IDENTIFICATION, ORIGIN AND GEOGRAPHIC SPREAD OF RARE EURASIAN ATM HAPLOTYPES

Olkova MV[1] ✉, Koshel SM[2], Ponomarev GYu[1], Alimov AA[1]

[1] Research Centre for Medical Genetics, Moscow, Russia
[2] Lomonosov Moscow State University, Moscow, Russia

The exploration of ATM gene haplotype evolutionary variability and its geographic dissemination across Eurasia enables the utilization of these data to predict the distribution of diseases associated with alleles comprising these haplotypes. The study aimed to identify the ATM gene haplotypes unique to Eurasia based on the panel of 28 polymorphic loci determining two major "yin-yang" haplotypes. The analysis was conducted using individual genotype data for 28 ATM gene loci across 41 groups of indigenous Eurasian populations. Haplotypes were identified using the Haploview 4.1 software. Phylogenetic analysis was performed in RStudio using the pegas and geneHapR packages. Cartographic visualization of ATM haplotype frequencies was implemented using the GeneGeo cartographic package. The H7j haplotype was identified in indigenous peoples of North Eurasia, which suggests its localized origin. The maximum H7i frequency was reported for Finno-Ugric peoples of the north and indigenous peoples of the northern Eurasia. The highest H8j haplotype frequency is observed in populations of southern China and Japan, which probably indicates the place of its origin. The remaining haplotypes are rare, these are typical for small geographic regions: H6j is sporadically distributed across Asia, H8i is localized to Transcaucasia, and H9i is found in two meta-groups of Eastern European peoples. The data obtained suggest the presence of the ATM gene haplotypes specific for the population of Eurasia, emphasize the contribution of intragenic recombination in shaping the gene haplotype diversity, and highlight the importance of the geographic and demographic context for preservation and spread of distinct genetic variants.

Keywords: ATM, haplotypes, yin-yang, Eurasia, phylogeny, gene geography

Funding: the study was conducted within the framework of the State Assignment for the Research Centre for Medical Genetics.

Acknowledgments: the authors would like to thank DNA sample donors and Biobank of North Eurasia for biomaterial provided for the study.

Author contribution: Olkova MV — data acquisition and statistical processing, phylogenetic analysis, cartographic analysis, manuscript writing; Koshel SM — cartographic analysis; Ponomarev GYu — statistical data processing; Alimov AA — manuscript writing, research management.

Compliance with ethical standards: the study was approved by the Ethics Commitee of the Research Centre for Medical Genetics (protocol No. 1 date 29 June 2020). All the subjects submitted the informed consent; the data were anonymized.

✉ **Correspondence should be addressed:** Marina V. Olkova
Moskvorechye 1, Moscow, 115522, Russia; genetics@inbox.ru

Received: 03.04.2026 **Accepted:** 30.04.2026 **Published online:** 23.05.2026

DOI: 10.24075/brsmu.2026.023

ИДЕНТИФИКАЦИЯ, ПРОИСХОЖДЕНИЕ И ГЕОГРАФИЧЕСКОЕ РАСПРОСТРАНЕНИЕ РЕДКИХ ЕВРАЗИЙСКИХ ГАПЛОТИПОВ ГЕНА ATM

М. В. Олькова[1]✉, С. М. Кошель[2], Г. Ю. Пономарев[1], А. А. Алимов[1]

[1] Медико-генетический научный центр имени Н. П. Бочкова, Москва, Россия
[2] Московский государственный университет имени М. В. Ломоносова, Москва, Россия

Изучение эволюционной изменчивости и географической распространенности гаплотипов гена ATM в пределах Евразии дает возможность использовать эти данные при прогнозировании распространения заболеваний, ассоциированных с входящими в состав гаплотипов аллелями. Целью работы было провести поиск уникальных для территории Евразии гаплотипов гена ATM на основе панели 28 полиморфных локусов, определяющих два основных «инь-ян»-гаплотипа. Исследование выполнено на основе данных об индивидуальных генотипах 28 локусов гена ATM в 41 группе коренных народов Евразии. Идентификацию гаплотипов осуществляли с помощью программного обеспечения Haploview 4.1. Филогенетический анализ выполнен в среде RStudio посредством пакетов pegas и geneHapR. Картографическая визуализация частот гаплотипов гена ATM реализована с использованием картографического пакета GeneGeo. Гаплотип H7j идентифицирован у автохтонных народов Северной Евразии, что свидетельствует о его локализованном происхождении. Для H7i выявлен частотный максимум среди финно-угорских народов севера и коренных народов северных территорий Евразии. Максимальная распространенность гаплотипа H8j наблюдается у населения юга Китая и Японии, что, вероятно, указывает на место его возникновения. Оставшиеся гаплотипы являются редкими и характерны для небольших географических регионов: H6j встречается у жителей Азии, H8i — Закавказья, H9i — в двух метагруппах народов Восточной Европы. Полученные данные свидетельствуют о наличии специфичных для народонаселения Евразии гаплотипов гена ATM, подчеркивают вклад внутригенной рекомбинации в формировании гаплотипического разнообразия гена и акцентируют важность географического и демографического контекста в сохранении и распространении отдельных генетических вариантов.

Ключевые слова: ATM, гаплотипы, инь-ян, Евразия, филогения, геногеография

Финансирование: исследование выполнено в рамках Государственного задания для ФГБНУ «Медико-генетический научный центр имени академика Н. П. Бочкова».

Благодарности: авторы благодарят доноров образцов ДНК и «Биобанк Северной Евразии» за предоставленный материал для исследования.

Вклад авторов: М. В. Олькова — сбор и статистическая обработка данных, филогенетический анализ, картографический анализ, подготовка текста статьи; С. М. Кошель — картографический анализ; Г. Ю. Пономарев — статистическая обработка данных; А. А. Алимов — подготовка текста статьи, руководство исследованием.

Соблюдение этических стандартов: исследование одобрено этическим комитетом ФГБНУ «Медико-генетический научный центр имени академика Н. П. Бочкова» (протокол № 1 от 29 июня 2020 г.). Все участники исследования подписали информированное согласие; данные деперсонифицированы.

✉ **Для корреспонденции:** Марина Викторовна Олькова
ул. Москворечье, д. 1, г. Москва, 115522, Россия; genetics@inbox.ru

Статья получена: 03.04.2026 **Статья принята к печати:** 30.04.2026 **Опубликована онлайн:** 23.05.2026

DOI: 10.24075/vrgmu.2026.023

The ATM (ataxia-telangiectasia mutated) gene product, a master regulator of the cell cycle and DNA repair machinery, constitutes a critical sentinel of genomic stability. Homozygous or compound heterozygous mutations in the gene ATM result in the development of ataxia-telangiectasia, a severe hereditary syndrome characterized by progressive neurological deficits, immunodeficiency, and increased predisposition to malignant neoplasms [1, 2]. Intriguingly, even heterozygous carriers of pathogenic ATM variants exhibit a statistically significant increase in susceptibility to certain cancers (most notably breast cancer) underscoring the profound clinical and societal imperative to decipher the genetic architecture of this locus[3–5]. Contemporary investigations into ATM mutations, now actively pursued across diverse populations, simultaneously target two distinct cohorts: cancer patients, to elucidate the contribution of pathogenic ATM alleles to tumorigenesis, and healthy individuals, to gauge carrier prevalence and identify subclinical cancer risk, thereby refining preventive strategies [6–12].

The recent genetic study involving the analysis of 3876 individuals, representing 46 combined groups of peoples of Eurasia and Africa, revealed a long linkage disequilibrium block consisting of 28 high-frequency polymorphic loci evenly distributed within the ATM gene between the 5' and 3' untranslated regions [4]. Through linkage analysis, eleven distinct ATM haplotypes were resolved, phylogenetically diverging into two branches that share a common ancestral origin: six haplotypes constitute branch I, and five belong to branch J. The two most prevalent haplotypes within these branches — H6i and H5j — carry opposing allelic variants, designated as "yin-yang" haplotypes [13–15]. These haplotypes are predominant across Eurasia, reaching a cumulative frequency of 93.7% [16]. According to modern concepts, branches I and J of the "yin-yang" ATM haplotypes originated under conditions of long isolation of the groups of *Homo sapiens* in Africa in the early stages of human evolution [14]. Subsequent migratory events facilitated the haplotype distribution observed in modern indigenous African populations. Nonetheless, the dissemination and frequency of ATM haplotypes in Eurasia exhibit substantial divergence from those documented on the African continent. In particular, only 5 currently known ATM out of 11 have been found in the indigenous population of Eurasia, while the whole spectrum is preserved in the indigenous population of Africa. These differences are likely to result from genetic drift and reflect migration processes that occurred during the human dispersal out of Africa. Thorough analysis of the genotypes formed on the basis of ATM hapolotypes has shown that a heterozygous genotype for "yin-yang" haplotypes predominates in the groups of indigenous peoples of Eurasia, which can be due to the effects of balancing selection. High frequency of heterozygotes for the H6i and H5j haplotypes carrying conflicting alleles suggests the possibility of recombination between "yin" and "yang" haplotypes with subsequent consolidation of recombinant forms in isolated groups. At the same time, "reverse" mutations in key markers defining the canonical "yin" and "yang" haplotypes are possible, with their subsequent fixation and frequency increase due to stochastic genetic drift. Such mutant and recombinant forms, carrying novel allelic combinations, may possess unique functional properties impacting ATM gene expression levels, ATM mRNA stability, and the kinetic activity of the ATM kinase protein. The study aimed to conduct a targeted search for the ATM haplotypes unique for Eurasia based on the panel of 28 polymorphic loci determining two major "yin-yang" haplotypes, as well as to conduct phylogenetic analysis of their origin and assess their geographic spread across the studied population of Eurasia.

METHODS

Sources of data on individual genotypes

The analysis was based on the data on genotypes of 3372 individuals representing 41 combined groups of indigenous peoples of Eurasia formed on the principle of common origin. The data were acquired from two independent sources:

1) 1975 DNA samples from 27 meta-groups of peoples living in Russia and neighboring countries were provided by Biobank of North Eurasia. DNA samples from unrelated individuals were included in the sample, provided self-identification of them and their ancestors with a particular ethnic group throughout at least three generations. The size of distinct samples varied between 41 and 129 individuals. Allelic variants of 28 polymorphic ATM gene markers were isolated from the data set by whole-genome genotyping using the Infinium Omni5Exome BeadChip Kit (Illumina Inc.).

2) Genotypes of 1397 individuals belonging to 14 groups of indigenous populations of Eurasia were obtained from the archive of the 1000 Genomes Project Phase 3 (Ensembl). The size of distinct samples was 85–113 individuals.

The list indigenous population meta-groups, abbreviated names of those, and sample size values are provided in Table 1.

Testing the quality of the DNA sample and polymorphic ATM gene loci typing

Inadequately typed samples were excluded based on the standard threshold GenCall score of 0.15. Calculation of the marker and DNA sample genotyping percentage and testing for Hardy–Weinberg equilibrium for all ATM loci were performed using the snpStats package in RStudio. Polymorphic markers and samples with the typing quality below 95% were excluded from the study.

Statistical data processing

Statistical processing and analysis of the data were performed in the RStudio integrated development environment (version 2025.09.1+401, Posit, USA) for R (version 4.5.1), as well as using Microsoft Excel (Microsoft Corporation, USA).

Search for and identification of ATM haplotypes

The haplotype structure was determined and the haplotype frequency was calculated using the Haploview 4.1 software tool (Broad Institute, USA).

Haplotypes were determined by the confidence interval method (Gabriel graph-based algorithm), provided that 95% of the pairs of polymorphic loci within the block showed the value $D' > 0.8$. Genotype frequency in the studied groups was calculated based on the reconstructed ATM haplotypes using the plyr package [17] for R.

Phylogenetic analysis

Phylogenetic analysis was conducted using the pegas [18] and geneHapR [19] R-packages in accordance with the principle of parsimony. Input files contained the data on haplotype frequencies in groups multiplied by a factor 1000.

Cartographic analysis

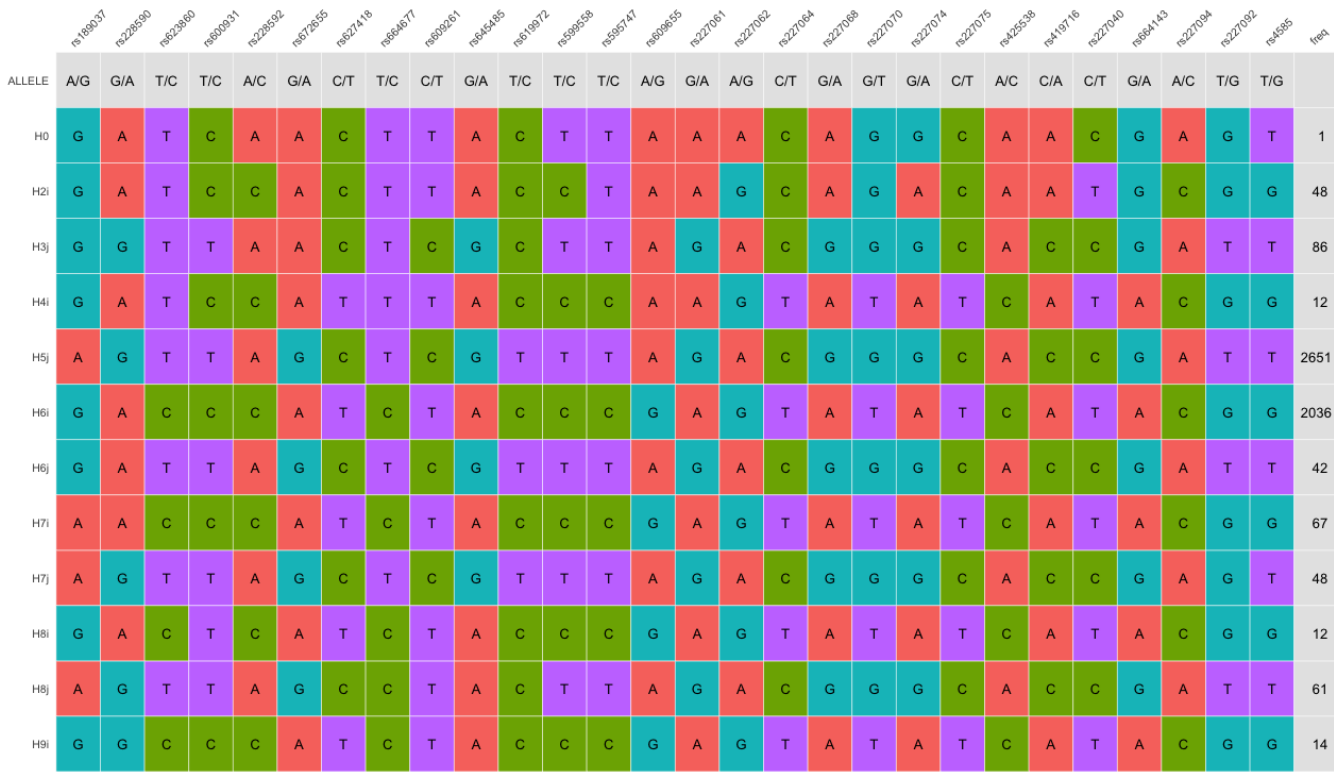
Cartographic visualization of the ATM gene haplotype and genotype frequencies in each groups was performed

Table 1. List of meta-groups of indigenous peoples included in the study

Abbreviation	Meta-group of indigenous peoples	Sample size
ALT	Altai peoples	83
ANC	Ancestral population	–
BASH	Bashkir	44
BEB	Bengali	86
BKHY	Buryats, Khamnigans, Yakuts	59
BRNW	Belarusians and Russians of the Northwest of the European Part of Russia	61
CDX	Dai from Xishuangbanna, China	93
CHB	Han from Beijing, China	103
CHKI	Chukchi, Koryaks, Itelmens	75
CHS	Southern Han, China	105
ECAU	Peoples of the North-Eastern Caucasus	123
ESN	Esan from Nigeria	99
ETRC	Peoples of Eastern Transcaucasia	87
FE	Peoples of the Far East	104
FIN	Finns	99
GBR	British	91
GIH	Gujaratis (India)	103
GWD	Gambians from the Western Division, Gambia	113
IBS	Spaniards	107
ITU	Telugu (India)	102
JPT	Japanese	104
KHMN	Khanty, Mansi, Nenets	63
KHV	Vietnamese	99
KKUN	Kazakhs, Karakalpaks, Uighurs, Nogais	53
KLMN	Mongols (other groups) and Kalmyks	79
KOUD	Komi and Udmurts	84
KV	Karelians and Vepsians	59
LWK	Luhya, Kenya	99
MCH	Mari and Chuvash	53
MKH	Mongols (Khalkha)	49
MOR	Mordva	41
MSL	Mende, Sierra Leone	85
PJL	Punjabis, Pakistan	96
RN	Russians of the northern European part of Russia	84
RNAR	Russians of the northern Arkhangelsk region	75
RSE	Russians of the southeast of European Russia	62
ST	Siberian Tatars	68
STU	Tamils of Sri Lanka	102
TAD	Tajiks, Pamiris, Yagnobis	72
TAT	Tatars	60
TSI	Italians	107
TUTO	Tuvans and Tofalar People	63
URSW	Ukranians and Russians of the southwest of European Russia	129
UZTK	Uzbeks, Turkmens, Kyrgyz	98
WCAU	Peoples of the Northwest Caucasus	91
WTRC	Peoples of Western Transcaucasia	56
YRI	Yoruba, Nigeria	108

using the original GeneGeo cartographic software package [20] developed under supervision of E.V. Balanovska and O.P. Balanovsky. Frequencies for the entire mapped territory were calculated by the average weighted interpolation method with the weighting function inversely proportional to the cube of the distance and a radius of influence of 4000 km using the following formula:

$$f(\boldsymbol{p}) = \frac{\sum f_i w_i}{\sum w_i}, \quad w_i = |\boldsymbol{p} - \boldsymbol{p}_i|^{-3}$$



	rs189037	rs228590	rs623860	rs600931	rs228592	rs672655	rs627418	rs664677	rs609261	rs645485	rs619972	rs599558	rs595747	rs606555	rs227061	rs227062	rs227064	rs227068	rs227070	rs227074	rs227075	rs425538	rs419716	rs227040	rs664143	rs227094	rs227092	rs4585	freq
ALLELE	A/G	G/A	T/C	T/C	A/C	G/A	C/T	T/C	C/T	G/A	T/C	T/C	T/C	A/G	G/A	A/G	C/T	G/A	G/T	G/A	C/T	A/C	C/A	C/T	G/A	A/C	T/G	T/G	
H0	G	A	T	C	A	A	C	T	T	A	C	T	T	A	A	A	C	A	G	G	C	A	A	C	G	A	G	T	1
H2i	G	A	T	C	C	A	C	T	T	A	C	C	T	A	A	G	C	A	G	A	C	A	A	T	G	C	G	G	48
H3j	G	G	T	T	A	A	C	T	C	G	C	T	T	A	G	A	C	G	G	G	C	A	C	C	G	A	T	T	86
H4i	G	A	T	C	C	A	T	T	T	A	C	C	C	A	A	G	T	A	T	A	T	C	A	T	A	C	G	G	12
H5j	A	G	T	T	A	G	C	T	C	G	T	T	T	A	G	A	C	G	G	G	C	A	C	C	G	A	T	T	2651
H6i	G	A	C	C	C	A	T	C	T	A	C	C	C	G	A	G	T	A	T	A	T	C	A	T	A	C	G	G	2036
H6j	G	A	T	T	A	G	C	T	C	G	T	T	T	A	G	A	C	G	G	G	C	A	C	C	G	A	T	T	42
H7i	A	A	C	C	C	A	T	C	T	A	C	C	C	G	A	G	T	A	T	A	T	C	A	T	A	C	G	G	67
H7j	A	G	T	T	A	G	C	T	C	G	T	T	T	A	G	A	C	G	G	G	C	A	C	C	G	A	G	T	48
H8i	G	A	C	T	C	A	T	C	T	A	C	C	C	G	A	G	T	A	T	A	T	C	A	T	A	C	G	G	12
H8j	A	G	T	T	A	G	C	C	T	A	C	T	T	A	G	A	C	G	G	G	C	A	C	C	G	A	T	T	61
H9i	G	G	C	C	C	A	T	C	T	A	C	C	C	G	A	G	T	A	T	A	T	C	A	T	A	C	G	G	14

Fig. 1. ATM haplotypes found in Eurasia

where p — arbitrary point of the mapped territory, $f(p)$ — frequency value at point p, p_i — meta-group geographic localization (interpolation nodes), f_i — frequency value at point p_i, w_i — ith point weight, and summation is extended over all points pi falling into the circle with the radius of 4000 cm and center in point p. In order to avoid the effects of map projection distortions on the calculations, during interpolation all distances were calculated on a sphere. The Winkel tripel projection with the central meridian 60° east longitude was used to create the maps.

RESULTS

The study was based on the data on individual genotypes of 28 ATM gene loci in 41 groups of indigenous peoples of Eurasia. All polymorphic markers within the meta-groups conformed to the Hardy–Weinberg equilibrium. Six unique haplotypes were identified (Fig. 1–4). Among these, two haplotypes — H6j and H8j — were most plausibly derived from recombination between the predominant "yin-yang" haplotypes of branches J and I: H5j and H6i (Fig. 3A, 4A). H6j and H8j were found in groups of people of Asian origin only. The other four haplotypes (H7j, H7i, H8i, H9i) differ from the canonical "yin" and "yang" by single nucleotide substitutions and, therefore, can result from both single point mutations and the recombination event (Table 2).

In northeastern Europe, two region-specific haplotypes belonging to two different branches were found: H7i (derived from branch I) and H7j (derived from branch J). The H7i haplotype is also found in East Asia and among the population of Finland, In contrast, the H7j haplotype has, to date, been identified exclusively among the indigenous populations of Northern Eurasia and is not found elsewhere.

Cartographic analysis has shown that the H7i haplotype (Fig. 4B) differing from H6i by the substitution of G for A in rs189037 is likely to emerge in the coastal regions of northern Eurasia among peoples of Finno-Ugric origin. Later it spread mainly to the east (all the way to the Pacific coast) and south (to the Alpine-Himalayan belt), and to a lesser extent to the west. The H7i haplotype maximum frequency has been reported for the meta-group “Khanty and Mansi” (12.5%), as well as for the geographically close group “Komi and Udmurts” (7.5%). The frequency ~5.5% is observed in meta-groups “Tuvans and Tofalar People” and “Chukchi, Koryaks, Itelmens”. Homozygous carriers of this haplotype (frequency 1–2%) were found in the region of its geographic maximum in representatives of the same meta-groups (“Khanty and Mansi”, “Komi and Udmurts”).

The H8j haplotype (Fig. 3D) is distinguished from the H5j haplotype, in which the listed below loci occupy the neighboring positions 8–11, by four nucleotide substitutions (T→C, C→T, G→A, T→C) in the loci rs664677, rs609261, rs645485, and rs619972. These substitutions correspond to alleles of the H6i haplotype that is more common in East Asia. The highest H8j frequency is observed in the Han population of the southern coast of China and in Japan: ~ 6%. Cartographic analysis revealed a gradient spread of H8j across all studied peoples of East Asia, as well as across North Eurasian meta-groups of Asian origin (“Khalkha Mongols”, “Other Groups of Mongols and Kalmyks”, “Altaians”, “Siberian Tatars”, and “Kazakhs, Karakalpaks, Uighurs, Nogais”), in which the haplotype frequency was 2–4%.

The H7i and H8j haplotype spatial distribution characterized by the clearly visible local frequency maximum and progressive frequency decrease with increasing distance from the maximum is highly likely to suggest the founder effect and long isolation of ancestral carriers of these haplotypes. The combination of relatively reported frequencies in the maxima and wide geographic spread is in favor of the ancient origin of these haplotypes.

The H7j haplotype (Fig. 3C) resulting from the substitution of T for G in the rs227092 locus of the H5j haplotype can be characterized as “Eastern European”, since its frequency maxima are observed in Mordovia and meta-group “Belarusians and

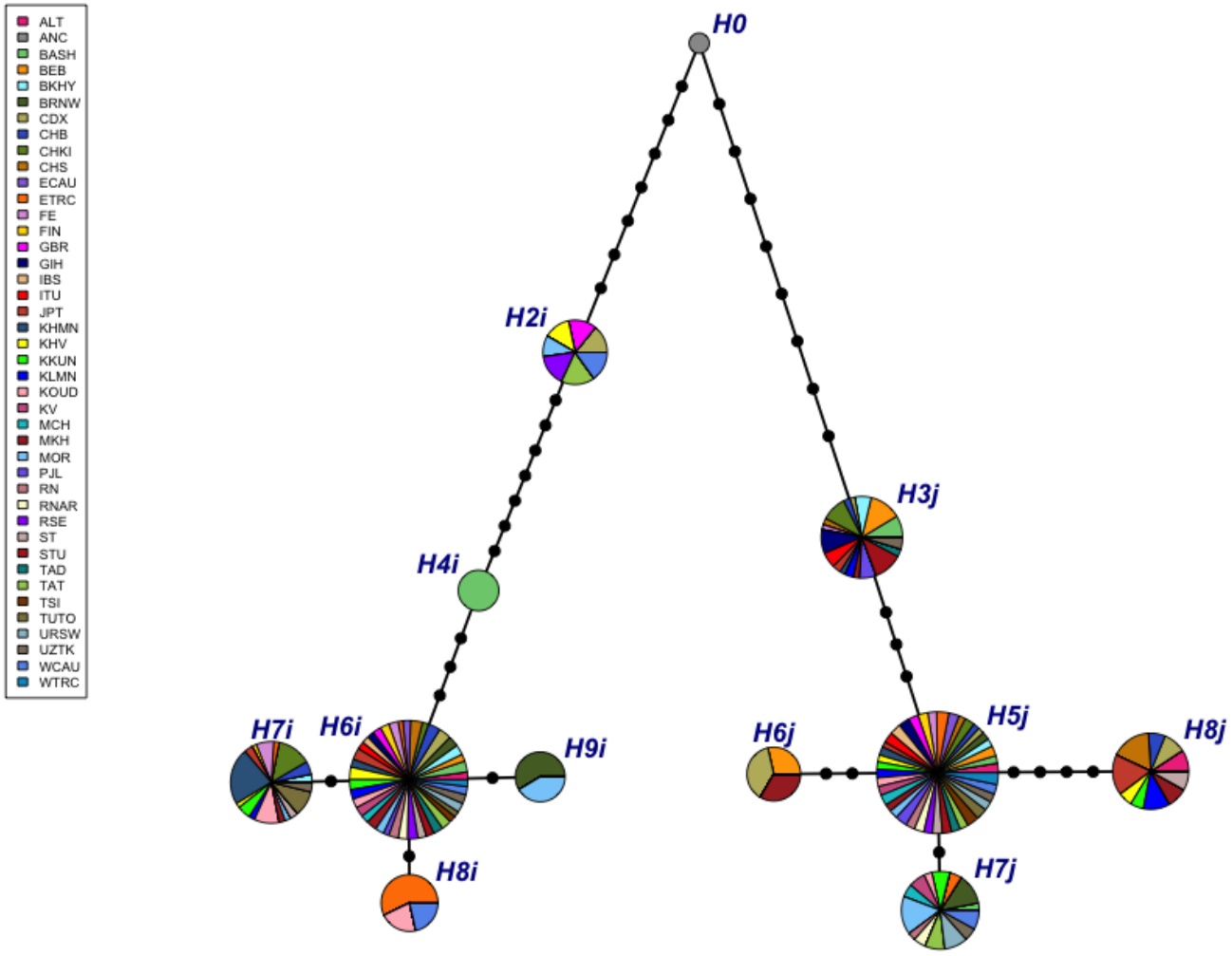


Fig. 2. Phylogenetic network of "Eurasian" ATM haplotypes

Russians of the Northwest of the European Part of Russia" (6.2 and 5%, respectively). H7j is widely spread among indigenous peoples of the east of Europe, all the way to Urals and Caucasus.

The rather rare H8i haplotype (Fig. 4c) resulting from the substitution of cytosine for thymine in the rs600931 position is found mainly in the population of Caucasus (frequency 1–3%), which can indicate its local spread under the conditions of isolation in the mountains. H8i was also found in one representative of the "Komi and Udmurts" meta-group.

The rare "Asian" H6j haplotype (frequency ~1–2%) is found in the Dai people (southwestern China), Khalkha Mongols from Mongolia, and people from Bangladesh (Fig. 3B). It results from the substitution of A and G alleles (typical for H5j) for alternative alleles G and A (typical for H6i) in the neighboring linked loci rs189037 and rs228590.

The H9i haplotype (Fig. 4D) was found only in individual representatives of two Eastern European groups of peoples: "Belarusians and Russians of the Northwest of the European Part of Russia" and "Mordva" (frequency 1–2%). H9i is distinguished by the substitution of adenine for guanine in the rs228590 position of the H6i haplotype.

DISCUSSION

In prior investigations, we established that the ATM gene features a conserved linkage disequilibrium block comprising 28 extra-exonic loci, which constitute a distinct haplotype. The population-based analysis of samples from Eurasia and Africa has shown that the haplotype allelic variants are distributed in accordance with the "yin-yang" concept. Phylogenetic reconstruction of the 11 identified ATM haplotypes confirmed the existence of two independent evolutionary lineages diverging from an ancestral haplotype. A high frequency of heterozygotes carrying these conflicting allelic haplotypes was ubiquitously reported throughout Eurasia, providing the impetus for this study to elucidate possible combinations of "yin" and "yang" haplotypes and to identify rare Eurasian ATM haplotypes. These were discovered during the study.

The cartographic analysis conducted has shown that unique ATM haplotypes are concentrated primarily in peripheral and often isolated regions, which emphasizes the role of demographic factors in their preservation.

In summary, we can say that haplotypes as the groups of genetic markers inherited together represent "genetic signatures" reflecting the history of peoples and their relationships over millennia. Investigation of their spread enables the reconstruction of ancient settlement paths and identification of the events associated with changes in gene flows that shaped the contemporary genetic landscape of Eurasia. The reported ATM haplotype divergence manifested by division into "yin" and "yang" variants of the common ancestral type suggest that there was an ancient evolutionary event, which left a clear mark on the genetic structure of the modern peoples of the continent. Presumably, such divergence is maintained by the effects of balancing selection, which, however, requires further study.

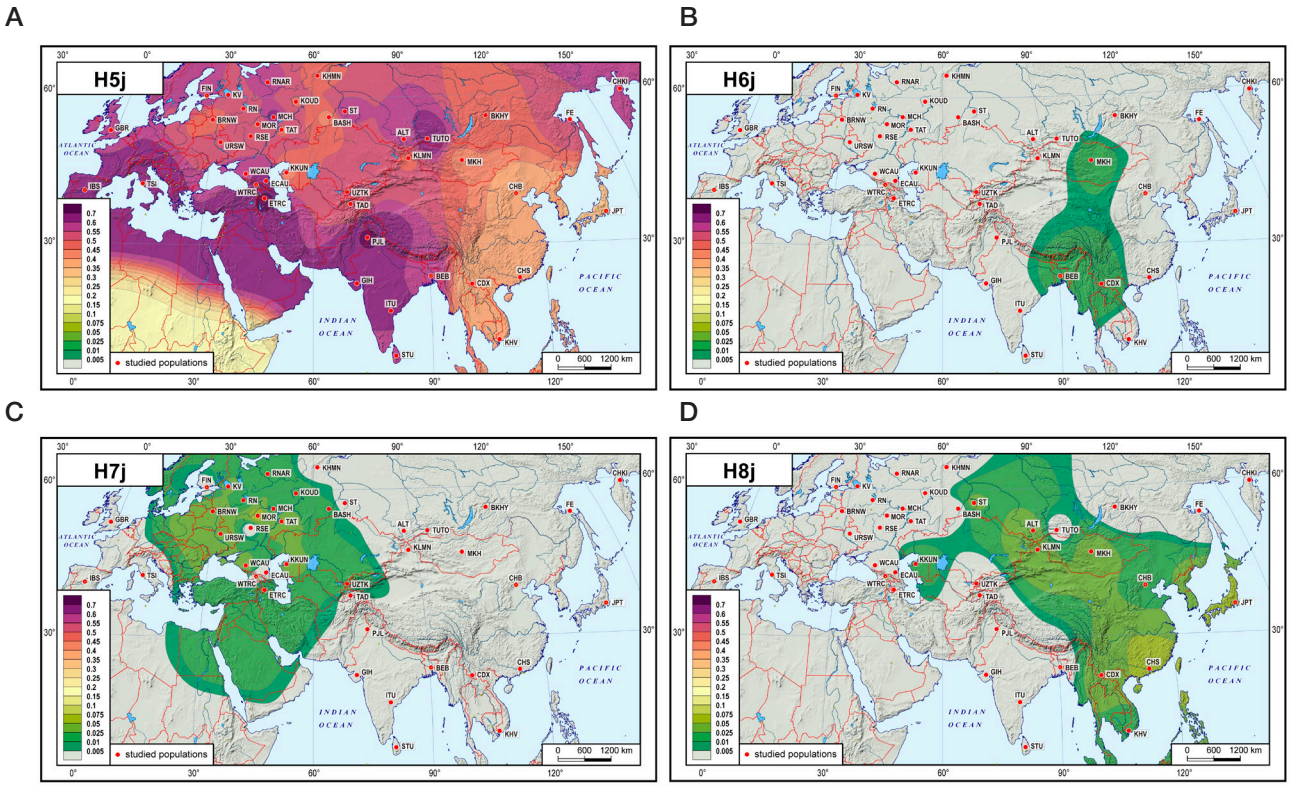


Fig. 3. Geographical distribution of ATM haplotypes, branch J

It is critically important that the recombination points identified in this study are localized in the ATM gene non-coding regions, adjacent to the regulatory sequences and polymorphic loci associated with predisposition to malignant neoplasms. This means that the ATM gene protein product expression variation may result from specific combinations of allelic variants being part of the haplotype in a certain individual. Therefore, the research focused on ATM haplotypes is of considerable interest not only for elucidating evolutionary processes, but also for medical genetics, in particular, for assessing individual risks of cancer and developing personalized prevention and treatment strategies.

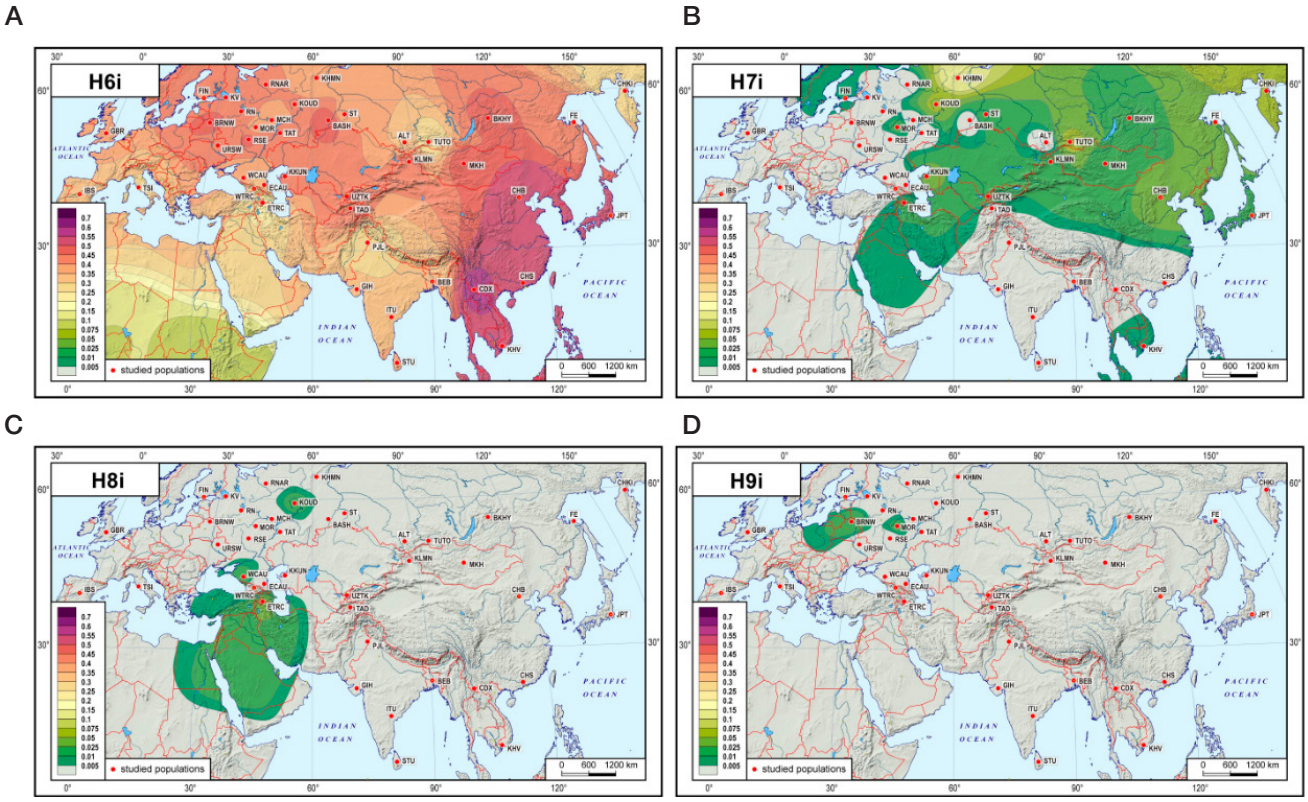


Fig. 4. Geographical distribution of ATM haplotypes, branch I

Table 2. Comparison of the allelic composition of the "yin-yang" and "Eurasian" ATM haplotypes

Haploblock loci	Similarity to "yin-yang" haplotypes					
	H6i/H7i	H6i/H8i	H6i/H9i	H5j/H6j	H5j/H7j	H5j/H8j
rs189037	RL	+	+	RL	+	+
rs228590	+	+	RL	RL	+	+
rs623860	+	+	+	+	+	+
rs600931	+	RL	+	+	+	+
rs228592	+	+	+	+	+	+
rs672655	+	+	+	+	+	+
rs627418	+	+	+	+	+	+
rs664677	+	+	+	+	+	RL
rs609261	+	+	+	+	+	RL
rs645485	+	+	+	+	+	RL
rs619972	+	+	+	+	+	RL
rs599558	+	+	+	+	+	+
rs595747	+	+	+	+	+	+
rs609655	+	+	+	+	+	+
rs227061	+	+	+	+	+	+
rs227062	+	+	+	+	+	+
rs227064	+	+	+	+	+	+
rs227068	+	+	+	+	+	+
rs227070	+	+	+	+	+	+
rs227074	+	+	+	+	+	+
rs227075	+	+	+	+	+	+
rs425538	+	+	+	+	+	+
rs419716	+	+	+	+	+	+
rs227040	+	+	+	+	+	+
rs664143	+	+	+	+	+	+
rs227094	+	+	+	+	+	+
rs227092	+	+	+	+	RL	+
rs4585	+	+	+	+	+	+

Note: RL — presumably recombinant or mutant locus. + — alleles matched in the "yin-yang" haplotype and the compared one. In the column header on the left — "yin-yang" haplotype, to the branch of which the haplotype compared specified through a slash on the right belongs.

CONCLUSIONS

The study reveals a pronounced genetic and geographic specificity of ATM haplotypes among indigenous populations of Eurasia. The unique haplotypes identified — H6j, H7j, H8j, H7i, H8i, and H9i — exhibit distinct geographic localization and peak frequency in peripheral, often demographically isolated regions, corroborating the pivotal roles of intragenic recombination, genetic drift, and historical-demographic processes in shaping and preserving regional genetic profiles. The clinical and genetic significance of these findings is underscored by the fact that potential recombination breakpoints reside within non-coding gene regions adjacent to regulatory sequences and polymorphic loci associated with predisposition to malignant neoplasms. This architecture suggests that specific haplotypes may modulate ATM gene expression, thereby influencing individual cancer risk. Consequently, the study of ATM haplotype architecture holds dual value: it deepens the understanding of evolutionary mechanisms underlying genetic diversity formation and offers promising avenues for personalized medicine, particularly in cancer risk stratification and targeted prevention strategies. Future research should prioritize functional validation of specific variants within these haplotypes, elucidation of evolutionary mechanisms maintaining "yin-yang" dimorphism and integration of indigenous genotype frequency data into clinical models for predictive and personalized therapy.

References

1. Amirifar P, Ranjouri MR, Lavin M, Abolhassani H, Yazdani R, Aghamohammadi A. Ataxia-telangiectasia: epidemiology, pathogenesis, clinical phenotype, diagnosis, prognosis and management. Expert Review of Clinical Immunology. 2020; 16 (9): 859–71. DOI: 10.1080/1744666X.2020.1810570.
2. Asekretova TV, Anderzhanova LH, Leonteva ME, Rodina YuA, Panferova AV, Aleksenko MYu, i dr. Harakteristika klinicheskih i laboratornyh proyavlenij v gruppe pacientov s sindromom ataksii-teleangiektazii. Voprosy gematologii/onkologii i immunopatologii v pediatrii. 2022; 21 (3): 47–55. DOI: 10.24287/1726-1708-2022-

21-3-47–55. Russian.
3. Lee SA, Lee KM, Lee SJ, Yoo KY, Park SK, Noh DY, et al. Antioxidant Vitamins Intake, Ataxia Telangiectasia Mutated (ATM) Genetic Polymorphisms, and Breast Cancer Risk. Nutrition and Cancer. 2010; 62 (8): 1087–94. DOI:10.1080/01635581.2010.492088.
4. Wang S, Zhang Y, Chen M, Wang Y, Feng Y, Xu Z, et al. Association of genetic variants in ATR-CHEK1 and ATM-CHEK2 pathway genes with risk of colorectal cancer in a Chinese population. Oncotarget. 2018; 9 (42): 26616–24. DOI: 10.18632/oncotarget.24299.
5. Mou J, Hu T, Wang Z, Chen W, Wang Y, Zhang W. ATM gene polymorphisms are associated with poor prognosis of non-small cell lung cancer receiving radiation therapy. Aging. 2020; 12 (8): 7465–79. DOI: 10.18632/aging.103094 PubMed PMID: 32329754.
6. Suspitsin E, Sokolenko A, Bizin I, Tumakova A, Guseva M, Sokolova N, et al. ATM mutation spectrum in Russian children with ataxia-telangiectasia. European Journal of Medical Genetics. 2020; 63 (1): 103630. DOI: 10.1016/j.ejmg.2019.02.003.
7. Podralska M, Dzikiewicz-Krawczyk A, Mosor M, Żurawek M, Iżykowska K, Słomski R, et al. The most frequent Polish ATM mutations are not susceptibility factors for tobacco-related cancers. Arch Med Sci. 2021; 17 (5): 1158–63. DOI: 10.5114/aoms.2020.94155.
8. Stankovic T, Kidd AMJ, Sutcliffe A, McGuire GM, Robinson P, Weber P, et al. ATM Mutations and Phenotypes in Ataxia-Telangiectasia Families in the British Isles: Expression of Mutant ATM and the Risk of Leukemia, Lymphoma, and Breast Cancer. The American Journal of Human Genetics. 1998; 62 (2): 334–45. DOI: 10.1086/301706. PubMed PMID: 9463314.
9. Fukunaga H, Taki Y, Prise KM. Diversity of ATM gene variants: a population-based genome data analysis for precision medicine. Hum Genomics. 2019; 13: 38. DOI: 10.1186/s40246-019-0234-2. PubMed PMID: 31443742; PubMed Central PMCID: PMC6708157.
10. Bu R, Siraj AK, Al-Rasheed M, Iqbal K, Azam S, Qadri Z, et al. Identification and characterization of ATM founder mutation in BRCA-negative breast cancer patients of Arab ethnicity. Sci Rep. 2023; 13 (1): 20924. DOI: 10.1038/s41598-023-48231-0.
11. Dalal Amandi AR, Jabbarpour N, Shiva S, Bonyadi M. Identification of Two Novel Pathogenic Variants of the ATM Gene in the Iranian-Azeri Turkish Ethnic Group by Applying Whole ExomeSequencing. CG. 2023; 24 (6): 345–53. DOI: 10.2174/0113892029268949231104165301.
12. Moslemi M, Vafaei M, Khani P, Soheili M, Nedaeinia R, Manian M, et al. The prevalence of ataxia telangiectasia mutated (ATM) variants in patients with breast cancer patients: a systematic review and meta-analysis. Cancer Cell Int. 2021; 21: 474. DOI: 10.1186/s12935-021-02172-8. PubMed PMID: 34493284; PubMed Central PMCID: PMC8424893.
13. Zhang J, Rowe WL, Clark AG, Buetow KH. Genomewide Distribution of High-Frequency, Completely Mismatching SNP Haplotype Pairs Observed To Be Common across Human Populations. The American Journal of Human Genetics. 2003; 73 (5): 1073–81. DOI: 10.1086/379154.
14. Dutta R, Mainsah J, Yatskiv Y, Chakrabortty S, Brennan P, Khuder B, et al. Intricacies in arrangement of SNP haplotypes suggest “Great Admixture” that created modern humans. BMC Genomics. 2017; 18 (1): 433. DOI: 10.1186/s12864-017-3776-5.
15. Curtis D, Amos W. The human genome harbours widespread exclusive yin yang haplotypes. Eur J Hum Genet. 2024; 32 (6): 691–6. DOI: 10.1038/s41431-023-01399-5.
16. Olkova MV, Koshel SM, Ponomarev GYu, Alimov AA. Phylogenetic relationships of ATM polymorphic loci haplotypes in Eurasian and African populations | bioRxiv [Internet]. [cited 2026 Mar 20]. Available from: https://www.biorxiv.org/content/10.1101/2025.05.10.652950v3.full.
17. Wickham H. The Split-Apply-Combine Strategy for Data Analysis. Journal of Statistical Software. 2011; 40 (1): 1–29. https://www.jstatsoft.org/v40/i01/.
18. Paradis E. PEGAS: an R package for population genetics with an integrated–modular approach. Bioinformatics. 2010; 26: 419–20. DOI: 10.1093/bioinformatics/btp696.
19. Zhang R, Jia G, Diao X. GeneHapR: an R package for gene haplotypic statistics and visualization. BMC Bioinformatics. 2023; 24 (1): 199. Available from: https://doi.org/10.1186/s12859-023-05318-9
20. Koshel SM. Geoinformacionnye tekhnologii v genogeografii. Sovremennaya geograficheskaya kartografiya. 2012; s. 158–166. Russian.

Литература

1. Amirifar P, Ranjouri MR, Lavin M, Abolhassani H, Yazdani R, Aghamohammadi A. Ataxia-telangiectasia: epidemiology, pathogenesis, clinical phenotype, diagnosis, prognosis and management. Expert Review of Clinical Immunology. 2020; 16 (9): 859–71. DOI: 10.1080/1744666X.2020.1810570.
2. Асекретова Т. В., Андержанова Л. Х., Леонтьева М. Е., Родина Ю. А., Панферова А. В., Алексенко М. Ю., и др. Характеристика клинических и лабораторных проявлений в группе пациентов с синдромом атаксии-телеангиэктазии. Вопросы гематологии/онкологии и иммунопатологии в педиатрии. 2022; 21 (3): 47–55. DOI: 10.24287/1726-1708-2022-21-3-47-55.
3. Lee SA, Lee KM, Lee SJ, Yoo KY, Park SK, Noh DY, et al. Antioxidant Vitamins Intake, Ataxia Telangiectasia Mutated (ATM) Genetic Polymorphisms, and Breast Cancer Risk. Nutrition and Cancer. 2010; 62 (8): 1087–94. DOI:10.1080/01635581.2010.492088.
4. Wang S, Zhang Y, Chen M, Wang Y, Feng Y, Xu Z, et al. Association of genetic variants in ATR-CHEK1 and ATM-CHEK2 pathway genes with risk of colorectal cancer in a Chinese population. Oncotarget. 2018; 9 (42): 26616–24. DOI: 10.18632/oncotarget.24299.
5. Mou J, Hu T, Wang Z, Chen W, Wang Y, Zhang W. ATM gene polymorphisms are associated with poor prognosis of non-small cell lung cancer receiving radiation therapy. Aging. 2020; 12 (8): 7465–79. DOI: 10.18632/aging.103094 PubMed PMID: 32329754.
6. Suspitsin E, Sokolenko A, Bizin I, Tumakova A, Guseva M, Sokolova N, et al. ATM mutation spectrum in Russian children with ataxia-telangiectasia. European Journal of Medical Genetics. 2020; 63 (1): 103630. DOI: 10.1016/j.ejmg.2019.02.003.
7. Podralska M, Dzikiewicz-Krawczyk A, Mosor M, Żurawek M, Iżykowska K, Słomski R, et al. The most frequent Polish ATM mutations are not susceptibility factors for tobacco-related cancers. Arch Med Sci. 2021; 17 (5): 1158–63. DOI: 10.5114/aoms.2020.94155.
8. Stankovic T, Kidd AMJ, Sutcliffe A, McGuire GM, Robinson P, Weber P, et al. ATM Mutations and Phenotypes in Ataxia-Telangiectasia Families in the British Isles: Expression of Mutant ATM and the Risk of Leukemia, Lymphoma, and Breast Cancer. The American Journal of Human Genetics. 1998; 62 (2): 334–45. DOI: 10.1086/301706. PubMed PMID: 9463314.
9. Fukunaga H, Taki Y, Prise KM. Diversity of ATM gene variants: a population-based genome data analysis for precision medicine. Hum Genomics. 2019; 13: 38. DOI: 10.1186/s40246-019-0234-2. PubMed PMID: 31443742; PubMed Central PMCID: PMC6708157.
10. Bu R, Siraj AK, Al-Rasheed M, Iqbal K, Azam S, Qadri Z, et al. Identification and characterization of ATM founder mutation in BRCA-negative breast cancer patients of Arab ethnicity. Sci Rep. 2023; 13 (1): 20924. DOI: 10.1038/s41598-023-48231-0.
11. Dalal Amandi AR, Jabbarpour N, Shiva S, Bonyadi M. Identification of Two Novel Pathogenic Variants of the ATM Gene in the Iranian-Azeri Turkish Ethnic Group by Applying Whole ExomeSequencing. CG. 2023; 24 (6): 345–53. DOI: 10.2174/0113892029268949231104165301.
12. Moslemi M, Vafaei M, Khani P, Soheili M, Nedaeinia R, Manian M, et al. The prevalence of ataxia telangiectasia mutated (ATM) variants in patients with breast cancer patients: a systematic review and meta-analysis. Cancer Cell Int. 2021; 21: 474. DOI: 10.1186/s12935-021-02172-8. PubMed PMID: 34493284; PubMed Central PMCID: PMC8424893.
13. Zhang J, Rowe WL, Clark AG, Buetow KH. Genomewide Distribution of High-Frequency, Completely Mismatching SNP

Haplotype Pairs Observed To Be Common across Human Populations. The American Journal of Human Genetics. 2003; 73 (5): 1073–81. DOI: 10.1086/379154.
14. Dutta R, Mainsah J, Yatskiv Y, Chakrabortty S, Brennan P, Khuder B, et al. Intricacies in arrangement of SNP haplotypes suggest "Great Admixture" that created modern humans. BMC Genomics. 2017; 18 (1): 433. DOI: 10.1186/s12864-017-3776-5.
15. Curtis D, Amos W. The human genome harbours widespread exclusive yin yang haplotypes. Eur J Hum Genet. 2024; 32 (6): 691–6. DOI: 10.1038/s41431-023-01399-5.
16. Olkova MV, Koshel SM, Ponomarev GYu, Alimov AA. Phylogenetic relationships of ATM polymorphic loci haplotypes in Eurasian and African populations | bioRxiv [Internet]. [cited 2026 Mar 20]. Available from: https://www.biorxiv.org/content/10.1101/2025.05.10.652950v3.full.
17. Wickham H. The Split-Apply-Combine Strategy for Data Analysis. Journal of Statistical Software. 2011; 40 (1): 1–29. https://www.jstatsoft.org/v40/i01/.
18. Paradis E. PEGAS: an R package for population genetics with an integrated–modular approach. Bioinformatics. 2010; 26: 419–20. DOI: 10.1093/bioinformatics/btp696.
19. Zhang R, Jia G, Diao X. GeneHapR: an R package for gene haplotypic statistics and visualization. BMC Bioinformatics. 2023; 24 (1): 199. Available from: https://doi.org/10.1186/s12859-023-05318-9
20. Кошель С. М. Геоинформационные технологии в геногеографии. Современная географическая картография. 2012; с. 158–166.