Lactase Deficiency in Russia: Multiethnic Genetic Study

Elena Kovalenko
Genotek Ltd.  https://orcid.org/0000-0001-5678-6557
Ekaterina Vergasova
Genotek Ltd.
Olesya Shoshina
Danone
Iaroslav Popov
Genotek Ltd.
Anna Ilinskaya
Genotek Ltd.
Anna Kim
Genotek Ltd.
Nikolay Plotnikov
Genotek Ltd.
Isroel Barenbaum
Genotek Ltd.
Artem Elmuratov
Genotek Ltd.
Valery Ilinsky
Genotek Ltd.
Olesya Volokh
Danone
Alexander Rakitko (✉ rakitko@genotek.ru)
Genotek Ltd.

Article

Keywords:

Posted Date: September 30th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1859111/v1
License: © This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Version of Record: A version of this preprint was published at European Journal of Clinical Nutrition on June 13th, 2023. See the published version at https://doi.org/10.1038/s41430-023-01294-8.
Abstract

Background

Lactase persistence — the ability to digest lactose through adulthood — is closely related to evolutionary adaptations and has affected many populations since the beginning of cattle breeding. Nevertheless, the contrast initial phenotype, lactase non-persistence or adult lactase deficiency, is still affecting large numbers of people worldwide.

Methods

We performed the largest multiethnic genetic study of lactase deficiency on 24439 people in Russia to date. The percent of each population group was estimated according to the local ancestry inference results. Additionally, we calculated frequencies of rs4988235 GG genotype in Russian regions using the information of current location and birthplace data in client's questionnaire.

Results

It turned out that among all studied population groups the frequency of GG genotype in rs4988235 was higher than in average in the European populations. In particular, the prevalence of lactase deficiency genotype in the East Slavs group was 42.8% (95% CI: 42.1–43.4%). We also investigated the regional prevalence of lactase deficiency by current place of residence.

Conclusions

Our study emphasizes the diagnostic significance of genetic testing, i.e. specifically for lactose intolerance parameter, as well as the scale of the problem of lactase deficiency in Russia which needs to be addressed by healthcare and food industry.

Introduction

Lactose is one of the most common disaccharides consumed by humans in milk and dairy products. The lactose content ranges from 0.1 to more than 7 g per 100 mL, depending on the source of milk, category of dairy products, fat content in the product and the type of starter used at the fermentation stage, due to which the final sugar concentration is significantly reduced [1–3]. The digestion of lactose occurs in the small intestine in the presence of β-galactosidase enzyme called lactase-phlorizin hydrolase (LPH) or lactase in short, the activity of which determines the ability of the body to successfully break down milk sugar into glucose and galactose monomers [4].
The lactase enzyme is encoded by the LCT gene located on chromosome 2. Its upstream regulatory region MCM6 is responsible for modulation LCT expression through epigenetic modifications. Polymorphisms in the MCM6, particularly rs182549 (−22018G/A) and rs4988235 (13910C/T), are associated with high age-related methylation level in both MCM6 and LCT regions which lead to lactase non-persistence by silencing LCT [5–6]. Such decreased lactase activity generally occurs after the weaning phase and is considered to be normal "wild type" and is widespread around two-thirds of the world's population with wide diversity in different countries [7–10]. The ability to digest lactose in adulthood is provided by maintenance of high level of lactase activity due to continued post weaning expression of the lactase gene, and results from at least 5 different mutations in a lactase regulatory region within an intron of MCM6 [12]. It is assumed that lactase persistence (LP) alleles have undergone a positive selection simultaneously with the spread of agriculture [11–13]. The average frequency of lactase persistence is higher in northern Europe, moderate in southern Europe, the Middle East and low in some Asian and African communities [11, 14–15]. Their evolutionary advantage provides access to nutrient-dense dairy products and liquids sources, which is critical for the populations living in hot, arid environments [14].

Decreased lactase activity in adulthood virtually leads to lactose malabsorption, which is related to the amount of mis-absorbed lactose. Taking a limited amount of lactose does not cause recognizable symptoms in people with lactose maldigestion: the majority of subjects with lactose maldigestion will tolerate up to 12 g of lactose as a single dose with no or minor symptoms or more in case of distribution of lactose-contained products to multiple doses throughout a day [16, 17]. For this very reason, worldwide rates of lactase deficiency and lactose malabsorption may be much higher than those of lactose intolerance. Symptoms of intolerance are associated with the amount of lactose consumed at a time and may not occur in people with lactase deficiency when consuming small doses of lactose. Lactose intolerance depends not only on the lactase genotype but also on the lactose dose, age, microbiota, specifics and sensitivity of the GI tract and ability to lactose digestion [18, 19]. In large intestine non-cleaved lactose disaccharides lead to a violation of water absorption and are fermented by the intestinal microbiota with the release of gases (H₂, CO₂, CH₄) and other metabolites. These factors negatively affect the function of the gastrointestinal (GI) tract [20, 21] as well as person wellbeing. Thus, a person may face digestive problems (diarrhea, abdominal pain, bloating, nausea) after taking lactose-containing food [21]. To avoid the digestive symptoms but keep the dairy in diet the solution can be the lactose-free and low lactose dairy [22]. This leads to a decrease of dairy products consumption, meanwhile they are a key source of calcium in a balanced diet, as well as an important source of protein and set of micronutrients [23, 24]. Most guidelines and public health organizations suggest consuming dairy products up to three servings per day to ensure adequate nutrient intakes and optimal bone health [25–30].

During the last decades the set of works has been dedicated to analyzing the various national management strategies [17], symptoms evaluations analysis [31] and the global and region-focused situation [7, 8]. Specifically, for Russia it was concluded that figures may overestimate the prevalence of
deficiency due to the unrepresentative samples in relation to the total population in the country [7]. Indeed, the Russian Federation is a country of more than 146 million people and at least 190 ethnic groups, which tend to be genetically diverse and may be more similar to people of European descent or almost indistinguishable from Asians. The population is unevenly distributed throughout the country, which is divided into 85 federal subjects or regions. Evaluation of accurate regional patterns of lactase deficiency is important to guide management of GI symptoms as well as a crucial driver for lactose-free and low-lactose products dairy market growth.

To evaluate the actual level of lactase deficiency in Russia here we focus on genotype-related hypolactasia. We aimed to compare frequency of GG genotype in rs4988235 (13910C/T) in regulatory region (MCM6) of the LCT lactase enzyme gene (in populations living in Russia and differences between Russian regions) which is in strong linkage disequilibrium (D’=0.9985, R² = 0.9825) with another lactase non-persistence CC genotype in rs182549 (−22018G/A). We also analyzed the prevalence of lactose intolerance phenotypic manifestations based on previously published data with the frequency of the presence of lactase persistence genotype which determines the genetic risk of lactase deficiency manifestation for the Russian population.

**Materials & Methods**

**Study populations**

In our study we analyzed genetic data of 40164 individuals from the database of Genotek, the Russian consumer genetics and research company ([www.genotek.ru](http://www.genotek.ru)). The Genotek clients included in our analyses provided informed consent for their data to be used for research purposes and responded to online questionnaire. The current research was approved by the Genotek Ethics Committee (protocol №8 "Lactose intolerance in the Russian population") and performed in accordance with the Declaration of Helsinki. DNA extraction and genotyping were performed on saliva samples that were genotyped on Illumina Infinium Global Screening Array v.1-v.3 microarrays (~ 650 000 SNPs). We also included uploaded to Genotek website samples from 23andMe genetic company ([www.23andme.com](http://www.23andme.com)) which were performed on 23andMe v5 custom microarray chip. All 23andMe users, who shared their data with us, signed informed consent for their data to be used for research purposes and responded to the online questionnaire. All samples in Genotek cohort were processed in batches (192–576 samples per batch). The GenomeStudio software (Illumina, San Diego, CA) and manually created cluster files were used to cluster the raw signals and call the genotypes. SNPs with call rate < 0.9 within the batch were removed. Then we removed the calls on the Y chromosome for women and the heterozygous calls on the X chromosome for men. We calculated the MCD (maximum cluster distance) metric for each genotype as the maximum (over all probes for SNP) of the normalized distances between individual signal and the corresponding cluster center in the plane with coordinates Norm Theta and Norm R. We considered all heterogeneous genotypes with MCD > 4 as missing calls. Finally, we removed samples with the call rate < 0.9 after all previous steps.
Statistical analysis of genotypes was performed using R scripts. Confidence intervals were constructed using the Hmisc package and binconf (alpha = 0.05).

**Ancestry estimation**

An in-house algorithm for local ancestry inference was used to assign the ancestry label to each individual. Our algorithm firstly used Positional Burrows-Wheeler Transform (PBWT) [32] for initial population prediction and then a custom Hidden Markov Model for phasing error correction and smoothing. We used the reference panel with 17559 genomes with known ancestry. Each individual from the reference panel corresponds to one of the 101 populations. All populations were clustered into 23 population groups based on their genetic similarity.

For each individual in the current study, we estimated the percent of each population group according to the local ancestry inference results. We assigned the population group label to people whose maximum percentage of that population group exceeds 70% and included for further analysis. For assuring enough samples for frequency estimation, we selected only those population groups in which there were 100 or more people. After such filtering seven population groups remained: Central Asia, East Slavs, Ashkenazi Jews, North Caucasus, Siberia, Volga-Ural region, Western Asia. The total number of populations in all groups was 56. The full list of all regions and populations used in current research is presented in Table S2. In order to show the adequacy of our approach and to visually validate the selected clusters, we constructed a PCA plot (see below).

Only representatives of selected population groups were considered for further analysis. Among them 29559 individuals were genotyped on Illumina GSA microarray and 835 — on 23andMe v5. All subsequent data preprocessing was performed using BCFtools (version bcftools 1.12-75-g5329f29) and PLINK (version v1.90b6.10). Sex chromosomes have been removed. Then all SNPs with missing calls in more than 5% of samples (plink --geno 0.05) and all samples with missing calls in more than 3% of SNPs (plink --mind 0.03) were filtered out. After this filtration 29104 samples remained.

All missing cases were simply replaced with homozygous major allele (plink–fill-missing-a2). To identify biases caused by distinct platforms, a basic $\chi^2$ test of associations between SNPs and chip type was performed. We identified 11023 SNPs passing the p-value threshold 0.00001 responsible for the major difference (plink --assoc–pfilter 0.00001). All found SNPs were removed from consideration.

For the proper analysis, we needed to filter out all people with at least one close relative (1st or 2nd degree) in the cohort. For this stage, we excluded all variants with minor allele frequency below 0.01 (plink –maf 0.01). Then the pairwise comparison of samples and PI-HAT calculation was made with plink –genome. Thus, 4635 samples with at least one close relative (PI-HAT > 0.125) were removed from the data.

Eventually, we obtained data from 24469 samples and pruned it (plink–indep 50 5 2). To the pruned data, we applied the principal component analysis (PCA) in PLINK. A PCA-plot was constructed from the
principal components and the HDBSCAN clustering algorithm was run.

After filtration by the presence of two SNPs (rs4988235 and rs182549) in the MCM6 gene 30 individuals were excluded, thus 24439 people remained, of which 509 individuals were genotyped on 23andMe v5 microarray type, 1036 — on Illumina GSA v.1, 4788 — on Illumina GSA v.2.0 and 18106 — on Illumina GSA v.3.0.

**Evaluation of frequencies in different Russian regions**

To calculate frequencies of rs4988235 GG genotype (which is in strong linkage disequilibrium with rs182549 CC genotype) in Russian regions we used the information of current location and birthplace data in client’s questionnaire, as well as information about the birthplace of close relatives (parents and grandparents). Since we aimed to focus primarily on Russian residents, we filtered out not only people born abroad, but also all people whose ancestors were also born abroad. Thus, the data for the frequency estimation included 11325 individuals. Finally, 10622 East Slavs clients were selected from this subcohort.

In the last decades the intensity of migration has increased significantly [33]. Since we aimed to study the prevalence of genetic variants of lactase deficiency in the context of population differences, we focus on information about the ancestral birthplaces of an individual rather than on the current location. Thus, if a grandparent was born in a region, such observation was given a weight of 1; if a mother or father was born in a region, the observation was given a weight of 0.5; for the individual himself, the observation was given a weight of 0.25. Hence for each region we calculated the total weight by summarizing all weights. If the total weight was less than 13 (e.g., 52 individuals, each with a minimum weight of 0.25) and there were less than 20 individuals in the region, then such a region was not included in further analysis. Thus, 69 Russian regions were included in research for estimating the rs4988235 GG genotype among all samples and 66 for only East Slavs. The list of Russian regions used in current research can be found in Tables S4-S5. Data for the map plotting were downloaded from gadm.org (version 3.6) and then were visualized in R using package ggplot2 version 3.3.5.

**Results**

To study the prevalence of lactase deficiency in the Russian population, we evaluated the frequencies of rs4988235 and rs182549 genotypes among 40164 individuals. Taking into account that the GG genotype in rs4988235 is in strong linkage disequilibrium with CC genotype in rs182549 (Tables S4-S5) we concentrated on the first one as it is more likely to be used in lactose intolerance research. Since the Russian population is multi-ethnic, we aimed to estimate the frequencies in groups of populations genetically distinct from each other. For this purpose, the proportion of each population group in the human genome was estimated using the LAI algorithm. People with evidence of recent admixture events were excluded from consideration. After QC filtering and exclusion of admixed individuals (see Methods), the final sample consisted of 29559 people, each assigned to one of the 7 population groups: Central Asia, East Slavs, Ashkenazi Jews, North Caucasus, Siberia, Volga-Ural region and Western Asia.
We constructed a PCA plot, clustering people in the filtered 56 populations according to their ethnicity (Fig. 1).

We obtained a single map of 7 population groups using genetic distance among population pairs. The distribution of populations into clusters was as expected. A total of 24469 people were included in the following research after the PCA plot was built (12410 female, 12059 male). The final cohort remained 24439 individuals with the presence of two lactase-related SNPs (rs4988235 and rs182549). The characteristics of the final cohort are summarized in Table 1.

Among the population groups, the highest prevalence of GG genotypes was observed in Western Asia and Siberia, as well as among Central Asia (96.3%, 88.7%, and 83.8%, respectively) (Fig. 2, Table S3).

We observed the lowest percentage of GG genotype (42.8%; 95% CI: 42.1–43.4%) among population groups in the East Slavs group, which consists of Russians, Belarusians and Ukrainians.

Among Russian regions, the highest percentage of people with lactase deficiency was found in North Ossetia (83.2%), slightly lower in the Chechen Republic (77.8%) and in Dagestan (74.3%) (Fig. 2b, S4). High prevalence of GG genotype was also detected in Buryatia (65%), Sakha Republic (63.4%) and Tomsk Oblast (60.5%).

The lowest percentage of people with the genotype associated with lactase deficiency was observed in Zabaykalsky Krai (22.8%). Among the top million cities and regions, the percentage of the GG genotype was as follows in Table 2:

It should be noted that these values were obtained not by ethnicity or ancestry estimation but by current place of residence, hence may differ from the values inherent in the natives of the region.

Among East Slavs living in different Russian regions, the lowest lactase deficiency prevalence was in Udmurtia (20.4%), Zabaykalsky Krai (23.0%) and Ivanovo Oblast (26.2%). The highest prevalence was detected in Buryatia (61.5%), Sakha (Yakutia) Republic and Tomsk Oblast (60.0% both) — Fig. S1, Table S5.

**Discussion**

The prevalence of LP genotype in Europe remained low until the Bronze Age [34, 35]. There are at least 2 hypotheses of the origin of lactase persistence. The first one is explained by the appearance of LP in Europeans after mass migration of the eastern steppe populations associated with the Yamnaya culture; the second one is explained by the appearance of LP initially in the pastoral steppe population when farming replaced hunting and gathering in the Middle East and then brought to Western Europe at the beginning of Corded Ware culture [36]. Each of these hypotheses has strong arguments for their existence; nevertheless, one of the recent study on the evolutionary origin of the LP genotype concludes that the increase in the percentage of LP in the European population was unlikely related to steppe movements and their genotypes, and began no earlier than after 3000 BP [15]. Other studies showed that
the large-scale migrations of steppe nomads were connected precisely with the beginning of milk consumption and domestication of horses: the advantage of a permanent source of protein combined with the potential epicenter of horse domestication allowed steppe people to begin large-scale migrations of their Pontic-Caspian steppe across Eurasia [37]. In addition to genetic variation in humans, high genetic diversity has been observed in cows from such regions where dairy farming is well practiced and humans are lactose persistent [38]. Recently, it has been suggested that the ability to digest milk is better correlated with the prevalence of fermentation rather than with the level of cattle production [34]. This prompts the puzzling question of the need for positive genetic selection of lactase persistence considering that its prevalence is lower than non-persistence.

The geographic specificity of the Caucasus primarily affected the genetic gap between the East European Plain and the Caucasus. There is a number of genetic and linguistic evidence indicating predominantly a Southern, Middle Eastern origin of the Caucasus populations in the process of Neolithic migrations from the steppe and subsequent genetic drift in a geographically restricted landscape [39, 40]. It is possible to speculate that the spread of lactase persistence from Eastern Europe could not pass the geographical barrier but remained as the main routes of ancient gene flow and mixing from the Anatolian side, where cattle breeding originated. It is worth to mention the cultural peculiarities of this region, which may have been both the cause and the consequence of the high prevalence of lactase deficiency.

The Siberian cluster of populations deserves special attention. The vast majority of Siberian populations, even from the far north, are genetically of southern origin and multilayered complex admixture, as emphasized by several studies [41–44]. In particular, it is assumed that the Sakha (Yakut) people are descended from a common ancestral population from the Lake Baikal region, genetically and culturally close to Mongols and Buryats until they were displaced by the latter [43]. The northward migration can be explained by the arrival of horse-breeding peoples from the Asian steppes at the beginning of the first millennium BC, at a time when reindeer breeding flourished in the Eastern Sayan region [42, 45]. Thus, the high percentage of lactase deficiency in Siberian populations, in addition to the genetic component, can be explained by the following factors: big practice in milk fermentation practices (the lactose content of mare's milk is 1.5 times higher than of cow's one [46]); later domestication and the presence of admixture with steppe herders have been suggested as possible explanations. Another explanation could be the fact that reindeer milk contains the lowest concentration of lactose among other herd animals (less than 3 g) and there was no need to digest large amounts of lactose [24].

Our results indicate that the prevalence of lactase deficiency in East Slavs populations (42.8%; 95% CI: 42.1–43.4%) is higher than that for European populations on average (28%; 95% CI: 19–37%) [7]. Moreover, the individuals from the North Caucasus are more susceptible to lactase deficiency than in Central Russia, which generally coincides with the frequency of GG genotype from north to south, as well as the history of cattle breeding. The moderate percentage of lactase deficiency across the territory of the Russian Federation persisted even after we analyzed its prevalence only among the East Slavs.
We obtained unique data of lactase deficiency genotypes among Ashkenazi Jews, since the previous statistics for this population were limited, outdated, and did not use genetic affiliation with Ashkenazi Jews. Our statistics turned out to be similar to previously published data (60–80%) [19, 47–48]. The prevalence of the lactase deficiency genotype is comparable with that of the North Caucasus and Central Asia. However, according to mtDNA data [49–50], the ancestors of most recent Ashkenazi Jews lived in Southern Europe and not in the Middle East or the Caucasus. They likely originated as a result of the migration of Mediterranean Jews northward to Germany and later underwent expansion eastward.

According to the results of the earlier published meta-analysis the prevalence of lactase deficiency among the Russian population was 61% (95% CI from 59 to 64%) [7]. Indeed, most of the studies conducted on the territory of Russia are either limited to a narrow cohort of people of a certain age (for example, students from 17 to 26 years old) [8] or a specific region of the analysis [51–52], or the assessment was carried out in small ethnic groups of the Russian population [53–55]. Thus, according to the data of the rs4988235 prevalence [8, 53–55, 58] among Russians, the frequency of lactase deficiency occurs in 35–54% of cases (Table S1). The prevalence of lactase deficiency based on analysis of rs4988235 is 42.8% among East Slavs populations which lies in the range revealed previously. Depending on the ethnic groups, the frequency of rs4988235 GG genotype varied from 28 to 94% of cases (see Table S1). As it mentioned before, generally investigations were carried out within particular ethnic group. For some populations values of lactase deficiency correlates with our estimations. For instance, values are different slightly within 5% in the case of Bashkortostan, Buryatia, Perm oblast, Kostroma oblast, Dagestan, Mordovia, Arkhangelsk oblast. It possibly could be explained by the fact that ethnicity is predominant for that region (as Buryats in Buryatia and Mokshas in Mordovia) or according to the chosen approach based on the assumption of ancestors it is a high probability of the interethnic marriages. In the vast majority of cases ethnic estimation is much higher than revealed value of lactase deficiency for the region where this ethnic group lives. This tendency is relevant for Sami (the Kola Peninsula) (48% vs 33.8% in Murmansk), Mari people (74% vs 56.6% in Mari El), Udmurts (55 or 40% vs 30.9%), Russians in Kursk (54% vs 37.9%), Russians in Rostov (54% vs 43.85%), Komi (70% vs 30.9% in Komi republic), Erzas (57% vs 40.7% in Mordovia). As it mentioned before, generally investigations were carried out within particular ethnic group. However, the aim of the current study was estimating the average prevalence of lactase deficiency in the regions based on the historical aspects of person's ancestry considering the fact of multiculturalism in Russia. It is known that the frequency of the genotype associated with lactase deficiency is affected by migration and interethnic marriages. For example, the prevalence of adult-type hypolactasia (rs4988235) among Nenets who had four Nenets grandparents was 90%. In cases of three, two, and one grandfather or grandmother of Nenets origin, the frequency of the genotype decreases sharply: 72%, 60%, and 28%, respectively [54]. Thus, accurate national estimates depend to a large extent on the proper representation of all population groups in a country.

Despite the high prevalence of lactase deficiency in Russia, there is a trend of consumption of dairy products. In 2020 dairy consumption reached 272 kg per capita [56]. Nevertheless, the annual amount of dairy products in Russia is still insufficient according to the Ministry of Health of the Russian Federation (340 kg per capita) [57]. This could potentially be due to the fact that people have gastrointestinal
symptoms after lactose products intake, thus they more frequently exclude dairy products from the diet and replace them with lactose-free alternatives [8, 58].

Genotyping for lactase deficiency has shown excellent correlation with the lactose hydrogen breath test (H\textsubscript{2} test), measurement of glucose and urinary galactose/creatinine levels, concluding that the analysis of the lactase non-persistence variants can be considered as accurate and convenient test for predicting the presence of this disease in a patient with suspected lactose malabsorption [22, 59–63].

Our cohort study genetic result highlights sizeable lactase deficiency across Russian regions and can testify in favour of the healthcare system to pay more attention to this problem as well as for the food system to rebalance the dairy market towards low lactose and lactose-free products. The analysis of the lactase non-persistence variants can be considered as an accurate and convenient test for predicting the presence of this disease in a patient with suspected lactose malabsorption therefore preventing nutritional imbalance.

**Conclusion**

We conducted the most extensive study of lactase deficiency genotype prevalence in the Russian population. We compared frequencies in different populations and in different regions thus obtaining almost complete regional coverage. Our results are in accord with previous observations from the past decades and can be explained by historical ethnic migration and the cultural characteristics of the Russian regions. The geographical features of residence and climate, animal species diversity and ancestral migration played a major role in the use of lactose-contained products in these regions.

From the practical standpoint, public awareness of the actual lactase deficiency conditions may positively impact on nutritional behaviour and status of the population, as well as supporting the growth of the market of lactose-free and low-lactose products. These features highlight the importance of such studies from scientific, epidemiological and economic standpoint.

**Declarations**

**Competing Interests statement**

The authors do not have any competing interests.

**Data Availability Statement**

For Genotek dataset, the user agreement (available at https://www.genotek.ru) states that disclosure of individual-level genetic information and/or self-reported Information to third parties for research purposes will not occur without explicit consent, and the consent was not obtained from the individuals. Due to the user agreement the individual level cannot be made directly available, and the dataset could pose a threat
to confidentiality. Data have to be accessed indirectly via Genotek Ltd, https://www.genotek.ru/. Data requests should be sent to the Genotek Ltd at info@genotek.ru.

References


34. Segurel, L. *et al.* Why and when was lactase persistence selected for? Insights from Central Asian herders and ancient DNA. *PLOS Biology* **18**, e3000742 (2020).


### Tables

Tables 1 and 2 are available in the Supplementary Files section.

### Figures
Figure 1

PCA-plot

Figure 2

a) The distribution of lactase deficiency by rs4988235 genotypes among all population groups. The histogram is sorted from the lowest to the highest percentage of GG genotype (blue) of the 24439 people.

b) The weighted frequency distribution of lactase deficiency by rs4988235 GG genotype among all population groups in Russian regions (11325 people). Higher percentages of GG genotype prevalence are
represented in deeper blue. No data (grey) displays regions which were excluded from the analysis because of insufficient data (See Evaluation of frequencies in different Russian regions in Materials and Methods). For the weighted frequency distribution of GG genotype in East Slavs only throughout Russia see Fig. S1.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- TableS1.xlsx
- TableS2.xlsx
- TableS3.xlsx
- TableS4.xlsx
- TableS5.xlsx
- figs1.jpg
- Table1.xlsx
- Table2.xlsx