ISSN 1022-7954, Russian Journal of Genetics, 2024, Vol. 60, No. 2, pp. 259–264. © Pleiades Publishing, Inc., 2024.

SHORT COMMUNICATIONS

Searching for Ethnospecific Risk Markers of Paranoid Schizophrenia in Bashkirs Based on the Results from Genome-Wide Association Study

A. E. Gareeva[a, b, *]

[a] *Institute of Biochemistry and Genetics, Ufa Federal Research Center, Russian Academy of Sciences, Ufa, 450054 Russia*
[b] *Bashkir State Medical University, Ufa, 450008 Russia*
**e-mail: annagareeva@yandex.ru*

Received July 5, 2023; revised August 21, 2023; accepted August 24, 2023

Abstract—It is now known that schizophrenia is a multifactorial disease in which both genetic and environmental factors play a role. In recent years, mainly through the use of genome-wide association studies (GWAS), many molecular genetic processes that increase susceptibility to schizophrenia have been identified. The objective of this study was to examine genetic factors associated with the risk of developing schizophrenia by conducting a genome-wide association study (GWAS) in Bashkirs from the Republic of Bashkortostan. The studied sample consisted of 139 patients with paranoid schizophrenia and 204 healthy individuals. Genome-wide genotyping of DNA samples was carried using a PsychChip biochip, which included 610000 single nucleotide polymorphisms (SNPs).

Keywords: genetics, schizophrenia, genome-wide association study, ethnicity, ethnospecific markers, Republic of Bashkortostan, Psychiatric Genomics Consortium (PGC)

DOI: 10.1134/S1022795424020030

Schizophrenia is a severe mental disorder. Pooled lifetime prevalence of schizophrenia is approximately 1%. High morbidity and mortality make schizophrenia a major public health problem [1]. The heritability estimate of schizophrenia is 81%, which points to the important role of the genetic component in the pathogenesis of this disease [1]. Genome-wide association studies (GWAS) have improved our understanding of the contribution of specific genetic factors to the risk of schizophrenia. GWAS have provided compelling evidence for the important role of common genetic variants in determining individual background susceptibility to schizophrenia [2]. To date, several large-scale GWAS have been conducted in different populations around the world, and hundreds of polymorphic risk loci for schizophrenia have been identified [3–6]. The largest GWAS to date identified 287 independent polymorphic risk loci for schizophrenia [6].

To identify ethnospecific genetic risk factors for developing paranoid schizophrenia, in the present study, a genome-wide association study was conducted in Bashkirs from the Republic of Bashkortostan (Fig. 1). The study involved 139 patients (70 males, 69 females) of Bashkir ethnicity with the diagnosis of paranoid schizophrenia (PSz), code F20.0, according to the International Classification of Diseases 10th Revision (ICD-10), being treated at the Republican Clinical Psychiatric Hospital No. 1 of the Ministry of Health of the Republic of Bashkortostan. The mean age of the patients was 24.9 ± 8.9 years. The mean age of onset of the disease was 22.4 ± 7.3 years. Information on ethnicity up to the third generation was obtained by interview. The control group consisted of 204 healthy individuals (108 males, 96 females) of the same age group, who were not under regular check-up by either psychiatrist or narcologist and denied a family history of mental illness. The mean age of healthy donors was 32.4 ± 12.4 years.

Genome-wide genotyping of DNA samples was conducted with the Illumina Human 610-QuadPsychChip biochip, which included 610000 single nucleotide polymorphisms (SNPs). Genome-wide association study assessing the SNP associations was performed using the PLINK 2.0 software package [7]. A detailed description of the genome-wide association study was reported previously [8]. To reduce the type I error, the FDR-BH (False Discovery Rate Bengamini-Hochberg) correction for multiple comparisons was applied [9].

GWAS performed in individuals of Bashkir ethnicity revealed the most pronounced differences between patients with PS and the control group at polymorphic loci located in the 1p36.13 chromosomal region, which, according to the results from a number of genome-wide studies, is associated with the risk of

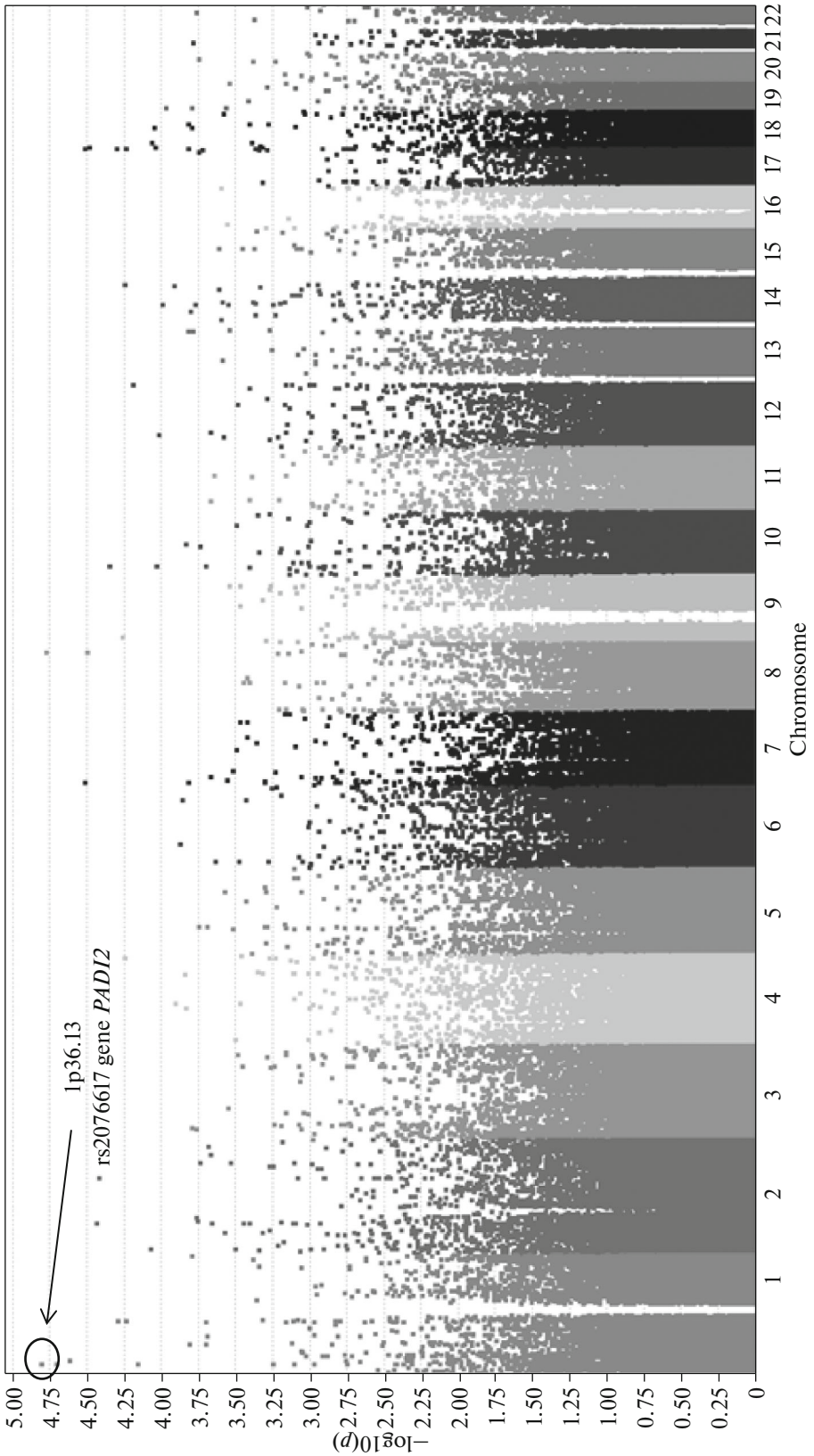

Fig. 1. Graphic representation of the results from genome-wide association study of 395832 SNPs with paranoid schizophrenia in Bashkirs (Manhattan plot). The *X* axis shows chromosomal locations of SNPs, and the *Y* axis shows negative decimal logarithm of *p*-values.

Table 1. Single nucleotide polymorphisms located in the 1p36.13 region and associated with paranoid schizophrenia in Bashkirs

Gene	rs no.	SNP	Allele 1	Frequency of allele 1 patients, %	Frequency of allele 1 control, %	p	OR	p_{fdr}
PADI2	rs2076617	g.17409017G>A	*A*	0.2662	0.4338	1.53E-05	0.472	0.768
PADI2	rs2016693	g.17397704A>C	*A*	0.2806	0.4485	1.95E-05	0.484	0.768
PADI2	rs2057096	g.17405809G>A	*G*	0.3309	0.4902	6.86E-05	0.522	0.806
PADI2	rs2057094	g.17405949C>T	*C*	0.3309	0.4902	6.86E-05	0.522	0.806
PADI1	rs3003406	g.17557133A>C	*C*	0.4209	0.2966	5.59E-04	1.833	0.814
PADI2	rs2076598	g.17395521G>A	*G*	0.3345	0.4681	7.73E-04	0.582	0.814
PADI1	rs11203339	g.17560972C>T	*T*	0.3669	0.25	1.06E-03	1.774	0.817
PADI1	rs4268393	g.17559196T>C	*C*	0.2842	0.1814	1.74E-03	1.809	0.824
PADI2	rs2076614	g.17413459G>A	*A*	0.2482	0.3529	3.69E-03	0.597	0.824
SDHB	rs4920653	g.17366871T>C	*C*	0.2806	0.3922	3.74E-03	0.618	0.824
PADI1	rs114209578	g.17541929C>A	*A*	0.2518	0.07108	0.011	0.336	0.849

developing schizophrenia (Fig. 1, Table 1) [10, 11]. In a number of previous studies, an association of the 1p36.13 chromosomal region with the development of schizophrenia was revealed [10, 11]. For instance, linkage analysis with twelve endophenotypes during a large-scale GWAS revealed an association of the 1p36.13 chromosomal region (*PAX7*, *UBR4*, *ALDH4A1*, *NBL1*, *HTR6*, *EPHA8*, *EPHB2*) with emotion recognition test that reached a genome-wide level of statistical significance with a LOD score of 3.5 (1p36) in 1004 patients with schizophrenia [12, 13].

The strongest association with PS was found for the rs2076617 polymorphism, located in the *PADI2* gene (p = 1.53E-05) (Table 1). The *PADI2* gene encodes peptidyl arginine deiminase type II enzyme, which is the member of the family of enzymes called peptidyl arginine deiminases, and consists of 16 exons, encompassing about 53 kb of genomic DNA, the length of its mRNA is 4363 bp [14]. The cleavage of arginine to citrulline is a process catalyzed by the peptidyl arginine deiminase enzyme (PAD), which converts the amino acid arginine to citrulline. During the modification process, the positively charged NH_2 group is removed with the addition of oxygen. Cyclic citrullinated proteins are present in the synovial tissue of patients with rheumatoid arthritis. The *PADI2* gene is widely expressed in the central nervous system, including neurons, glial cells, astrocytes, microglia, and oligodendrocytes. The deregulated expression of *PADI2* causes aberrant citrullination of glial fibrillary acidic protein (GFAP), eventually leading to the occurrence of neurological diseases [14].

In recent years, abnormal activation of PAD family proteins was found to be associated with the accumulation of large amounts of citrullinated proteins in patients with different neurodegenerative diseases, such as Alzheimer's disease, Parkinson's disease, multiple sclerosis, and Huntington's disease, suggesting that excessive citrullination of proteins led to the development of these [15] and other neuropsychiatric diseases with neurodegeneration [16]. Citrullination, that is the deamination of peptidyl arginine residues to peptidyl citrulline, was shown to be implicated in the etiology of a number of diseases. In multiple sclerosis, citrullination is thought to be the leading factor in the disease pathogenesis through hypercitrullination and destabilization of myelin. As such, inhibition of citrullination was suggested as a therapeutic strategy for multiple sclerosis [17]. On the contrary, A.M. Falcao et al. showed that citrullination by peptidylarginine deiminase 2 (PADI2) promoted normal oligodendrocyte differentiation, myelination, and motor function. This research group identified several targets for PADI2, including myelin and chromatin-related proteins, implicating PADI2 in epigenomic regulation. They also found that PADI2 inhibition and knockdown affected chromatin accessibility and prevented the upregulation of oligodendrocyte differentiation genes. Moreover, mice lacking PADI2 displayed motor dysfunction and a decreased number of myelinated axons in the corpus callosum. The data of A.M. Falcao et al. indicate that citrullination contributes to proper oligodendrocyte lineage progression and myelination [17]. Based on these findings, it can be suggested that the *PADI2* gene may be a candidate gene for schizophrenia, since oligodendrocyte and myelin dysfunction is known to lead to changes in synaptic formation and function, which can lead to cognitive dysfunction, a core symptom of schizophrenia [6].

According to the data from the 1000 Genomes Project, the frequency of the *rs2076617*A* allele in world populations varies, constituting 26.3% in Europeans (CEU); 39.5%, in Africans (AFR); and 57.3%, in Chinese (CHB) (https//www.ensembl.org/Homo_sapi-

Table 2. Allele and genotype frequency distributions of polymorphisms located in the 1p36.13 chromosomal region in the samples of patients with paranoid schizophrenia and control group of Bashkirs

Genotype/ Allele	Patients		Control		p	p_{fdr}	OR (CI95%)
	n	$p \pm$ spCI (%)	n	$p \pm$ sp CI (%)			
				rs2076617			
A/A	11	7.91 ± 2.29 4.02–13.72	37	18.14 ± 2.7 13.1–24.12	7.4E-03	0.841	0.39 (0.19–0.79)
A/G	52	37.41 ± 4.1 29.36–46.01	103	50.49 ± 3.5 43.42–57.54	0.017	0.862	0.59 (0.38–0.92)
G/G	76	54.68 ± 4.22 46.02–63.13	64	31.37 ± 3.25 25.07–38.22	1.6E-05	0.752	2.64 (1.69–4.12)
A	74	26.62 ± 2.65 21.52–32.23	177	43.38 ± 2.45 38.51–48.35	1.53E-05	0.864	0.47 (0.34–0.65)
G	204	73.38 ± 2.65 67.77–78.48	231	56.62 ± 2.45 51.65–61.49	1.53E-05	0.864	2.11 (1.52–2.94)
				rs2016693			
A/A	11	7.91 ± 2.29 4.02–13.72	42	20.59 ± 2.83 15.26–26.79	1.4E-05	0.988	0.33 (0.16–0.67)
A/C	56	40.29 ± 4.16 32.06–48.94	99	48.53 ± 3.5 41.49–55.61	0.132	0.931	
C/C	72	51.8 ± 4.24 43.17–60.35	63	30.88 ± 3.23 24.62–37.71	9.9E-05	0.822	2.41 (1.54–3.76)
A	78	28.06 ± 2.69 22.86–33.73	183	44.85 ± 2.46 39.96–49.82	1.95E-05	0.689	0.48 (0.35–0.67)
C	200	71.94 ± 2.69 66.27–77.14	225	55.15 ± 2.46 50.18–60.04	1.95E-05	0.689	2.09 (1.51–2.9)
				rs2057096			
G/G	16	11.51 ± 2.71 6.72–18.02	50	24.51 ± 3.01 18.77–31	2.7E-03	0.838	0.4 (0.22–0.74)
G/A	60	43.17 ± 4.2 34.8–51.83	100	49.02 ± 3.5 41.97–56.1	0.286	0.969	
A/A	63	45.32 ± 4.22 36.87–53.98	54	26.47 ± 3.09 20.55–33.08	3.0E-04	0.803	2.3 (1.46–3.63)
G	92	33.09 ± 2.82 27.59–38.96	200	49.02 ± 2.47 44.07–53.98	6.86E-05	0.744	0.51 (0.37–0.7)
A	186	66.91 ± 2.82 61.04 –72.41	208	50.98 ± 2.47 46.02–55.93	6.86E-05	0.744	1.94 (1.41–2.66)
				rs2057094			
C/C	16	11.51 ± 2.71 6.72–18.02	50	24.51 ± 3.01 18.77–31	2.7E-03	0.837	0.4 (0.22–0.74)
C/T	60	43.17 ± 4.2 34.8–51.83	100	49.02 ± 3.5 41.97–56.1	0.286	0.969	
T/T	63	45.32 ± 4.22 36.87–53.98	54	26.47 ± 3.09 20.55–33.08	3.0E-04	0.822	2.3 (1.46–3.63)
C	92	33.09 ± 2.82 27.59–38.96	200	49.02 ± 2.47 44.07–53.98	6.86E-05	0.717	0.51 (0.37–0.7)
T	186	66.91 ± 2.82 61.04–72.41	208	50.98 ± 2.47 46.02–55.93	6.86E-05	0.717	1.94 (1.41–2.66)

ens/Variation/Population?db=core;r=1:170820221-7-083022;v=rs2076617;vdb=variation;vf=1478939).

In the present study, the *rs2076617*A/A* homozygous genotype of the *PADI2* gene in patients with paranoid schizophrenia was found to be rare (7.91% of cases), while in controls it was detected more often (18.14% of cases) (p = 7.4E-03; OR = 0.39; CI95% 0.19–0.79). The *rs2076617*A/G* heterozygous genotype was also less common in patients (37.41%) than in healthy controls (50.49%) (p = 0.017; OR = 0.59; CI95% 0.38–0.92). The frequency of the *rs2076617*G/G* genotype of the *PADI2* gene was significantly higher in patients with PSz than in the control group (54.68 and 31.37%, respectively) (p = 1.6E-05; OR = 2.64; CI95% 1.69–4.12). Analysis of the allele frequency distribution of this polymorphic locus showed that the frequency of the *rs2076617*A* allele in patients with PSz was lower (26.62%) than in controls (43.38%). The *rs2076617*G* allele was detected in 73.38% of cases in patients with PSz, and in 56.62% of cases in healthy controls. Odds ratio for the *rs2076617*A* allele was 0.47 (CI95% 0.34–0.65), p = 1.53E–0.5; for the *rs2076617*G* allele, 2.11 (CI95% 0.34–0.65). However, after the introduction of the FDR-BH correction for multiple comparisons, these differences were appeared to be not statistically significant (*rs2076617*A/A* p_{fdr} = 0.841, *rs2076617*A/G* p_{fdr} = 0.862, *rs2076617*G/G* p_{fdr} = 0.752, *rs2076617*A* p_{fdr} = 0.864, *rs2076617*G* p_{fdr} = 0.864), Table 2.

In the sample of Bashkir patients and controls examined, the rs2016693 single nucleotide polymorphism showed strong association with PSz (Table 2). In patients, the frequency of the *rs2016693*A/A* homozygous genotype (7.91%) was significantly lower than that in the control group (20.59%) (P = 1.4E-05; OR = 0.33; CI95% 0.16–0.67). The *rs2016693*C/C* genotype was more frequent in patients (51.8%) than in controls (30.88%) (p = 9.9E-05; OR = 2.41; CI95% 1.54–3.76). The frequency of the *rs2016693*A* allele in the group of healthy controls was significantly higher (44.85%) than in patients (28.06%) (p = 1.95E-05; OR = 0.48; CI95% 0.35–0.67). The frequency of the *rs2016693*C* allele in patients (71.94%) exceeded its frequency in the control group, where it was 55.15% (OR = 2.09; CI95% 1.51–2.9). However, after the introduction of the FDR-BH correction, the differences in the genotype and allele frequency distributions at this polymorphic locus (rs2016693) turned out to be not statistically significant (Table 2).

Analysis of genotype and allele frequency distributions of the rs2057096 and rs2057094 SNPs showed identical frequency values for these polymorphic loci, and therefore, the results from the association study of only the rs2057096 SNP will be presented in detail below (Table 2). In the studied sample of patients and controls of Bashkir ethnicity, the rs2057096 SNP showed strong association with PSz (Table 1). In patients with paranoid schizophrenia, the frequency of the *rs2057096*G/G* homozygous genotype (11.51%) was significantly lower than that in the control group (24.51%) (p = 2.7E-03; OR = 0.41; CI95% 0.22–0.74). The *rs2057096*A/A* genotype was more frequent in patients with PS (45.32%) than in the control group (26.47%) (p = 3.0E-04; OR = 2.3; CI95% 1.46–3.63). The frequency of the *rs2057096*G* allele in the group of healthy controls was significantly higher (49.02%) than in patients with PSz (33.09%) (p = 6.86E-05; OR = 0.51; CI95% 0.37–0.7). The frequency of the *rs2057096*A* allele in patients with PSz (66.91%) exceeded its frequency in the control group, where it was 50.98% (OR = 1.94; CI95% 1.41–2.66). However, after the introduction of the FDR-BH correction, the differences in the genotype and allele frequency distributions at the rs2057096 polymorphic locus turned out to be not statistically significant (Table 2).

Thus, genome-wide association study showed the absence of association of paranoid schizophrenia in individuals of Bashkir ethnicity with the rs2076617 SNP of the *PADI2* gene, located in the 1p36.13 region, despite the available literature data demonstrating the association of the 1p36.13 chromosomal region [10–13] and the *PADI2* gene [16] with the development of schizophrenia in different populations. These differences may be associated both with an insufficient sample size for this type of research and indicate interpopulation differences in the formation of hereditary predisposition to paranoid schizophrenia.

ACKNOWLEDGMENTS

The author thanks the academic staff of the Division of Psychological Medicine and Clinical Neurosciences, Cardiff University, Cardiff (UK), M. O'Donovan, V. Escott-Price, M. Owen, and G. Leonenko for advice on data generation and analysis and participation in the project. Special thanks to the Director of the Institute of Biochemistry and Genetics, Ufa Federal Research Center, Russian Academy of Sciences, prof. E.R. Khusnutdinova for scientific guidance.

FUNDING

This work was supported by ongoing institutional funding. No additional grants to carry out or direct this particular research were obtained.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

This article does not contain any studies involving live animals.

All studies were conducted in accordance with the principles of biomedical ethics as outlined in the 1964 Declaration of Helsinki and its later amendments. They were also approved by the Ethics Committee of the Institute of Biochemistry and Genetics, Ufa Federal Research Center,

Russian Academy of Sciences, protocol No. 4 dated March 27, 2009.

Each participant in the study provided a voluntary written informed consent after receiving an explanation of the potential risks and benefits, as well as the nature of the upcoming study.

CONFLICT OF INTEREST

The author of this work declares that she has no conflicts of interest.

REFERENCES

1. Wang, J., Liu, J., Li, S., et al., Genetic regulatory and biological implications of the 10q24.32 schizophrenia risk locus, *Brain*, 2023. vol. 146, no. 4, pp.1403–1419. https://doi.org/10.1093/brain/awac352
2. Dennison, C.A., Legge, S.E., Pardiñas, A.F., and Walters, J., Genome-wide association studies in schizophrenia: recent advances, challenges and future perspective, *Schizophr. Res.*, 2020, vol. 217, pp. 4–12. https://doi.org/10.1016/j.schres.2019.10.048
3. O'Donovan, M.C., Craddock, N., Norton, N., et al., Molecular genetics of schizophrenia collaboration. identification of loci associated with schizophrenia by genome-wide association and follow-up, *Nat. Genet.*, 2008, vol. 40, no. 9, pp. 1053–1055. PMID: 18677311. https://doi.org/10.1038/ng.201
4. Ripke, S., Neale, B.M., Corvin, A., et al., Biological insights from 108 schizophrenia-associated genetic loci, *Nature*, 2014, vol. 511, no. 7510, pp. 421–427. https://doi.org/10.1038/nature13595
5. Lam, M., Chen, C.Y., Li, Z., et al., Comparative genetic architectures of schizophrenia in East Asian and European populations, *Nat. Genet.*, 2019, vol. 51, no. 12, pp. 1670–1678. https://doi.org/10.1038/s41588-019-0512-x
6. Trubetskoy, V., Pardiñas, A.F., Qi, T., et al., Mapping genomic loci implicates genes and synaptic biology in schizophrenia, *Nature*, 2022, vol. 604, no. 7906, pp. 502–508. https://doi.org/10.1038/s41586-022-04434-5
7. Purcell, S., Neale, B., Todd-Brown, K., et al., PLINK: a toolset for whole-genome association and population-based linkage analysis, *Am. J. Hum. Genet.*, 2007, vol. 81, no. 3, pp. 559–575. https://doi.org/10.1086/519795
8. Gareeva, A.E., Genome-wide association study of the risk of schizophrenia in the Republic of Bashkortostan, *Russ. J. Genet.*, 2023, vol. 59, no. 8, pp. 831–839. https://doi.org/10.1134/S1022795423080070
9. Benjamini, Y., Drai, D., Elmer, G., et al., Controlling the false discovery rate in behavior genetics research, *Behav. Brain Res.*, 2001, vol. 125, nos. 1–2, pp. 279–284. https://doi.org/10.1016/s0166-4328(01)00297-2
10. Abecasis, G.R., Burt, R.A., Hall, D., et al., Genome-wide scan in families with schizophrenia from the founder population of Afrikaners reveals evidence for linkage and uniparental disomy on chromosome 1, *Am. J. Hum. Genet.*, 2004, vol. 74, no. 3, pp. 403–417. https://doi.org/10.1086/381713
11. Escamilla, M.A., Ontiveros, A., Nicolini, H., et al., A genome-wide scan for schizophrenia and psychosis susceptibility loci in families of Mexican and Central American ancestry, *Am. J. Med. Genet.: Neuropsychiatr. Genet.*, 2007, vol. 144B, no. 2, pp. 193–199. https://doi.org/10.1002/ajmg.b.30411
12. Greenwood, T.A., Swerdlow, N.R., Gur, R.E., et al., Genome-wide linkage analyses of 12 endophenotypes for schizophrenia from the Consortium on the Genetics of Schizophrenia, *Am. J. Psychiatry*, 2013, vol. 170, no. 5, pp. 521–532. https://doi.org/10.1176/appi.ajp.2012.12020186
13. Greenwood, T.A., Lazzeroni, L.C., Calkins, M.E., et al., Genetic assessment of additional endophenotypes from the Consortium on the Genetics of Schizophrenia Family Study, *Schizophr. Res.*, 2016, vol. 170, no. 1, pp. 30–40. https://doi.org/10.1016/j.schres.2015.11.008
14. Wang, L., Chen, H., Tang, J., et al., Peptidylarginine deiminase and Alzheimer's disease, *J. Alzheimer's Dis.*, 2022, vol. 85, no. 2, pp. 473–484. https://doi.org/10.3233/JAD-215302
15. Bradford, C.M., Ramos, I., Cross, A.K., et al., Localisation of citrullinated proteins in normal appearing white matter and lesions in the central nervous system in multiple sclerosis, *J. Neuroimmunol.*, 2014, vol. 273, nos. 1–2, pp. 85–95. https://doi.org/10.1016/j.jneuroim.2014.05.007
16. Watanabe, Y., Nunokawa, A., Kaneko, N., et al., A two-stage case-control association study of PADI2 with schizophrenia, *J. Hum. Genet.*, 2009, vol. 54, no. 7, pp. 430–432. https://doi.org/10.1038/jhg.2009.52
17. Falcão, A.M., Meijer, M., Scaglione, A., et al., PAD2-mediated citrullination contributes to efficient oligodendrocyte differentiation and myelination, *Cell Rep.*, 2019, vol. 27, no. 4, pp. 1090–1102. https://doi.org/10.1016/j.celrep.2019.03.108

Translated by N. Maleeva