

The structure of Haplotypes and Diplotypes in the *PNPLA3* gene in the Yakut Population

Khariton A. Kurtanov, PhD^{1*}; Nadezhda I. Pavlova, PhD¹; Aleksandra T. Diakonova¹;
Lyubovy A. Sydykova, PhD²; Tuyara N. Aleksandrova¹; Yigulana P. Borisova¹;
Vladimir V. Dodokhov, PhD³

¹*Yakut Science Center of Complex Medical Problems*

²*M. K. Ammosov North-Eastern Federal University*; ³*Yakut State Agricultural Academy*
Yakutsk, the Republic of Sakha (Yakutia), Russia

Abstract

Background: The first GWAS searching for such genetic factors identified the *PNPLA3* gene as a major genetic determinant for the predisposition to nonalcoholic fatty liver disease in Hispanic, African American, and European American populations, according to liver fat contents, a finding that was subsequently confirmed by liver biopsy in Europeans and Asians. The aim of our research was to study the distribution of alleles, genotypes, haplotypes and diplotypes of polymorphic variants of the *PNPLA3* gene (rs2294918 and rs738409) in Yakuts.

Methods and Results: The *PNPLA3* SNPs (rs2294918 and rs738409) were analyzed by PCR-RFLP reaction. The *PNPLA3* rs738409 SNP in the Yakut population is characterized by a high frequency of the risk G allele (72%). According to the *PNPLA3* rs2294918 SNP, which suppresses the negative effect of rs738409, the protective *PNPLA3* (rs2294918) A allele was found only in 10.7% of study subjects.

Analysis of the distribution of the frequency of genotypes in the studied sample of Yakuts showed the predominance of the carriage of the *PNPLA3* rs738409 GG genotype (57.3%) and the *PNPLA3* rs2294918 GG genotype (80.7%). The frequency of the *PNPLA3* (rs2294918) AA and AG genotypes was 2.0% and 17.3%, respectively. The Yakuts often have two diplotypes [GG]-[GG] and [CG]-[GG]. Both diplotypes carry the *PNPLA3* rs738409 G allele (45.3% and 28%) and do not carry the *PNPLA3* rs2294918 A allele. The high frequency of the [GG]-[GG] and [CG]-[GG] diplotypes in Yakuts (45.3% and 25%, respectively), carrying mutant alleles G (rs738409) and not carrying the A allele (rs2294918), indicates that these diplotypes were probably adaptively favorable to the Yakuts.

Conclusion: The analysis of haplotypes and diplotypes based on the markers rs738409 and rs2294918 of the *PNPLA3* gene may contribute to future new biomarkers for the diagnosis of nonalcoholic fatty liver disease and nonalcoholic steatohepatitis, as well as provide fundamental knowledge on human adaptation to cold. (**International Journal of Biomedicine. 2020;10(4):433-437.**)

Key Words: haplotype • diplotype • Yakut • *PNPLA3* • rs738409 • rs2294918

Abbreviations

NAFLD, nonalcoholic fatty liver disease; **PNPLA3**, patatin-like phospholipase domain-containing protein 3; **GWAS**, genome-wide association study; **SNP**, single nucleotide polymorphism; **HWE**, Hardy–Weinberg Equilibrium; **NASH**, nonalcoholic steatohepatitis.

Introduction

Maintaining body temperature has been relevant for almost the entire history of mankind, especially in the northern regions of the planet. But in the modern world, even in the

north, people live all year round in warm houses, dress in warm clothes, and are exposed to the minimum effect of cold on the body. The hottest organ in our body is the liver. It heats up from 37.8°C to 38.5°C. This difference is due to the tasks that it performs. The liver is the central chemical laboratory in the

body, responsible for the production of digestive bile, filtration of the blood, and the processing of raw materials supplied with food into the necessary chemical elements for the work of other organs. However, in addition to the liver, there are two other most intense heat sources: muscles and brown adipose tissue. Brown fat is a major site of immobile thermogenesis, but other tissues, especially muscle and liver, can help adapt to the cold. Simcox et al.⁽¹⁾ demonstrated in mice exposed to cold that the acylcarnitines produced by the liver are required to maintain thermogenesis. Maintaining body weight requires balancing energy from food intake with energy expenditure.

In vitro studies have shown that PNPLA3 acts as an acyltransferase, which catalyzes the conversion of lysophosphatidic acid to phosphatidic acid, while the p.I148M variant reduces the enzyme activity.^(2,3) At the same time, the physiological function of PNPLA3 is not fully understood. It remains unclear, for example, what is the role of changes in enzyme activity in the mechanism that forms hepatic steatosis.^(4,5)

The first GWAS searching for such genetic factors identified the *PNPLA3* gene as a major genetic determinant for the predisposition to NAFLD in Hispanic, African American, and European American populations, according to liver fat contents,⁽⁶⁾ a finding that was subsequently confirmed by liver biopsy in Europeans and Asians. The most prominent variant is the *PNPLA3* rs738409 [G], which is a nonsynonymous substitution of cytosine for guanine (C>G) that changes codon 148 from encoding isoleucine (I) to methionine (M) (I>M, I148M).^(6,7)

Recently completed studies have shown that PNPLA3 can participate in the remodeling of lipid droplets in hepatocytes, leading to the accumulation of p.I148M variant protein on the surface of these droplets by disrupting ubiquitination (a decrease in the amount of ubiquitinated PNPLA3 against the background of inhibition of the bortezomib proteasome makes it difficult to release triglyceride triglycerides) and reducing the hydrolysis of triglycerides by lipases.^(8,9) This hypothesis supports the conclusion that the mutant rs2294918 polymorphism of the *PNPLA3* gene decreases the expression of PNPLA3 and may weaken the effect of p.I148M, thereby reducing the risk of developing steatosis and steatohepatitis. At the same time, the *PNPLA3* E434K (the rs2294918 G>A polymorphism) variant does not affect the functional ability of PNPLA3 to reduce the content of intracellular fat in hepatocytes.⁽¹⁰⁾

It is also interesting that the *PNPLA3* p.I148M variant is present in stellate cells, and the mutant allele activates stellate cells regardless of its role in hepatocytes.⁽¹¹⁾ In addition to inducing an accumulation of triglycerides in hepatocytes, the *PNPLA3* p.I148M variant increases the risk of progression of fatty liver disease.⁽¹²⁾ It was also shown that the *PNPLA3* p.I148M variant was associated with a higher level of alanine aminotransferase.⁽⁶⁾ The combination of the *PNPLA3* p.I148M variant, an increased level of aspartate aminotransferase, and a high level of fasting insulin became an accurate predictor of histologically confirmed, nonalcoholic steatohepatitis in patients with NAFLD.⁽¹³⁾ Later, it was found that the *PNPLA3* p.I148M variant is associated with the development of nonalcoholic steatohepatitis, liver fibrosis, and hepatocellular carcinoma.⁽¹⁴⁾

We previously found a high frequency of the *PNPLA3* (rs738409) [G] allele in the Yakut population (73%).⁽¹⁵⁾ The accumulation of triglycerides in hepatocytes, associated with the *PNPLA3* p.I148M variant, was probably an adaptation to a cold climate; this accumulation is not needed in the modern world, but leads to NAFLD among the Yakut population.

The aim of our research was to study the distribution of alleles, genotypes, haplotypes and diplotypes of polymorphic variants of the *PNPLA3* gene (rs2294918 and rs738409) in Yakuts.

Materials and Methods

The study of the *PNPLA3* SNPs (rs2294918 and rs738409) was carried out in the Department of Molecular Genetics at 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 150 DNA samples from healthy volunteers were examined. The inclusion criteria for the study were Yakuts by ethnicity, living in Yakutia, without liver damage by chronic viral hepatitis. Exclusion criteria: autoimmune hepatitis, primary biliary cholangitis, primary sclerosing cholangitis, hereditary hemochromatosis, Wilson-Konovalov disease, and alcohol abuse (>30g/l). For comparison, data from the 1000 Genomes project were used for the populations of Peruvians (n=85), Mexicans (n=64), Europeans from Utah in the United States (n=99), Japanese (n=104) and Yoruba tribes (n=108).⁽¹⁶⁾

Genomic DNA samples were isolated from peripheral blood lymphocytes using a commercial DNA kit, Excel Biotech (Yakutsk). The *PNPLA3* SNPs (rs2294918 and rs738409) were analyzed by PCR-RFLP reaction. The gene region containing the polymorphic variant was amplified with standard pairs of primers produced by Biotech-Industry LLC (Moscow, Russia). The reaction mixture (25 µl) for PCR contained of forward and reverse primer (Moscow, Russia), Dream Taq PCR master mix (12.5 µl), deionized water (9.5 µl), and DNA in the amount of 100 µg/ml (1 µl). A mixture for RFLP (20 µL) consisted of amplicate (7 µL), deionized water (10.9 µL), restriction buffer (2 µL), and the corresponding restriction endonucleases (2 e.a.).

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. RFLP products were detected by horizontal electrophoresis in 4% agarose gel stained with ethidium bromide using a standard Tris-acetate buffer at 120V for 1 hour.⁽¹⁵⁾

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 (or the participant's parent/guardian).

Statistical analysis was performed using Microsoft Excel 2010. The correspondence of the distributions of genotypes to the expected values at HWE and comparison of the frequencies of allelic variants/genotypes were performed using the chi-square test. Possible haplotype variants were calculated using the Punnett square (Table 1). The haplotype frequency was determined using the EM algorithm. Linkage

disequilibrium (LD) between SNP pairs was calculated using Lewontin's coefficient D' and Pearson's r^2 coefficient. Using the solid spine LD algorithm, a block structure was built ($D' > 0.75$). Haploview 4.2 software was used to construct the *PNPLA3* haplotypes and frequencies based on genotyping data and to perform association tests between alleles and haplotypes of the *PNPLA3* gene.⁽¹⁷⁾

Table 1.

Possible configurations (haplotypes) at two loci of the *PNPLA3* gene (shown in the Punnett square)

<i>PNPLA3</i>	Locus (SNP)	rs738409		
Locus (SNP)	Genotype	CC	CG	GG
rs2294918	AA	CA	CA, GA	GA
	AG	CA, CG	GA, CA, CG, GG	GG, GA
	GG	CG	GG, CG	GG

Results and Discussion

The *PNPLA3* rs738409 SNP in the Yakut population is characterized by a high frequency of the risk G allele (72%) (Table 2). The data from open sources of the 1000 Genomes project⁽¹⁰⁾ reveal that the *PNPLA3* (rs738409) [G] allele is also found with a high frequency in populations of Central and South

America (Peruvians - 71.8%, Mexicans - 55.5%, Colombians - 41%). Attention is drawn to the fact that the higher the G allele frequency, the higher the percentage of indigenous people in the population. For example, in the Puerto Rican population, where the indigenous Indian population is 11%, the frequency of the G allele is 31.7%. In Europeans, the frequency of the G allele is on average 22.6%. Among Asians, the Japanese have a high frequency of the G allele (42.3%). The owners of the lowest frequency of the G allele are Africans (8.6%-17.2%)

Analysis of the distribution of the frequency of genotypes of the *PNPLA3* rs738409 SNP, according to the open database of the 1000 Genomes project,⁽¹⁰⁾ showed that the G allele is found with a high frequency in all the world populations. For example, the average frequency of the G allele in African populations is 89.6%, in the populations of Central and South America - 78.8%, in the populations of East Asians (Chinese, Japanese, and Vietnamese) - 81.8%, among Europeans - 62.9%, and in South Asians (Indians and Pakistanis) - 77.2%.

According to the *PNPLA3* rs2294918 SNP, which suppresses the negative effect of rs738409, the protective the *PNPLA3* (rs2294918) A allele was found only in 10.7% of study subjects. According to the 1000 Genomes project,⁽¹⁰⁾ the *PNPLA3* (rs2294918) A allele is more common in Europeans than in other populations (32.3%).

Table 2.

Frequency of variants and distribution of the missense mutations in the *PNPLA3* gene in the Yakut population and in the populations from the "1000 Genomes" project

Population	SNP	NC	MAF	Protein Change	Risk Allele	ObsH	PredH	HWE <i>P</i> -value
YKT (n=150)	rs738409	G:C	0.720 (G)	I148M	G (148M)	0.293	0.403	0.0019
	rs2294918	G:A	0.107 (A)	E434K	G (434E)	0.173	0.191	0.4386
PEL (n=85)	rs738409	G:C	0.718 (G)	I148M	G (148M)	0.376	0.405	0.6473
	rs2294918	G:A	0.100 (A)	E434K	G (434E)	0.153	0.180	0.3655
MXL (n=64)	rs738409	G:C	0.555 (G)	I148M	G (148M)	0.422	0.494	0.3260
	rs2294918	G:A	0.172 (A)	E434K	G (434E)	0.312	0.285	0.8301
JPT (n=104)	rs738409	G:C	0.418 (G)	I148M	G (148M)	0.394	0.487	0.0736
	rs2294918	G:A	0.087 (A)	E434K	G (434E)	0.135	0.158	0.3183
CEU (n=99)	rs738409	G:C	0.217 (G)	I148M	G (148M)	0.354	0.340	0.9869
	rs2294918	G:A	0.323 (A)	E434K	G (434E)	0.438	0.438	0.8946
YRI (n=108)	rs738409	G:C	0.116 (G)	I148M	G (148M)	0.213	0.205	1.0
	rs2294918	G:A	0.097 (A)	E434K	G (434E)	0.157	0.176	0.5046

Abbreviations: NC, nucleotide substitution; MAF, minor allele frequency; ObsH, observed heterozygosity; PredH, predicted heterozygosity; YKT, Yakuts from Yakutia, Russia; PEL, Peruvians from Lima, Peru; MXL, Mexican Ancestry from Los Angeles, USA; JPT, Japanese in Tokyo, Japan; CEU, Utah Residents (CEPH) with Northern and Western European Ancestry; YRI, Yoruba in Ibadan, Nigeria.

A weak LD was observed between the two SNPs ($D' = 0.096$, $r^2 = 0.003$ in Yakuts; $D' = 0.48$, $r^2 = 0.003$ in Yoruba) (Fig. 1). In other samples, strong linkage was observed ($D' = 1$, $r^2 = 0.282$ in Peruvians; $D' = 1$, $r^2 = 0.068$ in Japanese; $D' = 1$, $r^2 = 0.132$ in Europeans; $D' = 0.884$, $r^2 = 0.202$ in Mexicans).

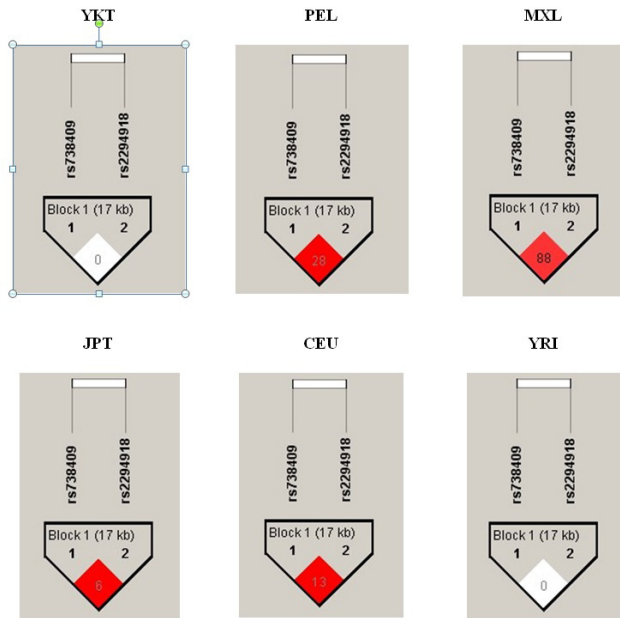


Fig. 1. LD between the *PNPLA3* SNPs

The color of the cell indicates the adhesion strength between SNPs: red - strong link ($D' = 1$, $LOD > 2$), white - weak link ($D' < 1$, $LOD < 2$).

Abbreviations: YKT, Yakuts from Yakutia, Russia; PEL, Peruvians from Lima, Peru; MXL, Mexican Ancestry from Los Angeles, USA; JPT, Japanese in Tokyo, Japan; CEU, Utah Residents (CEPH) with Northern and Western European Ancestry; YRI, Yoruba in Ibadan, Nigeria

Analysis of the distribution of the frequency of genotypes in the studied sample of Yakuts showed the predominance of the carriage of the *PNPLA3* rs738409 GG genotype (57.3%). The *PNPLA3* (rs2294918) AA and AG genotypes carrying the protective A allele are more common in European populations (11.1% and 42.4%, respectively) (Table 3).

Table 5.

Distribution of diplotypes of two SNP markers of the *PNPLA3* gene in the Yakut population and in the populations from the «1000 Genomes» project

№	Genotype / SNP		Diplotype	Diplotype frequency					
	rs738409	rs2294918		YKT (n=150)	PEL (n=85)	MXL (n=64)	JPT (n=104)	CEU (n=99)	YRI (n=108)
1	GG	GG	[GG][GG]	0.453	0.529	0.328	0.221	0.040	0.009
2	CG	GG	[CG][GG]	0.280	0.247	0.234	0.327	0.182	0.176
3	GG	AG	[GG][AG]	0.107	0	0.016	0	0	0
4	CG	AG	[CG][AG]	0.013	0.129	0.188	0.067	0.172	0.037
5	GG	AA	[GG][AA]	0.013	0	0	0	0	0
6	CC	AA	[CC][AA]	0.007	0.024	0.016	0.019	0.111	0.019
7	CC	AG	[CC][AG]	0.053	0.024	0.109	0.067	0.253	0.120
8	CC	GG	[CC][GG]	0.073	0.047	0.109	0.298	0.242	0.639
9	CG	AA	[CG][AA]	0	0	0	0	0	0

Abbreviations: see Fig.1

Table 3.

Distribution of the *PNPLA3* rs2294918 and rs738409 genotypes in Yakuts and in the populations from the «1000 Genomes» project

SNP	Genotype	Genotype frequency					
		YKT n=150	PEL n=85	MXL n=64	JPT n=104	CEU n=99	YRI n=108
rs738409	GG	0.573	0.529	0.344	0.231	0.040	0.009
	GC	0.293	0.376	0.422	0.385	0.354	0.213
	CC	0.133	0.094	0.234	0.385	0.606	0.778
rs2294918	GG	0.807	0.824	0.672	0.846	0.465	0.824
	GA	0.173	0.153	0.313	0.135	0.424	0.157
	AA	0.020	0.024	0.016	0.019	0.111	0.019

Abbreviations: see Fig.1

The frequency distribution of the *PNPLA3* gene haplotypes for two SNPs (rs738409, rs2294918) based on all detected variants is presented in Table 4. We identified two main haplotypes, the frequency of which was > 0.1 . One of the most common haplotypes carries variant G (148M), the other carries variant C (148I), and both carry the same variant G (434E). In other words, the more common two haplotypes carry the G (434E) allele, while the A (434K) protective allele does not occur in those found in the two main haplotypes. Protective allele A (434K) is shared by both rare haplotypes. Haplotype G-A (148M-434K) was not found in samples of Peruvians, Japanese, Europeans, and the Yoruba tribe. This haplotype (G-A) was found in Yakuts and Mexicans (6.9% and 1.1%, respectively).

Table 4.

Frequency of I148M-E434K haplotypes in the Yakut population and in the populations from the «1000 Genomes» project

№	Haplotype rs738409-rs2294918	Protein	Haplotype frequency					
			YKT n=150	PEL n=85	MXL n=64	JPT n=104	CEU n=99	YRI n=108
1	G-G	148M-434E	0.651	0.718	0.544	0.495	0.217	0.110
2	C-G	148I-434E	0.243	0.182	0.285	0.495	0.460	0.793
3	G-A	148M-434K	0.069	0	0.011	0	0	0
4	C-A	148I-434K	0.037	0.100	0.161	0.087	0.323	0.091

Abbreviations: see Fig.1

The distribution of diplotype frequencies for two SNPs (rs738409 and rs2294918) of the *PNPLA3* gene based on all detected variants is presented in Table 5. Eight diplotypes out of nine possible variants were found. The Yakuts often have two diplotypes [GG]-[GG] and [CG]-[GG]. Both diplotypes carry the *PNPLA3* rs738409 G allele (45.3% and 28%) and do not carry the *PNPLA3* rs2294918 A allele. The same distribution of diplotype frequencies is observed among Peruvians (52.9% and 24.7%), Mexicans (32.8% and 23.4%) and Japanese (22.1% and 32.7%). Diplotypes carrying the A protective allele are found at a low frequency (Table 4). The [GG]-[AA] and [CG]-[AA] diplotypes are absent in the population samples, with the exception of the [GG]-[AA] diplotype found in the Yakuts (1.3%). Among the seven discovered diplotypes among the Yoruba tribe, [CC]-[GG] is most common (63.9%). This diplotype does not carry the pathological allele G (rs738409) and does not carry the protective A allele (rs2294918).

Thus, the high frequency of the [GG]-[GG] and [CG]-[GG] diplotypes in Yakuts (45.3% and 25%, respectively), carrying mutant alleles G (rs738409) and not carrying the A allele (rs2294918), indicates that these diplotypes were probably adaptively favorable to the Yakuts. The high frequency of diplotypes [GG]-[GG] and [CG]-[GG] among the indigenous population of America may be due to the settling of the American continent through the Bering Strait by the ancestors of the American Indians.

In conclusion, the analysis of haplotypes and diplotypes based on the markers rs738409 and rs2294918 of the *PNPLA3* gene may contribute to future new biomarkers for the diagnosis of NAFLD and NASH, as well as provide fundamental knowledge on human adaptation to cold.

Competing Interests

The authors declare that they have no competing interests.

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).

References

1. Simcox J, Geoghegan G, Maschek JA, Bensard CL, Pasquali M, et al. Global Analysis of Plasma Lipids Identifies Liver-Derived Acylcarnitines as a Fuel Source for Brown Fat Thermogenesis. *Cell Metab.* 2017 Sep 5;26(3):509-522.e6. doi: 10.1016/j.cmet.2017.08.006.
2. Kumari M, Schoiswohl G, Chitruja C, Paar M, Cornaciu I, Rangrez AY, et al. Adiponutrin functions as a nutritionally regulated lysophosphatidic acid acyltransferase. *Cell Metab.* 2012 May 2;15(5):691-702. doi: 10.1016/j.cmet.2012.04.008.
3. Pingitore P, Pirazzi C, Mancina RM, Motta BM, Indiveri C, Pujia A, Montalcini T, et al. Recombinant PNPLA3 protein shows triglyceride hydrolase activity and its I148M mutation results in loss of function. *Biochim Biophys Acta.* 2014 Apr 4;1841(4):574-80. doi: 10.1016/j.bbali.2013.12.006.
4. Basantani MK, Sitnick MT, Cai L, Brenner DS, Gardner NP, Li JZ, et al. Pnpla3/Adiponutrin deficiency in mice does not contribute to fatty liver disease or metabolic syndrome. *J Lipid Res.* 2011 Feb;52(2):318-29. doi: 10.1194/jlr.M011205.
5. Chen W, Chang B, Li L, Chan L. Patatin-like phospholipase domain-containing 3/adiponutrin deficiency in mice is not associated with fatty liver disease. *Hepatology.* 2010 Sep;52(3):1134-42. doi: 10.1002/hep.23812.
6. Romeo S, Kozlitina J, Xing C, Pertsemlidis A, Cox D, Pennacchio LA, et al. Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease. *Nat Genet.* 2008 Dec;40(12):1461-5. doi: 10.1038/ng.257.
7. Xu R, Tao A, Zhang S, Deng Y, Chen G. Association between patatin-like phospholipase domain containing 3 gene (PNPLA3) polymorphisms and nonalcoholic fatty liver disease: a HuGE review and meta-analysis. *Sci Rep.* 2015 Mar 20;5:9284. doi: 10.1038/srep09284.
8. He S, McPhaul C, Li JZ, Garuti R, Kinch L, Grishin NV, et al. A sequence variation (I148M) in PNPLA3 associated with nonalcoholic fatty liver disease disrupts triglyceride hydrolysis. *J Biol Chem.* 2010 Feb 26;285(9):6706-15. doi: 10.1074/jbc.M109.064501.
9. Smagris E, BasuRay S, Li J, Huang Y, Lai KM, Gromada J, et al. Pnpla3I148M knockin mice accumulate PNPLA3 on lipid droplets and develop hepatic steatosis. *Hepatology.* 2015 Jan;61(1):108-18. doi: 10.1002/hep.27242.
10. Donati B, Motta BM, Pingitore P, Meroni M, Pietrelli A, Alisi A, et al. The rs2294918 E434K variant modulates patatin-like phospholipase domain-containing 3 expression and liver damage. *Hepatology.* 2016 Mar;63(3):787-98. doi: 10.1002/hep.28370.
11. Bruschi FV, Claudel T, Tardelli M, Caligiuri A, Stulnig TM, Marra F, et al. The PNPLA3 I148M variant modulates the fibrogenic phenotype of human hepatic stellate cells. *Hepatology.* 2017 Jun;65(6):1875-1890. doi: 10.1002/hep.29041.
12. Trépo E, Nahon P, Bontempi G, Valenti L, Falletti E, Nischalke HD, et al. Association between the PNPLA3 (rs738409 C>G) variant and hepatocellular carcinoma: Evidence from a meta-analysis of individual participant data. *Hepatology.* 2014 Jun;59(6):2170-7. doi: 10.1002/hep.26767.
13. Hyysalo J, Männistö VT, Zhou Y, Arola J, Kärjä V, Leivonen M, et al. A population-based study on the prevalence of NASH using scores validated against liver histology. *J Hepatol.* 2014 Apr;60(4):839-46. doi: 10.1016/j.jhep.2013.12.009.
14. Valenti L, Rumi M, Galmozzi E, Aghemo A, Del Menico B, De Nicola S, et al. Patatin-like phospholipase domain-containing 3 I148M polymorphism, steatosis, and liver damage in chronic hepatitis C. *Hepatology.* 2011 Mar;53(3):791-9. doi: 10.1002/hep.24123.
15. Kurtanov KA, Sydykova LA, Pavlova NI, Filippova NP, Dodokhov VV, Apsolikhova GA, et al. [Polymorphism of the adiponutrin gene (PNPLA3) in the indigenous inhabitants of the Republic of Sakha (Yakutia) with type 2 diabetes mellitus]. *Almanac of Clinical Medicine.* 2018;46 (3):258-263. doi: 10.18786/2072-0505-2018-46-3-258-263. [Article in Russian].
16. GSR and the 1000 Genomes Project [Internet]. Available from: <https://www.internationalgenome.org/>
17. HAPLOVIEW v.4.2 [Internet]. Available from: <https://www.broadinstitute.org/haploview/version-42-15-september-2009>

*Corresponding author: Khariton A. Kurtanov, PhD. Yakut Science Center of Complex Medical Problems, Yakutsk, the Republic of Sakha (Yakutia), Russia. E-mail: hariton_kurtanov@mail.ru