International Journal of Biomedicine 12(1) (2022) 155-159
http://dx.doi.org/10.21103/Article12(1)_OA19

ORIGINAL ARTICLE

Population Genetics

INTERNATIONAL JOURNAL OF BIOMEDICINE

Frequency of the Risk A Allele of rs17713054 Localized in the 3p21.31 COVID-19 Risk Locus in the Yakut Population

Nadezhda I. Pavlova[1*], PhD; Aleksey A. Bochurov[1]; Vladislav A. Alekseev[1]; Aleksandra T. Diakonova[1]; Vladimir V. Dodokhov[2], PhD; Khariton A. Kurtanov[3], PhD

[1]*Yakut Science Center of Complex Medical Problems*
[2]*Arctic State Agrotechnological University*
[3]*Institute for Biological Problems of Cryolithozone of SB of RAS*
Yakutsk, the Republic of Sakha (Yakutia), Russia

Abstract

Background: Genome-wide association studies identified the region of chromosome 3p21.31 as having the strongest association with the severe COVID-19 and susceptibility to SARS-CoV-2 infection. The aim of our study was to investigate the frequency of the risk A allele of rs17713054 localized in the 3p21.31 COVID-19 risk locus in Yakuts.

***Methods and Results*:** A total of 382 DNA samples from healthy Yakut volunteers (184 men and 198 women; the average age of 41.8±0.05 years) were examined. Our results show that the frequency of the risk A allele of the rs17713054 SNP in the Yakut population occurs at a frequency of 2% and generally corresponds to the frequency of East Asian populations (from 0% to 2%), geographically close to the Yakuts and belonging to the same Mongoloid race. (**International Journal of Biomedicine. 2022;12(1):155-159.**)

Key Words: SARS-CoV-2 • rs17713054 • epithelial-mesenchymal transition • leucine zipper transcription factor like 1

For citation: Pavlova NI, Bochurov AA, Alekseev VA, Diakonova AT, Dodokhov VV, Kurtanov KhA. Frequency of the Risk A Allele of rs17713054 Localized in the 3p21.31 COVID-19 Risk Locus in the Yakut Population. International Journal of Biomedicine. 2022;12(1):155-159. doi:10.21103/Article12(1)_OA19

Abbreviations

EMT, epithelial-mesenchymal transition; **GWAS**, genome-wide association study; **SNP**, single nucleotide polymorphism; **LZTFL1**, leucine zipper transcription factor like 1.

Introduction

The COVID-19 pandemic has killed millions of people worldwide. The predominant cause of death is pneumonia and severe acute respiratory distress syndrome.[1] However, COVID-19 can cause multiple organ failure due to cytokine release, micro- and macrovascular thrombosis, endothelial injury, acute kidney injury, and myocarditis.[2-4] A study of post-mortem lung biopsies from COVID-19 demonstrated widespread epithelial dysfunction with signs of EMT.[5,6]

Two large GWASs [7,8] identified the region of chromosome 3p21.31 as having the strongest association with the severe COVID-19; another study identified this locus as conferring susceptibility to SARS-CoV-2 infection.[9] GWASs have found that the 3p21.31 region is associated with a two-fold increase in the risk of respiratory failure. The influence of the 3p21.31 locus on the early epithelial response may contribute to susceptibility to SARS-CoV-2 infection.[9]

Downes et al.[10] identified the gain-of-function risk A allele of an SNP, rs17713054G>A, as a probable causative variant for the respiratory failure. The authors showed that

***Corresponding author:** Nadezhda I. Pavlova, PhD. Yakut Science Center of Complex Medical Problems. Yakutsk, the Republic of Sakha (Yakutia), Russia. E-mail: solnishko_84@inbox.ru*

the rs17713054-affected enhancer upregulates the interacting gene, leucine zipper transcription factor like 1 (*LZTFL1*). "This gene encodes a ubiquitously expressed protein that localizes to the cytoplasm. This protein interacts with Bardet-Biedl Syndrome (BBS) proteins and, through its interaction with BBS protein complexes, regulates protein trafficking to the ciliary membrane. Nonsense mutations in this gene cause a form of Bardet-Biedl Syndrome; a ciliopathy characterized in part by polydactyly, obesity, cognitive impairment, hypogonadism, and kidney failure. This gene may also function as a tumor suppressor; possibly by interacting with E-cadherin and the actin cytoskeleton and thereby regulating the transition of epithelial cells to mesenchymal cells."[11] The transmembrane protein E-cadherin is one of the critical components for epithelial integrityIn addition.[12] Thus, LZTFL1 is involved in the trafficking of numerous signaling molecules. [13-17]

For the 3p21.31 COVID-19 risk locus, a higher risk is associated with increased expression of the *LZTFL1* gene, a known inhibitor of EMT. Higher levels of the *LZTFL1* gene may delay the positive effects of an acute EMT response by blocking a reduction in the levels of ACE2 and TMPRSS2 and/or by slowing EMT-driven tissue repair.

The *LZTFL1* gene is widely expressed in pulmonary epithelial cells, including ciliated epithelial cells, which have been identified as one of the main cellular targets for SARS-CoV-2 infection.[18] The *LZTFL1* gene was identified as the gene having strong 3C contacts with the rs17713054 enhancer and lung expression quantitative trait loci (eQTL). *LZTFL1* is the most likely direct regulatory target of the rs17713054-containing epithelial–endothelial–fibroblast enhancer.

The aim of our study was to investigate the frequency of the risk A allele of rs17713054 localized in the 3p21.31 COVID-19 risk locus in Yakuts.

Materials and Methods

The study was carried out in the Department of Molecular Genetics at the Yakut Scientific Center for Complex Medical Problems (YSC CMP). For the study, we used DNA samples from the collection of biomaterials of the YSC CMP (Project "The Genome of Yakutia"; No. USE_507512).

A total of 382 DNA samples from healthy volunteers (184 men and 198 women; the average age of 41.8±0.05 years) were examined. The inclusion criteria for the study were Yakuts by ethnicity, living in Yakutia.

The study was approved by the Ethics Committee of the Yakut Science Center of Complex Medical Problems (YSC CMP). Written informed consent was obtained from each research participant.

Genomic DNA was isolated using the standard phenol-chloroform extraction method from frozen whole blood. Genotyping of DNA samples was carried out by analyzing PCR products - amplification of specific regions of the genome, followed by an analysis of restriction fragment length polymorphism. Primer selection was performed using the National Center for Biotechnology Information (NCBI) primer design tool, Primer-BLAST (forward primer: 5'-TGTCTGATTTTAAAGAAGTTTGGGT-3' and reverse primer: 5'- GGAGCAGAGCCCGTCATTAT-3').The sequence of the study site for the primer selection matrix and the primer specificity test were taken from the UCSC Genome Browser database (GRCh38/hg38). The primers were synthesized by Lumiprobe RUS Ltd (Moscow, Russia). The reaction mixture (25 µl) for PCR contained of forward and reverse primer (1.5µl (10 pmol/µl) of each oligonucleotide primer) (Moscow, Russia), ddH2O (13 µl), 10xPCR buffer (2.5 µl), 25 mM MgCl2 (2.5 µl), 2.5 mM dNTP Mix (2.5 µl), Taq polymerase (0.5 µl (1.5 units)), and DNA (1 µl). PCR was performed in an MJ Mini Gradient Thermal Cycler (BioRad). The temperature-time regime for carrying out the amplification of a given nucleotide sequence with the restriction enzyme used and the lengths of the restriction fragments is presented in Table 1. PCR products were detected by horizontal electrophoresis in a 2% agarose gel plate with the addition of ethidium bromide, a specific intercalating fluorescent DNA (RNA) dye, using a standard Tris-acetate buffer at a field voltage of ~ 20V/cm for 30 minutes.

Table 1.

PCR conditions for rs17713054

Temperature protocol for PCR			Restriction enzyme	Interpretation of the PCR-RFLP result bp)
Temperature	Time	Cycles		
95 °C	4 min	1	Zrm I	AA – 210; AG – 210, 125, 85; GG – 125, 85.
94 °C	30 sec	35		
64 °C	30 sec			
72 °C	1 min			
72 °C	10 min	1		

***bp**, base pair*

After PCR, the amplificate was subjected to restriction using ZrmI endonuclease (OOO SibEnzim, Novosibirsk) for 16 hours at 37°C. Restriction endonuclease selection was performed in a NEBcutter V2.0. RFLP products were detected by horizontal electrophoresis in a 3% agarose gel stained with ethidium bromide using a standard Tris-acetate buffer at 120V for 1 hour. Restriction products were visualized under UV light using a gel-documenting system (Vilber Lourmat, France) (Figure 1).

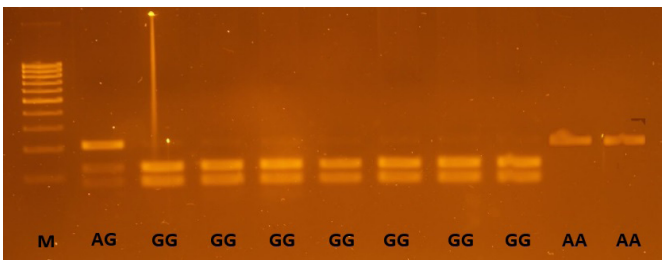

Fig. 1. Electrophoretogram: RFLP products in a 3% agarose gel.

M - marker PUC19 / + Msp I., AG – 210, 125, 85 bp; GG – 125, 85 bp; AA – 210 bp.

Statistical analysis was performed using Microsoft Excel 2010.

Results and Discussion

In our study, the rs17713054 SNP was analyzed in the Yakut population in comparison with the populations of East and South Asia, Africa, America, and Europe.[19] In the Yakut (YKT) population, the frequency of the risk A allele and the G allele of the rs17713054 was 2% and 98%, respectively (Table 2). The genotype distribution was as follows: AG=3.7%, AA=0.5%, GG=95.8%. When compared with an open database of populations from the 1000 Genomes

Table 2.

The frequency of genotypes and alleles of the rs17713054 SNP in the Yakut population and 1000 Genomes Project Phase 3 (population and subpopulations)

Group	ID	Name	Geographic region	Count	Allele freq.		Genotype freq.			Chi-sq.	*P*
					A	G	AA	AG	GG		
All Yakuts	YKT	Yakuts	North-Eastern Siberia	382	0.020	0.980	0.005	0.037	0.958	15.813	0.000
Yakut men				184	0.010	0.990	0	0.027	0.973	0.035	0.852
Yakut women				198	0.033	0.967	0.011	0.045	0.944	15.992	0.000
All 1000 Genomes Project Phase 3 individuals				2504	0.082	0.918	0,019	0.126	0.855	64.538	0.000
East Asian, EAS				504	0.005	0.995	0	0.010	0.990	0.013	0.911
East Asian sub-populations	CDX	Dai	East Asia	93	0.005	0.995	0	0.011	0.989	0.003	0.958
	CHB	Northern Han	East Asia	103	0	1	0	0	1	-	-
	CHS	Southern Han	East Asia	105	0	1	0	0	1	-	-
	JPT	Japanese	East Asia	104	0	1	0	0	1	-	-
	KHV	Kinh	East Asia	99	0.020	0.980	0	0.040	0.960	0.042	0.837
South Asian, SAS				489	0.296	0.704	0,094	0.403	0.503	0.514	0.473
South Asian sub-populations	BEB	Bengali	Central/South Asia	86	0.378	0.622	0.128	0.500	0.372	0.346	0.557
	GIH	Gujarati	Central/South Asia	103	0.267	0.733	0.078	0.379	0.544	0.110	0.741
	ITU	Telugu	Central/South Asia	102	0.289	0.711	0.098	0.382	0.520	0.500	0.479
	PJL	Punjabi	Central/South Asia	96	0.297	0.703	0.094	0.406	0.500	0.069	0.792
	STU	Sri Lankan Tamil	Central/South Asia	102	0.260	0.740	0.078	0.363	0.559	0.330	0.566
European, EUR				503	0.081	0.919	0	0.161	0.839	3.857	0.050
European sub-populations	CEU	European American	Europe	99	0.081	0.919	0	0.162	0.838	0.765	0.382
	FIN	Finnish	Europe	99	0.101	0.899	0	0.202	0.798	1.250	0.264
	GBR	British	Europe	91	0.071	0.929	0	0.143	0.857	0.538	0.463
	IBS	Iberian	Europe	107	0.047	0.953	0	0.093	0.907	0.257	0.612
	TSI	Toscani	Europe	107	0.103	0.897	0	0.206	0.794	1.405	0.236
American, AMR				347	0.043	0.957	0,003	0.081	0.916	0.208	0.648
American sub-populations	CLM	Colombian	Admixed	94	0.064	0.936	0.011	0.106	0.883	1.134	0.287
	MXL	Mexican American	Admixed	64	0.039	0.961	0	0.078	0.922	0.106	0.745
	PEL	Peruvian	Admixed	85	0.029	0.971	0	0.059	0.941	0.078	0.780
	PUR	Puerto Rican	Admixed	104	0.038	0.962	0	0.077	0.923	0.166	0.683
African, AFR				661	0.004	0.996	0	0.008	0.992	0.010	0.922
African sub-populations	ACB	Afro-Caribbean	Admixed	96	0.016	0.984	0	0.031	0.969	0.024	0.876
	ASW	African American	Admixed	61	0.008	0.992	0	0.016	0.984	0.004	0.949
	ESN	Esan	Africa	99	0	1	0	0	1	-	-
	GWD	Gambian	Africa	113	0.004	0.996	0	0.009	0.991	0.002	0.962
	LWK	Luhya	Africa	99	0	1	0	0	1	-	-
	MSL	Mende	Africa	85	0	1	0	0	1	-	-
	YRI	Yoruban	Africa	108	0	1	0	0	1	-	-

ID, population and subpopulation identifier; P - HWE p-value. A probability value of P<0.05 was considered statistically significant.

Project, the Yakuts were similar to the Kinh people of northern Vietnam (KHV subpopulation: A=2%; AA=0 and AG=4%) and the African subpopulation in Barbados (ACB: A=1.6%; AA=0 and AG=3.1%) (Table 2). In the African population (AFR), the risk A allele was absent, except for the aforementioned African subpopulation from the state of Barbados in the Caribbean (ACB) and African Americans from Southwest US (ASW: A=0.8%; AA=0 and AG=1.6%). Probably, in the ACB and ASW populations, the risk A allele is present due to cross-breeding with Europeans. In American populations (AMR), the risk A allele occurs at a frequency of 4% (AA=0.3% and AG=8.1%). Europeans rank second in the world in terms of the prevalence of the risk A allele (EUR: A=8%; AA=0 and AG=16.1%). Among Europeans, a high frequency is observed in Italians (TSI: A=10.3%; AA=0 and AG=20.6%) and Finns (FIN: A=10.1%; AA=0 and AG=20.2%).

A very high frequency of the risk A allele and allele-carrying A genotypes was observed in the South Asian population (SAS: A=30%; AA=9.4% and AG=40.4%), compared with all other considered populations (Table 2). In all the South Asian subpopulations presented, the frequency of the minor allele was high (26%-37.8%), but it should be noted that Bengalis have the highest frequency of the A allele of all analyzed populations (37.8%). The heterozygous AG genotype occurs in 50% of individuals in the Bengal population and 40.6% in the Punjabi population living in Pakistan and the Punjab state of northern India (PJL).

A nationwide cohort study of 29 million adults in England performed by Nafilyan et al.[20] analyzed ethnic differences for the risk of COVID-19-related death in the first and second waves of the pandemic in England. In the first wave, all ethnic minority groups had a higher risk of COVID-19-related death than the White British population. In the second wave, the risk of COVID-19-related death remained elevated for Pakistani (ASMR: 339.9 [95% CI: 303.7-376.2] and 166.8 [141.7-191.9] deaths per 100,000 population in men and women) and Bangladeshi (318.7 [247.4-390.1] and 127.1 [91.1-171.3] in men and women) people, but not for people from Black ethnic groups. Despite a reduction in the increased risk of mortality from COVID-19, after adjusting for socio-demographic characteristics and health status, the risk was significantly higher in people of Bangladeshi and Pakistani origin in both the first and second waves.

In conclusion, our results show that the frequency of the risk A allele of the rs17713054 SNP in the Yakut population occurs at a frequency of 2% and generally corresponds to the frequency of East Asian populations (from 0% to 2%), geographically close to the Yakuts and belonging to the same Mongoloid race. Probably, the risk A allele is present in the Yakut population due to ancient and/or recent miscegenation with Europeans.

Sources of Funding

The study was carried out within the framework of the project "The study of the genetic structure and burden of hereditary pathology of populations of the Republic of Sakha (Yakutia), using the Unique scientific equipment (USE) "The Genome of Yakutia" (Reg. No. USU_507512).

Competing interests

The authors declare that they have no competing interests.

References

1. Marini JJ, Hotchkiss JR, Broccard AF. Bench-to-bedside review: microvascular and airspace linkage in ventilator-induced lung injury. Crit Care. 2003 Dec;7(6):435-44. doi: 10.1186/cc2392.
2. Levi M, Thachil J, Iba T, Levy JH. Coagulation abnormalities and thrombosis in patients with COVID-19. Lancet Haematol. 2020 Jun;7(6):e438-e440. doi: 10.1016/S2352-3026(20)30145-9.
3. Varga Z, Flammer AJ, Steiger P, Haberecker M, Andermatt R, Zinkernagel AS, et al. Endothelial cell infection and endotheliitis in COVID-19. Lancet. 2020 May 2;395(10234):1417-1418. doi: 10.1016/S0140-6736(20)30937-5.
4. Ackermann M, Verleden SE, Kuehnel M, Haverich A, Welte T, Laenger F, et al. Pulmonary Vascular Endothelialitis, Thrombosis, and Angiogenesis in Covid-19. N Engl J Med. 2020 Jul 9;383(2):120-128. doi: 10.1056/NEJMoa2015432.
5. He J, Cai S, Feng H, Cai B, Lin L, Mai Y, et al. Single-cell analysis reveals bronchoalveolar epithelial dysfunction in COVID-19 patients. Protein Cell. 2020 Sep;11(9):680-687. doi: 10.1007/s13238-020-00752-4.
6. Borczuk AC, Salvatore SP, Seshan SV, Patel SS, Bussel JB, Mostyka M, et al. COVID-19 pulmonary pathology: a multi-institutional autopsy cohort from Italy and New York City. Mod Pathol. 2020 Nov;33(11):2156-2168. doi: 10.1038/s41379-020-00661-1.
7. Severe Covid-19 GWAS Group, Ellinghaus D, Degenhardt F, Bujanda L, Buti M, Albillos A, Invernizzi P, et al. Genomewide Association Study of Severe Covid-19 with Respiratory Failure. N Engl J Med. 2020 Oct 15;383(16):1522-1534. doi: 10.1056/NEJMoa2020283.
8. Pairo-Castineira E, Clohisey S, Klaric L, Bretherick AD, Rawlik K, Pasko D, et al.; GenOMICC Investigators; ISARIC4C Investigators; COVID-19 Human Genetics Initiative; 23andMe Investigators; BRACOVID Investigators; Gen-COVID Investigators, Shen X, Ponting CP, Fawkes A, Tenesa A, Caulfield M, Scott R, Rowan K, et al. Genetic mechanisms of critical illness in COVID-19. Nature. 2021 Mar;591(7848):92-98. doi: 10.1038/s41586-020-03065-y.
9. COVID-19 Host Genetics Initiative. Mapping the human genetic architecture of COVID-19. Nature. 2021 Dec;600(7889):472-477. doi: 10.1038/s41586-021-03767-x.
10. Downes DJ, Cross AR, Hua P, Roberts N, Schwessinger R, Cutler AJ, et al.; COvid-19 Multi-omics Blood ATlas (COMBAT) Consortium, Gill DR, Hyde SC, Knight JC, Todd JA, Sansom SN, Issa F, Davies JOJ, Hughes JR. Identification of LZTFL1 as a candidate effector gene at a COVID-19 risk locus. Nat Genet. 2021 Nov;53(11):1606-1615. doi: 10.1038/s41588-021-00955-3.
11. NCBI Resources. LZTFL1 leucine zipper transcription factor like 1 [Homo sapiens (human)]. Available from: https://www.ncbi.nlm.nih.gov/gene/54585

12. Ramirez Moreno M, Stempor PA, Bulgakova NA. Interactions and Feedbacks in E-Cadherin Transcriptional Regulation. Front Cell Dev Biol. 2021 Jun 28;9:701175. doi: 10.3389/fcell.2021.701175.
13. Wei Q, Zhou W, Wang W, Gao B, Wang L, Cao J, Liu ZP. Tumor-suppressive functions of leucine zipper transcription factor-like 1. Cancer Res. 2010 Apr 1;70(7):2942-50. doi: 10.1158/0008-5472.CAN-09-3826.
14. Promchan K, Natarajan V. Leucine zipper transcription factor-like 1 binds adaptor protein complex-1 and 2 and participates in trafficking of transferrin receptor 1. PLoS One. 2020 Jan 2;15(1):e0226298. doi: 10.1371/journal.pone.0226298.
15. Starks RD, Beyer AM, Guo DF, Boland L, Zhang Q, Sheffield VC, Rahmouni K. Regulation of Insulin Receptor Trafficking by Bardet Biedl Syndrome Proteins. PLoS Genet. 2015 Jun 23;11(6):e1005311. doi: 10.1371/journal.pgen.1005311.
16. Wei Q, Gu YF, Zhang QJ, Yu H, Peng Y, Williams KW, et al. Lztfl1/BBS17 controls energy homeostasis by regulating the leptin signaling in the hypothalamic neurons. J Mol Cell Biol. 2018 Oct 1;10(5):402-410. doi: 10.1093/jmcb/mjy022.
17. Seo S, Zhang Q, Bugge K, Breslow DK, Searby CC, Nachury MV, Sheffield VC. A novel protein LZTFL1 regulates ciliary trafficking of the BBSome and Smoothened. PLoS Genet. 2011 Nov;7(11):e1002358. doi: 10.1371/journal.pgen.1002358.
18. Ravindra NG, Alfajaro MM, Gasque V, Huston NC, Wan H, Szigeti-Buck K, et al. Single-cell longitudinal analysis of SARS-CoV-2 infection in human airway epithelium identifies target cells, alterations in gene expression, and cell state changes. PLoS Biol. 2021 Mar 17;19(3):e3001143. doi: 10.1371/journal.pbio.3001143.
19. Reference SNP (rs) Report (rs1773054). Available from: https://www.ncbi.nlm.nih.gov/snp/rs17713054?horizontal_tab=true
20. Nafilyan V, Islam N, Mathur R, Ayoubkhani D, Banerjee A, Glickman M, Humberstone B, Diamond I, Khunti K. Ethnic differences in COVID-19 mortality during the first two waves of the Coronavirus Pandemic: a nationwide cohort study of 29 million adults in England. Eur J Epidemiol. 2021 Jun;36(6):605-617. doi: 10.1007/s10654-021-00765-1.