 5, pp. 743–452.
6. Crawford, D.C. and Nickerson, D.A., Definition and Clinical Importance of Haplotypes, *Annu. Rev. Med.*, 2005, vol. 56, pp. 303–320.
7. Zhang, K., Qin, Z., Chen, T., et al., HapBlock: Haplotype Block Partitioning and Tag SNP Selection Software Using a Set of Dynamic Programming Algorithms, *Bioinformatics*, 2005, vol. 21, pp. 131–134.
8. Slatkin, M., Linkage Disequilibrium—Understanding the Evolutionary Past and Mapping the Medical Future, *Genetics*, 2008, vol. 9, pp. 477–485.

9. The International HapMap Consortium: A Haplotype Map of the Human Genome, *Nature*, 2005, vol. 437, pp. 1299–1320.
10. Frosst, P., Blom, H.J., Milos, R., et al., A Candidate Genetic Risk Factor for Vascular Disease: A Common Mutation Methylenetetrahydrofolate Reductase, *Nat. Genet.*, 1995, no. 10, pp. 111–113.
11. Friedman, G., Goldschmidt, N., and Friedlander, Y., A Common Mutation *A1298C* in Human Methylenetetrahydrofolate Reductase Gene: Association with Plasma Total Homocysteine and Folate Concentrations, *J. Nutr.*, 1999, no. 129, pp. 1656–1661.
12. Trifonova, E.A., Spiridonova, M.G., Puzyrev, V.P., et al., Haplotypes Structure of Methylenetetrahydrofolate Reductase Locus: Population Specificity and Association with Coronary Atherosclerosis, *Med. Genet.*, 2009, vol. 79, no. 1, pp. 39–47.
13. Glantz, S., *Medico-Biological Statistics*, New York: McGraw-Hill, 1994.
14. Barrett, J.C., Fry, B., Maller, J., and Daly, M.J., Haploview: Analysis and Visualization of LD and Haplotype Maps, *Bioinformatics*, 2005, no. 21, pp. 263–265.
15. Watterson, G.A., On the Number of Segregating Sites in Genetical Models without Recombination, *Theor. Popul. Biol.*, 1975, vol. 7, pp. 256–276.
16. Spiroski, I., Kedev, S., Antov, S., et al., Association of Methylenetetrahydrofolate Reductase (*MTHFR-677* and *MTHFR-1298*) Genetic Polymorphisms with Occlusive Artery Disease and Deep Venous Thrombosis in Macedonians, *Croat. Med. J.*, 2008, vol. 49, pp. 39–49.
17. Mayor-Olea, A., Callejon, G., Palomares, A.R., et al., Human Genetic Selection on the *MTHFR* 677C>T Polymorphism, *BMC Med. Genet.*, 2008, vol. 9, p. 104.
18. Rosenberg, N., Murata, M., Ikeda, Y., et al., The Frequent 5,10-Methylenetetrahydrofolate Reductase C677T Polymorphism Is Associated with a Common Haplotype in Whites, Japanese, and Africans, *Am. J. Hum. Genet.*, 2002, vol. 70, pp. 758–762.
19. Mao, R., Fan, Y., Chen, F., et al., Methylenetetrahydrofolate Reductase Gene Polymorphisms in 13 Chinese Ethnic Populations, *Cell Biochem. Funct.*, 2008, vol. 26, pp. 352–358.
20. Melo, S.S., Persuhn, D.C., Meirelles, N., et al., G1793A Polymorphisms in the Methylenetetrahydrofolate Gene: Effect of Folic Acid on Homocysteine Levels, *Mol. Nutr. Food Res.*, 2006, vol. 50, pp. 769–774.
21. Martin, Y.N., Salavaggione, O.E., and Eckloff, B.W., Human Methylenetetrahydrofolate Reductase Pharmacogenomics: Gene Resequencing and Functional Genomics, *Pharmacogenet. Genomics*, 2006, vol. 16, pp. 265–277.
22. Freitas, A.I., Mendona, I., and Guerra, G., Methylenetetrahydrofolate Reductase Gene, Homocysteine and Coronary Artery Disease: The *A1298C* Polymorphism Does Matter. Inferences from a Case Study (Madeira, Portugal), *Thromb. Res.*, 2008, vol. 122, no. 5, pp. 648–656.
23. Sarecka-Hujar, B., Żak, I., and Krauze, J., Carrier-State of Two or Three Polymorphic Variants of *MTHFR, IL-6* and *ICAM1* Genes Increases the Risk of Coronary Artery Disease, *Kardiol. Polska*, 2008, vol. 66, no. 12, pp. 1369–1277.
24. Szczeklik, A., Sanak, M., Jankowski, M., et al., Mutation *A1298C* of Methylenetetrahydrofolate Reductase: Risk for Early Coronary Disease not Associated with Hyperhomocysteinemia, *Am. J. Med. Genet.*, 2001, vol. 101, pp. 36–39.
25. Zak, I., Niemiec, P., Sarecka, B., et al., Carrier-State of *D* Allele in *ACE* Gene Insertion/Deletion Polymorphism Is Associated with Coronary Artery Disease, in Contrast to the C677T Transition in the *MTHFR* Gene, *Acta Biochim. Polonica*, 2003, vol. 50, no. 2, pp. 527–534.
26. Girelli, D., Martinelli, N., Pizzolo, F., et al., The Interaction between *MTHFR* 677 C ⟶ T Genotype and Folate Status Is a Determinant of Coronary Atherosclerosis Risk, *J. Nutr.*, 2003, vol. 133, pp. 1281–1285.
27. Morita, H., Kurihara, H., Tsubaki, S., et al., Methylenetetrahydrofolate Reductase Gene Polymorphism and Ischemic Stroke in Japanese, *Arterioscler. Thromb. Vasc. Biol.*, 1998, vol. 18, pp. 1465–1469.
28. Hsu, L.A., Ko, Y.L., Wang, S.M., et al., The *C677T* Mutation of the Methylenetetrahydrofolate Reductase Gene Is not Associated with the Risk of Coronary Artery Disease or Venous Thrombosis among Chinese in Taiwan, *Hum. Hered.*, 2001, vol. 51, nos. 1–2, pp. 41–45.
29. Klerk, M., Verhoef, P., Clarke, R., et al., *MTHFR* Studies Collaboration Group: *MTHFR* 677C ⟶ T Polymorphism and Risk of Coronary Heart Disease: A Meta-Analysis, *JAMA*, 2002, vol. 288, pp. 2023–2031.
30. Debette, S. and Markus, H.S., The Genetics of Cervical Artery Dissection: A Systematic Review, *Stroke*, 2009, vol. 40, no. 6, pp. 459–466.
31. Brattström, L., Wilcken, D.E., Ohrvik, J., and Brudin, L., Common Methylenetetrahydrofolate Reductase Gene Mutation Leads to Hyperhomocysteinemia but not to Vascular Disease: The Result of a Meta-Analysis, *Circulation*, 1998, vol. 98, pp. 2520–2526.
32. Wu, A.H. and Tsongalis, G.J., Correlation of Polymorphisms to Coagulation and Biochemical Risk Factors for Cardiovascular Diseases, *Am. J. Cardiol.*, 2001, vol. 15, no. 12, pp. 1361–1366.
33. Nakai, K., Itoh, C., Nakai, K., et al., Correlation between C677T *MTHFR* Gene Polymorphism, Plasma Homocysteine Levels and the Incidence of CAD, *Am. J. Cardiovasc. Drugs*, 2001, vol. 1, no. 5, pp. 353–361.
34. Trifonova, E.A., Eremina, E.R., Urnov, F.D., et al., The Genetic Diversity and Structure of Linkage Disequilibrium of the *MTHFR* Gene in Populations of Northern Eurasia, *Acta Nat.*, 2012, vol. 4, no. 1(12), pp. 79–96.
35. Poduri, A., Mukherjee, D., Sud, K., et al., *MTHFR* A1298C Polymorphism Is Associated with Cardiovascular Risk in End Stage Renal Disease in North Indians, *Mol. Cell Biochem.*, 2008, vol. 308, nos. 1–2, pp. 43–50.
36. Fung, M.M., Salem, R.M., and Lipkowitz, M.S., Methylenetetrahydrofolate Reductase (*MTHFR*) Polymorphism A1298C (Glu429Ala) Predicts Decline in Renal Function Over Time in the African-American Study of Kidney Disease and Hypertension (AASK)

Trial and Veterans Affairs Hypertension Cohort (VAHC), *Nephrol. Dial. Transplant.*, 2011, vol. 27, pp. 197–205.

37. Kolling, K., Ndrepepa, G., Koch, W., et al., Methylenetetrahydrofolate Reductase Gene C677T and A1298C Polymorphisms, Plasma Homocysteine, Folate, and Vitamin B12 Levels and the Extent of Coronary Artery Disease, *Am. J. Cardiol.*, 2004, vol. 93, pp. 1201–1206.
38. Schaid, D.J., Linkage Disequilibrium Testing When Linkage Phase Is Unknown, *Genetics*, 2004, vol. 166, pp. 505–512.
39. Zeng, D. and Lin, D.Y., Estimating Haplotype–Disease Associations with Pooled Genotype Data, *Genet. Epidemiol.*, 2005, vol. 28, no. 1, pp. 70–82.
40. Akey, J., Jin, L., and Xiong, M., Haplotypes vs Single Marker Linkage Disequilibrium Tests: What Do We Gain?, *Eur. J. Hum. Genet.*, 2001, vol. 9, no. 4, pp. 291–300.
41. Trifonova, E.A., Spiridonova, M.G., and Stepanov, V.A., Genetic Diversity and the Structure of Linkage Disequilibrium in the Methylenetetrahydrofolate Reductase Locus, *Russ. J. Genet.*, 2008, vol. 44, no. 10, pp. 1410–1419.
42. Kucher, A.N., Ivanova, O.F., Puzyrev, V.P., et al., Variability of Biochemical Gene Markers in Residents of the City of Tomsk, *Russ. J. Genet.*, 1994, vol. 30, no. 11, pp. 1530–1534.
43. Real, J.T., Martinez-Hervas, S., Garcia-Garcia, A.B., et al., Association of C677T Polymorphism in *MTHFR* Gene, High Homocysteine and Low HDL Cholesterol Plasma Values in Heterozygous Familial Hypercholesterolemia, *J. Atheroscler. Thromb.*, 2009, vol. 16, no. 6, pp. 815–820.
44. Kawamoto, R., Kohara, K., Oka, Y., et al., An Association of 5,10-Methylenetetrahydrofolate Reductase (*MTHFR*) Gene Polymorphism and Ischemic Stroke, *J. Stroke Cerebrovasc. Dis.*, 2005, vol. 14, no. 2, pp. 67–74.
45. Gil-Prieto, R., Hernndez, V., Cano, B., et al., Plasma Homocysteine in Adolescents Depends on the Interaction between Methylenetetrahydrofolate Reductase Genotype, Lipids and Folate: A Seroepidemiological Study, *Nutr. Metab.*, 2009, vol. 5, pp. 6–39.
46. Liao, D., Yang, X., and Wang, H., Hyperhomocysteinemia and High-Density Lipoprotein Metabolism in Cardiovascular Disease, *Clin. Chem. Lab. Med.*, 2007, vol. 45, no. 12, pp. 1652–1659.
47. Jakubowski, H., The Pathophysiological Hypothesis of Homocysteine Thiolactone-Mediated Vascular Disease, *J. Phys. Pharm.*, 2008, vol. 59, no. 19, pp. 155–167.