2025 • 29 • 5

Online ISSN 2500-3259

Сетевое издание

ВАВИЛОВСКИЙ ЖУРНАЛ ГЕНЕТИКИ И СЕЛЕКЦИИ

VAVILOV JOURNAL OF GENETICS AND BREEDING

Основан в 1997 г.
Периодичность 8 выпусков в год
doi 10.18699/vjgb-25-65

Учредители

Сибирское отделение Российской академии наук

Федеральное государственное бюджетное научное учреждение «Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук»

Межрегиональная общественная организация Вавиловское общество генетиков и селекционеров

Главный редактор

А.В. Кочетов – академик РАН, д-р биол. наук, профессор РАН (Россия)

Заместители главного редактора

Н.А. Колчанов – академик РАН, д-р биол. наук, профессор (Россия)
И.Н. Леонова – д-р биол. наук (Россия)
Н.Б. Рубцов – д-р биол. наук, профессор (Россия)
В.К. Шумный – академик РАН, д-р биол. наук, профессор (Россия)

Ответственный секретарь

Г.В. Орлова – канд. биол. наук (Россия)

Редакционная коллегия

Е.Е. Андронов – канд. биол. наук (Россия)
Ю.С. Аульченко – д-р биол. наук (Россия)
О.С. Афанасенко – академик РАН, д-р биол. наук (Россия)
Д.А. Афонников – д-р биол. наук, доцент (Россия)
Л.И. Афтанас – академик РАН, д-р мед. наук (Россия)
Л.А. Беспалова – академик РАН, д-р с.-х. наук (Россия)
А. Бёрнер – д-р наук (Германия)
Н.П. Бондарь – канд. биол. наук (Россия)
С.А. Боринская – д-р биол. наук (Россия)
П.М. Бородин – д-р биол. наук, проф. (Россия)
А.В. Васильев – чл.-кор. РАН, д-р биол. наук (Россия)
М.И. Воевода – академик РАН, д-р мед. наук (Россия)
Т.А. Гавриленко – д-р биол. наук (Россия)
И. Гроссе – д-р наук, проф. (Германия)
Н.Е. Грунтенко – д-р биол. наук (Россия)
С.А. Демаков – д-р биол. наук (Россия)
И.К. Захаров – д-р биол. наук, проф. (Россия)
И.А. Захаров-Гезехус – чл.-кор. РАН, д-р биол. наук (Россия)
С.Г. Инге-Вечтомов – академик РАН, д-р биол. наук (Россия)
А.В. Кильчевский – чл.-кор. НАНБ, д-р биол. наук (Беларусь)
С.В. Костров – чл.-кор. РАН, д-р хим. наук (Россия)
А.М. Кудрявцев – чл.-кор. РАН, д-р биол. наук (Россия)
И.Н. Лаврик – д-р биол. наук (Германия)
Д.М. Ларкин – канд. биол. наук (Великобритания)
Ж. Ле Гуи – д-р наук (Франция)
И.Н. Лебедев – чл.-кор. РАН, д-р биол. наук, проф. (Россия)
Л.А. Лутова – д-р биол. наук, проф. (Россия)
Б. Люгтенберг – д-р наук, проф. (Нидерланды)
В.Ю. Макеев – чл.-кор. РАН, д-р физ.-мат. наук (Россия)
В.И. Молодин – академик РАН, д-р ист. наук (Россия)
М.П. Мошкин – д-р биол. наук, проф. (Россия)
С.Р. Мурсалимов – канд. биол. наук (Россия)
Л.Ю. Новикова – д-р с.-х. наук (Россия)
Е.К. Потокина – д-р биол. наук (Россия)
В.П. Пузырев – академик РАН, д-р мед. наук (Россия)
Д.В. Пышный – чл.-кор. РАН, д-р хим. наук (Россия)
И.Б. Рогозин – канд. биол. наук (США)
А.О. Рувинский – д-р биол. наук, проф. (Австралия)
Е.Ю. Рыкова – д-р биол. наук (Россия)
Е.А. Салина – чл.-кор. РАН, д-р биол. наук, проф. (Россия)
В.А. Степанов – академик РАН, д-р биол. наук (Россия)
И.А. Тихонович – академик РАН, д-р биол. наук (Россия)
Е.К. Хлесткина – чл.-кор. РАН, д-р биол. наук, проф. РАН (Россия)
Э.К. Хуснутдинова – д-р биол. наук, проф. (Россия)
М. Чен – д-р биол. наук (Китайская Народная Республика)
Ю.Н. Шавруков – д-р биол. наук (Австралия)
Р.И. Шейко – чл.-кор. НАНБ, д-р с.-х. наук (Беларусь)
С.В. Шестаков – академик РАН, д-р биол. наук (Россия)
Н.К. Янковский – академик РАН, д-р биол. наук (Россия)

2025 • 29 • 5

Online ISSN 2500-3259

Online edition

VAVILOVSKII ZHURNAL GENETIKI I SELEKTSII

VAVILOV JOURNAL OF GENETICS AND BREEDING

Founded in 1997
Publication frequency: 8 issues a year
doi 10.18699/vjgb-25-65

Founders

Siberian Branch of the Russian Academy of Sciences

Federal Research Center Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences

The Vavilov Society of Geneticists and Breeders

Editor-in-Chief

A.V. Kochetov, Full Member of the Russian Academy of Sciences, Dr. Sci. (Biology), Professor of the RAS, Russia

Deputy Editor-in-Chief

N.A. Kolchanov, Full Member of the Russian Academy of Sciences, Dr. Sci. (Biology), Russia
I.N. Leonova, Dr. Sci. (Biology), Russia
N.B. Rubtsov, Professor, Dr. Sci. (Biology), Russia
V.K. Shumny, Full Member of the Russian Academy of Sciences, Dr. Sci. (Biology), Russia

Executive Secretary

G.V. Orlova, Cand. Sci. (Biology), Russia

Editorial board

O.S. Afanasenko, Full Member of the RAS, Dr. Sci. (Biology), Russia
D.A. Afonnikov, Associate Professor, Dr. Sci. (Biology), Russia
L.I. Aftanas, Full Member of the RAS, Dr. Sci. (Medicine), Russia
E.E. Andronov, Cand. Sci. (Biology), Russia
Yu.S. Aulchenko, Dr. Sci. (Biology), Russia
L.A. Bespalova, Full Member of the RAS, Dr. Sci. (Agricul.), Russia
N.P. Bondar, Cand. Sci. (Biology), Russia
S.A. Borinskaya, Dr. Sci. (Biology), Russia
P.M. Borodin, Professor, Dr. Sci. (Biology), Russia
A. Börner, Dr. Sci., Germany
M. Chen, Dr. Sci. (Biology), People's Republic of China
S.A. Demakov, Dr. Sci. (Biology), Russia
T.A. Gavrilenko, Dr. Sci. (Biology), Russia
I. Grosse, Professor, Dr. Sci., Germany
N.E. Gruntenko, Dr. Sci. (Biology), Russia
S.G. Inge-Vechtomov, Full Member of the RAS, Dr. Sci. (Biology), Russia
E.K. Khlestkina, Corr. Member of the RAS, Professor of the RAS, Dr. Sci. (Biology), Russia
E.K. Khusnutdinova, Professor, Dr. Sci. (Biology), Russia
A.V. Kilchevsky, Corr. Member of the NAS of Belarus, Dr. Sci. (Biology), Belarus
S.V. Kostrov, Corr. Member of the RAS, Dr. Sci. (Chemistry), Russia
A.M. Kudryavtsev, Corr. Member of the RAS, Dr. Sci. (Biology), Russia
D.M. Larkin, Cand. Sci. (Biology), Great Britain
I.N. Lavrik, Dr. Sci. (Biology), Germany
J. Le Gouis, Dr. Sci., France
I.N. Lebedev, Corr. Member of the RAS, Professor, Dr. Sci. (Biology), Russia
B. Lugtenberg, Professor, Dr. Sci., Netherlands
L.A. Lutova, Professor, Dr. Sci. (Biology), Russia
V.Yu. Makeev, Corr. Member of the RAS, Dr. Sci. (Physics and Mathem.), Russia
V.I. Molodin, Full Member of the RAS, Dr. Sci. (History), Russia
M.P. Moshkin, Professor, Dr. Sci. (Biology), Russia
S.R. Mursalimov, Cand. Sci. (Biology), Russia
L.Yu. Novikova, Dr. Sci. (Agricul.), Russia
E.K. Potokina, Dr. Sci. (Biology), Russia
V.P. Puzyrev, Full Member of the RAS, Dr. Sci. (Medicine), Russia
D.V. Pyshnyi, Corr. Member of the RAS, Dr. Sci. (Chemistry), Russia
I.B. Rogozin, Cand. Sci. (Biology), United States
A.O. Ruvinsky, Professor, Dr. Sci. (Biology), Australia
E.Y. Rykova, Dr. Sci. (Biology), Russia
E.A. Salina, Corr. Member of the RAS, Professor, Dr. Sci. (Biology), Russia
Y.N. Shavrukov, Dr. Sci. (Biology), Australia
R.I. Sheiko, Corr. Member of the NAS of Belarus, Dr. Sci. (Agricul.), Belarus
S.V. Shestakov, Full Member of the RAS, Dr. Sci. (Biology), Russia
V.A. Stepanov, Full Member of the RAS, Dr. Sci. (Biology), Russia
I.A. Tikhonovich, Full Member of the RAS, Dr. Sci. (Biology), Russia
A.V. Vasiliev, Corr. Member of the RAS, Dr. Sci. (Biology), Russia
M.I. Voevoda, Full Member of the RAS, Dr. Sci. (Medicine), Russia
N.K. Yankovsky, Full Member of the RAS, Dr. Sci. (Biology), Russia
I.K. Zakharov, Professor, Dr. Sci. (Biology), Russia
I.A. Zakharov-Gezekhus, Corr. Member of the RAS, Dr. Sci. (Biology), Russia

СОДЕРЖАНИЕ • 2025 • 29 • 5

© Сибирское отделение Российской академии наук, 2025
© Институт цитологии и генетики СО РАН, 2025
Вавиловский журнал генетики и селекции, 2025

VAVILOVSKII ZHURNAL GENETIKI I SELEKTSII • VAVILOV JOURNAL OF GENETICS AND BREEDING

CONTENTS • 2025 • 29 • 5

© Siberian Branch RAS, 2025
© Institute of Cytology and Genetics, SB RAS, 2025

Vavilov Journal of Genetics and Breeding. 2025;29(5):619-620
doi 10.18699/vjgb-25-66

Winter School on Cytogenomics

In the early 2000s, the possibilities of classical and molecular cytogenetics in the study of chromosomal abnormalities in humans were significantly expanded by the emerging technologies of genomic analysis. The use of chromosomal microarray analysis, various modifications of next-generation sequencing, chromosome conformation capture, and optical genome mapping marked the beginning of a new cytogenomic era in cytogenetics. The combination of cytogenetic and genomic technologies has opened up new perspectives in the diagnosis of complex and submicroscopic chromosomal aberrations. The Winter School on Cytogenomics, held in Tomsk on November 25–29, 2024, was devoted to discussing these current trends.

The event was organized by the Research Institute of Medical Genetics of the Tomsk National Research Medical Center of the Russian Academy of Sciences. It was dedicated to the memory of Corresponding Member of the Russian Academy of Medical Sciences, Professor Sergey Nazarenko – founder and first head of the Laboratory of Cytogenetics at the Institute, and also coincided to the 25th Anniversary of the Medical Genetics Division of the Siberian State Medical University. Lectures were delivered by Russian and foreign experts in the field of clinical cytogenetics and cytogenomics – N. Rubtsov (Novosibirsk), I. Lebedev (Tomsk), N. Shilova (Moscow), V. Chernykh (Moscow), V. Fishman (Novosibirsk), T. Liehr (Germany), M. Zamani-Esteki (Netherlands), P. Li (USA), J. Vermeesch (Belgium). Workshops were held on chromosomal microarray analysis (CMA), fluorescent *in situ* hybridization (FISH), quantitative real time PCR, as well as on the clinical interpretation of the CMA results. Within the framework of the School, a competition for talks by young scientists was organized, the participants of which and their colleagues had the opportunity to present their research on the pages of this Issue in the section “Medical cytogenomics”.

The Issue opens with an article by V.P. Pushkarev and co-authors “Molecular genetic study of triploidy and the hydatidiform mole in pregnancy loss: analysis of 10,000 consecutive cases”. By studying a significant samples of spontaneous abortions using quantitative fluorescent PCR, the frequency of complete hydatidiform mole was determined. This pathology arises due to genomic imprinting effects caused by abnormal combination of parental haploid genomes in the zygote – specifically two paternal genomes in the absence of a maternal one. The frequency was estimated as 0.11 % and it was close to the epidemiological data typical for European populations.

The article by A.S. Iakovleva and co-authors presents a case involving a combination of low-level mosaicism for trisomy 9 and uniparental disomy of the same chromosome. The use of a CMA and FISH made it possible to describe the mosaic karyotype in detail and demonstrate that such a combination of chromosomal abnormalities is a consequence of postzygotic trisomy rescue. It is noteworthy that trisomy in the zygote arose as a result of an error in the meiosis II, as indicated by the alternation of iso- and heterodisomy sites in uniparental disomy.

The article by A.E. Kopytova and co-authors presents a family case of Xq28 chromosome duplication. The area of revealed copy number gain overlaps with the region of chromosome Xq28 duplication syndrome. However, both brothers and their mother carry a rare, non-classical duplication. That is why the use of common algorithms for classifying the clinical significance of duplication defines it as a variant of unknown clinical significance. However, given the asymmetric inactivation of the X chromosome in a mother, healthy carrier of the duplication, the authors reasonably suggest considering this duplication as pathogenic.

The article by G.D. Moskvitin and co-authors presents a case of interstitial deletion 6p22.3-p24.3 in monozygotic twins from Yakutia with severe speech delay, intellectual disability and congenital malformations. The study highlights the importance of CMA in diagnosis of chromosomal disorders and discusses the challenges in establishing genephenotype correlations.

In their article, M.M. Antonova and co-authors examine the meiotic segregation features of the paracentric inversion inv(7)(q11.23q22), one of the most frequent in the human karyotype. FISH analysis of spermatozoa in the inversion carrier allowed to establish the predominant and almost equally proportions of gametes with inversion and intact chromosome 7. Recombinant chromosomes were noted in 0.7 % of gametes only, confirming the presence of crossing over in the inversion loop.

Finally, the article by A.V. Vozilova and co-authors describes the segregation of a balanced translocation t(3;10)(p25;p15) across seven family members spanning three generations. The structure of chromosomal rearrangement and the products of its meiotic segregation, including unbalanced translocations, were investigated using CMA and clinical exome sequencing. The mechanisms of formation and clinical features of segmental aneuploidy in the terminal regions of chromosomes 3 and 10 are discussed.

The articles in the section "Medical cytogenomics" of this Issue demonstrate the potential of current cytogenomic technologies in diagnosing of chromosomal disorders and reproductive abnormalities. It is important that these technologies are available in national research centers that have presented their results at the Winter School on Cytogenomics.

I.N. Lebedev, Corresponding Member
of the Russian Academy of Sciences,
Doctor of Biological Sciences,
Chairman of the Program Committee
of the Winter School on Cytogenomics,
Executive editor of the Issue

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2025;29(5):621-628

doi 10.18699/vjgb-25-67

Molecular genetic study of triploidy and the hydatidiform mole in pregnancy loss: analysis of 10,000 consecutive cases

V.P. Pushkarev [1, 2] ✉, A.S. Masycheva [1], E.A. Glazyrina [1], T.E. Serebrenikova [1], V.B. Chernykh [2]

[1] Medical Genetic Center LLC Progen, Moscow, Russia
[2] Bochkov Research Centre for Medical Genetics, Moscow, Russia
✉ v.p.pushkarev@gmail.com

Abstract. Approximately 10–15 % of clinically recognized pregnancies result in miscarriage, with chromosomal abnormalities identified in about 50 % of early pregnancy losses (PL). Triploidy accounts for approximately 12 % of all chromosomal abnormalities in miscarriages. The additional haploid set of chromosomes in triploidy may be of paternal (diandric triploidy) or maternal (digynic triploidy) origin. Diandric triploidy is associated with a partial hydatidiform mole (PHM), while pregnancies involving diploid embryos with two paternal genomes (and loss of the maternal nuclear genome) are the most common cause of a complete hydatidiform mole (CHM). The hydatidiform mole (HM) is the most prevalent form of gestational trophoblastic disease. Genotyping of products of conception (POC) is currently considered a reliable method for confirming HM and distinguishing its subtypes. The aim of this study was to use DNA genotyping of POCs to detect cases of triploidy, estimate the frequency of HM and its subtypes, and analyze the molecular and clinical characteristics of triploid pregnancies, CHM, and PHM in a Russian population. Between 2018 and 2024, a total of 10,000 consecutive PL cases were analyzed at the Medical Genetic Center Progen (Moscow). The main clinical indications included spontaneous miscarriage, missed miscarriage, and anembryonic pregnancy. DNA genotyping was performed using a five-color multiplex QF-PCR method, which included profiling of 26 autosomal STR markers, as well as DYS437, DXS6809, the *SRY* gene, and 30 markers from homologous regions located on different chromosomes. CHM was diagnosed based on the homozygosity of all STR markers. Triploidy was identified by analyzing peak area ratios of non-homozygous STR markers, which exhibited characteristic patterns of approximately 2:1 or 1:1:1. In our cohort, chromosomal abnormalities were identified in 58.8 % of all PL cases. Triploidy was detected in 8.3 % of the total sample, representing 14.3 % of all chromosomally abnormal POCs. Diandric triploidy accounted for 43 % of triploid cases. The prevalence of CHM was 0.11 %. The median age of women with triploidy was 32.1 years, and 27.9 years for those with CHM. Given the observed frequencies of PHM and CHM in our cohort, along with the relatively young maternal age associated with these conditions, enhancing current diagnostic protocols for HM – particularly through the incorporation of DNA genotyping of POCs – is essential for the effective prevention and timely diagnosis of post-molar malignant neoplasms in this population.

Key words: triploidy; hydatidiform mole (complete and partial); miscarriage; quantitative fluorescent PCR (QF-PCR); short tandem repeats (STR)

For citation: Pushkarev V.P., Masycheva A.S., Glazyrina E.A., Serebrenikova T.E., Chernykh V.B. Molecular genetic study of triploidy and the hydatidiform mole in pregnancy loss: analysis of 10,000 consecutive cases. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov J Genet Breed*. 2025;29(5):621-628. doi 10.18699/vjgb-25-67

Молекулярно-генетическое исследование триплоидии и пузырного заноса при невынашивании беременности: анализ 10 000 последовательных случаев

В.П. Пушкарев [1, 2] ✉, А.С. Масычева [1], Е.А. Глазырина [1], Т.Е. Серебреникова [1], В.Б. Черных [2]

[1] Медико-генетический центр «Проген», Москва, Россия
[2] Медико-генетический научный центр им. академика Н.П. Бочкова, Москва, Россия
✉ v.p.pushkarev@gmail.com

Аннотация. Из клинически признанных беременностей 10–15 % заканчиваются выкидышем, и около 50 % абортусов на ранних сроках беременности имеют хромосомные аномалии. Триплоидии составляют примерно 12 % от всех хромосомных аномалий абортусов. Дополнительный гаплоидный набор хромосом может быть отцовского (диандрическая триплоидия) или материнского происхождения (дигиническая триплоидия). Диандрическая триплоидия проявляется частичным пузырным заносом (ЧПЗ). Беременности диплоид-

© Pushkarev V.P., Masycheva A.S., Glazyrina E.A., Serebrenikova T.E., Chernykh V.B., 2025
This work is licensed under a Creative Commons Attribution 4.0 License

ными эмбрионами с двумя геномами отцовского происхождения (и потерей материнского ядерного генома) признаны наиболее частой причиной полного пузырного заноса (ППЗ). Пузырный занос (ПЗ) – это самый распространенный тип гестационной трофобластической болезни. Генотипирование абортусов в настоящее время рассматривается как надежный метод для подтверждения и дифференциальной диагностики подтипов ПЗ. Целью данного исследования было с помощью ДНК-генотипирования абортусов при невынашивании беременности (НБ) выявить случаи триплоидии, оценить частоту ПЗ, его подтипов, молекулярно-генетические и клинические особенности триплоидной беременности, ППЗ и ЧПЗ в российской популяции. С 2018 по 2024 г. в медико-генетическом центре «Проген» (Москва) были исследованы 10 000 последовательных случаев НБ. Основными направительными диагнозами являлись спонтанный выкидыш, неразвивающаяся беременность, анэмбриония. ДНК-генотипирование проводилось с помощью метода мультиплексной КФ-ПЦР, включавшего профилирование 26 аутосомных STR-маркеров, DYS437, DXS6809, *SRY* и 30 маркеров на гомологичных участках пар хромосом. Критерием ППЗ была гомозиготность всех STR-маркеров. Критерием триплоидии было соотношение площадей пиков всех негомозиготных STR-маркеров, близкое к 2:1 или 1:1:1. В нашей выборке из 10 000 случаев НБ аномальный кариотип абортусов был выявлен в 58.8 %, доля триплоидии составила 8.3 % от общего числа случаев или 14.3 % от абортусов с аномальным кариотипом. Доля диандрической триплоидии составила 43 %. Частота ППЗ была равна 0.11 %. Медианный возраст женщин с триплоидией был равен 32.1 года, с ППЗ – 27.9 года. Учитывая оцененную в нашей выборке частоту ЧПЗ и ППЗ и относительно молодой возраст женщин, у которых он встречался, необходимо совершенствовать имеющиеся методы диагностики ПЗ (включение ДНК-генотипирования) с целью адекватной профилактики и своевременной диагностики постпузырных злокачественных новообразований в данной возрастной группе.
Ключевые слова: триплоидия; пузырный занос (полный и частичный); невынашивание беременности; количественная флуоресцентная ПЦР (КФ-ПЦР); короткие тандемные повторы; short tandem repeats (STR)

Introduction

Ten to fifteen percent of clinically recognized pregnancies end in miscarriage, with approximately 50 % of early pregnancy losses (PL) attributed to chromosomal abnormalities (Soler et al., 2017; Essers et al., 2023). Triploidy accounts for about 12 % of all chromosomal abnormalities identified in spontaneous abortions (Jenderny, 2014; Soler et al., 2017).

Triploidy is a genetic anomaly in embryonic or fetal cells characterized by the presence of three haploid sets of chromosomes ($3n = 69$) instead of the normal diploid number. The additional haploid set may be of paternal (diandric triploidy) or maternal (digynic triploidy) origin. The parental origin significantly influences the phenotypic manifestations of triploid pregnancies and maternal complications. Diandric triploidy most commonly arises from the fertilization of an ovum by two sperm cells (dispermy), or less frequently by a diploid sperm, and typically results in the development of a partial hydatidiform mole (PHM) (Fig. 1*C*). According to the concept of postzygotic diploidization of triploid cells proposed by M.D. Golubovsky in 2003, a normal ovum fertilized by two sperm cells may give rise to all types of hydatidiform mole (HM), as well as to a fetus. Sporadic complete hydatidiform mole (CHM) develops following monospermic (85 % of cases) or dispermic (15 %) fertilization of an ovum in which the maternal chromosomes are lost or destroyed shortly after fertilization (Fig. 1*A*, *B*). The result of monospermic fertilization is an androgenetic diploid zygote formed by endoreplication of the paternal genome (Candelier, 2016). In 10–20 % of cases, recurrent CHM is associated with biallelic pathogenic variants in maternal-effect genes. The list of implicated genes is steadily growing and currently includes *NLRP7*, *KHDC3L*, *MEI1*, *TOP6BL*, and *REC114* (Murdoch et al., 2006; Parry et al., 2011; Nguyen et al., 2018).

The incidence of HM varies significantly across populations, ranging from 1–2 cases per 1,000 pregnancies in Europe and the USA to as high as 10 per 1,000 in India and Indonesia (Joyce et al., 2022). Both complete and partial HMs carry the potential for malignant transformation, with the risk of gestational trophoblastic neoplasia (GTN) being higher for CHM than for PHM (Joyce et al., 2022).

In clinical practice, the main diagnostic tools for HM are elevated serum levels of β-human chorionic gonadotropin (β-hCG) – often tens of times higher than in normal pregnancies – and ultrasonographic findings. A definitive diagnosis is established through histopathological examination. However, these methods have limitations, particularly in early PL (Fukunaga et al., 2005; Sazhenova et al., 2009; Buza, Hui, 2021).

Genotyping of products of conception (POC) is currently considered a reliable approach for the confirmation and differential diagnosis of HM subtypes (Furtado et al., 2013; Ronnett, 2018; Buza, Hui, 2021). Distinguishing molar from non-molar specimens and differentiating PHM from CHM is critical for estimating the risk of post-molar GTN, which varies by HM subtype and determines the length and intensity of clinical follow-up (Buza, Hui, 2021).

The aim of the present study was to identify cases of triploidy using DNA genotyping of POCs from PL, to assess the prevalence of HM and its subtypes, and to characterize the molecular genetic and clinical features of triploid pregnancies, CHM, and PHM in the Russian population.

Materials and methods

Between 2018 and 2024, a total of 10,000 consecutive cases of PL were analyzed in the laboratory of the Medical Genetics Center Progen (Moscow), with the majority of referrals originating from Moscow and the Moscow region. The primary clinical indications included spontaneous miscarriage, missed abortion, and anembryonic pregnancy. Informed consent was obtained from all patients.

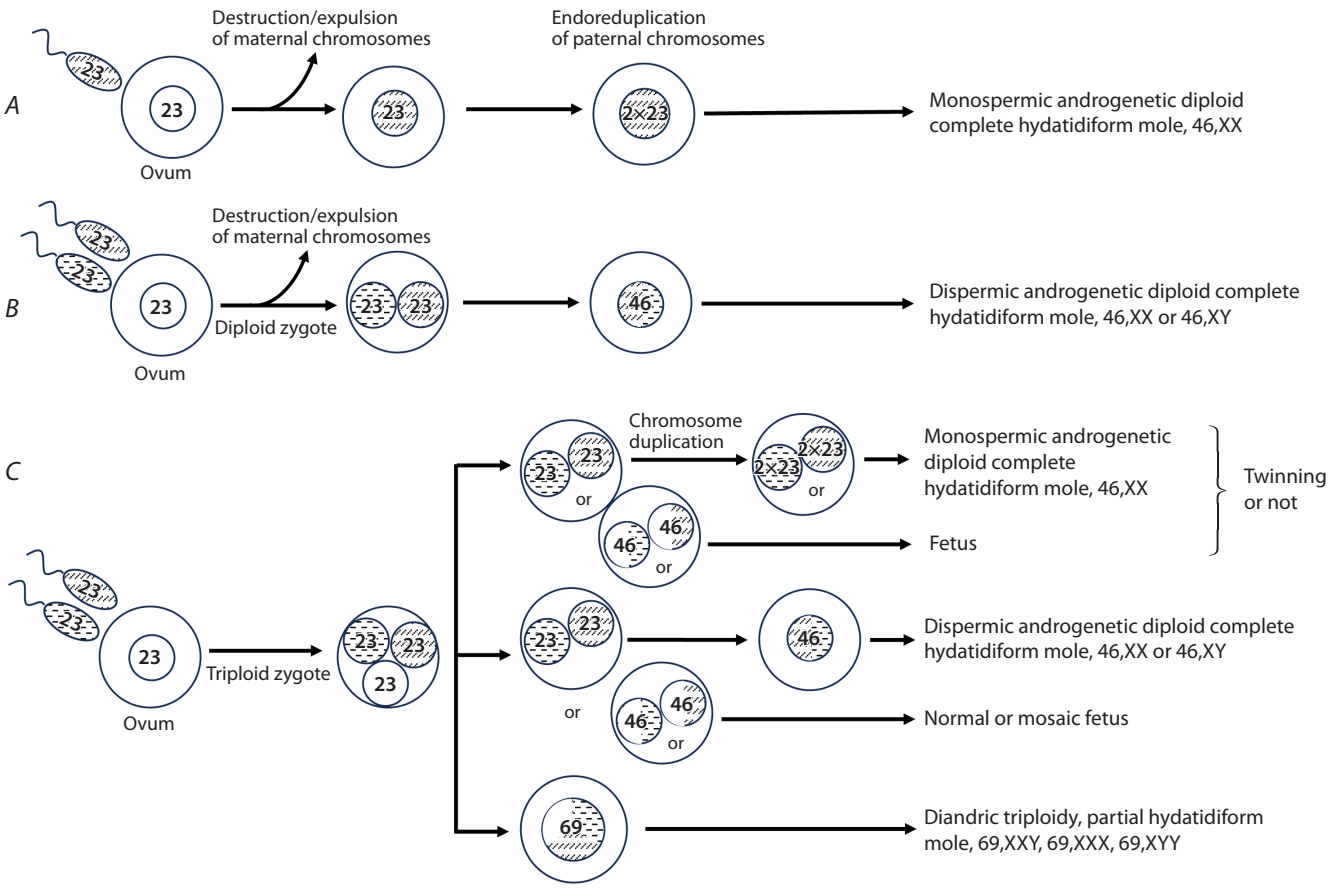


Fig. 1. The main mechanisms of sporadic hydatidiform mole (HM) development.

Complete hydatidiform mole develops after monospermic fertilization (85 % of cases) (*A*) or dispermic fertilization (15 % of cases) (*B*), where the maternal chromosomes were lost (or destroyed) immediately after conception. The result of the first scenario is an androgenetic diploid zygote with endoreplication of paternal chromosomes (*A*). *C* – post-zygotic diploidization of triploids (Golubovsky, 2003); a normal egg is fertilized by two sperm cells, resulting in a triploid zygote, which forms the basis for all types of HM and the fetus.

Chorionic villi, fetal membranes, and fetal tissues were examined as biological material. DNA genotyping was performed using five-color multiplex quantitative fluorescent polymerase chain reaction (QF-PCR), which included the profiling of 26 autosomal STR markers (D1S1656, D2S441, D3S1358, D4S2366, D4S2408, D5S818, D6S1017, D6S474, D7S820, D8S1179, D8S1115, D9S2157, D10S1248, D10S1435, THO1, D12S391, D13S317, D14S608, D15S659, D16S539, D18S535, D19S253, D20S482, D20S1082, D21S1412, D22S1045), a Y-STR marker (DYS437), an X-STR marker (DXS6809), the *SRY* gene, and 30 additional markers targeting homologous regions of different chromosome pairs. The selection criteria for STR markers included an expected heterozygosity of ≥0.7 and no more than 12 alleles in the Russian population (Smolyanitsky et al., 2004; Pesik et al., 2014; Zavarin et al., 2019).

PCR products were separated using a 3500 Genetic Analyzer (Thermo Fisher Scientific, USA), and electropherograms were analyzed with GeneMapper Software v5 (Thermo Fisher Scientific, USA). CHM was diagnosed based on homozygosity at all STR loci (Fig. 2), while triploidy was identified by peak area ratios of informative (heterozygous) STR markers approximating 2:1 or 1:1:1. The parental origin of triploidy (diandric or digynic) was determined by comparing the genotypes of the conceptus with those of the parents. The category "other chromosomal abnormalities" included autosomal monosomies and trisomies, sex chromosome aneuploidies, and complex karyotypic abnormalities. The "euploid karyotype" group included cases in which no chromosomal abnormalities were detected.

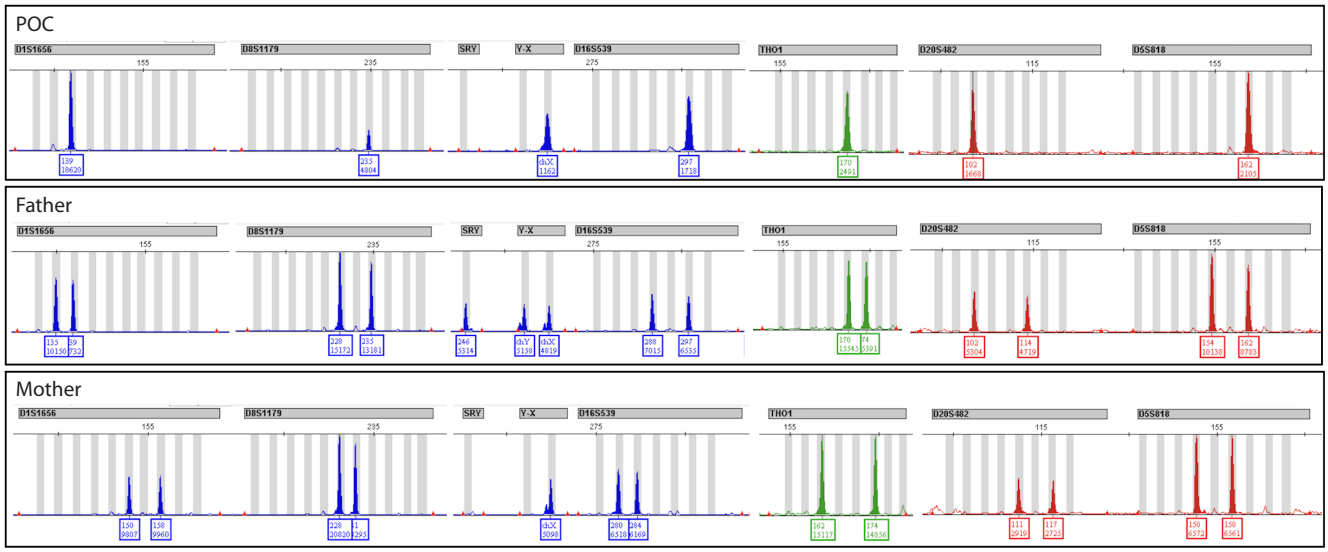
Statistical analyses were performed using R software (version 4.4.2).

Results

The median age of all women with PL was 34.6 years (interquartile range (IQR): 30.3–38.3 years). The median gestational age at PL was 7.5 embryonic weeks (IQR: 6.5–9). The summary of findings is presented in the Table.

Among 10,000 cases of PL, a normal (euploid) karyotype was identified in 4,122 samples (41.2 %; 95 % confidence interval (CI): 40.3–42.2). The median maternal age in this group was 33.5 years (IQR: 29.6–37.1), and the median gestational age at the time of PL was 7.5 weeks (IQR: 6.5–10).

An abnormal karyotype was detected in 5,878 cases (58.8 %; 95 % CI: 57.8–59.7). The median maternal age in this group was 35.4 years (IQR: 30.8–39.0), and the median



	D1S1656*	D8S1179	*SRY***	Y-X	D16S539	THO1	D20S482	D5S818
POC	139*	235	–	XX	297	170	102	162
Father	135–139	228–235	+	XY	288–297	170–174	102–114	154–162
Mother	150–158	228–231	–	XX	280–284	162–174	111–117	150–159

Note. * Length of alleles of informative STRs in nucleotides; ** Presence of *SRY* is indicated by a plus sign, absence by a minus sign. Shared alleles between the father's and POC's STR profiles are underlined. POC - product of conception.

Fig. 2. Identification of monospermic complete hydatidiform mole.

Results of the analysis of 10,000 cases of pregnancy loss

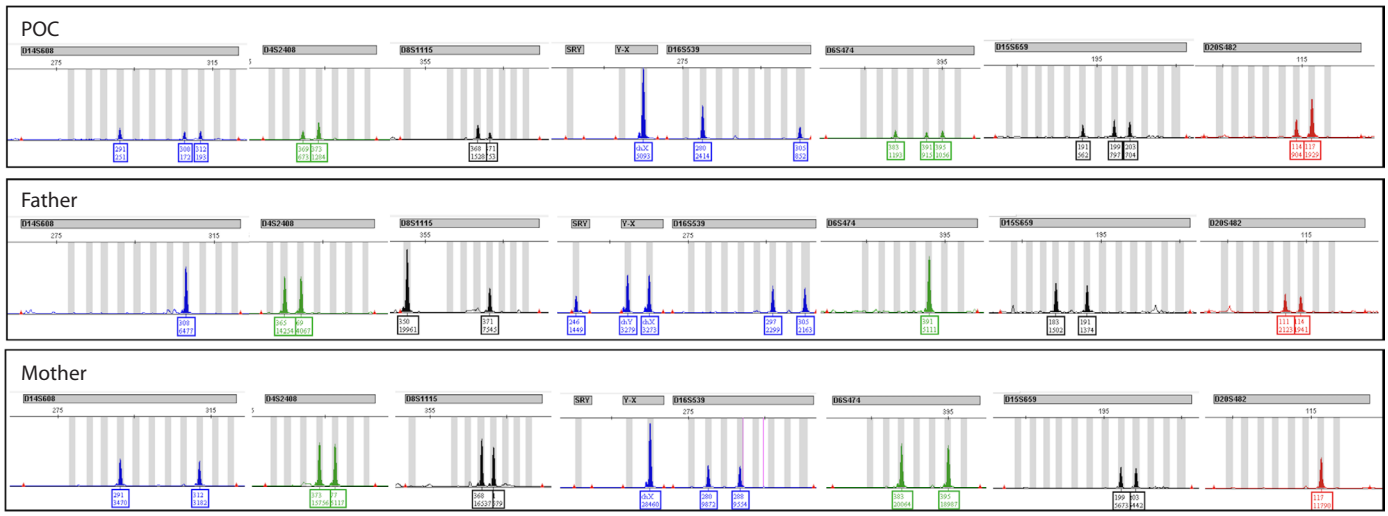
Indicator	Quantity	Proportion, % (95 % CI*)
Euploid karyotype	4,122	41.2 (40.3–42.2)
Abnormal karyotype, total	5,878	58.8 (57.8–59.7)
Autosomal, gonosomal aneuploidies, combined anomalies	5,038	50.4 (49.4–51.4)
Triploidy, total	829	8.30 (7.80–8.80)
69,XXY	448	4.50 (4.10–4.90)
69,XXX	363	3.60 (3.30–4.00)
69,XYY	13	0.13 (0.08–0.22)
68,XX	5	0.05 (0.02–0.12)
Diploid homozygous paternal genome	11	0.11 (0.06–0.20)

* 95 % confidence interval.

gestational age at miscarriage was 7.5 weeks (IQR: 6.5–9). According to the Mann–Whitney U-test, there was a statistically significant difference in maternal age between the euploid and aneuploid groups ($W = 8{,}233{,}198$, $p < 2.2 \times 10^{-16}$).

Triploidy was identified in 829 cases (8.3 % of all PL cases; 95 % CI: 7.8–8.8). The median maternal age in this group was 32.1 years (IQR: 28.2–35.8), and the median gestational age at the time of PL was 8 weeks (IQR: 7–9). The parental origin of triploidy was determined in 14 cases: digynic triploidy in eight cases (57 %) and diandric triploidy (partial hydatidiform mole, PHM) in six cases (43 %). An example of digynic triploidy is presented in Figure 3, and that of diandric triploidy, in Figure 4. Our observed digynic-to-diandric ratio is consistent with previous reports; for example, D. Massalska et al. (2021) identified diandric triploidy in 44.9 % of triploid cases.

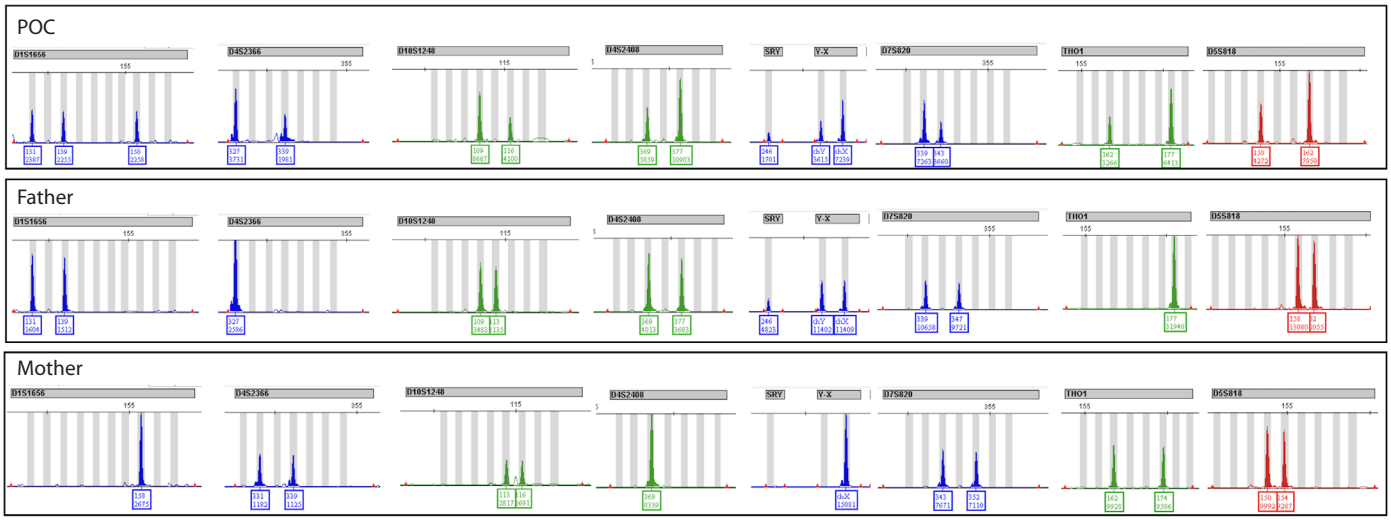
Complete hydatidiform mole was identified in 11 cases (0.11 % of all PL cases; 95 % CI: 0.06–0.20). Women in the CHM group were the youngest among all groups, with a



	D14S608*	D4S2408	D8S1115	*SRY***	Y-X	D16S539	D6S474	D15S659	D20S482
POC	291-308-312	369-373-373	368-368-371	–	XXX	280-280-305	383-391-395	191-199-203	114-117-117
Father	308	365-369	371	+	XY	297-305	391	183-191	111-114
Mother	291-312	373-377	368-371	–	XX	280-288	383-395	199-203	117

Note. * Length of alleles of informative STRs in nucleotides; ** Presence of *SRY* is indicated by a plus sign, absence by a minus sign. Shared alleles between the mother's and POC's STR profiles are underlined. POC - product of conception.

Fig. 3. Identification of triploidy and determination of its origin (digynic triploidy).



	D1S1656*	D4S2366	D10S1248	D4S2408	*SRY***	Y-X	D7S820	THO1	D5S818
POC	131-139-158	327-327-339	109-109-116	369-377-377	+	XXY	339-339-343	162-177-177	150-162-162
Father	131-139	327	109-113	369-377	+	XY	339-347	177	158-162
Mother	158	331-339	113-116	369	–	XX	343-352	162-174	150-154

Note. * Length of informative STR alleles is given in nucleotides; ** Presence of the *SRY* gene is indicated by a plus sign, absence by a minus sign. Shared alleles between the father's and POC's STR profiles are underlined. POC - product of conception.

Fig. 4. Identification of triploidy and determination of its origin (diandric triploidy).

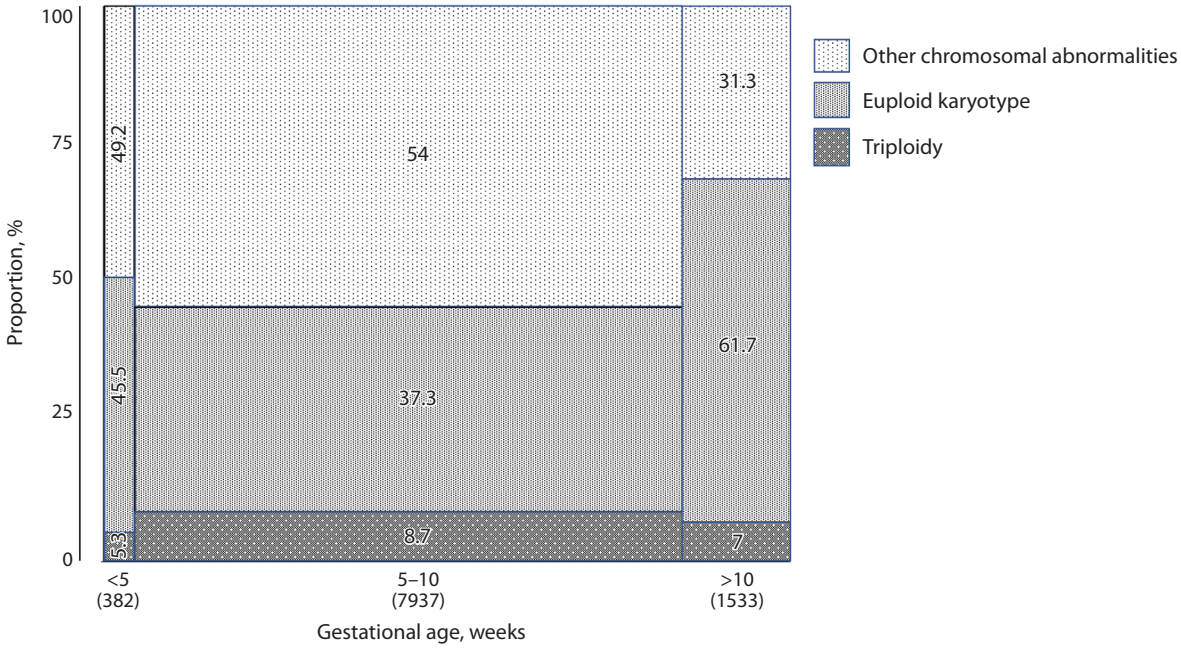


Fig. 5. Distribution of study results by gestational age at pregnancy loss.

The Y axis shows the proportion (%) and the X axis indicates gestational age at the time of pregnancy loss (in weeks). The total number of cases is shown in parentheses below each column. Gestational age was unavailable for 148 cases. The percentage labels on the bars indicate the proportion of each category among all cases at that specific gestational age. Bar width reflects the total number of cases at each gestational age. Other chromosomal abnormalities include autosomal and sex chromosome trisomies and monosomies, among others. Euploid karyotype refers to cases without chromosomal abnormalities. Triploidy indicates cases with a triploid karyotype.

median age of 27.9 years (IQR: 26.4–35.1). The median gestational age at the time of PL was 6.5 weeks (IQR: 6.5–7.5). In all 11 cases, homozygous STR profiles of the products of conception were observed, matching the paternal STR profile and differing from the maternal profile – consistent with a genome derived from a single sperm cell (Fig. 2).

The majority of PL cases (80.6 %; 95 % CI: 79.8–81.3) occurred between gestational weeks 5 and 10 (Fig. 5). PL cases occurred before 5 weeks of gestation in 3.9 % of cases (95 % CI: 3.5–4.3), and those after 10 weeks occurred in 15.6 % (95 % CI: 14.9–16.3). The highest proportion of triploid cases occurred between 5 to 10 weeks (8.7 % of all PL cases). Autosomal and sex chromosomal aneuploidies, as well as combined numerical chromosomal abnormalities, were also most common in this time window – 54 % of all PLs. Euploid POCs were more frequently observed after 10 gestational weeks (61.7 %).

Statistical significance of the frequency distribution across gestational age groups was assessed using Fisher's exact test. The comparisons yielded the following *p*-values: "<5 weeks" *vs.* "5–10 weeks": $p = 0.0012$; "<5 weeks" *vs.* ">10 weeks": $p = 8.6 \times 10^{-10}$; "5–10 weeks" *vs.* ">10 weeks": $p < 2.2 \times 10^{-16}$.

Discussion

In our cohort of 10,000 consecutive PL cases, molecular genetic analysis revealed chromosomal abnormalities in 58.8 % of samples, with triploidy accounting for 14.3 % of those with abnormal karyotypes. Similar proportions of chromosomal abnormalities and triploidy have been reported in studies analyzing chorionic villi from first-trimester miscarriages (Jenderny, 2014; Soler et al., 2017).

Approximately 80 % of PL cases occurred between gestational weeks 5 and 10 (Fig. 5). Classical clinical features of HM, such as vaginal bleeding and uterine enlargement, are rare at these early stages. This hampers the morphological differentiation between molar and non-molar tissue. It is estimated that 50 % of true PHM cases may be missed by routine histomorphology, with substantial inter- and intra-observer variability, even among experienced pathologists (Fukunaga et al., 2005; Hui et al., 2017). This may be due in part to the fact that trisomies involving chromosomes 7, 8, 13, 15, 16, 18, 21, and 22 can induce villous changes that mimic PHM (Buza, Hui, 2013; Gergely et al., 2024). These findings underscore the importance of DNA genotyping of POCs in differential diagnosis and in determining appropriate follow-up and prognosis for future pregnancies.

An increased incidence of CHM has been observed among women over 35 years of age and adolescent girls across different countries and ethnic groups (Hui et al., 2017). In our cohort, no bimodal distribution was observed; the CHM group ($n = 11$) had significantly younger maternal age than other groups. This could reflect sampling limitations or the low number of CHM cases detected.

Differentiating molar from non-molar pregnancies and distinguishing CHM from PHM is crucial for estimating the risk of post-molar gestational trophoblastic neoplasia, which varies by subtype and determines follow-up duration. CHM progresses to persistent/invasive mole in 15–20 % of cases

and to choriocarcinoma in 2–3 %, while the risks for PHM are lower (0.5–5 % and 0.015 %, respectively) (Buza, Hui, 2021; Ul'rich et al., 2024). DNA genotyping of POCs is recognized as a reliable diagnostic method for HM, with validated clinical sensitivity and specificity for both CHM and PHM (Furtado et al., 2013; Buza, Hui, 2021).

We found limited data on HM frequency among PLs in the Russian Federation. An analysis of statistical data on early reproductive losses in the Ryazan region from 2017 to 2021 revealed that "hydatidiform mole was diagnosed in exceptional cases." (Aleshkina, Konovalov, 2023). In various autonomous districts of the Tyumen region, during the period from 2016 to 2021, HM accounted for 0.11–0.17 % of pregnancies with abortive outcomes before 12 weeks of gestation (Mateykovich et al., 2023). The total number of triploid cases identified by us (n = 829), as well as the proportion of diandric triploidy (43 %), allow us to estimate the number of PHM cases at approximately 350 per 10,000 cases of PL, or 3.5 %. These findings highlight the need to revise current diagnostic approaches. Early and accurate diagnosis of HM is crucial for reducing complications and preserving fertility in young women, given the risk of progression to persistent trophoblastic disease.

A limitation of this study is the inability to identify recurrent HM, which is an autosomal recessive condition.

Conclusion

QF-PCR-based DNA genotyping of POCs reliably detects chromosomal abnormalities, including triploidy, CHM, and PHM. In our cohort of 10,000 PL cases, abnormal karyotypes were identified in 58.8 % of samples. Triploidy accounted for 8.3 % of all cases, or 14.3 % of those with abnormal karyotypes. The frequency of CHM was 0.11 %. The median maternal age in triploidy cases was 32.1 years (IQR: 28.2–35.8), while in CHM cases, it was 27.9 years (IQR: 26.4–35.1).

Given the observed frequency of both complete and partial HM in our cohort, as well as the relatively young age of the affected women, there is a pressing need to improve diagnostic protocols – particularly through the inclusion of DNA genotyping of POCs – to enable timely diagnosis and prevent post-molar malignant transformation in this age group.

References

Aleshkina O.S., Konovalov O.E. The dynamics of early reproductive losses in the Ryazan region. *Science of the Young* (*Eruditio Juvenium*). 2023;11(3):318-326. doi 10.23888/HMJ2023113318-326 (in Russian)

Buza N., Hui P. Partial hydatidiform mole: histologic parameters in correlation with DNA genotyping. *Int J Gynecol Pathol*. 2013;32(3):307-315. doi 10.1097/PGP.0b013e3182626011

Buza N., Hui P. Genotyping diagnosis of gestational trophoblastic disease: frontiers in precision medicine. *Mod Pathol*. 2021;34(9):1658-1672. doi 10.1038/s41379-021-00831-9

Candelier J.J. The hydatidiform mole. *Cell Adh Migr*. 2016;10(1-2):226-235. doi 10.1080/19336918.2015.1093275

Essers R., Lebedev I.N., Kurg A., Fonova E.A., Stevens S.J.C., Koeck R.M., von Rango U., … Paulussen A., Hoischen A., Brunner H.G., Salumets A., Zamani Esteki M. Prevalence of chromosomal alterations in first-trimester spontaneous pregnancy loss. *Nat Med*. 2023;29(12):3233-3242. doi 10.1038/s41591-023-02645-5

Fukunaga M., Katabuchi H., Nagasaka T., Mikami Y., Minamiguchi S., Lage J.M. Interobserver and intraobserver variability in the diagnosis of hydatidiform mole. *Am J Surg Pathol*. 2005;29(7):942-947. doi 10.1097/01.pas.0000157996.23059.c1

Furtado L.V., Paxton C.N., Jama M.A., Tripp S.R., Wilson A.R., Lyon E., Jarboe E.A., Thaker H.M., Geiersbach K.B. Diagnostic utility of microsatellite genotyping for molar pregnancy testing. *Arch Pathol Lab Med*. 2013;137(1):55-63. doi 10.5858/arpa.2012-0047-OA

Gergely L., Korbeľ M., Danihel Ľ., Repiská V., Tomka M., McCullough L., Priščáková P. Trisomy 16 mimicking hydatidiform mole. *Ceska Gynekol*. 2024;89(5):396-399. doi 10.48095/cccg2024396

Golubovsky M.D. Postzygotic diploidization of triploids as a source of unusual cases of mosaicism, chimerism and twinning. *Hum Reprod*. 2003;18(2):236-242. doi 10.1093/humrep/deg060

Hui P., Buza N., Murphy K.M., Ronnett B.M. Hydatidiform moles: genetic basis and precision diagnosis. *Annu Rev Pathol*. 2017;12:449-485. doi 10.1146/annurev-pathol-052016-100237

Jenderny J. Chromosome aberrations in a large series of spontaneous miscarriages in the German population and review of the literature. *Mol Cytogenet*. 2014;5(7):38. doi 10.1186/1755-8166-7-38

Joyce C.M., Fitzgerald B., McCarthy T.V., Coulter J., O'Donoghue K. Advances in the diagnosis and early management of gestational trophoblastic disease. *BMJ Med*. 2022;1(1):e000321. doi 10.1136/bmjmed-2022-000321

Massalska D., Ozdarska K., Roszkowski T., Bijok J., Kucińska-Chahwan A., Panek G.M., Zimowski J.G. Distribution of diandric and digynic triploidy depending on gestational age. *J Assist Reprod Genet*. 2021;38(9):2391-2395. doi 10.1007/s10815-021-02202-4

Mateykovich E.A., Novikova V.A., Radzinsky V.E. Intraterritorial differences in reproductive losses. *Meditsinskiy Sovet = Medical Council.* 2023;17(13):191-199. doi 10.21518/ms2023-252 (in Russian)

Murdoch S., Djuric U., Mazhar B., Seoud M., Khan R., Kuick R., Bagga R., Kircheisen R., Ao A., Ratti B., Hanash S., Rouleau G.A., Slim R. Mutations in *NALP7* cause recurrent hydatidiform moles and reproductive wastage in humans. *Nat Genet*. 2006;38(3):300-302. doi 10.1038/ng1740

Nguyen N.M.P., Ge Z.J., Reddy R., Fahiminiya S., Sauthier P., Bagga R., Sahin F.I., … Sahoo T., Ao A., Majewski J., Taketo T., Slim R. Causative mutations and mechanism of androgenetic hydatidiform moles. *Am J Hum Genet*. 2018;103(5):740-751. doi 10.1016/j.ajhg.2018.10.007

Parry D.A., Logan C.V., Hayward B.E., Shires M., Landolsi H., Diggle C., Carr I., … Malik S., Taylor G.R., Johnson C.A., Bonthron D.T., Sheridan E.G. Mutations causing familial biparental hydatidiform mole implicate *C6orf221* as a possible regulator of genomic imprinting in the human oocyte. *Am J Hum Genet*. 2011;89(3):451-458. doi 10.1016/j.ajhg.2011.08.002

Pesik V.Y., Fedunin A.A., Agdzhoyan A.T., Utevska O.M., Chukhraeva M.I., Evseeva I.V., Churnosov M.I., Lependina I.N., Bogunov Y.V., Bogunova A.A., Ignashkin M.A., Yankovsky N.K., Balanovska E.V., Orekhov V.A., Balanovsky O.P. Analysis of genetic diversity of Russian regional populations based on STR markers used in DNA identification. *Russian Journal of Genetics*. 2014;50(6):626-633. doi 10.1134/S1022795414060088

Ronnett B.M. Hydatidiform moles: ancillary techniques to refine diagnosis. *Arch Pathol Lab Med*. 2018;142(12):1485-1502. doi 10.5858/arpa.2018-0226-RA

Sazhenova E.A., Filippova M.O., Lebedev I.N. Epigenetic perspectives of hydatidiform mole: mechanisms of genomic imprinting disorders and issues of molecular genetic diagnostics. *Medical Genetics*. 2009;8(3):3-12 (in Russian)

Smolyanitsky A.G., Ivanov P.L., Kornienko I.V., Zamaraev V.S., Perepechina I.O., Komarovsky Yu.A., Pushkarev V.P., Khromov-Borisov N.N. Towards Russian reference population data on STR loci. *Int Congr Ser*. 2004;1261:242-244. doi 10.1016/S0531-5131(03)01610-8

Soler A., Morales C., Mademont-Soler I., Margarit E., Borrell A., Borobio V., Muñoz M., Sánchez A. Overview of chromosome abnormalities in first trimester miscarriages: a series of 1,011 consecutive chorionic villi sample karyotypes. *Cytogenet Genome Res*. 2017;152(2):81-89. doi 10.1159/000477707

Ul'rikh E.A., Rumyantsev A.A., Teletaeva G.M., Khokhlova S.V., Urmancheeva A.F., Tyulyandina A.S. Malignant trophoblastic tumors. *Malignant Tumors*. 2024;14(3s2-2):189-206. doi 10.18027/2224-5057-2024-14-3s2-1.2-07 (in Russian)

Zavarin V., Ilina V., Krassotkin Y., Makarova T., Sutiagina D., Semikhodskii A. Evaluation of sensitivity and specificity of sibship determination in the Caucasian population of the Russian Federation using the 23 STR loci VeriFiler panel. *Forensic Sci Int Genet Suppl Ser*. 2019;7(1):56-58. doi 10.1016/j.fsigss.2019.09.023

Conflict of interest. The authors declare no conflict of interest.
Received March 5, 2025. Revised May 12, 2025. Accepted May 20, 2025.

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2025;29(5):629-635

doi 10.18699/vjgb-25-68

A rare case of uniparental disomy 9 concomitant with low-level mosaicism for trisomy 9

A.S. Iakovleva, Zh.G. Markova, L.A. Bessonova, N.V. Shilova

Bochkov Research Centre for Medical Genetics, Moscow, Russia
shenkayter@icloud.com

Abstract. Uniparental disomy of chromosome 9, in combination with low-level mosaicism for chromosome 9, represents a rare chromosomal disorder. One of the mechanisms underlying the formation of uniparental disomy is the trisomy rescue, which concurrently results in low-level mosaicism. The diagnosis of mosaic aneuploidies poses significant challenges due to the limited sensitivity and resolution of conventional cytogenetic methods, which often fail to detect low-level mosaicism. Additionally, the variable distribution of cell lines within the patient's tissues, as well as the heterogeneity of samples derived from the same tissue, complicates the precise determination of the impact of mosaic trisomy on the phenotypic expression. Phenotypic manifestations associated with mosaic trisomy 9 are characterized by considerable variability. During the prenatal period, intrauterine growth restriction is frequently observed in cases of this chromosomal abnormality, although this finding is not pathognomonic for the condition. In liveborn infants with trisomy 9 mosaicism, characteristic phenotypic features may include craniofacial anomalies (such as micrognathia and ear malformations), scoliosis, low-set ears, feeding and respiratory difficulties, hip dysplasia, seizures, and developmental delays. To establish a diagnosis in a patient presenting with multiple dysembryogenic stigmata and psychomotor retardation, a comprehensive molecular cytogenetic analysis was conducted. This included high-resolution chromosomal microarray analysis (CMA) and fluorescence *in situ* hybridization (FISH) using targeted DNA probes. CMA identified regions of loss of heterozygosity (LOH) on chromosome 9, indicative of uniparental disomy, and suggested the presence of low-level mosaicism for trisomy 9. Subsequent FISH analysis of cultured lymphocytes, employing DNA probes specific to various regions of chromosome 9, confirmed the low-level mosaicism for trisomy 9. The results of our study are consistent with the idea that mosaicism for chromosome 9, particularly when combined with uniparental disomy, constitutes a complex genetic anomaly that can lead to a spectrum of phenotypic manifestations, including developmental delay, growth abnormalities, and behavioral anomalies. CMA and FISH are highly effective methods for the diagnosis of uniparental disomy and low-level mosaicism involving chromosome 9.

Key words: mosaicism; trisomy 9; uniparental disomy 9; chromosomal microarray analysis (CMA); FISH

For citation: Iakovleva A.S., Markova Zh.G., Bessonova L.A., Shilova N.V. A rare case of uniparental disomy 9 concomitant with low-level mosaicism for trisomy 9. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov J Genet Breed*. 2025;29(5):629-635. doi 10.18699/vjgb-25-68

Редкий случай однородительской дисомии хромосомы 9 в сочетании с мозаицизмом по трисомии 9

А.С. Яковлева, Ж.Г. Маркова, Л.А. Бессонова, Н.В. Шилова

Медико-генетический научный центр им. академика Н.П. Бочкова, Москва, Россия
shenkayter@icloud.com

Аннотация. Однородительская дисомия хромосомы 9 в сочетании с мозаицизмом низкого уровня хромосомы 9 – редкое хромосомное нарушение. Один из механизмов формирования однородительской дисомии – коррекция анеуплоидии, приводящая также к мозаицизму низкого уровня. Трудности диагностики мозаичных анеуплоидий связаны с ограничениями в чувствительности и разрешающей способности стандартных цитогенетических методов, затрудняющих выявление мозаицизма низкого уровня. Различное соотношение клеточных линий в тканях пациента или неоднородность образцов одной и той же ткани не позволяют однозначно определить влияние мозаичной трисомии на формирование фенотипа пациента. Фенотипические признаки мозаичной трисомии 9 (Т9) отличаются широкой вариабельностью. В пренатальном периоде при этой хромосомной патологии часто отмечается внутриутробная задержка развития. У живорожденных с мозаицизмом по Т9 могут наблюдаться характерные фенотипические особенности, такие как черепно-лицевые аномалии (микрогнатия, аномалии ушей), сколиоз, низко посаженные уши, дисплазия тазобедренного сустава, судороги и задержка развития, а также проблемы с кормлением и дыханием. С целью установления диагноза у пациента с множествен-

© Iakovleva A.S., Markova Zh.G., Bessonova L.A., Shilova N.V., 2025
This work is licensed under a Creative Commons Attribution 4.0 License

ными стигмами дисэмбриогенеза и задержкой психомоторного развития проведено комплексное исследование молекулярно-цитогенетическими методами, включающее хромосомный микроматричный анализ (ХМА) высокого разрешения и флуоресцентную гибридизацию *in situ* (FISH) с таргетными ДНК-зондами. ХМА показал наличие участков потери гетерозиготности на хромосоме 9, свидетельствующее об однородительской дисомии, и позволил предположить мозаицизм низкого уровня по Т9. Дополнительный FISH-анализ культивированных лимфоцитов с ДНК-зондами на различные районы хромосомы 9 позволил установить мозаицизм низкого уровня по Т9. Результаты нашего исследования согласуются с представлением о том, что мозаицизм по хромосоме 9 в сочетании с однородительской дисомией является сложной генетической аномалией, которая может привести к задержке развития, нарушению роста и особенностям поведения. ХМА и FISH – это эффективные методы диагностики однородительской дисомии и мозаицизма низкого уровня по трисомии хромосомы 9.
Ключевые слова: мозаицизм; трисомия хромосомы 9; однородительская дисомия хромосомы 9; хромосомный микроматричный анализ (ХМА); FISH

Introduction

Uniparental disomy (UPD) is a genetic anomaly in which both homologs of a chromosome or chromosomal segment are inherited exclusively from one parent in contrast to normal meiotic segregation, wherein a child inherits one homolog from the father and one from the mother. The term UPD was first introduced in 1980 (Engel, 1980) based on the observation that a partial diploid DNA sequence, and occasionally an entire pair of chromosomes may be inherited solely from one parent. The clinical significance of this phenomenon was not recognized until the 1990s when UPD was established as one of the possible causes of two neurodevelopmental syndromes – Prader–Willi and Angelman syndromes (Cassidy, Schwartz, 1998). UPD is subdivided into isodisomy where two identical copies of a chromosome are inherited from one parent, and heterodisomy where a pair of non-identical homologous chromosomes is inherited from one parent (Chen Q. et al., 2023). The frequency of UPD varies considerably among chromosomes (Eggermann et al., 2015). The highest UPD frequencies have been reported for acrocentric chromosomes 13, 14, 15, 21, and 22 due to their involvement in Robertsonian translocations – chromosomal rearrangements resulting from the fusion of the long arms and centromeres of two acrocentric chromosomes with the concomitant loss of material from their short arms.

For most chromosomes, UPD has no clinical consequences. However, chromosomes 6, 7, 8, 11, 13, 14, 15, and 20 contain genes subject to parent-of-origin-specific expression (imprinting). UPD involving these chromosomes may lead to the development of corresponding imprinting disorders as well as autosomal recessive diseases or X-linked recessive disorders in females. Rare inheritance of both sex chromosomes from a single parent may underlie transmission of X-linked disorders from father to son (Del Gaudio et al., 2020).

The clinical manifestations of UPD may range from varying degrees of intellectual disability and/or syndromes involving multiple congenital anomalies to an asymptomatic presentation. However, UPD involving autosomes is most frequently associated with intrauterine growth restriction, dysmorphic features, or multiple congenital malformations (Kotzot, 2002).

The estimated prevalence of UPD is approximately one case per 5,000 live births (Liehr, 2010). An analysis of data from more than 4 million individuals tested by the private genetics company 23andMe, and 431,094 participants from the Northern European UK Biobank, demonstrated that UPD affecting any chromosome (not only those harboring imprinted regions) occurs at a frequency of one in 2,000 births. Given that the 23andMe dataset primarily includes healthy individuals from the general population, this figure is considered a more representative estimate of the population-level frequency of UPD. Data from 23andMe further suggest that UPD of chromosomes lacking imprinted genes or genes associated with autosomal recessive disorders is often not associated with a pathological phenotype (Del Gaudio et al., 2020).

The most common mechanism leading to UPD is chromosomal nondisjunction during meiosis or mitosis. The principal mechanisms of UPD formation include monosomy rescue, trisomy rescue, mitotic error, and gamete complementation (Nakka et al., 2019). Aneuploidy correction occurs either through loss of the third chromosome (trisomy rescue) or duplication of a monosomic chromosome (monosomy rescue). Trisomy rescue may occur as a result of anaphase lag and can contribute to UPD formation.

Trisomy 9 (T9) is a rare chromosomal abnormality that can occur in either mosaic or non-mosaic forms (Cantú et al., 1996). The regular (non-mosaic) form of T9 is incompatible with live birth and is identified in 2.2–2.7 % of spontaneous abortions occurring in the first trimester of pregnancy (Benn, Grati, 2021). Nevertheless, postnatally diagnosed patients with mosaic T9 have been reported (Bruns, Campbell, 2015). In most individuals with mosaic T9, prenatal findings include intrauterine growth restriction or low fetal weight, oligohydramnios, placental insufficiency, premature rupture of membranes, and skeletal anomalies (Bruns, Campbell, 2015).

Postnatally mosaic T9 is typically characterized by multiorgan involvement, including craniofacial anomalies, malformations of the heart, genitourinary system, skeleton, and central nervous system, as well as abnormal ear mor-

phology, micrognathia, and hip dysplasia. Most reported patients also experience prenatal and perinatal complications related to respiration, growth, and feeding (Li M. et al., 2021). The severity and frequency of developmental anomalies and intellectual disability correlate with the proportion of trisomic cells in various tissues (Lee et al., 2018).

Herein we present a rare case of chromosome 9 loss of heterozygosity in combination with low-level mosaicism for T9 identified in a patient with psychomotor developmental delay and congenital anomalies, using a molecular cytogenetic approach.

Clinical description of the patient

The proband is a girl, 3 years and 7 months old at the time of examination, presenting with psychomotor developmental delay and feeding difficulties (does not chew solid food, experiences choking episodes). Parental ages at the time of birth were 37 years (mother) and 29 years (father). The child was born from the third pregnancy (I – miscarriage, II – maternal half-sister, 13 years old). Intrauterine growth restriction was diagnosed at 30 weeks of gestation. She was born with a birth weight of 2,657 g (10th percentile), length of 50 cm (10th percentile), head circumference of 32 cm (3rd percentile), and chest circumference of 31 cm (3rd percentile). Multiple minor anomalies (dysmorphic stigmata) of embryonic development were noted. Phenotype at the time of examination included: positional cranial deformation, enophthalmos, microtia, dysplastic auricular morphology, bilateral preauricular fistulas, congenital ptosis of the right upper eyelid, muscle hypotonia, bilateral mixed conductive hearing loss, and congenital dislocation of the right hip. Height: 89 cm (50th percentile), weight: 9.8 kg (<3rd percentile), head circumference: 45 cm (<3rd percentile).

Early psychomotor development: did not hold her head up, rolled from back to side by age one, does not sit independently; vocalizations present, reaches for objects and grasps them, makes spontaneous leg movements.

Karyotype: 46,XX – normal female.

Family history is negative for hereditary disorders.

Clinical exome sequencing (covering 6,640 genes), previously performed, revealed only variants of uncertain clinical significance (VUS).

Materials and methods

Genomic DNA was extracted from a peripheral venous blood sample collected in EDTA using the Gentra Puregene Blood Kit Plus (Qiagen, California) according to the manufacturer's protocol.

Chromosomal microarray analysis (CMA) was performed using the high-density CytoScan® HD Array Kit in accordance with the manufacturer's instructions (Affymetrix Inc., California, USA). Data were processed, analyzed, and normalized using Affymetrix Chromosome Analysis Suite (ChAS) version 4.0 with reference genome build NA33.1 (hg19).

Fluorescence *in situ* hybridization (FISH) was carried out according to the manufacturer's protocols on chromosome preparations obtained from 72-hour peripheral blood lymphocyte cultures. DNA probes targeting chromosome 9 were used: pericentromeric heterochromatin of chromosome 9 (SE 9 classical), and subtelomeric regions of the short and long arms of chromosome 9 (Sub Telomere 9pter, Sub Telomere 9qter) (KREATECH, Netherlands). The analysis was performed using an AxioImager M.1 epifluorescence microscope (Carl Zeiss) and the Isis digital image analysis software (MetaSystems).

Results

Chromosomal microarray analysis using high-density arrays with genotype-informative SNP probes revealed extensive regions of loss of heterozygosity (LOH) on chromosome 9. In parallel, the smooth signal indicated a slight shift toward increased copy number along the entire chromosome 9, suggesting the presence of low-level mosaicism for T9. The estimated level of T9 mosaicism ranged from 22 to 26 % in specific regions of the long arm of chromosome 9 (Fig. 1).

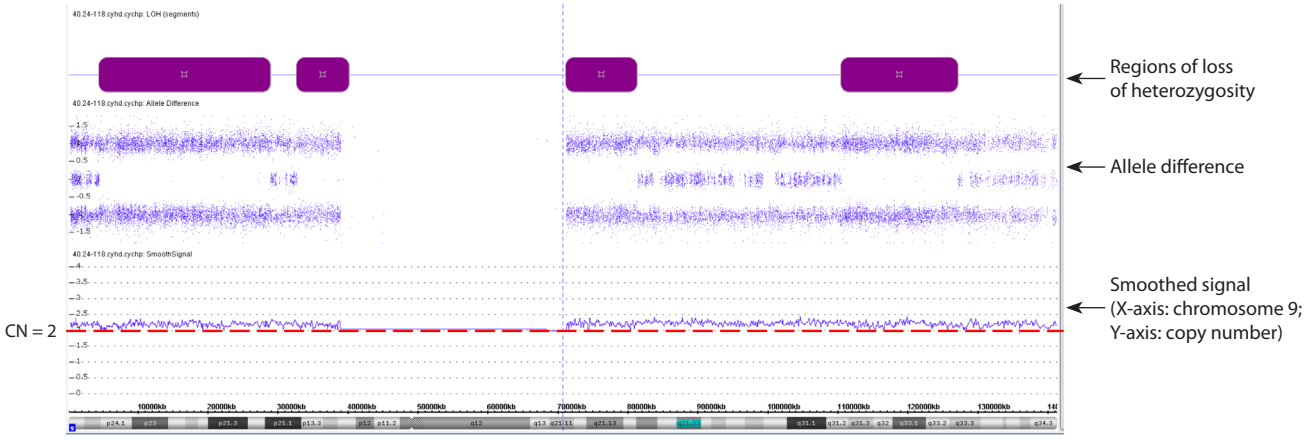


Fig. 1. CMA results.

Molecular karyotype of the patient (according to ISCN 2020): arr[GRCh37] 9p24.2p21.1(4284369_28756447)×2 hmz, 9p21.1p13.1(32561829_40087758)×2 hmz, 9q21.11q21.31(71013800_81233686)×2 hmz, 9q31.2q33.3(110291122_126976363)×2 hmz.

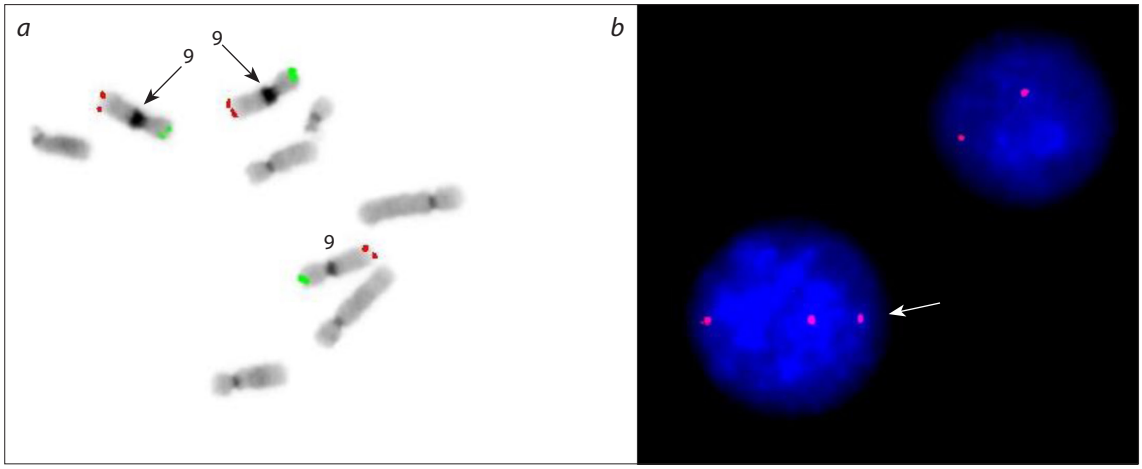


Fig. 2. FISH analysis with DNA probes for chromosome 9.

a – a metaphase spread hybridized with DNA probes targeting the subtelomeric regions of the short (green signal) and long (red signal) arms of chromosome 9 (inverted DAPI staining). Arrows indicate the chromosome 9 homologs each exhibiting identical pericentromeric heterochromatin block sizes; *b* – interphase nucleus hybridized with a DNA probe for the pericentromeric heterochromatin of the long arm of chromosome 9. The arrow indicates a nucleus with three hybridization signals corresponding to three copies of the 9q11.1 region.

Additional FISH analysis using DNA probes targeting the subtelomeric regions of the short and long arms of chromosome 9 confirmed the presence of a clone with trisomy involving the entire chromosome 9 (Fig. 2*a*). T9 was detected in six out of 100 metaphase spreads analyzed: ish 9(RH65569+, SHGC-149365+)×3[6/100]/9(RH65569+, SHGC-149365+)×2[94/100].

To assess the level of mosaicism, interphase FISH was performed using a DNA probe specific to the pericentromeric heterochromatin of the long arm of chromosome 9. Among 300 interphase nuclei analyzed, mosaicism with two distinct cell clones was identified: 6.7 % of cells exhibited three copies of the D9Z5 locus, and 93.3 % of cells contained two copies (Fig. 2*b*). FISH result: nuc ish(D9Z5×3)[20/300] (ISCN 2020).

Analysis of metaphase spreads revealed a polymorphism in the pericentromeric heterochromatin region of chromosome 9. In disomic cells, both homologs exhibited heterochromatin blocks of identical size. This heterochromatin block polymorphism supports the presence of uniparental disomy (UPD) for chromosome 9 (Liehr, 2010). Thus, molecular cytogenetic analysis confirmed low-level mosaicism for T9. It can be hypothesized that following the correction of T9 in an initially aneuploid embryo, two cell clones emerged during ontogenesis – one trisomic and the other disomic with UPD of chromosome 9. Parental material was unavailable for analysis, and therefore the parental origin of the UPD could not be determined.

Discussion

Chromosomal microarray analysis revealed four extended regions of homozygosity on chromosome 9, with a cumulative length of approximately 59 million base pairs, corresponding to 42 % of the total length of the chromosome. Based on the established theory that UPD can result from trisomy rescue, the observed LOH profile is more consistent with complete UPD of chromosome 9 than with segmental UPD. Since SNP array analysis is informative only for isodisomy, a mixed iso-/heterodisomy of chromosome 9 is likely (Fig. 3).

According to published molecular studies, extended regions of homozygosity on chromosome 9, detected by SNP arrays, most likely result from postzygotic trisomy rescue combined with mitotic recombination (Ma N. et al., 2023). This mechanism of UPD formation implies that low-level or cryptic mosaicism is likely present in many cases, particularly for chromosome 9 (Eggermann et al., 2018).

Trisomy 9 is most commonly detected during prenatal cytogenetic testing in extraembryonic tissues. When only a few abnormal cells are found in amniotic fluid cultures, regular T9 is frequently observed in chorionic villi, indicating confined placental mosaicism (CPM) (Ma N. et al., 2023). Chromosome 9 UPD is a rare genetic anomaly and is generally identified during prenatal diagnosis of mosaic trisomy 9. The first case demonstrating that trisomy rescue can result in a fetus with UPD of chromosome 9 was reported as early as 1992 (Willatt et al., 1992).

A particular diagnostic challenge in CPM involving trisomy is the risk of associated UPD, especially when the aneuploid chromosome harbors imprinted genes (Kotzot, 2002). Postzygotic correction of trisomy may occur in embryonic tissues, while the placenta remains partially or fully trisomic. The frequency of UPD in the setting of CPM is approximately 2 % (Malvestiti et al., 2015).

Current knowledge regarding the pathogenic mechanisms of UPD remains limited. UPD can affect phenotype through various pathways. The abnormal phenotype may result from imprinting, mosaicism for T9, or recessive mutations. In cases of isodisomy, the risk of monogenic disorders increases due to homozygosity for recessive alleles (Spence et al., 1988; Quan et al., 1997). For example,

UPD of chromosome 9 has been associated with Leigh syndrome (Tiranti et al., 1999; Xiao et al., 2019), cartilage-hair hypoplasia (Sulisalo et al., 1997), and amyotrophic lateral sclerosis (Yang et al., 2014). Whole-exome sequencing in our patient did not reveal pathogenic variants on chromosome 9, suggesting the absence of autosomal recessive disorders.

Another pathogenic effect of UPD may involve genomic imprinting which refers to parent-of-origin-specific gene expression. However, it has been shown that UPD of chromosome 9 does not result in an abnormal phenotype (Björck et al., 1999). Chromosome 9 harbors the imprinted gene *GLIS3*, which is expressed from the paternal allele, but this gene has not been linked to the clinical features observed in our patient.

Literature analysis shows that UPD of chromosome 9, whether maternal or paternal in origin, in the absence of trisomy, is not associated with developmental abnormalities or phenotypic deviations typically observed in T9 cases (see the Table).

Patients with mosaic T9 exhibit a broad spectrum of clinical manifestations affecting multiple organ systems. Craniofacial dysmorphism, cardiac anomalies, genitourinary defects, skeletal and central nervous system abnormalities are most frequently observed. Global developmental delay is also commonly reported (Li M. et al., 2021). The mild phenotypic features observed in our patient are most likely attributable to the low proportion of T9-positive cells. Birth weight of 2,657 g (10th percentile), length of 50 cm (10th percentile), head circumference of 32 cm (3rd percentile), and chest circumference of 31 cm (3rd percentile) indicate intrauterine growth restriction, likely associated with placental dysfunction potentially caused by confined placental mosaicism for T9. Since no prenatal cytogenetic testing was conducted in this case, the presence of mosaicism for T9 in both extraembryonic and embryonic tissues can only be presumed. Nevertheless, several reported cases indicate favorable pregnancy outcomes in instances of UPD of chromosome 9 combined with low-level mosaic T9 diagnosed prenatally (Chen Q. et al., 2023).

FISH-based assessment of mosaicism significantly improves the sensitivity for detecting low-level mosaicism. Unfortunately, in our case, peripheral blood was the only available tissue. It is possible that FISH analysis of additional tissues – derived from different germ layers – could provide insights into the distribution of abnormal cells and their tissue-specific phenotypic effects.

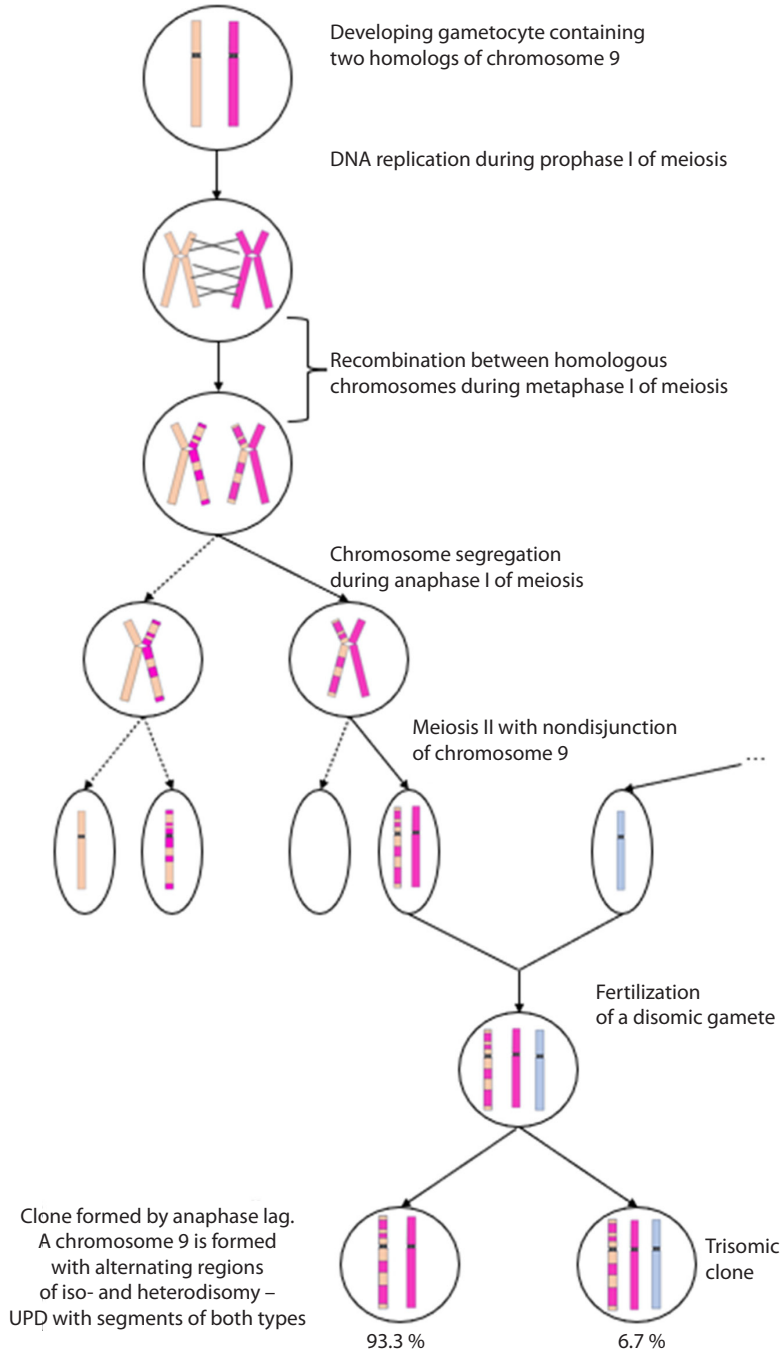


Fig. 3. Schematic representation of the formation of mixed uniparental iso-/heterodisomy of chromosome 9 with T9 mosaicism.

Conclusion

UPD of chromosome 9 is a rare genetic anomaly, it may have clinical significance due to imprinting defects and the manifestation of autosomal recessive disorders.

Published data on the correlation between clinical features, karyotype, and T9

No.	Age	UPD mat/pat	Karyotype	Sample	Clinical features	Reference
1	34 years	mat	46,XX,i(9)(p10),i(9)(q10)	Blood	Recurrent miscarriages	Björck et al., 1999
2	Fetus	mat	47,XX,+9[4]/46,XX[35]	Amniotic fluid	No abnormalities	Chen C.P. et al., 2022
			46,XX	Blood		
3	Fetus	mat	47,XX,+9[2]/46,XX[23]	Amniotic fluid	No abnormalities	Chen C.P. et al., 2023
			46,XX	Blood		
4	Fetus	mat	46,XY	Chorion	IUGR, spontaneous abortion at 12 weeks	Fritz et al., 2001
5	Fetus	pat	47,XXY[19]/46,XY[81]	Amniotic fluid	Normal at 3.5 years of age	Li D. et al., 2019
			47,XXY[17]/46,XY[83]	Cord blood		
6	Fetus	pat	47,XX,+9[30]/46,XX[70]	Maternal blood	IUGR, CP, global developmental delay, failure to thrive	Ma J. et al., 2015
	2 years		47,XX,+9[17]/46,XX[83]	Buccal epithelium, blood		

Note. IUGR – intrauterine growth restriction; CP – cerebral palsy.

This clinical case describing a patient with UPD of chromosome 9 combined with low-level mosaicism for trisomy 9 (T9) diagnosed using chromosomal microarray analysis and fluorescence *in situ* hybridization demonstrates the effectiveness of combining modern molecular genetic techniques.

The coexistence of UPD and T9 mosaicism may lead to variable phenotypic manifestations depending on the degree of mosaicism and the distribution of the trisomic clone across different tissues (Ma N. et al., 2023).

To establish a correlation between UPD of chromosome 9 and an abnormal phenotype, it is necessary to analyze additional tissue samples accessible for examination (e. g., skin fibroblasts, buccal epithelium, or urinary sediment cells) in order to assess the extent and tissue distribution of mosaicism. The phenomenon of UPD requires further investigation with a focus on the identification of specific genetic abnormalities and mosaicism patterns, which may allow for more accurate prognostic assessment.

References

Benn P., Grati F.R. Aneuploidy in first trimester chorionic villi and spontaneous abortions: windows into the origin and fate of aneuploidy through embryonic and fetal development. *Prenat Diagn*. 2021;41(5):519-524. doi 10.1002/pd.5795

Björck E.J., Anderlid B.M., Blennow E. Maternal isodisomy of chromosome 9 with no impact on the phenotype in a woman with two isochromosomes: i(9p) and i(9q). *Am J Med Genet.* 1999;87(1):49-52. doi 10.1002/(sici)1096-8628(19991105)87:1<49::aid-ajmg10>3.0.co;2-4

Bruns D.A., Campbell E. Twenty-five additional cases of trisomy 9 mosaic: birth information, medical conditions, and developmental status. *Am J Med Genet A*. 2015;167A(5):997-1007. doi 10.1002/ajmg.a.36977

Cantú E.S., Eicher D.J., Pai G.S., Donahue C.J., Harley R.A. Mosaic vs. nonmosaic trisomy 9: report of a liveborn infant evaluated by fluorescence in situ hybridization and review of the literature. *Am J Med Genet*. 1996;62(4):330-335. doi 10.1002/(SICI)1096-8628(19960424)62:4<330::AID-AJMG1>3.0.CO;2-V

Cassidy S.B., Schwartz S. Prader–Willi and Angelman syndromes. Disorders of genomic imprinting. *Medicine* (*Baltimore*). 1998;77(2):140-151. doi 10.1097/00005792-199803000-00005

Chen C.P., Chern S.R., Wu P.S., Chen S.W., Wu F.T., Chen L.F., Chen Y.Y., Wang W. Detection of maternal uniparental disomy 9 in association with low-level mosaic trisomy 9 at amniocentesis in a pregnancy associated with intrauterine growth restriction, abnormal first-trimester screening result (low PAPP-A and low PlGF), maternal preeclampsia and a favorable outcome. *Taiwan J Obstet Gynecol*. 2022;61(1):141-145. doi 10.1016/j.tjog.2021.11.024

Chen C.P., Ko T.M., Chen S.W., Chern S.R., Wu F.T., Pan Y.T., Pan C.W., Chen Y.Y., Wang W. Low-level mosaic trisomy 9 at amniocentesis associated with a positive non-invasive prenatal testing for trisomy 9, maternal uniparental disomy 9, intrauterine growth restriction and a favorable fetal outcome in a pregnancy. *Taiwan J Obstet Gynecol*. 2023;62(3):457-460. doi 10.1016/j.tjog.2023.03.008

Chen Q., Chen Y., Shi L., Tao Y., Li X., Zhu X., Yang Y., Xu W. Uniparental disomy: expanding the clinical and molecular phenotypes of whole chromosomes. *Front Genet*. 2023;14:1232059. doi 10.3389/fgene.2023.1232059

Del Gaudio D., Shinawi M., Astbury C., Tayeh M.K., Deak K.L., Raca G.; ACMG Laboratory Quality Assurance Committee. Diagnostic testing for uniparental disomy: a points to consider statement from the American College of Medical Genetics and Genomics (ACMG). *Genet Med*. 2020;22(7):1133-1141. doi 10.1038/s41436-020-0782-9

Eggermann T., Soellner L., Buiting K., Kotzot D. Mosaicism and uniparental disomy in prenatal diagnosis. *Trends Mol Med*. 2015;21(2):77-87. doi 10.1016/j.molmed.2014.11.010

Eggermann T., Mackay D.J.G., Tümer Z. Uniparental disomy and imprinting disorders. *OBM Genetics*. 2018;2(3):031. doi 10.21926/obm.genet.1803031

Engel E. Un nouveau concept genetique: la disomie uniparentale et son corollaire eventuel, l'isodisomie [A new genetic concept: the uniparental disomy and its potential effect, the isodisomy (author's transl.)]. *J Genet Hum*. 1980;28(1):11-22 (in French)

Fritz B., Aslan M., Kalscheuer V., Ramsing M., Saar K., Fuchs B., Rehder H. Low incidence of UPD in spontaneous abortions beyond the

5th gestational week. *Eur J Hum Genet*. 2001;9(12):910-916. doi 10.1038/sj.ejhg.5200741

Kotzot D. Review and meta-analysis of systematic searches for uniparental disomy (UPD) other than UPD 15. *Am J Med Genet*. 2002;111(4):366-375. doi 10.1002/ajmg.10569

Lee C.Y., Su H.J., Cheng Y.T., Ku Y.L., Ngo Y.G., Chen C.M., Ou Y.C., Lee M.C., Shaw S.W.S. Detection of fetal trisomy 9 mosaicism by noninvasive prenatal testing through maternal plasma DNA sequencing. *Taiwan J Obstet Gynecol*. 2018;57(4):594-597. doi 10.1016/j.tjog.2018.06.021

Li D., Wang Y., Zhao N., Chang L., Liu P., Tian C., Qiao J. A case report and mechanism analysis of a normal phenotype mosaic 47, XXY complicated by paternal iUPD (9) who had a normal PGD result. *BMC Med Genet*. 2019;20(1):172. doi 10.1186/s12881-019-0897-5

Li M., Glass J., Du X., Dubbs H., Harr M.H., Falk M., Smolarek T., Hopkin R.J., Zackai E., Sheppard S.E. Trisomy 9 mosaic syndrome: sixteen additional patients with new and/or less commonly reported features, literature review, and suggested clinical guidelines. *Am J Med Genet A*. 2021;185(8):2374-2383. doi 10.1002/ajmg.a.62251

Liehr T. Cytogenetic contribution to uniparental disomy (UPD). *Mol Cytogenet*. 2010;3:8. doi 10.1186/1755-8166-3-8

Ma J., Cram D.S., Zhang J., Shang L., Yang H., Pan H. Birth of a child with trisomy 9 mosaicism syndrome associated with paternal isodisomy 9: case of a positive noninvasive prenatal test result unconfirmed by invasive prenatal diagnosis. *Mol Cytogenet*. 2015;8:44. doi 10.1186/s13039-015-0145-4

Ma N., Zhu Z., Hu J., Pang J., Yang S., Liu J., Chen J., Tang W., Kuang H., Hu R., Li Z., Wang H., Peng Y., Xi H. Case report: detection of fetal trisomy 9 mosaicism by multiple genetic testing methods: report of two cases. *Front Genet*. 2023;14:1121121. doi 10.3389/fgene.2023.1121121

Malvestiti F., Agrati C., Grimi B., Pompilii E., Izzi C., Martinoni L., Gaetani E., Liuti M.R., Trotta A., Maggi F., Simoni G., Grati F.R. Interpreting mosaicism in chorionic villi: results of a monocentric series of 1001 mosaics in chorionic villi with follow-up amniocentesis. *Prenat Diagn*. 2015;35(11):1117-1127. doi 10.1002/pd.4656

Nakka P., Pattillo Smith S., O'Donnell-Luria A.H., McManus K.F.; 23andMe Research Team; Mountain J.L., Ramachandran S., Sathirapongsasuti J.F. Characterization of prevalence and health consequences of uniparental disomy in four million individuals from the general population. *Am J Hum Genet*. 2019;105(5):921-932. doi 10.1016/j.ajhg.2019.09.016

Quan F., Janas J., Toth-Fejel S., Johnson D.B., Wolford J.K., Popovich B.W. Uniparental disomy of the entire X chromosome in a female with Duchenne muscular dystrophy. *Am J Hum Genet*. 1997;60(1):160-165

Spence J.E., Perciaccante R.G., Greig G.M., Willard H.F., Ledbetter D.H., Hejtmancik J.F., Pollack M.S., O'Brien W.E., Beaudet A.L. Uniparental disomy as a mechanism for human genetic disease. *Am J Hum Genet*. 1988;42(2):217-226

Sulisalo T., Mäkitie O., Sistonen P., Ridanpää M., el-Rifai W., Ruuskanen O., de la Chapelle A., Kaitila I. Uniparental disomy in cartilage-hair hypoplasia. *Eur J Hum Genet*. 1997;5(1):35-42. doi 10.1007/BF03405875

Tiranti V., Lamantea E., Uziel G., Zeviani M., Gasparini P., Marzella R., Rocchi M., Fried M. Leigh syndrome transmitted by uniparental disomy of chromosome 9. *J Med Genet*. 1999;36(12):927-928

Willatt L.R., Davison B.C., Goudie D., Alexander J., Dyson H.M., Jenks P.E., Ferguson-Smith M.E. A male with trisomy 9 mosaicism and maternal uniparental disomy for chromosome 9 in the euploid cell line. *J Med Genet*. 1992;29(10):742-744. doi 10.1136/jmg.29.10.742

Xiao B., Wang L., Liu H., Fan Y., Xu Y., Sun Y., Qiu W. Uniparental isodisomy caused autosomal recessive diseases: NGS-based analysis allows the concurrent detection of homogenous variants and copy-neutral loss of heterozygosity. *Mol Genet Genomic Med*. 2019;7(10):e00945. doi 10.1002/mgg3.945

Yang Y., Muzny D.M., Xia F., Niu Z., Person R., Ding Y., Ward P., … Beaudet A.L., Lupski J.R., Plon S.E., Gibbs R.A., Eng C.M. Molecular findings among patients referred for clinical whole-exome sequencing. *JAMA*. 2014;312(18):1870-1879. doi 10.1001/jama.2014.14601

Conflict of interest. The authors declare no conflict of interest.

Received February 14, 2025. Revised June 11, 2025. Accepted June 16, 2025.

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2025;29(5):636-643

doi 10.18699/vjgb-25-69

A family case of a rare Xq28 duplication

A.E. Kopytova [1, 2] ✉, E.N. Tolmacheva [3], D.A. Emelina [4], O.S. Glotov [5], V.V. Miroshnikova [1, 2], T.S. Usenko [1, 2], O.Yu. Vasilyeva [3], E.D. Kasyanov [4], E.A. Fonova [3], I.V. Makarov [4, 6], A.D. Lobanov [7], G.E. Mazo [4], S.N. Pchelina [1, 2], I.N. Lebedev [3]

[1] Petersburg Nuclear Physics Institute named by B.P. Konstantinov of National Research Center "Kurchatov Institute", Gatchina, Leningradskaya Oblast, Russia
[2] Pavlov First St. Petersburg State Medical University, St. Petersburg, Russia
[3] Scientific Research Institute of Medical Genetics, Tomsk National Research Medical Center of the Russian Academy of Sciences, Tomsk, Russia
[4] V.M. Bekhterev National Research Medical Center for Psychiatry and Neurology of the Ministry of Health of the Russian Federation, St. Petersburg, Russia
[5] Federal Scientific and Clinical Center of Infectious Diseases of the Federal Medical and Biological Agency, St. Petersburg, Russia
[6] North-Western State Medical University named after I.I. Mechnikov, St. Petersburg, Russia
[7] Siberian State Medical University, Tomsk, Russia
✉ kopytovaalena@mail.ru

Abstract. Genetic factors contribute to the etiology of intellectual disability in 25–50 % of cases. Chromosomal abnormalities, such as microdeletions and microduplications, are the most significant genetic causes. We examined a family where two boys, aged 8 and 7, were diagnosed with mild intellectual disability. Using array-based comparative genomic hybridization, we detected a duplication of Xq28 in both brothers on the X chromosome inherited from a healthy mother with skewed (88 %) X-chromosome inactivation. The size of the rearrangement is 439.6 kilobases (kb). Eight genes are located in this region, including *F8*, *MTCP1*, *BRCC3*, *VBP1*, *RAB39B*, *CLIC2*, *FUNDC2*, and *CMC4*. This chromosomal region overlaps with the region of Xq28 duplication syndrome (OMIM 300815), characterized by intellectual disability, behavioral and psychiatric disorders, recurrent infections, atopic diseases, and specific facial features in affected male individuals. Whole-exome sequencing did not reveal pathogenic or likely pathogenic variants associated with neurodevelopmental disorders. These disorders have been previously linked to X-linked recessive single-nucleotide variants in *RAB39B* (OMIM 300271, 311510) and *CLIC2* (OMIM 300886). An assessment of the clinical significance of the identified duplication, using the AutoCNV internet resource and original data, allowed us to classify this variant as pathogenic. This implies that the identified duplication may be the cause of intellectual disability in patients.

Key words: Xq28 duplication syndrome; array-based comparative genomic hybridization; copy number variations (CNVs); intellectual disability; *RAB39B*; *CLIC2*

For citation: Kopytova A.E., Tolmacheva E.N., Emelina D.A., Glotov O.S., Miroshnikova V.V., Usenko T.S., Vasilyeva O.Yu., Kasyanov E.D., Fonova E.A., Makarov I.V., Lobanov A.D., Mazo G.E., Pchelina S.N., Lebedev I.N. A family case of a rare Xq28 duplication. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov J Genet Breed*. 2025;29(5): 636-643. doi 10.18699/vjgb-25-69

Funding. Chromosomal microarray analysis, X-chromosome inactivation status and real-time PCR were carried out with the support of the Russian Science Foundation grant No. 21-65-00017.
Collection and preparation of biological samples for cytogenetic research and whole-exome sequencing were carried out within the state assignment of the Ministry of Science and Higher Education of the Russian Federation (theme No. 1023031500037-7-1.6.8;1.6.1;1.6.2;1.6.3).
Whole-exome sequencing was performed with the support of the Federal Academic Leadership Program "Priority 2030" (Agreement 075-15-2023-132 dated 02/14/2023).

Acknowledgements. Molecular cytogenetic studies were performed using the equipment of the Medical Genomics Center for Collective Use on the basis of the Tomsk Scientific Research Center.

Семейный случай редкого варианта дупликации Xq28

А.Э. Копытова [1, 2] ✉, Е.Н. Толмачева [3], Д.А. Емелина [4], О.С. Глотов [5], В.В. Мирошникова [1, 2], Т.С. Усенко [1, 2], О.Ю. Васильева [3], Е.Д. Касьянов [4], Е.А. Фонова [3], И.В. Макаров [4, 6], А.Д. Лобанов [7], Г.Э. Мазо [4], С.Н. Пчелина [1, 2], И.Н. Лебедев [3]

[1] Петербургский институт ядерной физики им. Б.П. Константинова Национального исследовательского центра «Курчатовский институт», Гатчина, Ленинградская область, Россия
[2] Первый Санкт-Петербургский государственный медицинский университет им. академика И.П. Павлова, Санкт-Петербург, Россия

© Kopytova A.E., Tolmacheva E.N., Emelina D.A., Glotov O.S., Miroshnikova V.V., Usenko T.S., Vasilyeva O.Yu., Kasyanov E.D., Fonova E.A., Makarov I.V., Lobanov A.D., Mazo G.E., Pchelina S.N., Lebedev I.N., 2025

This work is licensed under a Creative Commons Attribution 4.0 License

[3] Научно-исследовательский институт медицинской генетики Томского национального исследовательского медицинского центра Российской академии наук, Томск, Россия
[4] Национальный исследовательский центр психиатрии и неврологии им. В.М. Бехтерева Министерства здравоохранения Российской Федерации, Санкт-Петербург, Россия
[5] Федеральный научно-клинический центр инфекционных болезней Федерального медико-биологического агентства, Санкт-Петербург, Россия
[6] Северо-Западный государственный медицинский университет им. И.И. Мечникова, Санкт-Петербург, Россия
[7] Сибирский государственный медицинский университет, Томск, Россия
kopytovaalena@mail.ru

Аннотация. К настоящему времени известно, что на долю генетических факторов, вносящих вклад в этиологию нарушения интеллектуального развития, приходится от 25 до 50 % случаев. Среди генетических причин наиболее существенную роль играют хромосомные аномалии, в том числе микроделеции и микродупликации. Нами обследована семья, в которой у двух мальчиков в возрасте 8 и 7 лет диагностирована легкая интеллектуальная недостаточность. С помощью матричной сравнительной геномной гибридизации у обоих братьев была обнаружена дупликация Xq28. Мать мальчиков является носительницей такой же дупликации с 88 % смещением инактивации Х-хромосомы. Размер перестройки составил 439.6 т. п. н. В данном регионе локализовано восемь генов (*F8*, *MTCP1*, *BRCC3*, *VBP1*, *RAB39B*, *CLIC2*, *FUNDC2*, *CMC4*). Рассматриваемый хромосомный регион перекрывается с областью синдрома дупликации Xq28 (OMIM 300815), характеризующегося интеллектуальной недостаточностью, поведенческими и психиатрическими нарушениями, рецидивирующими инфекциями, атопическими заболеваниями и характерными чертами лица у мужчин. Ранее описаны нарушения интеллектуального развития, обусловленные рецессивными однонуклеотидными вариантами в генах *RAB39B* (OMIM 300271, OMIM 311510) и *CLIC2* (OMIM 300886). Полноэкзомное секвенирование не выявило дополнительных патогенных и потенциально патогенных вариантов, ассоциированных с нарушениями интеллектуального развития. Оценка клинической значимости обнаруженной дупликации с помощью интернет-ресурса AutoCNV и собственных данных позволила классифицировать этот вариант как патогенный, что предполагает, что он может быть причиной интеллектуальной недостаточности у пациентов.
Ключевые слова: синдром дупликации Xq28; матричная сравнительная геномная гибридизация; вариации числа копий участков ДНК (CNV); интеллектуальная недостаточность; *RAB39B*; *CLIC2*

Introduction

Intellectual disability (ID) is a group of disorders characterized by limitations in both intellectual functioning and adaptive behavior (cognitive, speech, social abilities). According to the World Health Organization (2021), approximately 1–3 % of the population suffers from various forms of ID (Schalock et al., 2021). Genetic causes of ID are thought to be present in 25–50 % of cases (Lavrov et al., 2016).

Recent studies have shown that copy number variations (CNVs) are found in 15–25 % of patients with ID and/or multiple congenital anomalies (Iyer, Girirajan, 2015; Fedotov et al., 2024) and may play an important role in the etiology of ID. CNVs are changes in the number of copies of a specific DNA segment, such as microdeletions and microduplications, ranging from a few thousand base pairs to several megabases (Kearney et al., 2011). Xq28 duplication syndrome (OMIM #300815) is the most common cause of ID in men and has several variants depending on the genes involved and the extent of the duplication (Tolmacheva et al., 2022). The variant associated with increased copies of the region including the *RAB39B* and *CLIC2* genes is rare and has been described only in a few studies (El-Hattab et al., 2011; Vanmarsenille et al., 2014; Ballout et al., 2021). The manifestations of the disease phenotype are speculated to be the result of an increased dosage of two genes located in the duplicated segment: *RAB39B* and *CLIC2*. However, the underlying molecular mechanism remains largely unknown and the contribution of excessive *RAB39B* to the development of ID has yet to be confirmed (Wang Z. et al., 2023). It is necessary to describe new cases associated with increased doses of the *RAB39B* and *CLIC2* genes in order to clarify their role in the etiology of ID.

We presented a clinical case of two male ID patients with a rare Xq28 duplication. The aim of this study is to describe this duplication, which involves candidate genes *RAB39B* and *CLIC2* in order to better understand their effects and potential contribution to the ID phenotype.

Materials and methods

This study was approved by the Ethics Committee of the V.M. Bekhterev National Research Medical Center for Psychiatry and Neurology of the Russian Federation Ministry of Health (Protocol No. 3 dated 04/25/24). Written informed consent was obtained from parents for themselves and their children.

Peripheral blood samples from patients and their parents were obtained from the V.M. Bekhterev National Research Medical Center's Biobank.

The peripheral blood of patients and their relatives was collected in tubes containing EDTA for molecular genetic analyses. Genomic DNA was isolated from blood using phenol-chloroform extraction.

Chromosomal microarray analysis (array Comparative Genomic Hybridization (aCGH)) was performed using SurePrint G3 Human CGH 8×60K microarrays (Agilent Technologies, USA) according to the manufacturer's recommendations. Detection was performed using the SureScan Microarray Scanner (Agilent Technologies, USA). Data were obtained using the Scan software (version 9.1.1.1) and visualized with the Cytogenomics software (version 3.0.6.0). Interpretation of the clinical significance of CNVs was carried out in accordance with the American College of Medical Genetics and Genomics (ACMG), Clinical Genomics Resource

project and the Russian Society of Medical Geneticists (Brandt et al., 2020; Riggs et al., 2020; Lebedev et al., 2023), as well as the DGV, OMIM and DECIPHER databases. A detailed analysis of the clinical signs was conducted by reviewing literature data. The pathogenetic significance of the duplication was classified using the AutoCNV score (https://phoenix.bgi.com/autocnv/) and the assessment of the X-chromosome inactivation status (Tolmacheva et al., 2025). CNVs were classified as pathogenic if they had a total score of ≥0.99 according to a semi-quantitative scoring system (Lebedev et al., 2023).

To confirm a detected CNV in patients and determine its origin, we used real-time quantitative PCR with primers selected for exon 3 of the *CMC4* gene (F 5′-CTGTCATCC AAGAACTGCGTAA-3′, R 5′-TACTTTGATGCAGACTT CCGTG-3′).

X-chromosome inactivation status was determined based on the amplification of a highly polymorphic CAG repeat in the first exon of the androgen receptor (*AR*, Xq12) gene after DNA hydrolysis with the methyl-sensitive restriction endonuclease HpaII. PCR products were separated using fragment analysis. The degree of inactivation <80 % was considered a random pattern, and the degree of inactivation >80 % was considered a skewed X-chromosome inactivation (sXCI).

Whole-exome sequencing. The libraries were prepared using the KAPA HyperExome panel (Roche, USA), according to the manufacturer's protocol. The sequencing of the converted libraries (MGI Easy Universal Library Conversion Kit (App-A), MGI, China) was performed on a DNBSEQ-G50 NGS sequencer (MGI, China). After sequencing, FastQC was used for quality control to assess the raw sequence data (Andrews, 2020). The data obtained from sequencing experiments were aligned to the human reference genome, specifically the GRCh38 assembly, using an algorithm called Burrows–Wheeler Aligner (BWA v.0.7.17) (http://bio-bwa.sourceforge.net/). To eliminate possible duplication artifacts at the amplification stage, we used the GATK MarkDuplicates tool to identify and remove PCR duplicates (McKenna et al., 2010). After initial read mapping, the next steps involved recalibrating the quality scores of the reads and addressing potential biases in short insertion/deletion calls. This was achieved using the GATK Base Quality Score Recalibration (BQSR) tool and GATK's BaseRecalibrator and ApplyBQSR tools.

The search for variants was performed using GATK HaplotypeCaller, after which multilevel filtering was applied: low-quality variants were excluded (QUAL < 30, DP < 10), frequent variants were deleted (gnomAD_AF > 0.01) (https://gnomad.broadinstitute.org). The variants were annotated using ANNOVAR (Wang K. et al., 2010) and the refGene, ClinVar, gnomAD, and dbNSFP databases. The analysis of rare pathogenic variants was conducted in accordance with the criteria of the American College of Medical Genetics and Genomics (Richards et al., 2015) and the clinical significance of the variants was assessed using ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/). The predicted effect on the protein was evaluated using *in silico* tools such as SIFT (http://sift.jcvi.org), PolyPhen-2 (http://genetics.bwh.harvard.edu/wiki/pph2/about) and PROVEAN (http://provean.jcvi.org/index.php).

Results

A family with two patients, A. and I., born in 2015 and 2017, respectively, who have intellectual disabilities, consulted a psychiatrist at the Child Psychiatry Department of the V.M. Bekhterev National Research Medical Center located in St. Petersburg to clarify their diagnosis and select treatment. The patients were admitted to the hospital together with their mother to receive treatment.

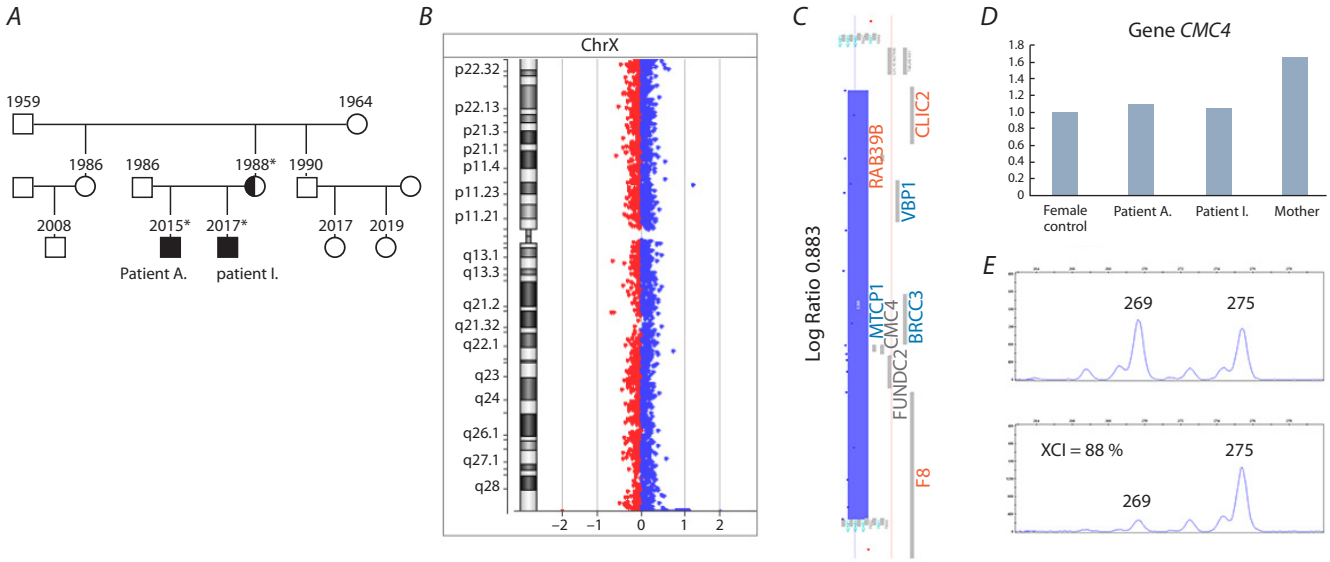
Patient A. An 8-year-old boy, born in 2015, Lezgin by nationality, has a family history of hereditary diseases. His younger brother has ID. He could sit since he was 7 months old, and walk since he was 1 year and 4 months old. Speech in the form of individual words began to appear around the age of 3.5. The perinatal period was burdened with complications, including a threat of pregnancy termination, anemia, and chronic fetal hypoxia. Urgent delivery was conducted by elective cesarean section at 41 weeks, birth weight was 3,500 grams (50th percentile), head circumference at birth was 35 cm (25th percentile), Apgar score was 7/7. During the neonatal period, the baby experienced prolonged jaundice and had feeding problems. He was also seen by an orthopedic specialist for diagnosis of pes valgus. Due to delays in speech development, the child was referred to a speech therapist. A speech delay of level III was identified, along with pseudobulbar dysarthria. The patient's height is 122 cm, weight is 24 kg at the time of examination.

By the time of treatment, his clinical picture showed signs of attention deficit hyperactivity disorder, aggression, tantrums, resistance to restrictions, lack of interest in studying, and sleep disturbance. He was consulted by a clinical psychologist for further evaluation. According to the results of the assessment, the psychologist identified an uneven intellectual development in the child, with a delay in verbal intelligence and difficulties with certain cognitive processes (attention, exhaustion of mental processes) of the organic type. During the work process, the boy required individual support due to his lack of self-organization skills and attention difficulties. In formal terms, according to the Wexler method, his verbal intellectual index (VIP) was 56 (in formal numerical terms corresponds to a mild level of underdevelopment), and his non-verbal intellectual index (NIP) was 94 (in formal numerical terms corresponds to the range of a low age norm).

Magnetic resonance imaging (MRI) of the brain did not reveal any evidence of neoplastic or demyelinating processes or focal changes in the brain tissue.

Video EEG monitoring of nighttime sleep revealed moderate changes in the bioelectric activity of the brain, with a predominance in the right frontocentral regions, increased excitability in deep structures at the diencephalic level and an increase in the phase of REM sleep II. However, no specific paroxysmal activity was recorded.

Based on the clinical presentation and hospital exams, a speech delay was diagnosed in combination with intellectual disability and specific learning difficulties. The patient also had problems with activity and attention at the time of admission to the study. At the time of enrollment in the study, the patient was taking tiapride to manage excitability and aggressive behavior.



The results of the molecular cytogenetic analysis of the family.

A – family history; * patients examined in this study. *B* – chromosome X profile from array-CGH of patient I. *C* – the blue bar indicates the region of duplication in the Xq28 chromosome of the patient and the genes within that region. *D* – Real-time PCR results (exon 3 of the *CMC4* gene). The X axis – the control female DNA and the DNA of the examined individuals; the Y axis – multiple change in the number of DNA copies. *E* – analysis of the X-chromosome inactivation status in a carrier of the Xq28 duplication.

Patient I. A 7-year-old boy, born in 2017, Lezgin by nationality, has a family history of hereditary diseases. His older brother has ID. He could sit since he was 7 months old, and walk since he was 1 year and 3 months old. Speech in the form of individual words began to appear around the age of 3. The perinatal period was burdened with complications, including a threat of pregnancy termination, anemia, and chronic fetal hypoxia. The mother gave birth by elective cesarean section on time, birth weight was 3,950 grams (50th–75th percentile), head circumference at birth was 37 cm (75th percentile), Apgar score was 7/7. The patient's height is 115 cm, weight is 20.5 kg at the time of examination. The boy's medical history includes pes valgus, chest wall deformities, enuresis and constipation. There were no visual or hearing impairments reported.

Cognitive impairments and attention deficit hyperactivity disorder were prominent in the clinical presentation. According to the results of psychological assessment, delayed speech and insufficient development of verbal and logical components of intellectual activity were identified. In terms of formal numbers, based on the Wexler assessment, the productivity of intellectual functioning (for preschoolers, WPPSI) was at a moderate level of underdevelopment (VIP = 56) and the normative level (NIP = 100).

MRI of the brain did not reveal any evidence of neoplastic or demyelinating processes or focal changes in the brain tissue.

Based on the clinical presentation and hospital exams, a mild intellectual disability with severe speech disorders, attention deficit hyperactivity disorder was diagnosed. At the time of enrollment in the study, the patient was receiving amitriptyline to manage attention disorders and hyperactivity.

The mother, born in 1988, Lezgin by nationality, has no known health problems. Her father experienced a severe heart attack and mother has an aggravated hernia. There were no medical or spontaneous abortions in the family's reproductive history. She has an older sister and a younger brother, who also have a son and two daughters, respectively. The family pedigree is shown in the Figure *A*. The clinical manifestations identified in the siblings were not observed in other relatives in the family. Genitals, skin, and appendages are normal, and the skeleton is free of pathology. Vision and hearing are without pathology. Speech is not impaired, speech development is timely. The gait is normal. There was one episode of an affective phase with psychosis after childbirth.

To clarify the causes of ID development in the family we performed a molecular cytogenetic study using array-based comparative genomic hybridization (see the Figure). We detected in both brothers a duplication of Xq28 on the X chromosome inherited from their healthy mother with skewed (88 %) X-chromosome inactivation. The size of the rearrangement is 439.6 kb. Eight genes are located in this region, including *F8*, *MTCP1*, *BRCC3*, *VBP1*, *RAB39B*, *CLIC2*, *FUNDC2*, and *CMC4*. The presence of CNV was confirmed in both siblings and their mother using real-time PCR.

Whole-exome sequencing was performed for both siblings to exclude other potentially pathogenic variants in the coding regions of genes that could contribute to the development of the disease. After analyzing the data from the exome sequence, no pathogenic or likely pathogenic variants were found that could explain the observed clinical picture.

Discussion

Clinical observations show that intellectual disability is more prevalent in males than in females (Mental retardation in children, 2024). X-linked intellectual disability (XLID) is known to contribute to a significant proportion of ID in males, accounting for approximately 10–15 % of cases (Tolmacheva et

al., 2022). To date, 114 different forms of XLID and 172 genes have been identified, variants in which can contribute to the development of the disorder (according to Greenwood Genetic Center, X-Linked Intellectual Disability) (Tolmacheva et al., 2025). Additionally, chromosomal microstructural rearrangements account for approximately 5 % of all cases of XLID (Bauters et al., 2008).

We examined a family where two boys, aged 8 and 7, were diagnosed with mild intellectual disability and had a 439.6 kb duplication on chromosome Xq28 inherited from a healthy mother. Eight genes were located in this region, including *F8*, *MTCP1*, *BRCC3*, *VBP1*, *RAB39B*, *CLIC2*, *FUNDC2*, and *CMC4*.

This chromosomal region overlaps with the region of Xq28 duplication syndrome (OMIM #300815). Xq28 duplication syndrome is a genetic condition linked to the X chromosome, causing ID and other neurodevelopmental issues. The syndrome is characterized by varying degrees of cognitive impairment, typically more pronounced in males. Affected individuals also experience a wide range of neurobehavioral abnormalities and facial dysmorphia (El-Hattab et al., 2011, 2015; Lannoy et al., 2013; Vanmarsenille et al., 2014; Voinova et al., 2015; Ballout et al., 2021). The main symptoms reported in patients with this syndrome are listed in the Table. In rare cases, duplication occurs *de novo*, but in most cases, affected boys inherit the distal duplication of the long arm of the X chromosome from their mothers. Heterozygous females did not show obvious clinical signs of the disease due to nonrandom X-chromosome inactivation (Amos-Landgraf et al., 2006; Lavrov et al., 2017; Tolmacheva et al., 2022). Sometimes mothers may have anxiety-depressive disorders, specific personality traits, speech difficulties, and seizures. We found skewed X-chromosome inactivation (88 %) in the mother with the Xq28 duplication, which may explain the lack of clinical symptoms of ID.

Xq28 region contains many sets of low-copy repeats (LCRs) in close proximity to each other, which render this region prone to non-allelic homologous recombination, which can lead to the formation of gametes with reciprocal microdeletions and microduplications (Vandewalle et al., 2009). The most frequently duplicated region includes the methyl-CpG-binding protein 2 (*MECP2*) gene, with a minimum duplication size of 0.2 million bp. Patients with *MECP2* duplications have severe ID, incurable seizures and recurrent infections. Duplications in the telomeric regions, including the GDP 1 dissociation inhibitor (*GDI1*) gene and the RAS-associated RAB39B protein (*RAB39B*) gene, are independently associated with ID (Tolmacheva et al., 2022). It has been noted that the severity of clinical symptoms in patients with duplications of the *GDI1* gene correlates with the number of copies of the gene (Vandewalle et al., 2009). It should be noted that in the clinical case we described, the duplicated region did not include the *MECP2* and *GDI1* genes, but included the *RAB39B* gene. The neurocognitive symptoms of the syndrome are speculated to be the result of an increased dosage of two genes located in the duplicated segment: *RAB39B* (OMIM #300271) and *CLIC2* (OMIM #300138), due to the identification of both loci within the smallest region of the overlap between the duplicated segments in all affected individuals with Xq28 duplication syndrome (Andersen et al., 2014; El-Hattab et al., 2015).

CLIC2 encodes a unique transmembrane chloride channel found in cardiac and skeletal muscle cells. This channel interacts with the ryanodine receptor 2 (RyR2) in modulating calcium release from the sarcoplasmic reticulum within skeletal and cardiac myocytes (Board et al., 2004; Meng et al., 2009). Pathogenic missense variants in the *CLIC2* gene are associated with a specific form of XLID (XLID 32, OMIM #300886) (Takano et al., 2012). The clinical manifestations of XLID 32 are presented in the Table. However, the effects of *CLIC2* duplication remain uncertain, quantitative expression analysis suggests no significant dosage sensitivity (Vanmarsenille et al., 2014).

Another candidate gene that may contribute to the phenotype is *RAB39B*. The *RAB39B* gene encodes a member of the Rab protein family, which are small GTPases involved in intracellular signaling proteins that coordinate vesicle trafficking during a variety of cellular processes, including neuronal development and signaling (Mignogna et al., 2015). Pathogenic missense variants in the *RAB39B* gene are associated with a specific form of XLID (XLID 72 (OMIM #300271)) and Weisman syndrome (OMIM #311510). The clinical manifestations of XLID 72 are presented in the Table. Loss of function mutations in *RAB39B* have been recently linked to early onset of Parkinson's disease (Wilson et al., 2014; Lesage et al., 2015). Four men with *RAB39B* duplication have been diagnosed with ID and behavioral disorders (Vanmarsenille et al., 2014). Additionally, overexpression of *RAB39B* in mouse primary hippocampal neurons demonstrated a significant reduction in neuronal branching and the number of synapses, resulting in impaired neuron development and synaptic dysfunction (Vanmarsenille et al., 2014). Neuronal overexpression of *RAB39B* impaired the recognition memory and the short-term working memory in mice and resulted in certain autism-like behaviors, including social novelty defect (Wang Z. et al., 2023).

Therefore, the pathogenic role of this aberration is unclear due to limited information in databases. *RAB39B* is a dose-sensitive gene, with evidence of haploinsufficiency (ClinGen DS, https://search.clinicalgenome.org/kb/gene-dosage/RAB39B). In this regard, based on the program for determining the pathogenic significance of CNVs (AutoCNV), the duplication of Xq28 containing this gene is assessed as a variant of uncertain clinical significance (with a score of 0). However, considering the segregation of inheritance in the family, item 5D should be selected in the AutoCNV program ("CNV is associated with a specific condition observed in the patient's family"). This gives a score of 0.45. In a study, 88 % skewed X-chromosome inactivation pattern was observed in a mother, which added 0.65 points (Tolmacheva et al., 2025). Therefore, the overall score for this CNV is 1.1, which allows us to interpret this variant as pathogenic.

The clinical manifestations of the clinical case we studied were similar to those reported in the literature. We identified common symptoms, which are characteristic of Xq28, XLID 32, and XLID 72 duplication syndromes. These include intellectual disability, impaired speech development, and

Comparison of patients clinical features with literature data

Manifestation	Xq28 duplication syndrome [a]	XLID 32 [b]	XLID 72 [c]	Patient A.	Patient I.
Neuropsychological development disorders					
ID	16/19	4/5	14/14	+	+
Attention deficit hyperactivity disorder	6/19	–	5/14	+	+
Aggression and irritability	6/19	–	1/14	+	–
Autism spectrum disorder	2/19		3/14	–	–
Delayed speech development	–	1/5	6/14	+	+
General developmental delay	–	2/5	4/14	–	–
Seizures	–	4/5	3/14	–	–
Depression, bipolar disorder, schizophrenia	3/19	–	–	–	–
Sleep disturbance	3/19	–	–	+	–
Cardiovascular abnormalities	–	2/5	–	–	–
Recurrent sinopulmonary infections					
Otitis media	8/19	–	–	–	–
Pneumonia	4/19	1/5	–	–	–
Upper respiratory tract infections	2/19	–	–	–	–
Atopic conditions					
Asthma	6/19	–	–	–	–
Allergic rhinitis	5/19	–	–	–	–
Eczema	2/19	–	–	–	–
Anthropometric abnormalities					
Obesity	5/19	–	1/14	–	–
Tall stature	3/19	–	–	–	–
Microcephaly	1/19	–	–	+	–
Dolichocephaly	–	–	2/14	–	–
Macrocephaly	–	–	6/14	–	+
Limb and/or digital abnormalities					
Clinodactyly	2/19	–	–	–	–
Preaxial polydactyly	1/19	–	–	–	–
Pes valgus	1/19	–	–	+	+
Facial dysmorphic features					
Tall forehead	11/19	–	–	–	–
Upper eyelid fullness	8/19	–	–	–	–
Broad nasal bridge	8/19	–	–	–	–
Thick lower lip	5/19	–	–	–	–
Long face	4/19	–	3/14	–	–
Large ears	4/19	4/5	5/14	–	–

Note. [a] According to (El-Hattab et al., 2011, 2015; Lannoy et al., 2013; Vanmarsenille et al., 2014; Ballout et al., 2021). [b] According to (Witham et al., 2011; Takano et al., 2012). [c] According to (Russo et al., 2000; Giannandrea et al., 2010).

attention deficit hyperactivity disorder. The medical history of both boys includes pes valgus. Previously described as a rare manifestation of Xq28 duplication syndrome, we have observed those in both boys in our clinical case. The older brother has sleep disturbance, which is typical for patients with Xq28 duplication syndrome. At the same time, we have identified unique symptoms in the younger brother that have not been previously described, such as chest wall deformities and enuresis.

Conclusion

By comparing the results of our molecular cytogenetic analysis with patient anamnesis data and information available in the literature, we have identified common clinical and phenotypic features (such as ID with other mental disorders and limb abnormalities) in boys with duplication of the Xq28 region, as well as in previously described patients with similar duplications, and in patients with ID, associated with variants in the *CLIC2* (XLID 32) and *RAB39B* (XLID 72) genes. Whole-exome sequencing did not reveal pathogenic and likely pathogenic variants associated with neurodevelopment disorders. The size of the rearrangement is 439.6 kb. Eight genes are located in this region, including *F8*, *MTCP1*, *BRCC3*, *VBP1*, *RAB39B*, *CLIC2*, *FUNDC2*, and *CMC4*. For the detected CNV, the total score according to the ACMG algorithm considering the X-chromosome inactivation status was 1.1. Based on the overall results, this variant may be interpreted as pathogenic, which may lead to clinical symptoms in patients. Based on the analysis of clinical cases reported in the literature, it is possible to assume that cognitive impairments may be associated with an increased expression of the *RAB39B* gene due to changes in the number of copies of this region.

References

Amos-Landgraf J.M., Cottle A., Plenge R.M., Friez M., Schwartz C.E., Longshore J., Willard H.F. X chromosome-inactivation patterns of 1,005 phenotypically unaffected females. *Am J Hum Genet.* 2006; 79(3):493-499. doi 10.1086/507565

Andersen E.F., Baldwin E.E., Ellingwood S., Smith R., Lamb A.N. Xq28 duplication overlapping the *int22h-1/int22h-2* region and including *RAB39B* and *CLIC2* in a family with intellectual and developmental disability. *Am J Med Genet A*. 2014;164A(7):1795-1801. doi 10.1002/AJMG.A.36524

Andrews S. A quality control tool for high throughput sequence data. 2020. Accessed: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Ballout R.A., El-Hattab A.W., Schaaf C.P., Cheung S.W. Xq28 duplication syndrome, int22h1/int22h2 mediated. *GeneReviews®*. 2021. https://www.ncbi.nlm.nih.gov/books/NBK349624/

Bauters M., Weuts A., Vandewalle J., Nevelsteen J., Marynen P., Van Esch H., Froyen G. Detection and validation of copy number variation in X-linked mental retardation. *Cytogenet Genome Res.* 2008; 123(1-4):44-53. doi 10.1159/000184691

Board P.G., Coggan M., Watson S., Gage P.W., Dulhunty A.F. CLIC-2 modulates cardiac ryanodine receptor Ca^{2+} release channels. *Int J Biochem Cell Biol.* 2004;36(8):1599-1612. doi 10.1016/j.biocel.2004.01.026

Brandt T., Sack L.M., Arjona D., Tan D., Mei H., Cui H., Gao H., … Vincent L.M., Reavey C., Lai A., Richard G., Meck J.M. Adapting ACMG/AMP sequence variant classification guidelines for single-gene copy number variants. *Genet Med.* 2020;22(2):336-344. doi 10.1038/S41436-019-0655-2

El-Hattab A.W., Fang P., Jin W., Hughes J.R., Gibson J.B., Patel G.S., Grange D.K., Manwaring L.P., Patel A., Stankiewicz P., Cheung S.W. *Int22h-1/int22h-2*-mediated Xq28 rearrangements: intellectual disability associated with duplications and in utero male lethality with deletions. *J Med Genet.* 2011;48(12):840-850. doi 10.1136/jmedgenet-2011-100125

El-Hattab A.W., Schaaf C.P., Fang P., Roeder E., Kimonis V.E., Church J.A., Patel A., Cheung S.W. Clinical characterization of *int22h1/int22h2*-mediated Xq28 duplication/deletion: new cases and literature review. *BMC Med Genet.* 2015;16:12. doi 10.1186/S12881-015-0157-2

Fedotov D.A., Kashevarova A.A., Lebedev I.N. CNVs in patients with neurodevelopmental disorders: meta-analysis. *Russ J Genet* 2024; 60(5):572-587. doi 10.1134/S1022795424700066

Giannandrea M., Bianchi V., Mignogna M.L., Sirri A., Carrabino S., D'Elia E., Vecellio M., … Gecz J., Van Esch H., Raynaud M., Chelly J., D'Adamo P. Mutations in the small GTPase gene RAB39B are responsible for X-linked mental retardation associated with autism, epilepsy, and macrocephaly. *Am J Hum Genet*. 2010; 86(2):185-195. doi 10.1016/j.ajhg.2010.01.011

Iyer J., Girirajan S. Gene discovery and functional assessment of rare copy-number variants in neurodevelopmental disorders. *Brief Funct Genomics*. 2015;14(5):315-328. doi 10.1093/bfgp/elv018

Kearney H.M., Thorland E.C., Brown K.K., Quintero-Rivera F., South S.T. American College of Medical Genetics standards and guidelines for interpretation and reporting of postnatal constitutional copy number variants. *Genet Med.* 2011;13(7):680-685. doi 10.1097/GIM.0b013e3182217a3a

Lannoy N., Grisart B., Eeckhoudt S., Verellen-Dumoulin C., Lambert C., Vikkula M., Hermans C. Intron 22 homologous regions are implicated in exons 1–22 duplications of the *F8* gene. *Eur J Hum Genet.* 2013;21(9):970-976. doi 10.1038/ejhg.2012.275

Lavrov A.V., Bannikov A.V., Chausheva A.I., Dadali E.L. Genetics of mental retardation. *Rossiyskiy Vestnik Perinatologii i Pediatrii = Russian Bulletin of Perinatology and Pediatrics*. 2016;61(6):13-20. doi 10.21508/1027-4065-2016-61-6-13-20 (in Russian)

Lebedev I.N., Shilova N.V., Iourov I.Yu., Malysheva O.V., Tveleneva A.A., Minzhenkova M.E., Markova Zh.G., Tolmacheva E.N., Kashevarova A.A. Guidelines of the Russian Society of Medical Geneticists for chromosomal microarray analysis. *Meditsinskaya Genetika = Medical Genetics*. 2023;22(10):3-47. doi 10.25557/2073-7998.2023.10.3-47 (in Russian)

Lesage S., Bras J., Cormier-Dequaire F., Condroyer C., Nicolas A., Darwent L., Guerreiro R., … Hardy J., Tison F., Singleton A., Brice A.; French Parkinson's Disease Genetics Study Group (PDG) and the International Parkinson's Disease Genomics Consortium (IPDGC). Loss-of-function mutations in *RAB39B* are associated with typical early-onset Parkinson disease. *Neurol Genet*. 2015;1(1):e9. doi 10.1212/NXG.0000000000000009

McKenna A., Hanna M., Banks E., Sivachenko A., Cibulskis K., Kernytsky A., Garimella K., Altshuler D., Gabriel S., Daly M., DePristo M.A. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010;20(9):1297-1303. doi 10.1101/GR.107524.110

Meng X., Wang G., Viero C., Wang Q., Mi W., Su X.D., Wagenknecht T., Williams A.J., Liu Z., Yin C.C. CLIC2-RyR1 interaction and structural characterization by cryo-electron microscopy. *J Mol Biol.* 2009;387(2):320-334. doi 10.1016/J.JMB.2009.01.059

Mental retardation in children: Rubricator for clinical recommendations. Russian Society of Psychiatrists. 2024. Available at https://cr.minzdrav.gov.ru/view-cr/676_2

Mignogna M.L., Giannandrea M., Gurgone A., Fanelli F., Raimondi F., Mapelli L., Bassani S., … Passafaro M., Gatti S., Esteban J.A., Huganir R., D'Adamo P. The intellectual disability protein RAB39B selectively regulates GluA2 trafficking to determine synaptic AMPAR composition. *Nat Commun.* 2015;6:6504. doi 10.1038/ncomms7504

OMIM: An Online Catalog of Human Genes and Genetic Disorders. http://www.omim.org/

Richards S., Aziz N., Bale S., Bick D., Das S., Gastier-Foster J., Grody W.W., Hegde M., Lyon E., Spector E., Voelkerding K., Rehm H.L. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med.* 2015;17(5):405-424. doi 10.1038/GIM.2015.30

Riggs E.R., Andersen E.F., Cherry A.M., Kantarci S., Kearney H., Patel A., Raca G., Ritter D.I., South S.T., Thorland E.C., Pineda-Alvarez D., Aradhya S., Martin C.L. Technical standards for the interpretation and reporting of constitutional copy-number variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics (ACMG) and the Clinical Genome Resource (ClinGen). *Genet Med.* 2020;22(2):245-257. doi 10.1038/S41436-019-0686-8

Russo S., Cogliati F., Cavalleri F., Cassitto M.G., Giglioli R., Toniolo D., Casari G., Larizza L. Mapping to distal Xq28 of nonspecific X-linked mental retardation MRX72: linkage analysis and clinical findings in a three-generation Sardinian family. *Am J Med Genet*. 2000;94(5):376-382. doi 10.1002/1096-8628(20001023)94:5<376::aid-ajmg6>3.0.co;2-a

Schalock R.L., Luckasson R., Tassé M.J. An overview of intellectual disability: definition, diagnosis, classification, and systems of supports (12th ed.). *Am J Intellect Dev Disabil.* 2021;126(6):439-442. doi 10.1352/1944-7558-126.6.439

Takano K., Liu D., Tarpey P., Gallant E., Lam A., Witham S., Alexov E., Chaubey A., Stevenson R.E., Schwartz C.E., Board P.G., Dulhunty A.F. An X-linked channelopathy with cardiomegaly due to a *CLIC2* mutation enhancing ryanodine receptor channel activity. *Hum Mol Genet.* 2012;21(20):4497-4507. doi 10.1093/hmg/ddS292

Tolmacheva E.N., Fonova E.A., Lebedev I.N. X-linked CNV in pathogenetics of intellectual disability. *Russ J Genet*. 2022;58(10):1193-1207. doi 10.1134/S102279542210009X

Tolmacheva E.N., Kashevarova A.A., Fonova E.A., Salyukova O.A., Seitova G.N., Nazarenko L.P., Agafonova A.A., … Vovk S.L., Fedotov D.A., Vasilyeva O.Y., Skryabin N.A., Lebedev I.N. Prevalence of CNVs on the X chromosome in patients with neurodevelopmental disorders. *Mol Cytogenet.* 2025;18(1):3. doi 10.1186/s13039-025-00703-w

Vandewalle J., Van Esch H., Govaerts K., Verbeeck J., Zweier C., Madrigal I., Mila M., … Spaich C., Rauch A., Fryns J.P., Marynen P., Froyen G. Dosage-dependent severity of the phenotype in patients with mental retardation due to a recurrent copy-number gain at Xq28 mediated by an unusual recombination. *Am J Hum Genet.* 2009;85(6):809-822. doi 10.1016/j.ajhg.2009.10.019

Vanmarsenille L., Giannandrea M., Fieremans N., Verbeeck J., Belet S., Raynaud M., Vogels A., Männik K., Õunap K., Jacqueline V., Briault S., Van Esch H., D'Adamo P., Froyen G. Increased dosage of RAB39B affects neuronal development and could explain the cognitive impairment in male patients with distal Xq28 copy number gains. *Hum Mutat*. 2014;35(3):377-383. doi 10.1002/humu.22497

Voinova V.Y., Vorsanova S.G., Yurov Y.B., Kolotiy A.D., Davidova Y.I., Demidova I.A., Novikov P.V., Iourov I.Y. Clinical and genetic characteristics of the X chromosome distal long arm microduplications encompassing the *MECP2* gene. *Zhurnal Nevrologii i Psikhiatrii Imeni S.S. Korsakova = S.S. Korsakov Journal of Neurology and Psychiatry*. 2015;115(10):10-16. doi 10.17116/jnevro201511510110-16 (in Russian)

Wang K., Li M., Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 2010;38(16):e164. doi 10.1093/nar/gkq603

Wang Z., Niu M., Zheng N., Meng J., Jiang Y., Yang D., Yao P., Yao T., Luo H., Xu H., Ge Y., Zhang Y.W., Zhang X. Increased level of RAB39B leads to neuronal dysfunction and behavioural changes in mice. *J Cell Mol Med.* 2023;27(9):1214-1226. doi 10.1111/jcmm.17704

Wilson G.R., Sim J.C.H., McLean C., Giannandrea M., Galea C.A., Riseley J.R., Stephenson S.E.M., … Kalscheuer V.M., D'Adamo P., Bahlo M., Amor D.J., Lockhart P.J. Mutations in *RAB39B* cause X-linked intellectual disability and early-onset Parkinson disease with α-synuclein pathology. *Am J Hum Genet.* 2014;95(6):729-735. doi 10.1016/j.ajhg.2014.10.015

Witham S., Takano K., Schwartz C., Alexov E. A missense mutation in *CLIC2* associated with intellectual disability is predicted by *in silico* modeling to affect protein stability and dynamics. *Proteins*. 2011;79(8):2444-2454. doi 10.1002/prot.23065

Conflict of interest. The authors declare no conflict of interest.
Received March 5, 2025. Revised April 21, 2025. Accepted April 24, 2025.

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2025;29(5):644-651

doi 10.18699/vjgb-25-70

A familial case of interstitial deletion of the short arm of chromosome 6p22.3-p24.3 in twins with severe delay in psychomotor and speech development

G.D. Moskvitin [1, 2] ✉, D.B. Kochkina [1, 2], E.E. Gurinova [1], D.A. Fedotov [3], L.V. Bekenieva [1, 2], A.A. Kashevarova [3], A.L. Sukhomyasova [1], I.N. Lebedev [3], N.R. Maximova [1]

[1] M.K. Ammosov North-Eastern Federal University, Yakutsk, Russia
[2] M.E. Nikolaev Republic Hospital No. 1 – National Center of Medicine, Yakutsk, Russia
[3] Research Institute of Medical Genetics, Tomsk National Research Medical Center of the Russian Academy of Sciences, Tomsk, Russia
✉ Moskvitin.gavriil@mail.ru

Abstract. Interstitial deletions of the short arm of chromosome 6 are even rarer than distal deletions of 6p24-pter, with an incidence rate of 1:1,000,000 (according to MalaCards, https://www.malacards.org/). These deletions are associated with developmental delays, autism spectrum disorders, congenital anomalies, and dysmorphic features. The objective of our study was to identify chromosomal abnormalities in twins from a Yakut family exhibiting severe psycho-speech developmental delays, intellectual disability combined with dysmorphisms, and congenital anomalies. In this paper, two new cases involving monozygotic twins from a Yakut family, who underwent array comparative genomic hybridization (aCGH), were reported. The diagnostic results revealed a rare interstitial deletion in the region 6p22.3-p24.3, measuring 7.5 Mb, which was subsequently confirmed using a conventional cytogenetics (GTG-banding) method. According to the cytogenetic analysis, the karyotypes of the parents were normal, indicating a *de novo* structural chromosomal rearrangement in the patients. Additionally, a comparative phenotypic analysis of these twins with each other and with other previously reported patients was performed; they were found to have overlapping deletions in the 6p22-p24 region. Furthermore, a literature review and an analysis of the gene content of the deleted region 6p22.3-p24.3 were conducted, and so was a discussion of the genotype-phenotype correlation. The results of the phenotypic analysis revealed both common and distinct dysmorphogenic features, including craniofacial dysmorphisms, deformities of the auricles, and abnormalities in the development of the upper and lower limbs, which are often mentioned in the literature. However, the analyzed data, both from the literature and our observations, showed that all patients lacked a common deleted region in the 6p22-p24 area, creating challenges in establishing an accurate diagnosis. The findings indicate the complexity of defining the minimally overlapping region responsible for the observed phenotypic and behavioral traits and highlight the importance of a systematic and multi-level approach to diagnosing severe psycho-speech developmental delays.

Key words: interstitial deletion 6p22.3-p24.3; intellectual disorders; psychomotor and speech delay; autism spectrum disorder; microarray comparative genomic hybridization

For citation: Moskvitin G.D., Kochkina D.B., Gurinova E.E., Fedotov D.A., Bekenieva L.V., Kashevarova A.A., Sukhomyasova A.L., Lebedev I.N., Maximova N.R. A familial case of interstitial deletion of the short arm of chromosome 6p22.3-p24.3 in twins with severe delay in psychomotor and speech development. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov J Genet Breed*. 2025;29(5):644-651. doi 10.18699/vjgb-25-70

Funding. The chromosomal microarray analysis was carried out with the support of grant RSF No. 21-65-00017, https://rscf.ru/en/project/21-65-00017/
Clinical and genealogical as well as cytogenetic studies of the patients were carried out at the expense of State Order of the Ministry of Science and Higher Education of the Russian Federation No. FSRG-2024-0001 "Genomics of the Arctic: diagnosis, prevention and treatment".

Acknowledgements. The equipment used in the research was from the Center for Collective Use "Medical Genomics" (Tomsk NRMC).
We express our appreciation to the colleagues from SAI RS(Ya) "RH No. 1 – NCoM named after M.E. Nikolaev" and their families for their help in collecting clinical and instrumental data.

Семейный случай интерстициальной делеции короткого плеча хромосомы 6p22.3-p24.3 у близнецов с грубой задержкой психо-речевого развития

Г.Д. Москвитин [1, 2] ✉, Д.Б. Кочкина [1, 2], Е.Е. Гуринова [1], Д.А. Федотов [3], Л.В. Бекенева [1, 2], А.А. Кашеварова [3], А.Л. Сухомясова [1], И.Н. Лебедев [3], Н.Р. Максимова [1]

© Moskvitin G.D., Kochkina D.B., Gurinova E.E., Fedotov D.A., Bekenieva L.V., Kashevarova A.A., Sukhomyasova A.L., Lebedev I.N., Maximova N.R., 2025
This work is licensed under a Creative Commons Attribution 4.0 License

[1] Северо-восточный федеральный университет им. М.К. Аммосова, Якутск, Россия
[2] Государственное автономное учреждение Республики Саха (Якутия) «Республиканская больница № 1 – Национальный центр медицины им. М.Е. Николаева», Якутск, Россия
[3] Научно-исследовательский институт медицинской генетики Томского национального исследовательского медицинского центра Российской академии наук, Томск, Россия
Moskvitin.gavriil@mail.ru

Аннотация. Интерстициальные делеции короткого плеча хромосомы 6 встречаются еще реже, чем дистальные делеции 6p24-pter с частотой 1:1 000 000 (по данным MalaCards, https://www.malacards.org/), и ассоциируются с задержками развития, расстройствами аутистического спектра, врожденными аномалиями, а также дисморфическими особенностями. Цель нашего исследования заключалась в поиске хромосомной патологии у близнецов из якутской семьи с грубой задержкой психо-речевого развития, умственной отсталостью в сочетании с дисморфиями и врожденными аномалиями. В этой работе мы сообщаем о двух новых пациентах – монозиготных близнецах из одной якутской семьи, которым была проведена микроматричная сравнительная геномная гибридизация (aCGH). В результате диагностики обнаружена редкая интерстициальная делеция в регионе 6p22.3-p24.3 размером 7.5 Мб, которая ретроспективно была подтверждена анализом GTG – дифференциального окрашивания хромосом. По данным цитогенетического исследования, кариотипы родителей были нормальными, что говорит о *de novo* структурной хромосомной перестройке у пациентов. Также мы выполнили сравнительный фенотипический анализ этих близнецов между собой и с другими ранее описанными в литературе пациентами, у которых были найдены перекрывающиеся делеции в регионе 6p22-p24. Кроме того, проведены обзор литературы и анализ генного состава делетированного региона 6p22.3-p24.3 с обсуждением корреляции генотип-фенотип. По результатам фенотипического анализа выявлены как общие, так и различные стигмы дизморфогенеза, такие как краниофациальные дисморфии, деформации ушных раковин и отклонения в развитии верхних и нижних конечностей, часто упоминаемые в литературе. Однако в проанализированных данных как в литературе, так и в наших наблюдениях у всех пациентов отсутствовал общий делетированный регион в области 6p22-p24, что создает трудности в установлении точного диагноза. Полученные результаты указывают на сложность однозначного определения минимально перекрывающегося региона, ответственного за наблюдаемые фенотипические и поведенческие особенности, и на важность последовательного и многоуровневого подхода к диагностике грубой задержки психо-речевого развития.

Ключевые слова: интерстициальная делеция 6p22.3-p24.3; интеллектуальные расстройства; задержка психо-речевого развития; расстройство аутистического спектра; микроматричная сравнительная геномная гибридизация

Introduction

The frequency of intellectual disorders (ID) in the world is 2–3 % (McKenzie et al., 2016); 1–3 % of children suffer from delayed psychomotor development combined with dysmorphia and congenital anomalies (Shaffer, 2005). It is known that the proportion of children with disabilities due to mental and behavioral disorders in Russia reaches 31 % (Freize et al., 2025). Genetic factors account for 17–47 % of the causes of intellectual disabilities (Moeschler, Shevell, 2006). Aneuploidies, large deletions and duplications, and unbalanced chromosomal translocations occur in 30–35 % of patients with intellectual disabilities and, as a rule, underlie syndromic forms of intellectual disability (Willemsen, Kleefstra, 2014).

Deletions affecting the distal part of the short arm of chromosome 6 are relatively rare. According to the MalaCards website (https://www.malacards.org/), the frequency of 6p24-pter chromosome deletion syndrome in the population is less than 1 per 1,000,000 people. Distal deletions of 6p24-pter are associated with developmental delay, brain malformations (including Dandy–Walker malformation, MIM 220200), anterior chamber abnormalities, hearing loss, ear abnormalities, micrognathia, and heart defects (Mirza et al., 2004). Patients with larger 6p23-pter deletions also have microcephaly, genital abnormalities, speech disorders, and delayed motor development (Plaja et al., 1994; Celestino-Soper et al., 2012). Interstitial deletions on 6p22-p24 are registered even less frequently and are usually associated with delayed psychomotor development and growth, hypotension, as well as a number of congenital anomalies, including hydrocephalus, microcephaly, structural eye abnormalities, hypertelorism, low-set and deformed ears, nasal anomalies, micrognathia, palate anomalies, short neck with folds on the skin, heart defects, kidneys and feet, abnormal genitals and abnormal fingers with nail hypoplasia (Plaja et al., 1994; Mirza et al., 2004; Celestino-Soper et al., 2012).

There are two reports in the scientific literature about interstitial deletion on chromosome 6p22.3-p24.3. In one of them, the authors used microarray comparative genomic hybridization (aCGH) to identify a ~5.4 Mb deletion on chromosome 6p22.3-p23 in a 15-year-old patient with intellectual disability and autism spectrum disorder (ASD) (Celestino-Soper et al., 2012). They suggest that the cause of developmental delay and ASD is related to the deletion of the *ATXN1*, *DTNBP1*, *JARID2*, and *NHLRC1* genes. The same article describes 17 more patients who had overlapping interstitial deletions on chromosome 6p22-p24. Most patients had neurological or behavioral abnormalities, including developmental and speech delays, ASD, attention deficit hyperactivity disorder (ADHD), repetitive movements, and various dysmorphic facial features.

Another article describes a rare case of interstitial deletion on the short arm of chromosome 6 in a fetus with multiple malformations, detected prenatally by the standard cytogenetic method of amniotic fluid at the 26th week of pregnancy. After termination of pregnancy, the authors eliminated the possibility of insertion of chromosome 6 material into any other chromosome using fluorescent *in situ* hybridization (FISH) with a

full-chromosome probe for chromosome 6 and subtelomeric 6p and 6q probes. Next, molecular karyotyping was performed using the aCGH method, which revealed a rare *de novo* interstitial deletion 6p22.3-p24.3 (Colmant et al., 2009).

In this study, two new twin patients from the same Yakut family who were diagnosed with a rare *de novo* interstitial deletion in the 7.5 Mb region 6p22.3-p24.3 are described. Based on the analysis of the previous data, as well as published materials, a comparative phenotypic analysis of these twins between themselves and with other patients with overlapping deletions in the 6p22-p24 region was conducted. A review of the literature and an analysis of the gene composition with a discussion of genotype and phenotype correlations were carried out.

The purpose of the research was to find a chromosomal pathology in twins from a Yakut family who have a severe delay in psycho-speech development and mental retardation.

Materials and methods

The research was approved by the Committee on Biomedical Ethics of the Scientific Research Institute of Medical Genetics of Tomsk National Research Medical Center (Protocol No. 15 dated 28.02.2023). Informed voluntary consent to participate in the research was received, signed by the parents of the study participants.

Clinical, genealogical and cytogenetic studies of the studied family were conducted on the basis of the Medical and Genetic Center of the State Autonomous Institution "RH No. 1 – NCoM named after M.E. Nikolaev" using the resources of the biocollection "DNA Bank of Congenital and Hereditary Pathology and Populations of the Republic of Sakha (Yakutia)".

Cytogenetic examination (karyotyping) was performed on peripheral blood lymphocytes of the patients with GTG-differential staining of chromosomes at the level of 550 bands according to generally accepted protocols under a light microscope.

Microarray comparative genomic hybridization (aCGH) was performed using SurePrint G3 Human CGH 8 × 60K microarray (Agilent Technologies, Santa Clara, California, USA) in accordance with the manufacturer's recommendations based on the Scientific Research Institute of Medical Genetics of the Tomsk National Research Medical Center. Labeling and hybridization of the patient's DNA and reference DNA (Human Reference DNA, Agilent Technologies) were performed using enzymatic labeling and hybridization protocols (v. 7.5, Agilent Technologies). Array images were obtained using the Agilent SureScan microarray scanner (Agilent Technologies). The data obtained were analyzed using the CytoGenomics (v. 5.3.0.14) software (Agilent Technologies) and publicly available databases of genomic variants: (DGV) (http://projects.tcag.ca/variation), MIM (https://omim.org/), DECIPHER (https://www.deciphergenomics.org/) ClinView Analytics (https://clinical-intelligence.org/services/clinview-analytics/). The aCGH results were analyzed in accordance with the recommendations of the American Collegium of Medical Genetics and Genomics (ACMG) (Riggs et al., 2020) and the Russian Society of Medical Geneticists (Lebedev et al., 2023).

Results

The patients, 7-year-old boys from a Yakut family, have been registered at the Medical and Genetic Center of the RH No. 1 – NCoM since 2021 at the age of four with a diagnosis of "Residual organic damage of the central nervous system with severe mental retardation. General speech underdevelopment of level 1. Cerebral palsy, mixed tetraparesis".

It is known from the medical history that the family had previously applied to the Medical and Genetic Center at the 30th week of pregnancy in connection with the carrying of monochoric diamniotic twins. Ultrasound examination of the fetuses revealed a number of changes: fetus No. 1 had edema of Warton's jelly, as well as a hydrocele; fetus No. 2 had polyhydramnios and congenital heart disease, including a defect of the interventricular septum and possibly an aortic defect, dilation of the pulmonary artery throughout, a narrow isthmus of the aorta with suspected aortic coarctation. Both fetuses had bradycardia.

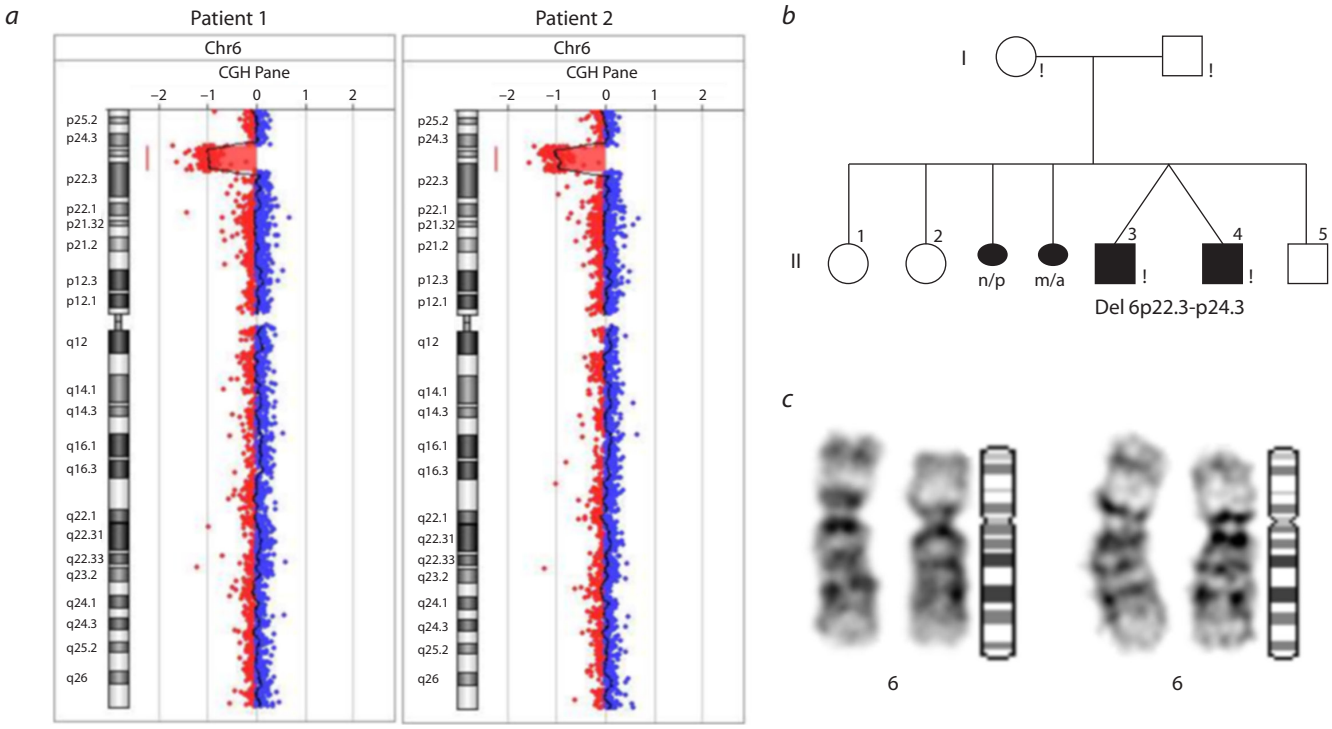
The obstetric and gynecological medical history of the mother is burdened: the first two pregnancies ended in childbirth on time, the third pregnancy ended in spontaneous miscarriage at the early stages, the fourth was terminated at the request of the mother, the sixth ended in childbirth on time. The patients were born from the fifth pregnancy that was threatened with early termination (see the Figure *b*). The delivery was performed by caesarean section at 36 weeks of gestation. The Apgar score was 5/7 for both children. The birth weight of patient 1 and patient 2 was 3,030 g (percentile 25.1; SDS 0.67) and 2,845 g (percentile 14.0; SDS 1.08), respectively, the height of both patients was 50 cm (percentile 52.4; SDS 0.06).

In terms of psychomotor development, both children began to hold their heads at the 2nd month, turn over at the 4th and 5th months, the first child started to sit at the 7th–8th months, the other one first sat at the 9th month. The children started walking with support from the age of one, but at some point both began to lose their acquired skills. The brothers resumed independent walking by the age of two. Among other things, both boys had a delay in speech development, the first words appeared closer to the age of 2 years. However, after the age of three, a regression in psycho-speech development was noted. There is currently no speech. They communicate with pointing gestures and facial expressions, and make inarticulate sounds; if necessary, they lead their relatives by the hand to the object of interest.

Based on the results of the examination and analysis of the phenotypic data of both boys, it is possible to identify both commonalities and differences in their phenotypic characteristics (Table 1).

In both patients, MRI of the brain with angiography revealed the signs of residual encephalopathy. Based on the examination and assessment of the mental status of the patients, a psychiatrist diagnosed them with "Other organic disorders of behavior and emotions with intellectual and mnestic decline and autistic-like behavior".

As a result of the aCGH analysis, a pathogenic deletion was detected in the p24.3-p22.3 region of chromosome 6 (arr[GRCh37] 6p24.3p22.3(10514204_17972394)x1; ISCN,



Standard and molecular cytogenetic study, ancestry of the studied family.

a – profile of aCGH chromosome 6 in patients 1 and 2; *b* – family ancestry; n/p – non-developing pregnancy; m/a – medical abortion; *c* – G-stained chromosomes 6 in patients 1 (left) and 2 (right).

2020) (see the Figure *a*). This chromosomal rearrangement has a length of 7.5 Mb and was detected in both patients (see the Figure *a*). 55 genes are localized in the 6p22.3-p24.3 deletion region, among which 11 are pathogenetically significant according to the MIM database (Table 2).

The standard cytogenetic examination of chromosomes was carried out retrospectively at the Medical and Genetic Center of the State Autonomous Institution "RH No. 1 – NCoM named after M.E. Nikolaev". As a result, interstitial deletion of 6p22-p24 was confirmed in both patients (see the Figure *c*). According to the cytogenetic study, the karyotypes of the parents were normal, indicating a *de novo* structural chromosomal rearrangement in the patients.

Discussion

The interstitial deletion 6p22.3-p24.3 found in Yakut patients in certain regions overlaps with the previously described interstitial deletions in the 6p22-p24 region presented in the scientific literature; however, they have different phenotypic and behavioral features. Search of the databases of diagnostic laboratories Medical Genetics Laboratories (MGL, https://med-gen.ru/en/) and Signature Genomic Laboratories (SGL, https://www.bionity.com/en/companies/18667/signature-genomic-laboratories-llc.html) and literary sources revealed 19 more overlapping interstitial deletions, which coincide with the deletion found in our patients with a diagnosis of "Delayed psycho-speech development, mental retardation and autism-like behavior" (Table 3).

Out of these 19 patients, 13 (patients 1, 2, 4–7, 9, 11–14, 16, 17) were also diagnosed with ASD and/or manifestations associated with ASD, including delayed speech development, ADHD, and behavioral abnormalities (Table 3) (Tuchman, Rapin, 2022; Goldstein, Schwebach, 2024). Some authors suggest (Celestino-Soper et al., 2012) that some of the following genes may be responsible for ASD traits: *ATXN1*, *JARID2*, *DTNBP1,* and *NHLRC1*. Some studies have shown that homozygous mice with *ATXN1* gene knockout exhibit similar aberrations to transgenic mice of the spinocerebellar ataxia type 1 (SCA1) model with polyglutamine expansions (Matilla et al., 1998; Crespo-Barreto et al., 2010). Despite the absence of ataxic symptoms characteristic of SCA1, as well as progressive cerebellar degeneration in knockout mice, both models showed abnormalities in spatial learning and memory, motor learning and coordination.

In addition, changes in the expression of genes associated with the functional activity of the cerebellum have been reported (Matilla et al., 1998; Crespo-Barreto et al., 2010). It should be noted that in a scientific article (Colmant et al., 2009) described in the literature, cerebellar hypoplasia was recorded in a fetus with deletion 6p22.3-p24.3. In addition, the meta-analysis has shown that single nucleotide polymorphic variants in the *ATXN1* gene are associated with IQ in patients with ADHD (Rizzi et al., 2011). A. Bremer and co-authors (2009) hypothesized that *ATXN1* haploinsufficiency may contribute to the learning difficulties observed in patients with a 6p22 deletion, which can be noted in our patients who have a deleted 6p22 region (Table 1).

Given the importance of haploinsufficiency for cognitive functions and associations with behavioral abnormalities in mouse models, P.B. Celestino-Soper and co-authors (2012)

Table 1. Phenotypic features in twins with deletion 6p22.3-p24.3

Phenotypic features	Patient 1	Patient 2
Protruding occiput	+	+
Narrow face	+	+
Wide eyebrows	+	+
Thick eyelashes	+	+
Divergent strabismus	+	+
Wide, flat bridge of the nose	+	+
Wide tip of the nose	+	+
Deep filter	+	+
Elongated lips with a wide cupid's bow	+	+
Drooping corners of the mouth	+	+
Macrotia	+	+
Darwin's tubercle	Left	–
Antihelix	Protruding from both sides	Protruding on the right, flattened on the left
A deformity of the curl that resembles a question mark in shape	Left auricle	Right auricle
"Fleshy" earlobes	+	+
Poorly developed subcutaneous fat tissue	+	+
Wide umbilical ring	+	+
Clinodactyly of 5 fingers	+	+
Knee joint area	Popliteal cord on the right, incomplete extension of the knee joints	The area of hyperkeratosis is dirty gray in color
Feet	Swollen feet, sandal gap, flat-valgus feet	Sandal-shaped gap, flat-valgus position of the feet, protruding metatarsophalangeal joint of the 5th toe of the right foot
Gait	Unsteady, wide base, occasional tiptoe walking	Unsteady, wide base, periodically moves on the outer surface of the feet
Trunk ataxia	+	–
Peculiarities of behavior during examination	Hyperactive, active, does not follow instructions, does not follow objects, does not respond to name	Sleepy, lethargic, turns away during examination, does not follow instructions, does not monitor objects

suggest that heterozygous deletions affecting *ATXN1* functionality may be associated with negative consequences of developmental delay and ASD, both in isolation and in combination with other gene deletions.

Deletions in the 6p.24 region are also associated with heart defects. The *EDN1* gene (located at 6p24.1) encodes the protein called endothelin-1. A study of the distribution of messenger RNA in various tissues revealed that it is distributed differently in brain and heart tissues. Endothelin has an effect on the central nervous system and on the excitability of neurons. Moreover, the *EDN1* gene is involved in both craniofacial and cardiac development (Bogani et al., 2005). Our patients have the same deletion region 6p24.1 as other 10 patients (2, 6–8, 10, 11, 14, 16, 17, 19), who also had congenital heart defects (Table 3), as well as the deleted *EDN1* gene. It should be noted that the scientific literature describes a mutation in the endothelin-1 gene that causes auriculocondylar syndrome (MIM 615706), as well as isolated "question mark ears" syndrome (MIM 612798) (Table 2). According to the phenotypic comparison, the patients described by us had similar deformities of the auricles: macrotia, a deformed notch of the outer curl, resembling a "question mark" in shape (Table 1).

The *JARID2* gene is expressed in both embryonic and adult human neurons (Berge-Lefranc et al., 1996) and can function as a transcriptional repressor (Toyoda et al., 2003);

Table 2. Characteristics of genes located in the region 6p22.3-p24.3 (coordinates: 6:10514204-17972394 at the GRCh37 assembly)

Gene symbol	Coordinates	Function of genes	Associated diseases	MIM
ATXN1	6:16299112-16761491	Binds to RNA and proteins; participates in transcription and developmental processes	Spinocerebellar ataxia type 1, AD (MIM 164400)	601556
DTNBP1	6:15522807-15663058	Organelle biogenesis; plays a role in neuronal function	Hermansky–Pudlak syndrome (MIM 203300); Schizophrenia, AR (MIM 181500)	607145
JARID2	6:15246069-15522042	Binds to DNA, chromatin and proteins; transcriptional repressor; plays a role in CNS development	Developmental delay with varying degrees of intellectual disability and dysmorphic facial features, AD (MIM 620098)	601594
CAP2	6:17393595-17557780	Binds to actin	Dilated cardiomyopathy, 2I, AR (MIM 620462)	618385
EDN1	6:12290361-12297194	Binds to signaling receptors; participates in hormonal activity	Auriculocondylar syndrome 3, AR (MIM 615706); Question mark ears, isolated, AD (MIM 612798)	131240
GCM2	6:10873223-10882041	DNA binding activity	Hyperparathyroidism 4, AD (MIM 617343); Isolated familial hypoparathyroidism 2, AD and AR (MIM 618883)	603716
MAK	6:10762723-10838553	Phosphorus-containing transferase activity and protein tyrosine kinase activity	Retinitis pigmentosa 62, AR (MIM 614181)	154235
PHACTR1	6:12716312-13290446	Binds to actin and protein phosphatase 1	Developmental and epileptic encephalopathy type 70, AD (MIM 618298)	608723
SYCP2L	6:10886831-10979320	An oocyte-specific gene product that localizes to centromeres at the dictyotene stage and regulates the survival of primary oocytes	Premature ovarian failure 24, AR (MIM 620840)	616799
GCNT2	6:10492223-10629368	Acetyl glucosaminyl transferase activator and N-acetyllactosaminide beta-1,6-N-acetylglucosaminyl transferase activator	[Blood group Ii], AD (MIM 110800); Adult i phenotype without cataract, AD (MIM 110800); Cataract 13 with phenotype i in adults, AR (MIM 116700)	600429
TBC1D7	6:13266542-13328583	Activates GTPases and binds to small GTPases; plays a role in regulation of cell growth and differentiation	Macrocephaly/megalencephaly syndrome, AR (MIM 612655)	612655

Note. AD – autosomal dominant inheritance; AR – autosomal recessive inheritance.
* Gene functions are given based on the Gene Ontology Annotation (http://www.ebi.ac.uk/GOA/).

its mouse homologue, jumonji (*Jmj*), is necessary for normal neural tube formation and heart development (Takahashi et al., 2004). Patients with a heterozygous deletion in the *JARID2* gene, which is assumed to lead to haploinsufficiency of the *JARID2* gene, had clinical manifestations of disorders of the nervous system (Barøy et al., 2013; Verberne et al., 2021) (Table 2). The described features in these patients, like in our twins, had characteristic features such as developmental delay, ASD, behavioral disorders, and minor facial phenotype features (Table 1).

In this study, 11 out of 21 patients, including ours, revealed various stigmas of dysmorphogenesis, including craniofacial dysmorphia, ear deformities, and abnormalities in the development of the upper and lower extremities, often mentioned in the literature. In addition, most of the patients had speech disorders and behavioral disorders. However, in the analyzed data, both in the literature and in the present observations, all patients lacked a common deleted region in the 6p22-p24 region, which makes it difficult to establish an accurate diagnosis.

Conclusion

In this research, two new cases of *de novo* interstitial deletion 6p22.3-p24.3 in monozygotic twins from the same Yakut family were analyzed. After studying the literature, the fact of the rare occurrence of such a deletion in the world was proven. A comparison of the phenotypic and behavioral features between our patients and patients previously described in the literature, who had overlapping deletions in the 6p22-p24 region, was made. In addition to a number of common phenotypic features, differences were found between all patients with deletion in the 6p22-p24 region, including our twins. The phenotypic manifestations caused by variants in certain areas of this region are likely to manifest with incomplete penetrance. This fact may indicate a variation in the severity of traits depending on the presence of modifying factors that

Table 3. Clinical features of patients with interstitial deletions in the 6p22-p24 region

Patients	Sex	Chr6 region	Coordinates (hg19)	Size (Mb)	Inheritance	Age (years/ months)	DD/ MR	DSD	RAS	ADHD	SP	SU	HT	CHD	CBD	DS
1	m	p22.3-p23	13662096-19042218	5.4	not mat	15 y	+	+	+	+	n/a	–	+	–	n/a	+
2	m	p22.3	16572367-17543199	1.0	mat	4 y	+	+	+	+	+	–	–	+	n/a	–
3	f	p22.3-p24.3	9621501-24218259	14.6	uk	1 m	n/a	n/a	n/a	n/a	n/a	n/a	n/a	+	–	+
4	m	p23-p24.3	10269968-13915223	3.6	uk	17 y	+	+	–	+	+	+	–	–	–	+
5	f	p22.3	16186391-21421705	5.2	dn	7 y	+	+	–	n/a	n/a	+	–	n/a	n/a	–
6	m	p22.3-p24.1	12058814-20896726	8.8	uk	3 y	+	+	n/a	+	n/a	+	+	+	+	+
7	m	p22.2-p25.2 or p21.33-p23	(2.3-4.2)–(25.2-27.0) or (13.4-15.2)–(30.4-32.1)	n/a	dn	3 y	+	+	n/a	n/a	n/a	n/a	+	+	n/a	+
8	m	p22.3-p24	(7.1-13.4)–(15.2-25.2)	n/a	dn	9 m	+	n/a	n/a	n/a	n/a	n/a	n/a	+	+	+
9	m	p22.1/p22.2-p23	14.4–21.6	n/a	uk	15 y	+	+	n/a	+	n/a	n/a	+	n/a	n/a	+
10	f	p22.3-p23/p24.1	11.9–18.7	n/a	uk	13 m	+	n/a	n/a	n/a	n/a	n/a	+	+	+	+
11	f	p22.3-p24.1	(13.0-14.0)–21.7	n/a	uk	34 m	+	+	n/a	n/a	+	n/a	+	+	+	+
12	m	p22.3-p24.1	10.0–15.8	n/a	not mat	20 y	+	n/a	n/a	n/a	+	n/a	+	–	–	+
13	m	p22.3-p24.2	10.0–18.7	n/a	dn	4 y	+	+	n/a	n/a	n/a	n/a	n/a	–	+	+
14	m	p24.2-p25.1	(4.2-6.1)–10.4-11.9)	n/a	dn	23m	n/a	+	n/a	n/a	n/a	n/a	n/a	+	–	+
15	m	p23	13889301-15153952	1.3	dn	n/i	n/a	n/a	+	n/a	n/a	n/a	n/a	n/a	n/a	n/a
16	f	p22.1-p23	14446670-27741682	13.3	dn	16 y	+	+	uk	+	n/a	n/a	n/a	+	–	+
17	f	p22.3	16132021-23152021	7.0	dn	4 y	+	+	–	–	n/a	n/a	–	+	–	+
18	n/i	p22.3	18829825-23576125	4.7	uk	n/i	+	n/a	+	n/a	n/a	n/a	+	n/a	n/a	+
19 (fetus)	m	p22.3-p24.4	n/a	15.2	dn	Week 26	n/a	n/a	–	n/a	n/a	n/a	n/a	+	+	+
20 (patient 1 of the twins (present study)	m	p22.3-p24.3	10514204-17972394	7.5	dn	7 y	+	+	Autistic-like behavior	+	+	–	–	+	–	+
21 (patient 2 of the twins (present study)	m	p22.3-p24.3	10514204-17972394	7.5	dn	7 y	+	+	Autistic-like behavior	+	+	–	–	+	–	+

Note. DD/MR – developmental delay, mental retardation; DSD – delayed speech development; RAS – autism spectrum disorder; ADHD – attention deficit hyperactivity disorder; SP – stereotypical behavior; SU – seizures; HT – hypotension; CHD – congenital heart defects; CBD – congenital brain defect; DS – dysmorphic signs; uk – unknown; n/a (i) – no data; mat – maternal type of inheritance; not mat – non-maternal type of inheritance; dn – *de novo*. The phenotypic data of patients (1–18) were taken from a scientific article by the author (Celestino-Soper et al., 2005), patient "fetus", from (Colmant, 2009); the lines highlighted in a darker shade are patients from the present examined family.

may be found in other alleles, regulatory elements, or genes located in different parts of the genome. This makes it difficult to unambiguously identify the minimally overlapping region responsible for the observed phenotypes, and indicates the importance of a consistent and multi-level approach to the diagnosis of severe delayed psycho-speech development.

References

Barøy T., Misceo D., Strømme P., Stray-Pedersen A., Holmgren A., Rødningen O.K., Blomhoff A., Helle J.R., Stormyr A., Tvedt B., Fannemel M., Frengen E. Haploinsufficiency of two histone modifier genes on 6p22.3, *ATXN1* and *JARID2*, is associated with intellectual disability. *Orphanet J Rare Dis*. 2013;8(1):3. doi 10.1186/1750-1172-8-3

Berge-Lefranc J.L., Jay P., Massacrier A., Cau P., Mattei M.G., Bauer S., Marsollier C., Berta P., Fontes M. Characterization of the human *jumonji* gene. *Hum Mol Genet*. 1996;5(10):1637-1641. doi 10.1093/hmg/5.10.1637

Bogani D., Willoughby C., Davies J., Kaur K., Mirza G., Paudyal A., Haines H., … Greenfield A., Denny P., Brown S.D., Ragoussis J., Arkell R.M. Dissecting the genetic complexity of human 6p deletion syndromes by using a region-specific, phenotype-driven mouse screen. *Proc Nat Acad Sci USA*. 2005;102(35):12477-12482. doi 10.1073/pnas.0500584102

Bremer A., Schoumans J., Nordenskjöld M., Anderlid B.M., Giacobini M. An interstitial deletion of 7.1 Mb in chromosome band 6p22.3 associated with developmental delay and dysmorphic features including heart defects, short neck, and eye abnormalities. *Eur J Med Genet*. 2009;52(5):358-362. doi 10.1016/j.ejmg.2009.06.002

Celestino-Soper P.B., Skinner C., Schroer R., Eng P., Shenai J., Nowaczyk M.M., Terespolsky D., … Stevenson R.E., Kang S.H., Cheung S.W., Beaudet A.L., Stankiewicz P. Deletions in chromosome 6p22.3-p24.3, including *ATXN1*, are associated with developmental delay and autism spectrum disorders. *Mol Cytogenet*. 2012;5:17. doi 10.1186/1755-8166-5-17

Colmant C., Brisset S., Tachdjian G., Gautier V., Ftouki M., Laroudie M., Druart L., Frydman R., Picone O. Interstitial deletion 6p22.3-p24.3 characterized by CGH array in a foetus with multiple malformations. *Prenat Diagn*. 2009;29(9):908-910. doi 10.1002/pd.2306

Crespo-Barreto J., Fryer J.D., Shaw C.A., Orr H.T., Zoghbi H.Y. Partial loss of ataxin-1 function contributes to transcriptional dysregulation in spinocerebellar ataxia type 1 pathogenesis. *PLoS Genet*. 2010;6(7):e1001021. doi 10.1371/journal.pgen.1001021

Freize V.V., Anokhina M.V., Malyshko L.V., Goncharenko A.Yu., Semenova N.V. Morbidity of the child population with mental disorders and behavioral disorders in the Russian Federation in 2018 and 2022. *Obozreniye Psikhiatrii i Meditsinskoy Psikhologii imeni V.M. Bekhtereva = V.M. Bekhterev Review of Psychiatry and Medical Psychology*. 2025;59(2):100-113. doi 10.31363/2313-7053-2025-2-986 (in Russian)

Goldstein S., Schwebach A.J. The comorbidity of pervasive developmental disorder and attention deficit hyperactivity disorder: results of a retrospective chart review. *J Autism Dev Disord*. 2004;34(3):329-339. doi 10.1023/b:jadd.0000029554.46570.68

Lebedev I.N., Shilova N.V., Yurov I.Yu., Malysheva O.V., Tveleneva A.A., Minzhenkova M.E., Markova Zh.G., Tolmacheva E.N., Kashevarova A.A. Guidelines of the Russian society of medical geneticists for chromosomal microarray analysis. *Meditsinskaya Genetika = Medical Genetics*. 2023;22(10):3-47. doi 10.25557/2073-7998.2023.10.3-47 (in Russian)

Lin R.J., Cherry A.M., Chen K.C., Lyons M., Hoyme H.E., Hudgins L. Terminal deletion of 6p results in a recognizable phenotype. *Am J Med Genet.* 2005;136(2):162-168. doi 10.1002/ajmg.a.30784

Matilla A., Roberson E.D., Banfi S., Morales J., Armstrong D.L., Burright E.N., Orr H.T., Sweatt J.D., Zoghbi H.Y., Matzuk M.M. Mice lacking ataxin-1 display learning deficits and decreased hippocampal paired-pulse facilitation. *J Neurosci*. 1998;18(14):5508-5516. doi 10.1523/JNEUROSCI.18-14-05508.1998

McKenzie K., Milton M., Smith G., Ouellette-Kuntz H. Systematic review of the prevalence and incidence of intellectual disabilities: current trends and issues. *Curr Dev Disord Rep*. 2016;3(2):104-115. doi 10.1007/s40474-016-0085-7

Mirza G., Williams R.R., Mohammed S., Clark R., Newbury-Ecob R., Baldinger S., Flinter F., Ragoussis J. Refined genotype-phenotype correlations in cases of chromosome 6p deletion syndromes. *Eur J Hum Genet*. 2004;12(9):718-728. doi 10.1038/sj.ejhg.5201194

Moeschler J.B., Shevell M.; American Academy of Pediatrics Committee on Genetics. Clinical genetic evaluation of the child with mental retardation or developmental delays. *Pediatrics*. 2006;117(6):2304-2316. doi 10.1542/peds.2006-1006

Plaja A., Vidal R., Soriano D., Bou X., Vendrell T., Mediano C., Pueyo J.M., Labraña X., Sarret E. Terminal deletion of 6p: report of a new case. *Ann Genet*. 1994;37(4):196-199

Riggs E.R., Andersen E.F., Cherry A.M., Kantarci S., Kearney H., Patel A., Raca G., Ritter D.I., South S.T., Thorland E.C., Pineda-Alvarez D., Aradhya S., Martin C.L. Technical standards for the interpretation and reporting of constitutional copy-number variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics (ACMG) and the Clinical Genome Resource (ClinGen). *Genet Med*. 2020;22(2):245-257. doi 10.1038/s41436-019-0686-8

Rizzi T.S., Arias-Vasquez A., Rommelse N., Kuntsi J., Anney R., Asherson P., Buitelaar J., … Steinhausen H.C., Taylor E., Faraone S.V., Franke B., Posthuma D. The *ATXN* and *TRIM3* genes are related to intelligence in an ADHD background: evidence from a large collaborative study totaling 4,963 subjects. *Am J Med Genet B Neuropsychiatr Genet*. 2011;156(2):145-157. doi 10.1002/ajmg.b.31149

Shaffer L.G. American College of Medical Genetics guideline on the cytogenetic evaluation of the individual with developmental delay or mental retardation. *Genet Med*. 2005;7(9):650-654. doi 10.1097/01.gim.0000186545.83160.1e

Takahashi M., Kojima M., Nakajima K., Suzuki-Migishima R., Motegi Y., Yokoyama M., Takeuchi T. Cardiac abnormalities cause early lethality of *jumonji* mutant mice. *Biochem Biophys Res Commun*. 2004;324(4):1319-1323. doi 10.1016/j.bbrc.2004.09.203

Toyoda M., Shirato H., Nakajima K., Kojima M., Takahashi M., Kubota M., Suzuki-Migishima R., Motegi Y., Yokoyama M., Takeuchi T. *jumonji* downregulates cardiac cell proliferation by repressing *cyclin* D1 expression. *Dev Cell*. 2003;5(1):85-97. doi 10.1016/s1534-5807(03)00189-8

Tuchman R., Rapin I. Epilepsy in autism. *Lancet Neurol*. 2002;1(6):352-358. doi 10.1016/s1474-4422(02)00160-6

Verberne E.A., Goh S., England J., van Ginkel M., Rafael-Croes L., Maas S., Polstra A., … Mannens M.M.A.M., Bakshi M., Mallette F.A., van Haelst M.M., Campeau P.M. *JARID2* haploinsufficiency is associated with a clinically distinct neurodevelopmental syndrome. *Genet Med*. 2021;23(2):374-383. doi 10.1038/s41436-020-00992-z

Willemsen M.H., Kleefstra T. Making headway with genetic diagnostics of intellectual disabilities. *Clin Genet*. 2014;85(2):101-110. doi 10.1111/cge.12244

Conflict of interest. The authors declare no conflict of interest.
Received February 18, 2025. Revised May 14, 2025. Accepted May 16, 2025.

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2025;29(5):652-657

doi 10.18699/vjgb-25-71

Study of the meiotic segregation of chromosome 7 with a paracentric inversion in spermatosoa of a heterozygous carrier

M.M. Antonova, D.A. Yurchenko, Zh.G. Markova, N.V. Shilova

Bochkov Research Centre for Medical Genetics, Moscow, Russia
masha-antonova-1999@yandex.ru

Abstract. A paracentric inversion (PAI) is a rare type of balanced intrachromosomal structural rearrangement. Heterozygotes for PAI are usually phenotypically normal, but the presence of the inversion may occasionally lead to synapsis and recombination disruptions during meiosis. PAI can be responsible for the production of recombinant chromosomes and unbalanced gametes. The risks associated with the birth of a child with chromosomal imbalances due to the generation of unbalanced crossover gametes is considered to be low. Nonetheless, viable offspring with intellectual disabilities and/or congenital abnormalities, as well as early miscarriages, stillbirth and infertility in heterozygous carriers of PAI have been described. Paracentric inversions may arise on various chromosomes. PAI with breakpoints on the long arm of chromosome 7 is among the most prevalent ones in humans. To assess the meiotic behavior of abnormal chromosome 7, as well as the empirical risk of producing gametes with recombinant chromosomes, the sperm FISH analysis of a male heterozygous carrier of inv(7)(q11.23q22) was performed. The percentage of recombinant sperms was 0.7 % and chromosomal imbalance was represented as reciprocal breakage products of a dicentric chromosome 7. Notably, spermatozoa with a dicentric chromosome 7 were not observed, which confirms its instability during meiosis I. Meiotic segregation analysis in the heterozygous carrier of inv(7)(q11.23q22) revealed a predominant formation of gametes containing either the inverted or the intact chromosome 7, occurring at frequencies of 52.2 and 47.8 %, respectively. This report is the first study providing a detailed description of meiotic segregation patterns of inv(7)(q11.23q22) by using a sperm FISH approach. Recombinant gamete formation confirms the occurrence of crossing-over within the inversion loop. Consequently, the individual risk of generating gametes (and subsequent zygotes) with chromosome 7 imbalance for this heterozygous carrier remains low.

Key words: paracentric inversion; chromosome 7; sperm FISH; meiotic segregation; sperm recombinant chromosomes

For citation: Antonova M.M., Yurchenko D.A., Markova Zh.G., Shilova N.V. Study of the meiotic segregation of chromosome 7 with paracentric inversion in spermatosoa of heterozygous carrier. *Vavilovskii Zhurnal Genetiki i Selektsii* = *Vavilov J Genet Breed*. 2025;29(5):652-657. doi 10.18699/vjgb-25-71

Исследование мейотической сегрегации хромосомы 7 с парацентрической инверсией в сперматозоидах гетерозиготного носителя

М.М. Антонова, Д.А. Юрченко, Ж.Г. Маркова, Н.В. Шилова

Медико-генетический научный центр им. академика Н.П. Бочкова, Москва, Россия
masha-antonova-1999@yandex.ru

Аннотация. Парацентрическая инверсия (ПаИ) – это редкая сбалансированная внутрихромосомная структурная перестройка. Хотя гетерозиготные носители ПаИ обычно не имеют клинически значимых аномалий фенотипа, факт присутствия в кариотипе хромосомы с инвертированным сегментом предопределяет проблемы синапсиса и рекомбинации в мейозе у таких индивидов и приводит к формированию рекомбинантных хромосом с хромосомным дисбалансом. Риск рождения больного ребенка для носителей ПаИ из-за производства несбалансированных гамет в результате мейотической рекомбинации считается низким. Однако были описаны случаи рождения ребенка с нарушением интеллектуального развития и/или пороками развития; случаи спонтанных абортов, бесплодия у носителей из-за классической рекомбинации в инвертированном хромосомном сегменте. ПаИ могут быть сформированы на различных хромосомах. Показано, что у человека одной из частых среди парацентрических инверсий является ПаИ с локализацией точек разрыва в длинном плече хромосомы 7. С целью оценки мейотического поведения хромосомы 7 с парацентрической инверсией в длинном плече и эмпирического риска формирования гамет с рекомбинантными хромосомами проведено молекулярно-цитогенетическое исследование клеток эякулята у мужчины – гетерозиготного носителя ПаИ хромосомы 7 – inv(7)(q11.23q22). Рекомбинантные хромосомы 7 обнаружены с частотой 0.7 % и в гаметах представлены суммарно реципрокными продуктами разрыва дицентрической хромосомы 7. Сперматозоиды с дицентрической хромосомой 7 не обна-

© Antonova M.M., Yurchenko D.A., Markova Zh.G., Shilova N.V., 2025
This work is licensed under a Creative Commons Attribution 4.0 License

ружены, что подтверждает факт нестабильности этой хромосомы в мейозе I у носителя данной парацентрической инверсии. Показано, что мейотическая сегрегация у гетерозиготного носителя inv(7)(q11.23q22) проходит с преимущественным формированием гамет с инвертированной и интактной хромосомой 7 с частотой 52.2 и 47.8 % соответственно. Впервые получены сведения о частоте формирования гамет с рекомбинантными хромосомами при мейотической сегрегации inv(7)(q11.23q22), что подтверждает факт наличия кроссинговера в инверсионной петле. Персонализированный риск формирования гамет (зигот) с дисбалансом материала хромосомы 7 у гетерозиготного носителя данной инверсии является низким.

Ключевые слова: парацентрическая инверсия; хромосома 7; FISH сперматозоидов; мейотическая сегрегация; рекомбинантные хромосомы

Introduction

Inversion is an intrachromosomal structural rearrangement in which two breaks occur, and the segment lying between the breakpoints rotates 180°. In paracentric inversions (PAI) of chromosomes, both breakpoints are located on the same arm of the same chromosome. Thus, the centromere is not involved in the rearrangement, and the rearranged chromosome consists of an inverted segment and two flanking, distal, non-inverted regions. PAI occurs with a frequency of 0.1–0.5 % (Gardner, Amor, 2018). Most often, PAI is found in chromosomes 1, 3, 5, 6, 7, and 11, with breakpoints localized at 3(p13p25), 6(p12p23), 6(p12p25), 7(q11q22), and 11(q21q23) (Pettenati et al., 1995). Heterozygous carriers of PAI do not exhibit clinically significant phenotypic abnormalities (Madan, 1995; Yang et al., 1997; Muss, Schwanitz, 2007). However, the presence of a chromosome with an inverted segment in the karyotype can lead to problems during meiotic segregation, resulting in the formation of gametes with recombinant chromosomes. This, in turn, may lead to zygotes with chromosomal imbalance and the birth of a child with chromosomal pathology. A key feature of synapsis and recombination in paracentric inversions during the pachytene stage of prophase I is the formation of an inversion loop (Fig. 1*a*).

Depending on the number of crossovers between a normal chromosome and its PAI homologue, various meiotic segregation outcomes are possible. If crossing-over occurs outside the inversion loop, no recombinant chromosomes will form. A single crossover within the inversion loop can lead to the formation of a recombinant dicentric chromosome and an acentric fragment (Fig. 1*b*-1) (Phelan et al., 1993; Anton et al., 2005). Cells containing an acentric fragment undergo apoptosis. The dicentric chromosome is unstable and may rupture during anaphase of meiosis I, resulting in gametes with abnormal chromosomes: one with an inverted duplication and an adjacent terminal deletion (inv dup del) and the other with a terminal deletion of the chromosome arm (Feldman et al., 1993; Mitchell et al., 1994) (Fig. 1*b*-2). The empirical risk of gametes with recombinant chromosomes can be assessed using FISH analysis of ejaculate cells (Bhatt et al., 2009; Balasar, Acar, 2020). In cases of classical segregation leading to an unstable dicentric chromosome, commercially available DNA probes targeting the centromeric and subtelomeric regions of the chromosome with PAI are sufficient for analysis.

Reports on meiotic segregation in inversion carriers show wide variability in the frequency of recombinant gametes, ranging from 0 to 38 % (Morel et al., 2007; Anton et al., 2005; Bhatt et al., 2009). This variability influences the reproductive outcomes for couples where one partner carries an inversion. For male heterozygous carriers of PAI, determining the frequency of abnormal gametes allows for personalized risk

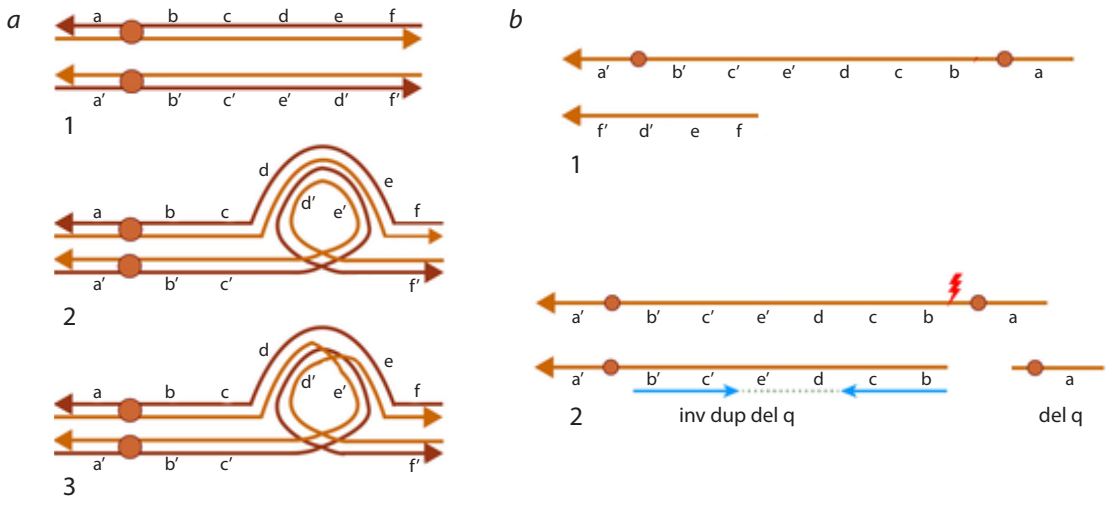


Fig. 1. Meiotic segregation of a chromosome with a PAI:

a – formation of an inversion loop in meiosis I: homologous chromosomes, the lower one has an inversion (1); formation of an inversion loop (2); crossing-over within the inversion loop (3). *b* – theoretically possible variants of gametes during meiotic segregation of PAI: dicentric chromosome and acentric fragment (1); break of the dicentric chromosome, forming an inverted duplication with an adjacent terminal deletion (inv dup del) and a chromosome with a terminal deletion of the arm (2). Adapted (Burssed et al., 2022).

assessment of having a child with chromosomal imbalance and improves medical and genetic counseling for the family.

The aim of our study was to evaluate meiotic segregation of chromosome 7 with paracentric inversion in ejaculate cells and determine the frequency of gametes with recombinant chromosomes.

Material and methods

The patient was a healthy 41-year-old man without clinical phenotypic abnormalities, enrolled in an assisted reproductive technology (ART) program for male infertility. Samples of peripheral venous blood and ejaculate were collected for analysis.

Cytogenetic study was performed on cultured peripheral blood lymphocytes according to a standard protocol (Cytogenetic Methods…, 2009). GTG-banding (550 bands) revealed the karyotype 46,XY,inv(7)(q11.23q22).

The inverted segment size relative to the q arm and the total length of chromosome 7 were calculated as 27.4 and 16.8 %, respectively.

Preparations from spermatozoa were obtained in accordance with a previously developed protocol (Tarlycheva et al., 2021).

FISH analysis of spermatozoa was performed using DNA probes on the centromeric region of chromosome7 (SE 7 (D7Z1), SpBlue), subtelomeric region of the long arm of chromosome 7 (Subtel 7q, SpRed), subtelomeric region of the long arm of chromosome 2 (Subtel 2q, SpGreen) as a control of ploidy and hybridization efficiency (Leica, Kreatech, Germany) according to the protocol of the manufacturing company. FISH analysis of peripheral blood lymphocytes was performed using locus-specific DNA probes on chromosome 7 labeled with various fluorochromes: ELN (7q11) (SpO)/7q22 (SpG) (Leica, Kreatech, Germany).

Hybridization signals were analyzed using an Axio Imager M.1 epifluorescence microscope (Carl Zeiss, Germany) and Isis software (MetaSystems, Germany).

Results

FISH analysis of peripheral blood lymphocytes confirmed PAI in the patient (Fig. 2).

To evaluate the frequency of gametes with recombinant and non-recombinant (normal and inverted) chromosome 7, FISH analysis of the patient's ejaculate cells was performed using a combination of DNA probes targeting the subtelomeric region of the long arm and the centromeric region of chromosome 7, as well as the subtelomeric region of the short arm of chromosome 2. In gametes with non-recombinant chromosomes, one blue, one red, and one green hybridization signal should be observed. In gametes with recombinant chromosomes – inv dup del(7q) or del(7q) – only one blue (from the centromeric region of chromosome 7) and one green (control) hybridization signal will be present, while the red hybridization signal will be absent, as all such chromosomes exhibit a terminal deletion of the long arm of chromosome 7. Meanwhile, gametes with a recombinant dicentric chromosome can be identified by the presence of two blue hybridization signals (corresponding to the centromeric region of chromosome 7) and one green control signal (Fig. 3).

The results of the frequency analysis of gametes with non-recombinant (normal and balanced) and recombinant chromosome 7 are presented in Table 1. During the analysis of 6,116 ejaculate cells, recombinant chromosome 7 was detected at a frequency of 0.7 %, and in mature germ cells (gametes), it was represented exclusively by reciprocal products of the breakage of a dicentric chromosome 7. Spermatozoa carrying the recombinant dicentric chromosome were not detected, confirming the instability of this chromosome during meiosis I in the carrier of this paracentric inversion.

To assess the frequency of gametes with intact and inverted chromosome 7, FISH analysis was performed using a combination of DNA probes targeting the q11 (red hybridization signal) and q22 (green hybridization signal) regions of chromosome 7. The distance between the hybridization signals

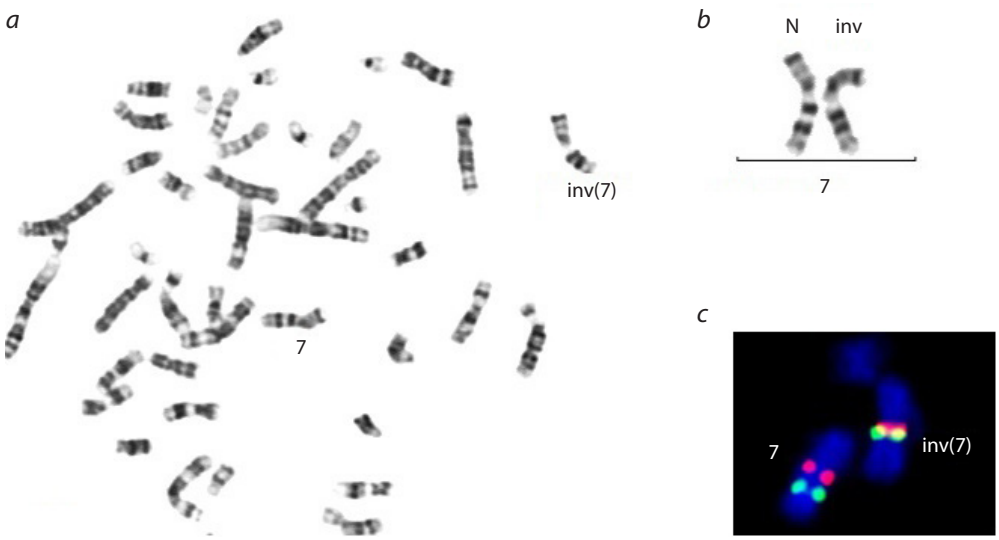


Fig. 2. Metaphase plate (*a*), fragment of the karyogram of the patient with inv(7)(q11.23q22) (*b*) and the result of hybridization with locus-specific DNA probes on chromosome 7 (*c*) – the convergence of hybridization signals from locus-specific DNA probes to regions 7q11 (red) and 7q22 (green) in one of the homologues of chromosome 7 indicates the presence of PAI.

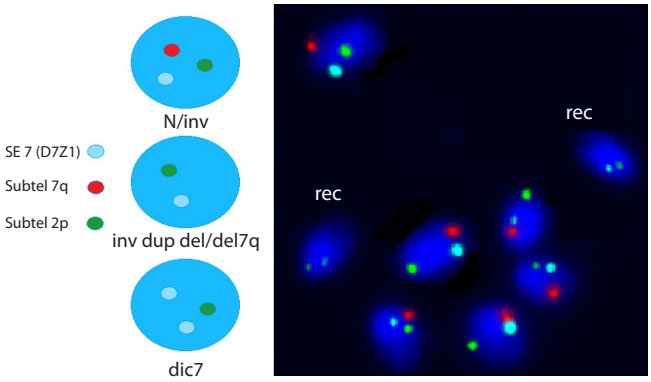


Fig. 3. Hybridization patterns expected in gametes due to crossing-over in the inversion loop in a male carrier of inv(7)(q11q22) and the result of FISH analysis performed on preparations from the ejaculate of a heterozygous inv(7)(q11q22) carrier to estimate the frequency of gametes with recombinant chromosome 7.

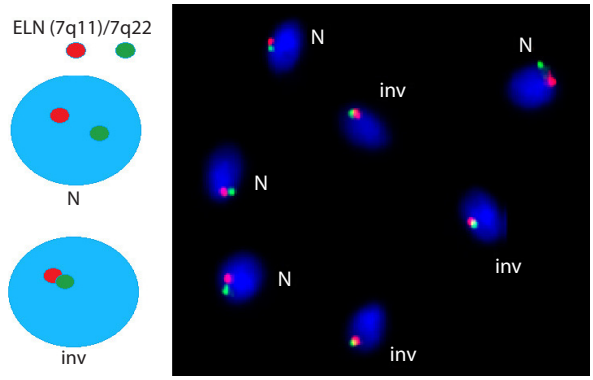


Fig. 4. Hybridization patterns enabling evaluation of gamete types in the absence of recombination within the inversion loop in a male carrier of inv(7)(q11.23q22), and the results of FISH analysis on ejaculate preparations from a heterozygous inv(7)(q11.23q22) carrier for assessing the frequency of gametes with intact and inverted chromosome 7.

Table 1. Frequency of gametes with non-recombinant and recombinant chromosome 7

DNA probes	N/inv		Inv dup del/del7q		Dicentric		Total cells
	Number of cells	%	Number of cells	%	Number of cells	%	
SE 7/Subtel 7q	6,074	99.3	42	0.7	0	0	6,116

Table 2. Frequency of gametes with intact and inverted chromosome 7

DNA probes	Inv		N		Total cells	Critical value of the *t*-value	*p*
	Number of cells	%	Number of cells	%			
ELN (7q11)/7q22	1,697	52.2	1,553	47.8	3,250	1.972	<0.05

allowed for the determination of whether chromosome 7 was intact or inverted. In the case of an inversion, the red and green hybridization signals appeared closer together. The hybridization results and possible signal patterns are presented in Figure 4.

The results of the analysis of the frequency of gametes with intact and inverted chromosome 7 are presented in Table 2. A total of 3,250 cells were analyzed, with the frequency of cells carrying inverted and intact chromosome 7 being 52.2 and 47.8 %, respectively.

Discussion

Constitutional chromosomal abnormalities are among the known genetic factors contributing to male infertility, increased risk of miscarriage, and the birth of children with developmental disorders. Paracentric inversions (PAIs) can not only disrupt meiosis and spermatogenesis but also lead to the formation of mature gametes with chromosomal imbalance due to the generation of recombinant chromosomes during male gametogenesis. The classic meiotic segregation scenario for PAIs involves crossing-over within the inversion loop, followed by the formation of a dicentric chromosome, its subsequent breakage, and the production of gametes carrying inv dup del and deleted chromosomes.

Multiple factors influence the formation of the inversion loop, including the size of the inverted segment. The risk of generating gametes with recombinant chromosomes depends on the likelihood of meiotic crossing-over occurring within the inversion loop. If the inversion is small, the probability of crossing-over within the inverted segment is low, as the number of crossover events appears to be proportional to chromosome length. Studies on the meiotic segregation of pericentric inversions have demonstrated that when the inverted segment constitutes <30 % of the chromosome length, recombinant gametes are not formed. If the inverted segment spans 30–50 % of the chromosome length, the frequency of recombinant gametes is <5 %, increasing to 20.5 % when the inverted segment exceeds 50 % (Morel et al., 2007). A positive correlation between the size of the inverted segment and the frequency of recombinant gametes has also been observed in the limited studies on the meiotic behavior of PAIs. For instance, an analysis of meiotic segregation patterns in blastocysts during preimplantation genetic testing of couples carrying PAIs revealed that the frequency of blastocysts with recombinant chromosomes increased with the size of the inverted segment, ranging from complete absence (when the inversion was <37.5 % of the chromosome length) to 12 % (for larger inversions) (Xie et al., 2019). Our previous research also demonstrated that in a heterozygous carrier of a polymorphic PAI in the short arm of chromosome 8 (with the inverted segment constituting 3.2 % of the chromosome length), the frequency of recombinant gametes was 0.03 % (Yurchenko et al., 2022).

Since only one chromosome arm is involved in the paracentric inversion and findings indicate that synapsis initiates distally on both arms in metacentric and submetacentric chromosomes but involves only one arm in acrocentric chromosomes (Brown et al., 1998), it was proposed to modify the evaluation criteria for PAIs. Instead of calculating the size of the inverted segment relative to the entire chromosome, it should be calculated relative to the length of the arm containing the inversion. S. Bhatt et al. demonstrated that when the PAI size is less than 50 % of the corresponding chromosome arm length, the percentage of recombinant spermatozoa ranges from 0 to 3.72 %, increasing to 10 % or more when the PAI exceeds 50 % of the arm length (Bhatt et al., 2014). In the present case of a heterozygous carrier of inv(7)(q11.23q22), where the inverted segment constitutes 16.8 % of chromosome 7 length and 27.4 % of its q-arm, the frequency of recombinant gametes was 0.7 %. These findings support the established correlation between the size of the inverted segment and recombination frequency in PAIs.

Limited studies on male gametogenesis in PAI carriers have reported an absence of recombinant chromosomes during meiotic segregation of inv(7)(q11q22) (Bhatt et al., 2009, 2014). The authors refer to an original study (Martin, 1986) in which meiotic segregation analysis was performed on pronuclear chromosomes obtained via *in vitro* penetration of spermatozoa from an inv(7)(q11q22) carrier into golden hamster (*Mesocricetus auratus*) oocytes. After analyzing 94 metaphase spreads, the authors concluded that no recombinant chromosome 7 was present (Martin et al., 1986).

In our analysis assessing the frequency of recombinant chromosome 7, hybridization patterns were examined in over 6,000 ejaculate cells. This allowed us to obtain reliable evidence of recombinant chromosomes in a heterozygous inv(7)(q11.23q22) carrier, contradicting previous findings.

The frequency of gametes with inverted chromosome 7 was statistically significantly different ($p < 0.05$) from that of gametes with intact chromosome 7. Thus, we suggest that heterozygous inv(7)(q11.23q22) carriers exhibit a preferential tendency to produce gametes with inverted chromosome 7 during meiotic segregation. However, drawing definitive conclusions is challenging due to the potential for random signal proximity, which could introduce systematic bias and overestimate the frequency of gametes with inverted chromosome 7.

Conclusion

A key aspect of genetic counseling for families carrying chromosomal rearrangements is assessing the risk of having children with chromosomal abnormalities caused by pathological segregation patterns during gametogenesis in the parent carrying the rearrangement. Determining the degree of genetic risk, along with the potential medical and social consequences of the anticipated chromosomal pathology, enables the development of personalized preventive strategies to avoid the birth of an affected child. FISH analysis of ejaculate cells is a specific method for studying the meiotic behavior of chromosomal abnormalities, including paracentric inversions. By identifying an effective combination of DNA probes for molecular cytogenetic analysis of male gametogenesis, it becomes possible to investigate segregation patterns and evaluate recombination events occurring during meiosis in carriers of chromosomal abnormalities. The assessment of the risk of having a child with chromosomal imbalance directly depends on understanding the frequency of recombinant gamete formation.

This study demonstrates that meiotic segregation of the paracentric inversion inv(7)(q11.23q22) in the long arm of chromosome 7 predominantly results in gametes carrying either an intact or inverted chromosome 7. For the first time, data on the frequency of recombinant gamete formation during meiotic segregation of inv(7)(q11q22) have been obtained, confirming the occurrence of crossing-over within the inversion loop. The personalized risk of producing gametes (or zygotes) with chromosomal imbalance in a heterozygous carrier of inv(7)(q11q22) is 0.7 %, which is considered low.

References

Anton E., Blanco J., Egozcue J., Vidal F. Sperm studies in heterozygote inversion carriers: a review. *Cytogenet Genome Res*. 2005;111(3-4):297-304. doi 10.1159/000086903

Balasar Ö., Acar H. Investigation of the interchromosomal effects in male carriers with structural chromosomal abnormalities using FISH. *Turk J Urol*. 2020;46(3):178-185. doi 10.5152/tud.2020.19255

Bhatt S., Moradkhani K., Mrasek K., Puechberty J., Manvelyan M., Hunstig F., Lefort G., Weise A., Lespinasse J., Sarda P., Liehr T., Hamamah S., Pellestor F. Breakpoint mapping and complete analysis of meiotic segregation patterns in three men heterozygous for paracentric inversions. *Eur J Hum Genet*. 2009;17(1):44-50. doi 10.1038/ejhg.2008.144

Bhatt S.S., Manvelyan M., Moradkhani K., Hunstig F., Mrasek K., Puechberty J., Lefort G., Sarda P., Weise A., Liehr T., Pellestor F. Inverted segment size and the presence of recombination hot spot clusters matter in sperm segregation analysis. *Cytogenet Genome Res*. 2014;142(2):145-149. doi 10.1159/000356142

Brown G.M., Leversha M., Hulten M., Ferguson-Smith M.A., Affara N.A., Furlong R.A. Genetic analysis of meiotic recombination in humans by use of sperm typing: reduced recombination within a heterozygous paracentric inversion of chromosome 9q32-q34.3. *Am J Hum Genet*. 1998;62(6):1484-92. doi 10.1086/301863

Burssed B., Zamariolli M., Bellucco F.T., Melaragno M.I. Mechanisms of structural chromosomal rearrangement formation. *Mol Cytogenet*. 2022;15(1):23. doi 10.1186/s13039-022-00600-6

Cytogenetic Methods of Diagnosis of Chromosomal Diseases. Methodical manual for doctors. Moscow, 2009 (in Russian)

Feldman G.L., Weiss L., Phelan M.C., Schroer R.J., Van Dyke D.L. Inverted duplication of 8p: ten new patients and review of the literature. *Am J Med Genet*. 1993;47(4):482-486. doi 10.1002/ajmg.1320470410

Gardner R.J. McKinlay, Amor D.J. Gardner and Sutherland's Chromosome Abnormalities and Genetic Counselling. Oxford Univ. Press, 2018. doi 10.1093/med/9780199329007.001.0001

Madan K. Paracentric inversions: a review. *Hum Genet*. 1995;96(5):503-515. doi 10.1007/BF00197403

Martin R.H. Sperm chromosome analysis in a man heterozygous for a paracentric inversion of chromosome 7 (q11q22). *Hum Genet*. 1986;73(2):97-100. doi 10.1007/BF00291594

Mitchell J.J., Vekemans M., Luscombe S., Hayden M., Weber B., Richter A., Sparkes R., Kojis T., Watters G., Der Kaloustian V.M. U-type exchange in a paracentric inversion as a possible mechanism of origin of an inverted tandem duplication of chromosome 8. *Am J Med Genet*. 1994;49(4):384-387. doi 10.1002/ajmg.1320490406

Morel F., Laudier B., Guérif F., Couet M.L., Royère D., Roux C., Bresson J.L., Amice V., De Braekeleer M., Douet-Guilbert N. Meiotic segregation analysis in spermatozoa of pericentric inversion carriers using fluorescence in-situ hybridization. *Hum Reprod*. 2007;22(1):136-141. doi 10.1093/humrep/del317

Muss B., Schwanitz G. Characterization of inversions as a type of structural chromosome aberration. *Int J Hum Genet*. 2007;7(2):141-161. doi 10.31901/24566330.2007/07.02.04

Pettenati M.J., Rao P.N., Phelan M.C., Grass F., Rao K.W., Cosper P., Carroll A.J., … Jackson-Cook C., Brown J., Schwartz S., Van Dyke D.L., Palmer C.G. Paracentric inversions in humans: a review of 446 paracentric inversions with presentation of 120 new cases. *Am J Med Genet*. 1995;55(2):171-187. doi 10.1002/ajmg.1320550207

Phelan M.C., Stevenson R.E., Anderson E.V., Jr. Recombinant chromosome 9 possibly derived from breakage and reunion of sister chromatids within a paracentric inversion loop. *Am J Med Genet*. 1993;46(3):304-308. doi 10.1002/ajmg.1320460313

Tarlycheva A.A., Markova Zh.G., Yurchenko D.A., Shilova N.V. Optimization of the sperm processing protocol for subsequent molecular cytogenetic studies. *Klinicheskaia Laboratornaia Diagnostika = Clinical Laboratory Diagnostics*. 2021;66(10):603-609. doi 10.51620/0869-2084-2021-66-10-603-609 (in Russian)

Xie P., Hu L., Tan Y., Gong F., Zhang S., Xiong B., Peng Y., Lu G.X., Lin G. Retrospective analysis of meiotic segregation pattern and interchromosomal effects in blastocysts from inversion preimplantation genetic testing cycles. *Fertil Steril*. 2019;112(2):336-342.e3. doi 10.1016/j.fertnstert.2019.03.041

Yang S.P., Bidichandani S.I., Figuera L.E., Juyal R.C., Saxon P.J., Baldini A., Patel P.I. Molecular analysis of deletion (17)(p11.2p11.2) in a family segregating a 17p paracentric inversion: implications for carriers of paracentric inversions. *Am J Hum Genet*. 1997;60(5):1184-1193

Yurchenko D.A., Tarlycheva A.A., Minzhenkova M.E., Markova Zh.G., Shilova N.V. Estimation of the frequency of recombination in the inversion loop in a carrier of polymorphic paracentric inversion 8p23.1. *Medicinskaya Genetika = Medical Genetics*. 2022; 21(12):60-63. doi 10.25557/2073-7998.2022.12.60-63 (in Russian)

Conflict of interest. The authors declare no conflict of interest.
Received February 14, 2025. Revised June 3, 2025. Accepted June 4, 2025.

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2025;29(5):658-665

doi 10.18699/vjgb-25-72

Familial translocation between chromosomes 3 and 10: meiotic segregation, diagnostics and clinical features of chromosomal imbalance

A.V. Vozilova [1], A.S. Tarasova [1], E.A. Ivanov [1], V.P. Pushkarev [1], N.I. Nalyotova [1], A.I. Pobedinskaya [1], A.S. Sabitova [1], N.V. Shilova [2]

[1] Chelyabinsk Regional Children's Clinical Hospital, Chelyabinsk, Russia
[2] Bochkov Research Centre for Medical Genetics, Moscow, Russia
vozilova_a@mail.ru

Abstract. Reciprocal translocations are the most common structural chromosomal rearrangements, occurring at a frequency of 0.08–0.3 % in the human population. The vast majority of carriers of reciprocal translocations are phenotypically normal, but have an increased risk of miscarriage or the birth of children with intellectual disabilities and multiple congenital abnormalities due to meiotic malsegregation of chromosomes involved in the translocation. This study presents a familial case of translocation involving the distal regions of the short arms of chromosomes 3 and 10, detected in seven family members across three generations. The investigation was prompted by the detection of a deletion 10p15 and a duplication 3p25 revealed through clinical exome sequencing in a proband exhibiting phenotypic abnormalities, which may correspond to der(10)t(3;10)(p25;p15). GTG cytogenetic study of the proband's family revealed that the mother, grandmother, aunt and brother – none of whom displayed any clinical or phenotypic manifestations – were carriers of a balanced chromosomal rearrangement, t(3;10)(p25;p15). By contrast, the karyotype of the proband's sibling – a girl with severe cognitive, neurological, and developmental abnormalities – was found to be 46,XX,der(3)t(3;10)(p25;p15)dmat. Molecular karyotyping facilitated further clarification of the chromosomal imbalance and the precise breakpoints on both chromosomes involved in the translocation. This study provides a detailed description of the clinical and phenotypic manifestations resulting from the presence of derivative chromosomes 3 and 10 in the karyotype. Additionally, it discusses the mechanisms underlying the formation of chromosomal imbalances in the family members with the abnormal phenotype, the relationship between the severity of clinical manifestations and changes in gene dosage due to chromosomal rearrangements, as well as potential preventive and rehabilitative measures aimed at reducing the risk of chromosomal pathology in the families with carriers of autosomal reciprocal translocations.

Key words: reciprocal translocations; meiotic segregation; genome imbalance; der(3); der(10); GTG-banded chromosomes; clinical exome sequencing; chromosomal microarray; 3p deletion syndrome

For citation: Vozilova A.V., Tarasova A.S., Ivanov E.A., Pushkarev V.P., Nalyotova N.I., Pobedinskaya A.I., Sabitova A.S., Shilova N.V. Familial translocation between chromosomes 3 and 10: meiotic segregation, diagnostics and clinical features of chromosomal imbalance. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov J Genet Breed*. 2025;29(5):658-665. doi 10.18699/vjgb-25-72

Семейная транслокация между хромосомами 3 и 10: мейотическая сегрегация, диагностика и клинические проявления хромосомного дисбаланса

А.В. Возилова [1], А.С. Тарасова [1], Е.А. Иванов [1], В.П. Пушкарев [1], Н.И. Налетова [1], А.И. Побединская [1], А.С. Сабитова [1], Н.В. Шилова [2]

[1] Челябинская областная детская клиническая больница, Челябинск, Россия
[2] Медико-генетический научный центр им. академика Н.П. Бочкова, Москва, Россия
vozilova_a@mail.ru

Аннотация. Реципрокные транслокации являются наиболее частой структурной хромосомной перестройкой и встречаются в популяции с частотой 0.08–0.3 %. Большинство носителей реципрокных транслокаций фенотипически нормальны, но имеют повышенный риск привычного невынашивания беременности или рождения детей с нарушением интеллектуального развития и множественными врожденными аномалиями и/или поро-

© Vozilova A.V., Tarasova A.S., Ivanov E.A., Pushkarev V.P., Nalyotova N.I., Pobedinskaya A.I., Sabitova A.S., Shilova N.V., 2025
This work is licensed under a Creative Commons Attribution 4.0 License

ками развития вследствие патологической мейотической сегрегации хромосом, вовлеченных в транслокацию. В данной работе представлен семейный случай транслокации между дистальными участками короткого плеча хромосомы 3 и короткого плеча хромосомы 10, когда перестройка между хромосомами 3 и 10 была обнаружена у семи членов семьи в трех поколениях. Поводом для обследования семьи стало выявление при клиническом секвенировании экзома делеции сегмента p15 хромосомы 10 и дупликации сегмента p25 хромосомы 3 у пробанда с аномалиями фенотипа, что может соответствовать der(10)t(3;10)(p25;p15). Цитогенетическое исследование (GTG-окрашивание хромосом) членов семьи показало, что мать, а также бабушка, тетя и сибс пробанда без клинических и фенотипических аномалий являются носителями сбалансированной хромосомной перестройки – t(3;10)(p25;p15). У сибса пробанда, девочки с тяжелыми когнитивными, неврологическими нарушениями и аномалиями развития, определен кариотип 46,XX,der(3)t(3;10)(p25;p15)dmat. Молекулярное кариотипирование позволило уточнить размер хромосомного дисбаланса и точки разрыва на обеих хромосомах, вовлеченных в транслокацию. В статье представлено описание клинико-фенотипических особенностей при наличии в кариотипе дериватных хромосом 3 и 10. Обсуждаются механизм формирования хромосомного дисбаланса у членов семьи с аномалиями фенотипа, зависимость тяжести клинических проявлений от размера и генного состава обнаруженных хромосомных перестроек, а также необходимые мероприятия, направленные на предупреждение рождения ребенка с хромосомной патологией в семьях носителей аутосомных реципрокных транслокаций.

Ключевые слова: реципрокная транслокация; мейотическая сегрегация; хромосомный дисбаланс; дериватная хромосома 3; дериватная хромосома 10; GTG-окрашивание хромосом; клиническое секвенирование экзома; хромосомный микроматричный анализ; 3p делеционный синдром

Introduction

Reciprocal translocations (RT) involve the reciprocal exchange of genetic material between two chromosomes, with a breakpoint occurring in each chromosome. Such exchanges can be balanced (if no chromosomal material is lost or gained) or unbalanced (if there is a net loss or gain of genetic material in one or both chromosomes). RT is one of the most common structural chromosomal abnormalities, with an estimated frequency of 1 in 500 to 1 in 625 newborns (Ogilvie, Scriven, 2002). The population frequency of balanced translocation carriers ranges from 0.08 to 0.3 % (Kochhar, Ghosh, 2013).

As a rule, carriers of reciprocal translocations (RT) are phenotypically normal. However, their reproductive potential is often compromised by an increased risk of infertility, recurrent miscarriage, or the birth of children with intellectual disabilities and multiple congenital anomalies. This risk arises from the high likelihood of chromosomal imbalance in the offspring, resulting from aberrant meiotic segregation in RT carriers (Hu et al., 2016).

Chromosome segregation patterns are established during meiosis I, and in rare cases, may also result from errors in meiosis II. In RT carriers, the formation of bivalents between non-homologous chromosomes involved in the translocation becomes impossible during prophase I. Instead, a quadrivalent structure forms, ensuring complete homosynapsis between the rearranged chromosomes. During gametogenesis, three segregation patterns are possible – 2:2, 3:1, and 4:0 – reflecting the distribution of chromosomes from the quadrivalent to daughter gametocytes. The predominant segregation pattern is largely determined by the quadrivalent configuration, which itself depends on the breakage–reunion points in the rearranged chromosomes (Gardner, Amor, 2018).

Of the 32 theoretically possible gamete combinations resulting from meiotic segregation in reciprocal translocation (RT) carriers, only two produce genetically balanced gametes: those containing either both non-rearranged chromosomes or both derivative chromosomes (alternate 2:2 segregation pattern). All the other segregation patterns result in gametes with chromosomal imbalance. The 2:2 malsegregation patterns include: adjacent-1 segregation – produces gametes with partial trisomy/monosomy of the translocated segment; adjacent-2 segregation – leads to partial trisomy/monosomy of the centric segment. In 3:1 segregation, gametes with 22 or 24 chromosomes are formed. Resulting zygotes contain 45 or 47 chromosomes. Zygotes with 47 chromosomes (trisomic) demonstrate the highest viability among unbalanced outcomes. In 4:0 segregation gametes receive either all four chromosomes or none from the quadrivalent. Resulting zygotes exhibit either double trisomies or double monosomies. These zygotes are uniformly nonviable (Shilova, 2016).

The viability of carriers and the severity of clinical manifestations in cases of chromosomal imbalances depend on three key factors: the size of the imbalanced region, its chromosomal location, specific genes involved in the affected regions. Notably, translocations with terminal breakpoints demonstrate a significantly increased frequency of embryos with chromosomal imbalance. The terminal location of breakpoints represents an independent risk factor resulting in the birth of viable offspring with multiple congenital anomalies, chromosomal imbalance. Statistical analysis reveals that carriers of RTs with at least one terminal breakpoint (0.2 of the size of the respective chromosome arm and less) have a 6-fold increased risk of producing viable offspring with these adverse outcomes compared to RTs without terminal breakpoints (Shilova, 2019). When chromosomal imbalances affect genes critical for embryonic development, developmental arrest typically occurs either during early embryogenesis or later in prenatal development (Beyer et al., 2019). In cases where the imbalance is compatible with continued in utero development, gestation typically results in the birth of a child with congenital malformations and/or developmental abnormalities (Shilova, 2016).

This study investigates the phenotypic and genetic consequences of meiotic segregation patterns in translocations

between chromosomes 3 and 10, specifically involving their terminal regions, across three generations of a single family. We present: a clinical case of 3p deletion syndrome resulting from genomic imbalance in a female, with concurrent cases of 10p15 deletion syndrome in male and female cousins.

Materials and methods

Proband III-1, a boy born in 2006, was first evaluated by a clinical geneticist at the Chelyabinsk Regional Children's Clinical Hospital in 2009. In 2018, his newborn sister (III-5) and parents (II-1, II-2) underwent cytogenetic analysis. Six additional family members, including the proband's brother (III-3), maternal aunt (II-4), her two daughters (III-7, III-10), as well as the maternal grandmother (I-1) and grandfather (I-2), were examined in 2024 (Fig. 1).

Cytogenetic analysis was performed on GTG-banded metaphase chromosome preparations (550-band resolution) obtained from PHA-stimulated peripheral blood T-lymphocytes, following a standardized cytogenetic protocol (Medical Genetics, 2022).

Molecular and cytogenetic investigations were performed in an external laboratory. High-resolution chromosomal microarray analysis (CMA) was conducted using the Affymetrix CytoScan HD oligonucleotide microarray platform, following the manufacturer's protocol (Affymetrix, USA). Data analysis was performed using the Chromosome Analysis Suite (ChAS) software (v4.0). Clinical exome sequencing (CES) was performed via next-generation sequencing (NGS) with paired-end reads. Sequencing data were processed by aligning reads to the human reference genome (GRCh38/hg38). The DECIPHER database was utilized to assess genes within the chromosomal imbalance region for haploinsufficiency and triplosensitivity effects.

Cytogenetic and CMA results were interpreted according to the International System for Human Cytogenomic Nomenclature (ISCN 2024).

All studies involving human participants complied with the ethical guidelines of the National Committee for Research Ethics and the Declaration of Helsinki (1964, with later amendments). Written informed consent was obtained from all participants or their legal guardians.

Results

Male proband (III-1), born in 2006, was the product of an uncomplicated first pregnancy and delivery. His birth parameters included: weight: 2,600 g (1st–2nd centile, ~3 %), length: 51 cm (4th–5th centile, ~50 %), Apgar scores: 7/8. The boy exhibited significant psychomotor delay: head control achieved at 3 months, independent sitting at 10–11 months, ambulation at 2.5 years (developed progressively stiff gait). The first genetic assessment (2009) at Chelyabinsk Regional Children's Clinical Hospital revealed normal male karyotype (46,XY) on GTG-cytogenetics. Despite recommendations for annual follow-up, the family was lost to genetic surveillance for 9 years.

The patient was re-evaluated at the same institution due to progressive neurological deterioration and admitted to neurology service (2023). At age 7, the patient experienced significant motor regression – lost ambulation capacity (currently only able to crawl), markedly limited expressive language (5-word vocabulary). A history of seizure-like episodes was characterized by ocular squeezing, risus sardonicus (sustained grimacing), perioral cyanosis, respiratory distress, myoclonic jerks of extremities. Clinical exome sequencing (CES) was performed in 2024 to investigate progressive neurodevelopmental regression, complex seizure disorder, suspected underlying genetic etiology.

CES revealed that proband III-1 carried a 2,993,266 bp deletion on the short arm of chromosome 10 (chr10:179763–3173029), encompassing 27 genes, including 10 protein-coding genes. Among these, three were OMIM-annotated: *ZMYND11* (associated with autosomal dominant intellectual developmental disorder 30 [AD]), *WDR37* (linked to neurooculo-cardiogenitourinary syndrome [AD]), *PITRM1* (implicated in autosomal recessive spinocerebellar ataxia 30 [AR]). Additionally, a 9,809,749 bp duplication was detected

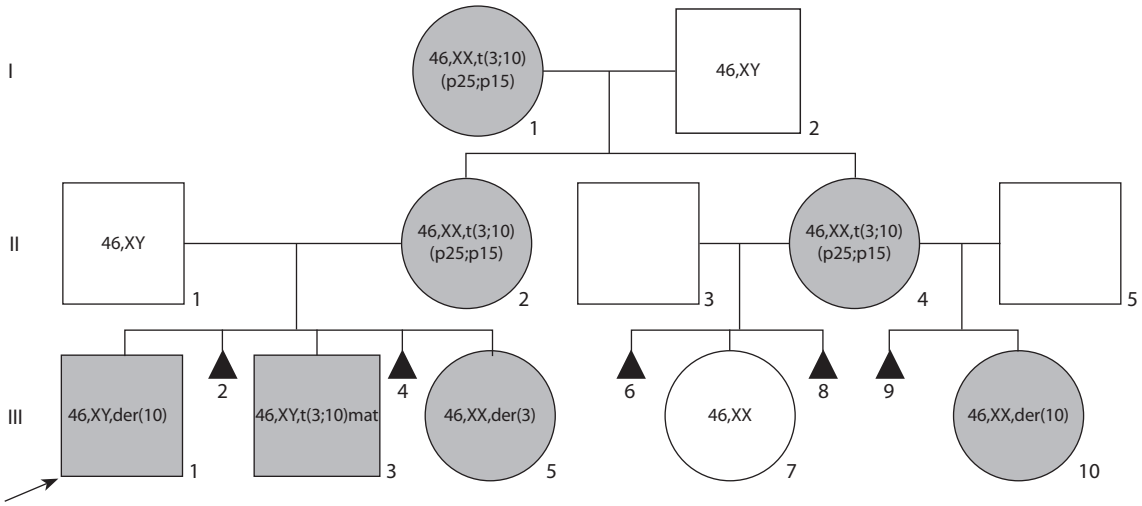


Fig. 1. Schematic representation of one and the same family members – carriers of the translocation t(3;10)(p25;p15) – examined across three generations (the arrow points to the proband).

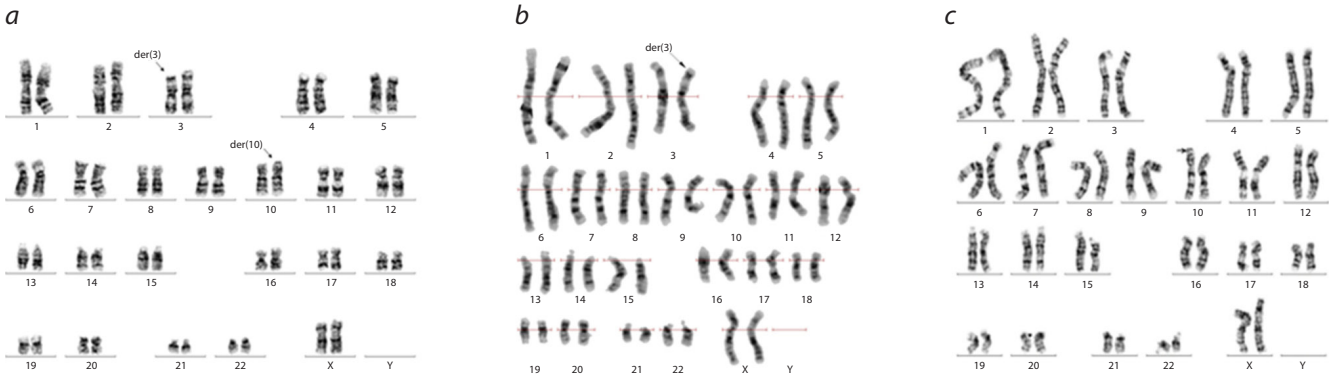


Fig. 2. Unbalanced and balanced variants of familial translocation between chromosomes 3 and 10.

a – karyotype of the mother II-2, reciprocal translocation t(3;10)(p25;p15); *b* – karyotype of the proband's sister III-5 46,XX,der(3)t(3;10)(p25;p15)dmat; *c* – karyotype of the proband's female cousin III-10 46,XX,der(10)t(3;10)(p25;p15)dmat). GTG-banding of the chromosomes.

on the short arm of chromosome 3 (chr3:2735965–12545714), spanning 115 genes (58 protein-coding ones). These findings imply that the proband carries a derivative chromosome 10 resulting from a translocation between chromosomes 3 and 10.

In 2018, the proband's parents sought medical genetic counseling for their three-month-old daughter (III-5). The girl was born from the mother's (II-2) fifth pregnancy, which occurred against a background of complicated obstetric-gynecological history. This was the mother's third delivery, resulting in a full-term cesarean section. The mother's previous pregnancies included: two live births – the proband (III-1, born in 2006) and a male sibling (III-3, born in 2012), two spontaneous abortions in early gestation (unknown etiology; no genetic or cytogenetic analysis was performed on the embryos). Prenatal findings in III-5 included marginal chorion presentation, intrauterine growth restriction (IUGR), congenital heart disease (CHD), musculoskeletal anomaly, polydactyly of the upper extremities. The gestation was complicated by chronic decompensated placental insufficiency and polyhydramnios. Birth parameters (III-5) were as follows: birth weight: 1,970 g (<1st centile, 3 %), length: 45 cm (1st centile, 3 %), head circumference: 35 cm (5th centile, 75 %), thoracic circumference: 30 cm (1st centile, 3 %). These anthropometric measurements indicate severe intrauterine growth restriction (IUGR). The Apgar score was 7. Clinical and phenotypic features included multiple congenital malformations, subclinical hyperthyroidism, iliac ectopia of the left kidney, flexion contractures of both thumbs, hip dysplasia, grade 2 cerebral ischemia, movement disorder suppression syndrome, ocular abnormalities: microphthalmia, microcornea. Congenital heart disease (CHD) included primum atrial septal defect (ASD), persistent left superior vena cava, patent foramen ovale (PFO), elongated Eustachian valve, grade 1 mitral regurgitation, group 1 pulmonary hypertension, circulatory failure (class 2A).

The cytogenetic investigation revealed derivative chromosome 3, karyotype 46,XX,der(3). Parental karyotyping clarified the origin of chromosomal anomaly: the mother (II-2) is the carrier of a reciprocal translocation – 46,XX,t(3;10)(p25;p15). The father (II-1) had a normal male karyotype – 46,XY. Thus, the karyotype of the proband's sister (III-5) was determined as 46,XX,der(3)t(3;10)(p25;p15)dmat (Fig. 2).

Molecular karyotyping of the proband's sister (III-5) precisely delineated the genomic imbalance and translocation breakpoints: arr[GRCh38] 3p26.3p25.2(11007_12547742)x1,10p15.3p15.2(45908_3500569)x3. Chromosome 3 deletion can be described as follows: size: 12,536,735 bp, genes affected: 132 (60 protein-coding ones), including 37 OMIM-annotated genes. Chromosome 10 microduplication was 3,454,661 bp in size, genes affected: 33 (11 protein-coding ones).

It can be concluded that the chromosomal imbalances observed in proband III-1 (karyotype: 46,XY,der(10)t(3;10)(p25;p15)mat), who inherited the derivative chromosome 10, and sibling III- 5, who inherited the derivative chromosome 3, result from meiotic adjacent-1 2:2 malsegregation of a maternal reciprocal translocation t(3;10)(p25;p15) (Fig. 3).

In 2024, the proband's parents sought genetic re-evaluation for their 12-year-old son (III-3) (born 2012), who exhibited no clinical abnormalities. G-banding revealed a balanced translocation, 46,XY,t(3;10)(p25;p15)mat, inherited from his mother (II-2).

The same year, the mother's sister (II-4) consulted medical geneticist with her two daughters: a phenotypically normal older daughter (III-7, born 2006) (46,XX) and a younger daughter (III-10, born 2017) presented with delayed motor development (gait instability, limited speech) and dysmorphic features. The karyotype of the younger girl showed an unbalanced derivative chromosome 10: 46,XX,der(10)t(3;10)(p25;p15)mat. Molecular karyotyping of the mother (II-2) revealed a balanced reciprocal translocation t(3;10)(p25;p15). Parental karyotyping traced the translocation to the grandmother (I-1), confirming a familial balanced translocation, 46,XX,t(3;10)(p25;p15) (Fig. 1).

Thus, in the presented family the daughters II-2, II-4 and the grandson III-3 inherited balanced translocation 46,XX,t(3;10)(p25;p15) from the grandmother I-1. As a result of meiotic adjacent-1 malsegregation proband III-1 and his cousin III-10

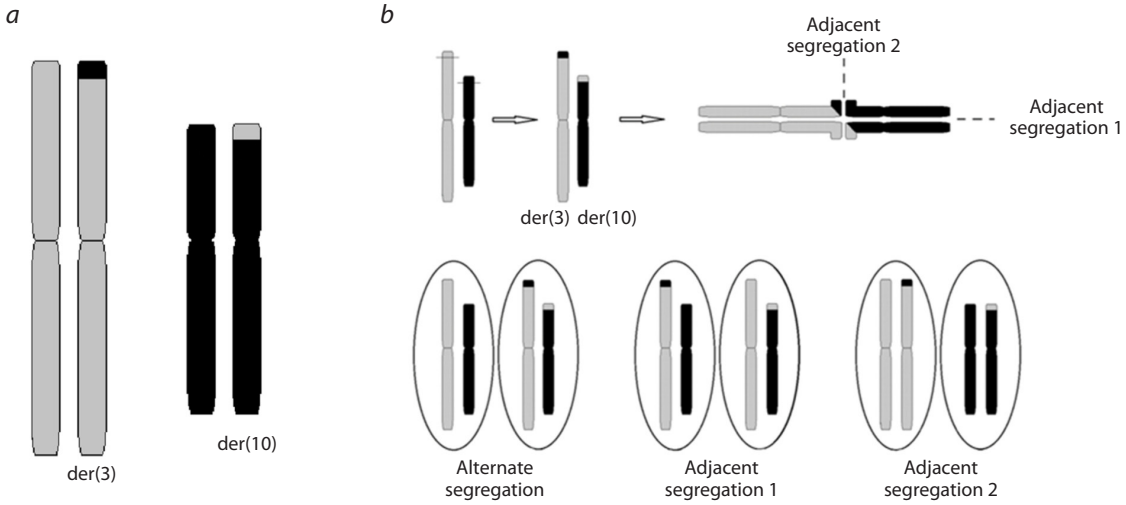


Fig. 3. Schematic representation of: *a* – reciprocal translocation formation between chromosomes 3 and 10, t(3;10)(p25;p15); *b* – variants of meiotic segregation patterns of chromosomes with translocation.

had unbalanced karyotype with derivative chromosome der(10) and the karyotype of proband's sister III-5 with the severe clinical manifestations contains derivative chromosome der(3) (Fig. 3).

Discussion

Chromosome banding analysis at a 550-band resolution per haploid set enables the detection of chromosomal abnormalities in 3–10 % of patients with intellectual disability. GTG-banding remains the gold standard for diagnosing chromosomal imbalances exceeding 8–10 Mb in size (Lebedev et al., 2023). However, as with any method relying on subjective interpretation, the accuracy of results depends heavily on the cytogeneticist's expertise, chromosome spreading quality, and banding clarity. Unfortunately, the derivative chromosome 10 (der(10)) in the proband (III-1) was not detected during initial cytogenetic analysis in 2009, significantly prolonging the "diagnostic odyssey" of this family.

As shown in Figures 1 and 3, four family members exhibited a balanced/normal karyotype, resulting from maternal alternate meiotic segregation of a familial reciprocal translocation t(3;10). In contrast, three individuals – the proband (III-1) and his cousin (III-10) (both with der(10)), as well as the proband's sister (III-5, with der(3)) – had unbalanced karyotypes due to adjacent-1 malsegregation (2:2 segregation). This suggests that balanced and unbalanced gametes (zygotes) are formed at near-equal frequencies (4:3) during meiotic segregation of this translocation. Literature analysis indicates that meiotic malsegregation occurs in ~30 % of reciprocal translocation cases, with adjacent-1 segregation being the most prevalent (~80 %). Adjacent-2 segregation is observed in ~13 % of cases, while tertiary (exchange) segregation accounts for ~7 % (Shilova, 2016).

Given that both mothers (II-2 and II-4) had a history of spontaneous abortions (the grandmother's obstetric history is unknown), we hypothesize that these pregnancy losses may represent embryos with chromosomal imbalances resulting from either meiotic adjacent-2 segregation of the familial translocation, or aneuploidy involving chromosomes 10 and 3 (which would be incompatible with live birth).

Presumably, in this reciprocal translocation, only conceptuses with chromosomal imbalances arising from adjacent-1 segregation (2:2) appear viable. Literature evidence suggests that when breakpoints occur in terminal chromosomal regions (as in our case), the likelihood of live-born children with multiple congenital anomalies and/or chromosomal imbalances increases sixfold (Shilova et al., 2019). In the present study, both breakpoints were located in terminal regions, which likely contributed to the birth of affected children with severe phenotypic manifestations.

The limited family pedigree does not permit a statistically robust assessment of the empirical recurrence risk for chromosomal imbalance in offspring. Nevertheless, our findings indicate a substantial observed risk (3/7, or 43 %) of unbalanced outcomes. This information is particularly relevant for the proband's brother (III-3), a balanced carrier of t(3;10), and should inform both clinical counseling and his future reproductive planning. Notably, among the grandmother's five grandchildren, only III-3 and one cousin (III-7) show no clinically significant phenotypic abnormalities. The other three (III-1, III-5, and III-10) exhibit developmental defects and multiple congenital anomalies resulting from genomic imbalance. While modern reproductive technologies (such as PGD) could mitigate this risk, the high probability of unbalanced segregation warrants heightened clinical vigilance regarding the reproductive choices of translocation carrier III-3. For instance, longitudinal follow-up of families with identified chromosomal rearrangements (diagnosed 15–34 years prior) revealed two significant findings: in four families, parents had no recollection of their children's previous cytogenetic diagnoses, and in four additional families (representing approximately 10 % of the study cohort), parents failed to comprehend the clinical implications of the karyotyping results (Bache et al., 2007).

Clinical and phenotypic manifestations in children with unbalanced genome formed during meiotic segregation t(3;10)(p25;15)

Physical features of 3p25-pter deletion (Malmgren et al., 2007; Fu et al., 2021)	**III-5 Proband's sister** Karyotype, CMA: 3p25-pter deletion, 10p15-pter duplication	**III-1 Proband** CES: 3p25-pter duplication, 10p15-pter deletion	**III-10 Proband's female cousin** Karyotype: 3p25-pter duplication, 10p15-pter deletion
Low birthweight	+	No	No
Developmental delay	+ No speech or its understanding	+ Says up to ten words, no self-care	+ Indistinct speech
Growth retardation/ abnormality	+	+	+
Locomotor activity	No Did not walk: knee-joint contracture, joints of the hand contracture	Independent walking from the age of 2.5 years up to the age of 7, currently only crawling	Independent walking, gait stiffness increases
Feeding problems	+	+	+
Hypotonia	+	+	+
Microcephaly	+	No (macrocephaly observed)	No (macrocephaly observed)
Blepharophimosis	+	No (ophthalmoptosis observed)	No
Ocular hypertelorism/ Hypertellrism	No	+	+
Flat nasal bridge	+	No	+
Polydactyly	+	No	No
Synophrys	+	No	No
Micrognathia	+	No	No
Low set ears	+ Plump earlobes/earlobe fullness	+	+

Note. Children with similar chromosomal imbalance are marked with a background.

Molecular characterization using chromosomal microarray analysis (CMA) and clinical exome sequencing (CES) in the proband (III-1) and his affected sister (III-5) enabled precise mapping of translocation breakpoints, identification of genes contained within the unbalanced chromosomal regions. Based on the karyotyping results, we hypothesized that the proband (III-1) and his female cousin (III-10) carried a genomic imbalance involving a 10p15-pter deletion and a 3p25-pter duplication, indicating the presence of a derivative chromosome 10 (der(10)). In contrast, chromosomal microarray analysis (CMA) of the proband's sister (III-5) identified a derivative chromosome 3 (der(3)) with a terminal 3p25-pter deletion (12.5 Mb) and a 10p15-pter duplication (3.5 Mb) (see the Table).

Although all patients with chromosomal imbalances in our cohort exhibited severe cognitive and physical impairments, detailed clinical evaluation revealed that the phenotypes of the two der(10) carriers were similar but different from that of the proband's sister (III-5) with der(3).

The terminal deletion of the 10p15-pter region was first reported by D. Elliott et al. (1970). To date, approximately 50 cases of 10p15-pter deletions with varying lengths have been described in the literature. The core clinical features associated with 10p15-pter deletions include cognitive impairment, behavioral abnormalities, speech delay, locomotor dysfunction, craniofacial dysmorphism, hypotonia, brain malformations, and seizures. These features are attributed to haploinsufficiency of the *ZMYND11* (OMIM 608668) and *DIP2C* (OMIM 611380) genes (DeScipio et al., 2012); both are located within the deleted region identified in proband III-1 and his female cousin III-10, who exhibited characteristic clinical manifestations. Notably, the phenotypic spectrum in these cases may reflect not only the 10p deletion but also the concurrent 3p25-pter duplication, which encompasses 13 protein-coding genes known to be triplosensitive. The patients require specialized neurorehabilitation, particularly the female cousin III-10, who demonstrates progressive locomotor deterioration similar to proband III-1.

In contrast, patient III-5 (the proband's sister) exhibited severe clinical features consistent with 3p deletion syndrome, further supporting the pathogenicity of the 3p25-pter deletion identified in her case (Verjaal, De Nef, 1978; Malmgren et al., 2007; Fu et al., 2021) (see the Table). Deletions of the terminal 3p region represent a rare chromosomal abnormality associated with characteristic phenotypic features, including microcephaly, ptosis, hypertelorism, and micrognathia. Affected individuals typically exhibit low birth weight, hypotonia, intellectual disability, developmental delay, delayed bone maturation, and renal anomalies. Congenital heart defects – particularly atrioventricular septal defects – occur in approximately one-third of cases (Martins et al., 2021). In our study, the proband's sister (III-5) had severe manifestations from birth, including low birth weight, complete absence of locomotor activity and speech, microcephaly, polydactyly, synophrys, and micrognathia. Notably, these features were absent in the proband (III-1) and his female cousin (III-10), highlighting the phenotypic divergence between the two genomic imbalances. Of particular interest is the contrasting cranial growth patterns observed: microcephaly in III-5 (with 3p25-pter deletion) versus macrocephaly in cases with 3p25-pter duplication. This reciprocal phenotype likely reflects dosage sensitivity of genes within this region, underscoring the critical role of gene copy number in neurodevelopment and craniofacial morphogenesis.

As previously noted, the 3p25-pter deletion region encompasses 132 genes, including 25 morbid genes associated with clinical phenotypes. Among these, haploinsufficiency of *SETD5*, *BRPF1*, *CRBN*, *ATG7*, *SLC6A11*, *GRM7*, and *ARPC4* has been linked to neurodevelopmental disorders and cognitive impairment. Notably, the region also includes *CHL1*, a candidate gene for nonspecific intellectual disability due to its high expression in the developing brain (Martins et al., 2021; Tsuboyama, Iqbal, 2021). The concurrent 10p15-pter duplication in this patient may further contribute to the abnormal phenotype, as this region contains two triplosensitive genes: *LARP4B* and *DIP2C*. Of these, *DIP2C* has been specifically associated with developmental delay and speech impairment, suggesting a potential additive or synergistic effect of the dual genomic imbalance.

Conclusion

The 14-year clinical odyssey of this three-generation family enabled the identification of carriers with both the balanced reciprocal translocation t(3;10) and derivative chromosomes resulting from adjacent-1 meiotic 2:2 malsegregation. Breakpoints were precisely mapped using high-resolution genomic techniques – chromosomal microarray analysis (CMA) and clinical exome sequencing (CES). Variability in the size and gene content of the imbalanced regions correlated with the severity of phenotypic abnormalities among affected individuals. Notably, patients sharing similar genomic imbalances exhibited comparable and progressively worsening clinical manifestations, underscoring the necessity for multidisciplinary care involving neurologists and rehabilitation specialists.

Chromosomal imbalances involving concurrent terminal deletions and duplications of non-homologous chromosomes typically arise from meiotic segregation errors in parental reciprocal translocations. Thus, parental karyotyping is essential to identify translocation carriers. Although the proband's brother (III-3) with the balanced translocation remains asymptomatic, genetic counseling is critical to inform him of his reproductive risks. Advanced preimplantation (PGT) and prenatal diagnostic methods can significantly reduce the likelihood of transmitting severe genomic disorders to his offspring.

References

Bache I., Brondum-Nielsen K., Tommerup N. Genetic counseling in adult carriers of a balanced chromosomal rearrangement scertained in childhood: experiences from a nationwide reexamination of translocation carriers. *Genet Med.* 2007;9(3):185-187. doi 10.1097/gim.0b013e3180314671

Beyer C.E., Lewis A., Willats E., Mullen J. Preimplantation genetic testing using Karyomapping for a paternally inherited reciprocal translocation: a case study. *J Assist Reprod Genet.* 2019;36(5):951-963. doi 10.1007/s10815-019-01413-0

DeScipio C., Conlin L., Rosenfeld J., Tepperberg J., Pasion R., Patel A., McDonald M., … Martinet D., Mason T., Clark D., Spinner N., Krantz I. Subtelomeric deletion of chromosome 10p15.3: clinical findings and molecular cytogenetic characterization. *Am J Med Genet A*. 2012;158A(9):2152-2161. doi 10.1002/ajmg.a.35574

Elliott D., Thomas G.H., Condron C.J., Khuri N., Richardson F. C-group chromosome abnormality (?10p-). Occurrence in a child with multiple malformations. *Am J Dis Child*. 1970;119(1):72-73. doi 10.1001/archpedi.1970.02100050074017

Fu J., Wang T., Fu Z., Li T., Zhang X., Zhao J., Yang G. Case report: a case report and literature review of 3p deletion syndrome. *Front Pediatr.* 2021;9:618059. doi 10.3389/fped.2021.618059

Gardner R.J.M., Amor D.J. Gardner and Sutherland's Chromosome Abnormalities and Genetic Counseling. Oxford Univ. Press, 2018

Hu L., Cheng D., Gong F., Lu C., Tan Y., Luo K., Wu X., He W., Xie P., Feng T., Yang K., Lu G., Lin G. Reciprocal translocation carrier diagnosis in preimplantation human embryos. *EBioMedicine*. 2016;14:139-147. doi 10.1016/j.ebiom.2016.11.007

ISCN 2024 – An International System for Human Cytogenomic Nomenclature (2024). *Cytogenet Genome Res*. 2024;164(Suppl.1):1-224. doi 10.1159/000538512

Kochhar P.K., Ghosh P. Reproductive outcome of couples with recurrent miscarriage and balanced chromosomal abnormalities. *J Obstet Gynaecol Res.* 2013;39(1):113-120. doi 10.1111/j.1447-0756.2012.01905.x

Lebedev I.N., Shilova N.V., Iourov I.Yu., Malysheva O.V., Tveleneva A.A., Minzhenkova M.E., Markova Zh.G., Tolmacheva E.N., Kashevarova A.A. Guidelines of the Russian Society of Medical Geneticists for chromosomal microarray analysis. *Medicinskaya Genetika = Medical Genetics*. 2023;22(10):3-47. doi 10.25557/2073-7998.2023.10.3-47 (in Russian)

Malmgren H., Sahlén S., Wide K., Lundvall M., Blennow E. Distal 3p deletion syndrome: detailed molecular cytogenetic and clinical characterization of three small distal deletions and review. *Am J Med Genet Part A.* 2007;143A(18):2143-2149. doi 10.1002/ajmg.a.31902

Martins M., Arantes R., Botelho P., Souto M., Moutinho O., Pinto Leite R. Familiar del 3p syndrome: the uncertainty of the prognosis. A case report. *Clin Case Rep.* 2021;9(4):2365-2368. doi 10.1002/ccr3.4036

Medical Genetics: a national guide. Moscow: GEOTAR-Media Publ., 2022 (in Russian)

Ogilvie C.M., Scriven P.N. Meiotic outcomes in reciprocal translocation carriers ascertained in 3-day human embryos. *Eur J Hum Genet.* 2002;10(12):801-806. doi 10.1038/sj.ejhg.5200895

Shilova N.V. Improving Approaches to the Diagnosis of Chromosomal Aberrations in the Context of Personalized Medicine. Extended abstract of the Dr. med. sci. diss. Moscow, 2016. Available at: https://viewer.rusneb.ru/ru/000199_000009_006659654?page=1&rotate=0&theme=white (in Russian)

Shilova N.V., Minzhenkova M.E., Antonenko V.G. Evaluation of risk factors for the birth of children with chromosomal imbalance in carriers of autosomal reciprocal translocations. *Russ J Genet.* 2019; 55(9):1103-1112. doi 10.1134/S1022795419090151

Tsuboyama M., Iqbal M.A. *CHL1* deletion is associated with cognitive and language disabilities – case report and review of literature. *Mol Genet Genomic Med*. 2021;9(7):e1725. doi 10.1002/mgg3.1725

Verjaal M., De Nef M.B. A patient with a partial deletion of the short arm of chromosome 3. *Am J Dis Child.* 1978;132(1):43-45. doi 10.1001/archpedi.1978.02120260045012

Conflict of interest. The authors declare no conflict of interest.
Received February 15, 2025. Revised May 23, 2025. Accepted July 2, 2025.

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2025;29(5):666-675

doi 10.18699/vjgb-25-73

Transposable elements as key regulators of placental development

M.A. Zhilkina, E.N. Tolmacheva, S.A. Vasilyev

Research Institute of Medical Genetics, Tomsk National Research Medical Center of the Russian Academy of Sciences, Tomsk, Russia
maria.amelkina@medgenetics.ru

Abstract. Transposable elements (TEs), comprising over one-third of the human genome, play a crucial role in its evolution, serving as a significant source of regulatory sequences. Under normal circumstances, their activity is tightly controlled by DNA methylation mechanisms; however, the effectiveness of this suppression varies substantially across tissues. The placenta, characterized by global hypomethylation, represents a unique environment where retroviruses and retrotransposons, typically silenced in somatic cells, gain the opportunity for activation. This distinct epigenetic landscape of the placenta allows transposons to participate in the regulation of genomic activity, influencing processes ranging from early embryogenesis to postnatal development. DNA hypomethylation in the placenta not only promotes TE mobilization, but also opens the possibility of using their components as independent genes and regulatory elements – promoters, enhancers, and other functional modules. These elements are involved in key aspects of placental development, including syncytiotrophoblast formation, extravillous trophoblast invasion, spiral artery remodeling, and endometrial decidualization. Importantly, TEs can serve as sources of alternative promoters for neighboring genes, and ancient mammalian transposons contain multiple transcription factor binding sites, enabling coordinated regulation of genes sharing a common function. Despite the growing interest in the role of transposable elements in placental development and function, many questions remain unanswered. In particular, the mechanisms of non-long terminal repeat (non-LTR) retrotransposon function during pregnancy remain poorly understood. A deep understanding of these processes is necessary to elucidate regulatory disorders in the placenta associated with major obstetric syndromes. This review examines the contribution of transposable elements to the functioning of the human genome, particularly their impact on gene expression, in the context of pregnancy and placental development.

Key words: transposable elements; retrotransposons; retroviruses; placenta development

For citation: Zhilkina M.A., Tolmacheva E.N., Vasilyev S.A. Transposable elements as key regulators of placental development. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov J Genet Breed*. 2025;29(5):666-675. doi 10.18699/vjgb-25-73

Funding. This work was supported by the Russian Science Foundation (project No. 24-25-00335).

Мобильные элементы как ключевые регуляторы развития плаценты

М.А. Жилкина, Е.Н. Толмачёва, С.А. Васильев

Научно-исследовательский институт медицинской генетики Томского национального исследовательского медицинского центра Российской академии наук, Томск, Россия
maria.amelkina@medgenetics.ru

Аннотация. Мобильные элементы (transposable elements, TE), составляющие свыше трети человеческого генома, играют ключевую роль в его эволюции, выступая важным источником регуляторных последовательностей. В норме их активность жестко контролируется механизмами метилирования ДНК, однако эффективность такого подавления существенно различается между тканями. Плацента, отличающаяся глобальным гипометилированием, представляет собой уникальную среду, где ретровирусы и ретротранспозоны, обычно молчащие в соматических клетках, получают возможность активации. Этот особый эпигенетический ландшафт плаценты позволяет транспозонам участвовать в регуляции геномной активности, оказывая влияние на процессы, протекающие от раннего эмбриогенеза до постнатального развития. Гипометилирование ДНК в плаценте не только способствует мобилизации TE, но и открывает возможность использования их компонентов в качестве самостоятельных генов и регуляторных элементов – промоторов, энхансеров и других функциональных модулей. Эти элементы вовлечены в ключевые аспекты плацентарного развития, включая формирование синцитиотрофобласта, инвазию вневорсинчатого трофобласта, ремоделирование спиральных артерий и децидуализацию эндометрия. Важно отметить, что TE могут служить источниками альтернативных промоторов для соседних генов, а древние транспозоны млекопитающих содержат множественные сайты связывания транскрипционных факторов, обеспечивая скоординированную регуляцию генов, объединенных общей функцией. Несмотря на растущий интерес к роли мобильных

© Zhilkina M.A., Tolmacheva E.N., Vasilyev S.A., 2025
This work is licensed under a Creative Commons Attribution 4.0 License

элементов в развитии и функционировании плаценты, многие вопросы остаются без ответа. В частности, малоизученными продолжают быть механизмы функционирования в ходе беременности ретротранспозонов, не содержащих длинных концевых повторов (non-LTR ретротранспозонов). Глубокое понимание этих процессов необходимо для прояснения нарушений регуляции в плаценте при больших акушерских синдромах. В данном обзоре рассматривается вклад мобильных элементов в функционирование генома человека, в частности их влияние на экспрессию генов, в контексте беременности и развития плаценты.
Ключевые слова: мобильные элементы; ретротранспозоны; ретровирусы; развитие плаценты

Introduction

Approximately 40 % of the mammalian genome is comprised of mobile genetic elements called “transposons” (TEs) (Chesnokova et al., 2022). At first glance, such an abundance of TEs in mammalian genomes seems paradoxical, given the potential risks associated with uncontrolled transposition (Doolittle, Sapienza, 1980). However, this coexistence reflects an ongoing evolutionary arms race between TEs and their hosts, resulting in a dynamic equilibrium. Although most mammalian TEs have been inactivated through mutations or transcriptional/post-transcriptional silencing, there are exceptions. Some TE/host interactions, initially driven by the need for TE replication, can be repurposed to perform important functions in host development or physiology.

In recent decades, it has become clear that such adaptation of mobile genetic element sequences to perform new functions in the host genome is a crucial step in their evolution. J. Brosius and S.J. Gould (1992) made a significant contribution to understanding this process, challenging the view of mobile genetic elements solely as “junk DNA” and proposing that TEs should be considered a source of evolutionary innovation through the mechanism of exaptation – the repurposing of existing genetic elements to perform new functions. It is important to note that while adaptation involves the refinement of features under the direct selection for their current function, exaptation describes the use of pre-existing traits for entirely new purposes (Brosius, Gould, 1992). As such, mobile genetic elements (such as transposons and retrotransposons) can take on biologically significant roles in gene regulation, formation of new functional elements of the genome or its structure (Chuong et al., 2016).

The reproductive strategy of placental mammals, characterized by intrauterine development and prolonged lactation, requires significant energetic and metabolic expenditures on the part of the maternal organism (Hamilton, Boyd, 1960). Under such substantial maternal costs, natural selection predictably favors the development of mechanisms for early elimination of non-viable embryos in the early stages of ontogenesis. From this perspective, the matter of the preservation of mobile elements in the genome, despite their potentially destructive effects and the strong action of selection, is of scientific interest.

The evolutionary persistence of TEs can be explained by their strategic integration into key processes that determine organism viability at critical stages of development. These fundamental processes include an activation of the embryonic genome, a successful embryo implantation, and placentation. The effectiveness of this strategy is supported by the large-scale invasion of TEs into mammalian genomes.

The features of epigenetic regulation in the placenta, such as global DNA hypomethylation and the presence of partially methylated domains of extended genomic regions with intermediate levels of methylation (Novakovic, Saffery, 2013), provide unique conditions for the activation of endogenous retroviruses and retrotransposons that are repressed in most somatic tissues (Honda, 2016). The brevity of existence and temporary nature of placenta as an organ further explains the specificity of its epigenome organization.

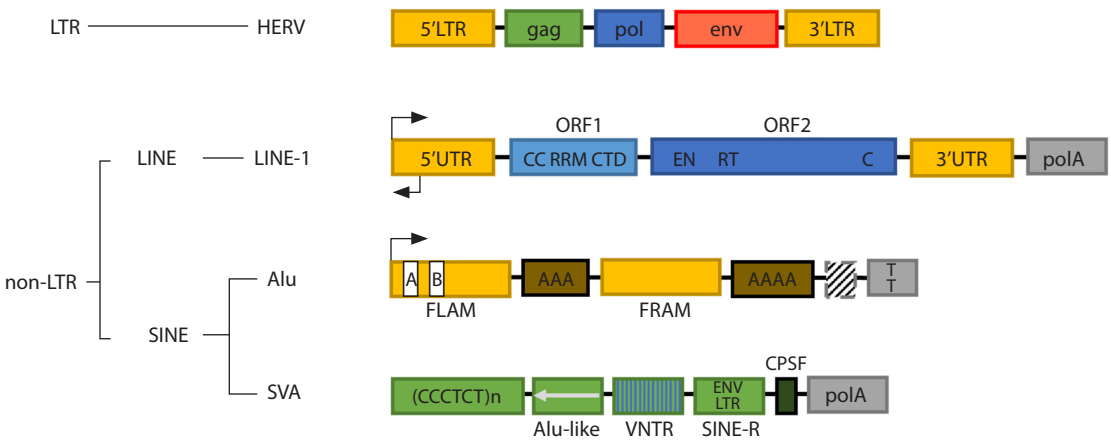
This review attempts to systematize current data on the functional significance of mobile elements in placental development and function.

Transposable elements in mammalian genomes

Mobile genetic elements are highly abundant in mammalian genomes. While previously considered “junk DNA”, their significant influence on host genome function is now a well-established fact. According to current data, approximately 50 % of the human genome is comprised of retrotransposons and DNA transposons (de Koning et al., 2011).

From the perspective of molecular transposition mechanisms, all transposable elements (TEs) are divided into two main classes (Wicker et al., 2007). The first class groups the elements called “retrotransposons”. For movement, these elements utilize an RNA intermediate followed by reverse transcription via a “copy-and-paste” mechanism while preserving the original sequence intact (Mustafin, 2018). The second class is represented by DNA transposons, which transpose genes without RNA involvement via a “cut-and-paste” mechanism (TIR and Cryptons) or through replicative transposition (Helitrons and Mavericks) (Mustafin, 2018). Retrotransposons, in turn, are classified into five orders based on their molecular organization, transposition mechanisms, and reverse transcriptase phylogeny: endogenous retroviruses (ERVs) with long terminal repeats (LTR-retrotransposons), LINE and SINE elements, DIRS-like elements, and Penelope-like elements (Wicker et al., 2007).

Typical ERVs contain three conserved coding domains (gag, env, pol) and are flanked by identical long terminal repeats (LTRs) on both sides (see the Figure). However, over the course of vertebrate evolution, most ERVs have acquired multiple mutations, resulting in the loss of their ability to fully express viral proteins (Johnson, 2019). The human-specific group of LTR-containing retrotransposons is commonly referred to as HERVs (human endogenous retroviruses).



LTR-containing and non-LTR retrotransposons in the human genome.

For HERV retrotransposons: long terminal repeat (LTR) (yellow blocks), and gag (green block), env (red block), pol (blue block) encoding domains. For LINE-1 retrotransposon: untranslated regions (UTR) (yellow blocks); sense and antisense internal promoters (black arrows); ORF1 includes a coiled-coil (CC) domain, RNA recognition motif (RRM) and C-terminal domain (CTD); ORF2 includes endonuclease (EN), reverse transcriptase (RT) and cysteine-rich (C) domains; poly(A) tail (polA follows 3′ UTR). For Alu: FLAM (free left Alu monomer); FRAM (free right Alu monomer); RNA polymerase III transcription start site (black arrow) and conserved cis-acting sequences required for transcription (white blocks A and B in the left Alu monomer); adenosine-rich fragment (brown block AAA between the left and right Alu monomers); terminal poly(A) tail (brown block AAAA); flanking genomic DNA of variable size (hatched gray block), followed by the pol III RNA termination signal (gray block TT). For human SVA: CCCTCT hexameric repeat; inverted Alu-like repeat (green block with a reverse arrow); GC-rich VNTR (hatched green block); SINE-R sequence homologous with HERV-K10 (ENV and LTR regions); cleavage and polyadenylation specificity factor (CPSF) binding site; terminal poly(A) tail (polA) (as per Lee et al., 2024).

Mobile genetic elements lacking long terminal repeats (non-LTR) are primarily represented by two classes: long interspersed nuclear elements (LINEs) and short interspersed nuclear elements (SINEs). LINE elements have a length of several thousand nucleotide pairs, whereas the size of SINEs usually does not exceed 600 base pairs (Kramerov, Vassetzky, 2011; Bourque, 2018). The fundamental difference between these groups lies in their transcription mechanisms: LINEs, like LTR-retrotransposons, are expressed by RNA polymerase II, while most SINEs are transcribed with the involvement of RNA polymerase III (Kramerov, Vassetzky, 2011).

SINE elements exhibit exceptionally high abundance in mammalian genomes, exceeding 100,000 copies. They replicate employing a retrotransposition mechanism based on the "copy-and-paste" sequential transcription into RNA, reverse transcription to form cDNA, and integration into new genomic loci. This process is entirely dependent on the enzymatic machinery encoded by LINE elements. In the human genome, the most prevalent SINE family is Alu – 300-nucleotide sequences that evolved from 7SL RNA (Lee et al., 2024).

LINE-1 elements, constituting approximately 17 % of the human genome (Chesnokova et al., 2022), have a complex structure. Full-length functional copies, of which there are approximately one thousand, contain untranslated regions (UTRs) necessary for transpositional activity. These elements include a 5′-UTR with a unique bidirectional promoter, two open reading frames (ORF1 and ORF2) encoding ORF1p and ORF2p proteins, and a 3′-UTR with a polyadenylation signal (see the Figure). Of particular interest is the organization of the LINE-1 promoter region, containing both a sense promoter regulating the expression of retrotransposition proteins and an antisense promoter (ASP) (Lee et al., 2024).

Regulation of retrotransposon activity in mammalian somatic cells is critical for maintaining genomic stability. Numerous studies confirm the key role of epigenetic mechanisms, particularly DNA methylation, in suppressing the potentially hazardous transpositional activity of these elements (Slotkin, Martienssen, 2007). This control mechanism act as an important protective barrier, preventing the development of genomic disorders and associated pathological conditions.

Unique epigenetic landscape of the placenta

The placenta is characterized by global DNA hypomethylation, which distinguishes its epigenetic profile from somatic tissues (Ehrlich et al., 1982). The average level of 5-methylcytosine in human placental tissue is 2.5–3 %, whereas in umbilical cord blood it reaches ~4 % (Price et al., 2012). The given epigenetic status is a key factor in regulating expression of genes, controlling placental growth and trophoblast functional activity (Robinson, Price, 2015).

The premise of placental hypomethylation is epigenetic reprogramming, a key feature of which in the zygote and embryo at the preimplantation stage of development is the loss of DNA methylation, such that the late morula/early blastocyst exhibits the lowest level of DNA methylation compared to any other period of ontogenesis. Subsequent

de novo methylation in the inner cell mass is accompanied by TE repression, while placenta-forming trophectoderm cells maintain a hypomethylated state of these elements (Price et al., 2012).

Although the functional role of the reduced level of genome methylation observed in the placenta is still not fully understood, studies show that it can activate the expression of mobile elements that is normally suppressed in other tissues (Macaulay et al., 2011). DNA methylation of HERV families in the placenta exhibits widely varying, but on average reduced levels compared to embryonic and adult tissues (Reiss et al., 2007). In contrast, the average DNA methylation index of Alu is similar in placental and fetal tissues (Price et al., 2012; Rondinone et al., 2021), and DNA methylation of the LINE-1 retrotransposon is reduced and more variable in the placenta compared to fetal tissues.

However, a decrease in TE methylation levels does not always lead to an increase in their transcriptional activity. For example, in a recent study, S. Lanciano et al. found that only a small number of copies of young L1s are activated upon decreased DNA methylation in the genome, whereas most hypomethylated L1 loci unexpectedly remain silent (Lanciano et al., 2024). The promoters of young active L1 elements are hypomethylated in human embryonic stem cells compared to differentiated cells, which partially explains their higher level of expression.

Hypermethylation of LINE-1 in the placenta has also been reported in some pregnancy pathologies. Hydatidiform mole is one cause of pregnancy loss and the most common type of gestational trophoblastic disease. In patients with hydatidiform mole, a twofold increase in LINE-1 methylation levels was observed throughout placental development and differentiation, whereas the level of overall genome methylation and other repeats remained the same in this pathology (Lou et al., 2020). In spontaneous abortions with aneuploidy, increased LINE-1 methylation was observed in extraembryonic tissues (Vasilyev et al., 2021). However, at the same time, LINE-1 is hypomethylated in extraembryonic tissues of spontaneous abortuses with a normal karyotype, which can lead to enhanced LINE-1 activation and subsequent mutational insertions (Lou et al., 2020).

An example of the influence of hypomethylation on the activity of mobile elements in the placenta is the hypomethylation of the AluY retrotransposon in the *KCNH5* locus. A differentially methylated region in the promoter region and first exon of transcript 1a of the *KCNH5* gene is of retrotransposon origin: 147 bp of the promoter and 162 bp of the exon evolved from a SINE-element of the AluY family. This element, which first appeared in the primate genome about 25–30 million years ago, has been preserved only in humans, great apes, and Old World monkeys, indicating its recent (on an evolutionary scale) integration. Hypomethylation of AluY in the placenta correlates with activation of an alternative *KCNH5* transcript, demonstrating how epigenetic modification of mobile elements can participate in tissue-specific gene regulation (Macaulay et al., 2011).

Functional exaptation of mobile genetic elements in the placenta

Low levels of DNA methylation in the placenta have facilitated the use of TE parts as functional regulatory sequences. In particular, TEs have been integrated into placenta-specific enhancers, alternative promoters, and other cis-regulatory elements, contributing to the evolutionary diversification of placental functions (Hoyt et al., 2022).

TE derivatives play an important role in various processes, including altering splicing patterns, enhancing recombination, forming enhancer and silencer regions, utilizing alternative promoters, and gene neofunctionalization (Brosius, 1999). The regulatory activity of TEs is apparent as early as the blastocyst stage and is maintained throughout mammalian prenatal development, including in the placenta. Certain integrated retroviral sequences have evolved into critically important regulatory elements, modulating the expression of neighboring genes or even forming novel gene loci (Johnson, 2019).

The genes *ERVW-1* (syncytin-1) and *ERVFRD-1* (syncytin-2) are classical examples of HERV elements that have undergone exaptation, acquiring placenta-specific functions (Macaulay et al., 2011). They have retained the ability to encode envelope (env) proteins, which typically mediate viral entry into cells (Nelson et al., 2003). However, in the placenta, these proteins have acquired a novel physiological function: they mediate the differentiation and fusion of cytotrophoblast cells, leading to the formation of the multinucleated syncytiotrophoblast (Pötgens et al., 2002).

The syncytin family, derived from HERVs, is a unique group of fusogenic proteins that play a crucial role in placental morphogenesis. Experimental data indicate that the surface SU domain of these proteins is essential for cell fusion, as evidenced by its inhibition with specific antibodies (Shimode, 2023).

In addition to cell fusion, syncytin-1 regulates critical functions such as proliferation and antiviral responses in trophoblast stem cells (West et al., 2022). Syncytin-2 contains a typical retroviral immunosuppressive env domain (Mangeney et al., 2007). Its expression in human cytotrophoblast cells suggests the involvement of this protein in establishing immunological tolerance during pregnancy, potentially through suppression of the maternal immune response to the fetus. Thus, former viral envelope proteins have been adapted to perform entirely new functions that are crucial for successful pregnancy.

The protein suppressin, the gene of which also originates from an ERV, performs opposing functions by inhibiting cell fusion. Suppressin has been identified in cultured human trophoblast cells and placental tissue samples. Suppressin utilizes ASCT2 as a receptor to inhibit syncytin-1-mediated fusion of cytotrophoblast cells (Sugimoto et al., 2013). In placental development, the balance of syncytin and suppressin gene expression determines the differentiation pathways of trophoblasts. It directs cells either towards fusion, forming the multinucleated syncytiotrophoblast, or towards invasion, forming the invasive trophoblast. Therefore, the regulation

of these two HERV-derived genes is critical for normal placental formation and function.

The imprinted genes *PEG10* (paternally expressed 10) and *PEG11/RTL1* (retrotransposon like 1), expressed from the paternal homolog, are also derived from ERVs. *PEG10* contains two overlapping open reading frames, the product of one of which features protease activity and plays an important role in the formation of fetal capillaries in mice (Clark et al., 2007). Both *PEG10* and *PEG11/RTL1* encode proteins that are highly homologous with the group-specific antigen and polymerase proteins of the *sushi-ichi* retrotransposon of the pufferfish genome, which belongs to the *Ty3/gypsy* family (Kim et al., 1994; Song et al., 1994). Functional studies in model organisms have demonstrated the key role of these genes in embryonic development.

In *PEG10* knockout mice, the labyrinthine and trabecular layers of the chorion are absent, accompanied by early embryonic lethality (Ono et al., 2006). Furthermore, CRISPR-Cas-induced deletion of *PEG10* in trophoblast stem cells led to impaired differentiation. Increased expression of the *PEG11* gene, or its deficiency, led to late embryonic lethality and neonatal death with damage to placental capillary networks in mice (Sekita et al., 2008; Kitazawa et al., 2017). These data highlight the fundamental significance of ERV-derived genes for ensuring normal placental development and successful pregnancy, demonstrating complex evolutionary mechanisms of exaptation of viral elements to perform critical physiological functions.

Mobile genetic elements as a source of placenta-specific enhancers

TE-related sequences are widespread throughout the human genome, found both within genes and in adjacent regulatory regions. According to Refseq data, 27.4 % of transcribed human DNA sequences have at least one transcript variant with insertions of TE sequences in untranslated regions (van de Lagemaat, 2003). Approximately 45 % of human enhancers are TE-derived (Simonti, 2017).

The function of enhancers is to regulate gene expression through the binding of transcription factors. In placental tissue, a significant prevalence of certain transposon classes is observed among placenta-specific enhancers. LTR-retrotransposons show the highest representation, followed by SINE, LINE, and DNA transposon elements (Sun et al., 2021). In humans, TE-derived enhancers are involved in context-specific gene regulation, including the expression of genes related to pregnancy, early embryonic development, and the formation of innate immunity (Modzelewski et al., 2022).

Human placental enhancers often overlap with specific families of endogenous retroviruses (ERVs), including MER21A, MER41A/B, and MER39B, typically associated with immune responses and placental function (Sun et al., 2021). MER41A/B elements create multiple binding sites for transcription factors, including ones located near the *FBN2* gene, which encodes the placenta-specific peptide hormone placentin, stimulating glucose secretion and trophoblast invasion (Yu et al., 2020; Sun et al., 2021). The MER41 family has six subfamilies, including A/B/C/D/E/G (Kojima, 2018). The evolutionary significance of these elements is underscored by their role in the formation of interferon-stimulated cis-regulatory elements that interact with the key transcription factors STAT1 and IRF1 (Schmid, Bucher, 2010; Chuong et al., 2016; Buttler, Chuong, 2022).

Another important example of TE-derived regulatory elements is LTR10A, acting as a powerful enhancer for essential placental genes, including *ENG* (Frost et al., 2023). The ENG protein plays a significant role in regulating trophoblast differentiation (Mano et al., 2011).

Leptin (LEP), encoded by one of the TE-regulated genes, performs multiple functions in early pregnancy. This hormone is involved in regulating implantation, trophoblast invasion, and placental angiogenesis, creating the necessary conditions for normal fetal development (Pérez-Pérez et al., 2018). In addition to its regulatory function, leptin promotes trophoblast proliferation and inhibits apoptotic processes (Magariños et al., 2007, Pérez-Pérez et al., 2008). *LEP* expression in the placenta is controlled by the transposon MER11 (Bi et al., 1997).

Of equal importance is the corticotropin-releasing hormone (*CRH*) gene, which regulates the duration of pregnancy. Its placental expression is controlled by the primate-specific element THE1B (Dunn-Fletcher et al., 2018).

LTR8B and MER11D elements, which are associated with the *PSG* gene cluster (Frost et al., 2023), encoding pregnancy-specific glycoproteins, exhibit notable evolutionary patterns. Their distribution among primates correlates with the type of placentation: from 6–24 genes in Old World monkeys to 1–7 genes in New World monkeys and complete absence in lemurs with epitheliochorial placentas (Zimmermann, Kammerer, 2021). Convergent evolution of this cluster in primates and mice (Rudert et al., 1989) suggests its important role in the development of the hemochorial placenta. These results suggest that integration of LTR8B elements prior to the expansion of the PSG cluster in humans was an important step that contributed to high expression of these genes in the trophoblast.

The function of *PSG* genes during pregnancy remains unclear. However, low levels of circulating *PSGs* are associated with recurrent pregnancy loss, fetal growth restriction, and preeclampsia (Towler et al., 1977; Arnold et al., 1999).

MER61D/E elements employ a specific regulatory mechanism, participating in the formation of binding sites for the transcription factor TP63 (Li et al., 2014). This factor, related to p53 (Riege et al., 2020), supports trophoblast proliferation, preventing premature differentiation. MER61 elements expand the TP63 binding network, participating in cellular stress responses (Su et al., 2015), thus highlighting the multifunctional nature of transposable elements with regards to placental development regulation.

Placenta-specific gene expression from transposable element promoters

Promoters formed from transposons represent an evolutionarily significant mechanism of coordinated gene regulation (Modzelewski et al., 2022). Such mechanisms are particu-

larly important in critical stages of embryonic development, requiring precise temporal and spatial organization of gene expression. Furthermore, transposon-derived regulatory elements enhance the reliability of genetic programs by creating redundancy in transcriptional factor interaction networks.

All orders of retrotransposons and DNA transposons can initiate the formation of chimeric transcripts in mammalian embryos, although their relative activity varies significantly among species and developmental stages (Oomen et al., 2025). The highest concentration of such transcripts is observed in oocytes and at the stage of embryonic genome activation, covering the period from the two-cell stage to the compacted morula stage (8–16 blastomeres). The integration of TE-derived promoter sequences into the host genome creates evolutionary prerequisites for the emergence of new gene expression patterns in various cell types, and also contributes to the generation of shortened or elongated protein isoforms, which ultimately can lead to gene neofunctionalization (Ashley et al., 2018).

Comparative analysis of the transcriptomes of preimplantation embryos of five placental mammal species (mice, pigs, cows, rabbits, and rhesus macaques) revealed species-specific features of transposon-mediated regulation. LTR elements predominate in mouse embryos (59 % of all TE-initiated transcripts), while LINE elements dominate in rabbits (40 %), and SINE elements, in rhesus macaques (42 %). Notably, SINE elements, despite their relatively recent evolutionary origin, demonstrate the ability to form chimeric transcripts in all studied species, although their number varies greatly: from 112 transcripts in cows to 3,910 in rhesus macaques (Oomen et al., 2025). These findings emphasize the important role transposon elements play in regulating early embryonic development in placental mammals.

The mechanisms of TE-mediated expression regulation significantly differ. Transposon-derived promoters can either fuse with canonical gene promoters or completely replace them, as well as function as alternative regulatory elements located in various positions relative to the transcription initiation site. LTR elements, which retain promoter activity in both the sense and antisense orientations, are of particular interest (van de Lagemaat et al., 2003).

A classic example of TE-mediated regulation is the *CYP19A1* gene, which encodes aromatase P450 and the expression of which in the placenta is controlled by the LTR promoter MER21A (van de Lagemaat et al., 2003). Another significant example is the pleiotrophin (*PTN*) gene, demonstrating tissue-specific alternative regulation: while one of its transcripts is expressed ubiquitously, including placental tissue, another variant, controlled by the 5′-LTR HERV-E, exhibits strict placenta-specificity (Reiss et al., 2007; Benson et al., 2009).

Another example is the *INSL4* gene, encoding a primate-specific peptide of insulin-like hormone and involved in the apoptosis of placental cells. Its expression is controlled by a HERV element, which likely determines the placenta-specific nature of expression in humans and other modern primates (Macaulay et al., 2011).

The LINE-1 antisense promoter can also initiate the formation of chimeric transcripts, in which 5′-antisense LINE-1 sequences are joined to exons of neighboring genes via splicing. However, the functional significance of LINE-1 element activation in the placenta is currently unclear. Previously, using bioinformatics methods, 988 putative LINE-1 chimeric transcripts were identified, with 911 of them being described for the first time (Criscione et al., 2016). Notably, the products of genes specific to neural tissue and placenta predominate among these transcripts, but experimental confirmation of these data has not yet been obtained.

An important step in the establishment and maintenance of pregnancy in many placental mammals is the differentiation (decidualization) of endometrial stromal fibroblasts into decidual stromal cells in response to progesterone. Decidualization triggers extensive reprogramming in the endometrium, leading to dramatic changes in gene expression, recruitment of immunosuppressive immune cells, vascular remodeling, and secretory transformation of uterine glands (Gellersen et al., 2007). In mammals, approximately 13 % of differentially expressed endometrial genes are located within 200 kb of the placental mammal-specific DNA transposon MER20 (Class II TE), which is thought to be involved in regulating key placental genetic networks, including cAMP-dependent signaling pathways in endometrial cells (Lynch et al., 2011). MER20 contains multiple binding sites for various transcription factors, and ancient mammalian transposons in general are enriched in hormone-sensitive regulatory elements that define endometrial cell identity (Lynch et al., 2015).

Furthermore, mapping of functionally active regions of the genome in decidual stromal cells showed that approximately 90 % of open chromatin regions, 58 % of enhancers, and 31 % of promoters overlapped with DNA regions derived from ancient TEs, most of which were specific to mammals or eutherians (Lynch et al., 2015).

Thus, the evolutionary domestication of transposable elements in the mammalian genome has led to the formation of a unique regulatory landscape necessary for the development and function of the placenta as an evolutionary novel organ. These changes have affected both the embryonic and maternal components of the placenta, providing complex mechanisms for their interaction.

Alternative mechanisms of mobile genetic element impact on placental development

In addition to the aforementioned mechanisms of placental development regulation, TEs perform other functions in mammals that, albeit not studied in detail with respect to placentation, play a significant role in early embryonic development, likewise characterized by genome hypomethylation.

Of particular interest is transposon-dependent alternative splicing, in which TEs can contain donor or acceptor splice sites, modifying the canonical pre-mRNA processing pathways and contributing to the emergence of new protein isoforms with unique functional properties (Modzelewski et al., 2022).

A vivid example of this phenomenon is the AluY element, which integrated into intron 6 of the *TBXT* gene in the hominoid ancestor's genome about 25 million years ago. The interaction of this element with the older AluSx1, located in the reverse orientation, leads to the formation of a hairpin structure in pre-mRNA, which excludes exon 6 from the mature mRNA. The resulting TBXTΔexon6 alternative isoform, specific to hominoids, correlates in time with the loss of a tail in this evolutionary lineage. Experimental expression of this isoform in mice leads to tail development disorders, confirming the key role of this TE-mediated modification in the evolution of primate morphology (Xia et al., 2024).

Transposon elements also have a significant impact on chromatin architecture through the formation of binding sites for the CTCF protein, which mediates the formation of topologically associated domains (TADs) (Rao et al., 2017). Approximately 20 % of species-specific TADs contain CTCF sites encoded by species-specific TEs (Choudhary et al., 2020). Although most TAD boundaries are evolutionarily conserved (Vietri et al., 2015), SINE elements show significant enrichment in these regions (Lu et al., 2020), apparently performing a stabilizing function for CTCF-site clusters (Kentepozidou et al., 2020). TEs can also contribute to the establishment of species-specific chromatin loops by introducing new CTCF anchor motifs (Choudhary, 2020).

Notably, some TADs in pluripotent stem cells are formed by an alternative mechanism dependent on the transcription of HERV-H elements (Santoni et al., 2012, Ohnuki et al., 2014). Likewise, in mouse embryos at the 2-cell stage, MERVL elements are not only the main source of promoters for controlling early embryonic development gene expression, but also contribute to the formation of domain boundaries (Kruse et al., 2019). Similar mechanisms may be involved in establishing specific chromatin conformations and in trophoblast stem cells, determining their differentiation potential.

TEs are actively involved in chromatin remodeling processes. It has been shown that the expression of LINE-1 plays an important role in chromatin organization during the activation of the mouse zygotic genome. Prolonged transcriptional activation of LINE-1 or premature transcriptional suppression of LINE-1 in mouse zygotes leads to developmental arrest. At the same time, this effect is not explained by the proteins encoded by LINE-1, but depends on the expression of non-coding RNA LINE-1 (Jachowicz et al., 2017). They act as a nuclear scaffold to recruit the Nucleolin and Kap1 proteins to suppress the Dux/MERVL transcription program at the two-cell stage and maintain the function of the pluripotency gene network in mouse embryonic stem cells (Percharde et al., 2018).

It is also notable that HERV-K elements, which have been relatively recently introduced into the human genome (Belshaw et al., 1999), are actively transcribed during normal human embryogenesis, starting from the eight-cell stage to the preimplantation blastocyst stage. Similar mechanisms may regulate the balance between proliferation and differentiation of trophoblast cells, determining the correct formation of placental structures.

An important evolutionary mechanism is TE-mediated recombination, often occurring between species-specific Alu and LTR elements, which can lead to gene duplication with subsequent neofunctionalization. A classic example is the duplication of the growth hormone gene in catarrhine primates, caused by recombination between Alu elements (Barsh et al., 1983). Growth hormone from duplicated genes is expressed in the placenta and interacts with growth hormone and prolactin receptors in placental tissues (Haig, 2008). This example demonstrates how TE-mediated genomic rearrangements can directly affect placental physiology and evolution.

Thus, numerous studies confirm that the early development of human embryos occurs with the active participation of retroviral and retrotransposon transcripts (Grow et al., 2015), which emphasizes the fundamental role of these elements in the formation and regulation of embryogenesis in mammals. Similar mechanisms likely operate in the placenta, making mobile elements important participants in the formation and function of this unique organ.

Conclusion

The evolutionary persistence of transposons in mammalian genomes is partly driven by their strategic integration into critical stages of early placental development. It is pivotal that this integration provides TEs with a dual selective advantage: guaranteed vertical transmission via incorporation into genes essential for implantation and placentation (where their elimination leads to embryonic lethality), and access to a unique epigenetic niche. Global hypomethylation and the presence of partially methylated domains in the placenta create an environment permissive to limited retrotransposon activity without catastrophic consequences for the host organism.

Of critical significance is the synergy between the transient nature of the placenta and TE replicative strategies: the ephemeral existence of this organ mitigates the long-term risks of uncontrolled transposition, while simultaneously providing a unique opportunity for functional testing of novel mobile element insertions. Partially methylated domains in the trophoblast genome serve as a molecular platform for exaptation, where potentially beneficial new regulatory mechanisms (such as providing alternative promoters or enhancers) are selectively fixed in the genome. Such dynamics transforms an initially parasitic relationship into a symbiotic one, where TEs are guaranteed replication and transmission, and the host benefits from a source of evolutionary innovation for regulating placental development.

References

Arnold L.L., Doherty T.M., Flor A.W., Simon J.A., Chou J.Y., Chan W.Y., Mansfield B.C. Pregnancy-specific glycoprotein gene expression in recurrent aborters: a potential correlation to interleukin-10 expression. *Am J Reprod Immunol*. 1999;41(3):174-182. doi 10.1111/j.1600-0897.1999.tb00530.x

Ashley J., Cordy B., Lucia D., Fradkin L.G., Budnik V., Thomson T. Retrovirus-like gag protein arc1 binds RNA and traffics across sy-

naptic boutons. *Cell*. 2018;172(1-2):262-274.e11. doi 10.1016/j.cell.2017.12.022

Barsh G.S., Seeburg P.H., Gelinas R.E. The human growth hormone gene family: structure and evolution of the chromosomal locus. *Nucleic Acids Res*. 1983;11(12):3939-3958. doi 10.1093/nar/11. 12.3939

Belshaw P.J., Walsh C.T., Stachelhaus T. Aminoacyl-CoAs as probes of condensation domain selectivity in nonribosomal peptide synthesis. *Science*. 1999;284(5413):486-489. doi 10.1126/science.284.5413.486

Benson D.A., Karsch-Mizrachi I., Lipman D.J., Ostell J., Sayers E.W. GenBank. *Nucleic Acids Res*. 2009;37(D1):D26-D31. doi 10.1093/nar/gkn723

Bi S., Gavrilova O., Gong D.W., Mason M.M., Reitman M. Identification of a placental enhancer for the human leptin gene. *J Biol Chem*. 1997;272(48):30583-30588. doi 10.1074/jbc.272.48.30583

Bourque G., Burns K.H., Gehring M., Gorbunova V., Seluanov A., Hammell M., Imbeault M., Izsvák Z., Levin H.L., Macfarlan T.S., Mager D.L., Feschotte C. Ten things you should know about transposable elements. *Genome Biol*. 2018;19(1):199. doi 10.1186/s13059-018-1577-z

Brosius J.P. Analyses and interventions: anthropological engagements with environmentalism. *Curr Anthropol*. 1999;40(3):277-309. doi 10.1086/200019

Brosius J., Gould S.J. On "genomenclature": a comprehensive (and respectful) taxonomy for pseudogenes and other "junk DNA". *Proc Natl Acad Sci USA*. 1992;89(22):10706-10710. doi 10.1073/pnas.89.22.10706

Buttler C.A., Chuong E.B. Emerging roles for endogenous retroviruses in immune epigenetic regulation. *Immunol Rev*. 2022;305(1):165-178. doi 10.1111/imr.13042

Chesnokova E., Beletskiy A., Kolosov P. The role of transposable elements of the human genome in neuronal function and pathology. *Int J Mol Sci*. 2022;23(10):5847. doi 10.3390/ijms23105847

Choudhary M.N.K., Friedman R.Z., Wang J.T., Jang H.S., Zhuo X., Wang T. Co-opted transposons help perpetuate conserved higher-order chromosomal structures. *Genome Biol*. 2020;21(1):16. doi 10.1186/s13059-019-1916-8

Chuong E.B., Elde N.C., Feschotte C. Regulatory evolution of innate immunity through co-option of endogenous retroviruses. *Science*. 2016;351(6277):1083-1087. doi 10.1126/science.aad5497

Clark M.B., Jänicke M., Gottesbühren U., Legge M., Poole E.S., Warren P.T. Mammalian gene *PEG10* expresses two reading frames by high efficiency –1 frameshifting in embryonic-associated tissues. *J Biol Chem*. 2007;282(52):37359-37369. doi 10.1074/jbc.M705676200

Criscione S.W., Theodosakis N., Micevic G., Cornish T.C., Burns K.H., Neretti N., Rodić N. Genome-wide characterization of human L1 antisense promoter-driven transcripts. *BMC Genomics*. 2016;17:463. doi 10.1186/s12864-016-2800-5

de Koning A.P.J., Gu W., Castoe T.A., Batzer M.A., Pollock D.D. Repetitive elements may comprise over two-thirds of the human genome. *PLoS Genet*. 2011;7(12):e1002384. doi 10.1371/journal.pgen.1002384

Doolittle W.F., Sapienza C. Selfish genes, the phenotype paradigm and genome evolution. *Nature*. 1980;284(5757):601-603. doi 10.1038/284601a0

Dunn-Fletcher C.E., Muglia L.M., Pavlicev M., Wolf G., Sun M.A., Hu Y.C., Huffman E., … Swaggart K.A., Lamm K.Y.B., Jones H., Macfarlan T.S., Muglia L.J. Anthropoid primate-specific retroviral element THE1B controls expression of *CRH* in placenta and alters gestation length. *PLoS Biol*. 2018;16:e2006337. doi 10.1371/journal.pbio.2006337

Ehrlich M., Gama-Sosa M.A., Huang L.-H., Midgett R.M., Kuo K.C., McCune R.A., Gehrke C. Amount and distribution of 5-methylcytosine in human DNA from different types of tissues or cells. *Nucleic Acids Res*. 1982;10(8):2709-2721. doi 10.1093/nar/10.8.2709

Frost J.M., Amante S.M., Okae H., Jones E.M., Ashley B., Lewis R.M., Cleal J.K., Caley M.P., Arima T., Maffucci T., Branco M.R. Regulation of human trophoblast gene expression by endogenous retroviruses. *Nat Struct Mol Biol*. 2023;30(4):527-538. doi 10.1038/s41594-023-00960-6

Gellersen B., Brosens I.A., Brosens J.J. Decidualization of the human endometrium: mechanisms, functions, and clinical perspectives. *Semin Reprod Med*. 2007;25(6):445-453. doi 10.1055/s-2007-991042

Grow E.J., Flynn R.A., Chavez S.L., Bayless N.L., Wossidlo M., Wesche D.J., Martin L., Ware C.B., Blish C.A., Chang H.Y., Pera R.A., Wysocka J. Intrinsic retroviral reactivation in human preimplantation embryos and pluripotent cells. *Nature*. 2015;522(7555):221-225. doi 10.1038/nature14308

Haig D. Placental growth hormone-related proteins and prolactin-related proteins. *Placenta*. 2008;29:S36-S41. doi 10.1016/j.placenta.2007.09.010

Hamilton W.J., Boyd J.D. Development of the human placenta in the first three months of gestation. *J Anat*. 1960;94(3):297-328

Honda T. Links between human LINE-1 retrotransposons and hepatitis virus-related hepatocellular carcinoma. *Front Chem*. 2016;4:21. doi 10.3389/fchem.2016.00021

Hoyt S.J., Storer J.M., Hartley G.A., Grady P.G.S., Gershman A., de Lima L.G., Limouse C., … Eichler E.E., Phillippy A.M., Timp W., Miga K.H., O'Neill R.J. From telomere to telomere: the transcriptional and epigenetic state of human repeat elements. *Science*. 2022;376(6588):eabk3112. doi 10.1126/science.abk3112

Jachowicz J.W., Bing X., Pontabry J., Bošković A., Rando O.J., Torres-Padilla M.-E. LINE-1 activation after fertilization regulates global chromatin accessibility in the early mouse embryo. *Nat Genet*. 2017;49(10):1502-1510. doi 10.1038/ng.3945

Johnson W.E. Origins and evolutionary consequences of ancient endogenous retroviruses. *Nat Rev Microbiol*. 2019;17(6):355-370. doi 10.1038/s41579-019-0189-2

Kentepozidou E., Aitken S.J., Feig C., Stefflova K., Ibarra-Soria X., Odom D.T., Roller M., Flicek P. Clustered CTCF binding is an evolutionary mechanism to maintain topologically associating domains. *Genome Biol*. 2020;21(1):5. doi 10.1186/s13059-019-1894-x

Kim A., Terzian C., Santamaria P., Pelisson A., Purd'homme N., Bucheton A. Retroviruses in invertebrates: the gypsy retrotransposon is apparently an infectious retrovirus of Drosophila melanogaster. *Proc Natl Acad Sci USA*. 1994;91(4):1285-1289. doi 10.1073/pnas.91.4.1285

Kitazawa M., Tamura M., Kaneko-Ishino T., Ishino F. Severe damage to the placental fetal capillary network causes mid-to late fetal lethality and reductionin placental size in *Peg11/Rtl1* KO mice. *Genes Cells*. 2017;22(2):174-188. doi 10.1111/gtc.12465

Kojima K.K. Human transposable elements in Repbase: genomic footprints from fish to humans. *Mob DNA*. 2018;9:2. doi 10.1186/s13100-017-0107-y

Kramerov D.A., Vassetzky N.S. SINEs. *Wiley Interdiscip Rev RNA*. 2011;2(6):772-786. doi 10.1002/wrna.91

Kruse K., Díaz N., Enriquez-Gasca R., Gaume X., Torres-Padilla M.-E., Vaquerizas J.M. Transposable elements drive reorganisation of 3D chromatin during early embryogenesis. *BioRxiv*. 2019. doi 10.1101/523712

Lanciano S., Philippe C., Sarkar A., Pratella D., Domrane C., Doucet A.J., van Essen D., Saccani S., Ferry L., Defossez P.A., Cristofari G. Locus-level L1 DNA methylation profiling reveals the epigenetic and transcriptional interplay between L1s and their integration sites. *Cell Genom*. 2024;4(2):100498. doi 10.1016/j.xgen.2024.100498

Lee M., Ahmad S.F., Xu J. Regulation and function of transposable elements in cancer genomes. *Cell Mol Life Sci*. 2024;81(1):157. doi 10.1007/s00018-024-05195-2

Li Y., Moretto-Zita M., Leon-Garcia S., Parast M.M. p63 inhibits extravillous trophoblast migration and maintains cells in a cytotropho-

blast stem cell-like state. *Am J Pathol*. 2014;184(12):3332-3343. doi 10.1016/j.ajpath.2014.08.006

Lou C., Goodier J.L., Qiang R.A. Potential new mechanism for pregnancy loss: considering the role of LINE-1 retrotransposons in early spontaneous miscarriage. *Reprod Biol Endocrinol*. 2020;18(1):6. doi 10.1186/s12958-020-0564-x

Lu J.Y., Shao W., Chang L., Yin Y., Li T., Zhang H., Hong Y., … Liu W., Yan P., Ramalho-Santos M., Sun Y., Shen X. Genomic repeats categorize genes with distinct functions for orchestrated regulation. *Cell Rep*. 2020;30(10):3296-3311. doi 10.1016/j.celrep.2020.02.048

Lynch V.J., Leclerc R.D., May G., Wagner G.P. Transposon-mediated rewiring of gene regulatory networks contributed to the evolution of pregnancy in mammals. *Nat Genet*. 2011;43(11):1154-1159. doi 10.1038/ng.917

Lynch V.J., Nnamani M.C., Kapusta A., Brayer K., Plaza S.L., Mazur E.C., Emera D., … Young S.L., Lieb J.D., DeMayo F.J., Feschotte C., Wagner G.P. Ancient transposable elements transformed the uterine regulatory landscape and transcriptome during the evolution of mammalian pregnancy. *Cell Rep*. 2015;10(4):551-561. doi 10.1016/j.celrep.2014.12.052

Macaulay E.C., Weeks R.J., Andrews S., Morison I.M. Hypomethylation of functional retrotransposon-derived genes in the human placenta. *Mamm Genome*. 2011;22(11-12):722-735. doi 10.1007/s00335-011-9355-1

Magariños M.P., Sánchez-Margalet V., Kotler M., Calvo J.C., Varone C.L. Leptin promotes cell proliferation and survival of trophoblastic cells. *Biol Reprod*. 2007;76(2):203-210. doi 10.1095/biolreprod.106.051391

Mangeney M., Renard M., Schlecht-Louf G., Bouallaga I., Heidmann O., Letzelter C., Richaud A., Ducos B., Heidmann T. Placental syncytins: genetic disjunction between the fusogenic and immunosuppressive activity of retroviral envelope proteins. *Proc Natl Acad Sci USA*. 2007;104(51):20534-20539. doi 10.1073/pnas.0707873105

Mano Y., Kotani T., Shibata K., Matsumura H., Tsuda H., Sumigama S., Yamamoto E., Iwase A., Senga T., Kikkawa F. The loss of endoglin promotes the invasion of extravillous trophoblasts. *Endocrinology*. 2011;152(11):4386-4394. doi 10.1210/en.2011-1088

Modzelewski A.J., Chong J.G., Wang T., He L. Mammalian genome innovation through transposon domestication. *Nat Cell Biol*. 2022;24(9):1332-1340. doi 10.1038/s41556-022-00970-4

Mustafin R.N. Functional dualism of transposon transcripts in evolution of eukaryotic genomes. *Russ J Dev Biol*. 2018;49:339-355. doi 10.1134/S1062360418070019

Nelson P.N., Carnegie P.R., Martin J., Davari Ejtehadi H., Hooley P., Roden D., Rowland-Jones S., Warren P., Astley J., Murray P.G. Demystified. Human endogenous retroviruses. *Mol Pathol*. 2003;56(1):11-18. doi 10.1136/mp.56.1.11

Novakovic B., Saffery R. Placental pseudo-malignancy from a DNA methylation perspective: unanswered questions and future directions. *Front Genet*. 2013;4:285. doi 10.3389/fgene.2013.00285

Ohnuki M., Tanabe K., Sutou K., Takahashi K. Dynamic regulation of human endogenous retroviruses mediates factor-induced reprogramming and differentiation potential. *Proc Natl Acad Sci USA*. 2014;111(34):12426-12431. doi 10.1073/pnas.1413299111

Ono R., Nakamura K., Inoue K., Naruse M., Usami T., Wakisaka-Saito N., Hino T., … Miki H., Kohda T., Ogura A., Yokoyama M., Kaneko-Ishino T. Deletion of Peg10, an imprinted gene acquired from a retrotransposon, causes early embryonic lethality. *Nat Genet*. 2006;38(1):101-106. doi 10.1038/ng1699

Oomen M.E., Rodriguez-Terrones D., Kurome M., Zakhartchenko V., Mottes L., Simmet K., Noll C., … Savatier P., Göke J., Wolf E., Kaessmann H., Torres-Padilla M. An atlas of transcription initiation reveals regulatory principles of gene and transposable element expression in early mammalian development. *Cell*. 2025;188(4):1156-1174.e20. doi 10.1016/j.cell.2024.12.013

Percharde M., Lin C.-J., Yin Y., Guan J., Peixoto G.A., Bulut-Karslioglu A., Biechele S., Huang B., Shen X., Ramalho-Santos M. A LINE1-nucleolin partnership regulates early development and ESC identity. *Cell*. 2018;174(2):391-405.e19. doi 10.1016/j.cell.2018.05.043

Pérez-Pérez A., Toro A., Vilariño-García T., Maymó J., Guadix P., Dueñas J., Fernández-Sánchez M., Varone C., Sánchez-Margalet V. Leptin action in normal and pathological pregnancies. *J Cell Mol Med*. 2018;22(2):716-727. doi 10.1111/jcmm.13369

Pötgens A.J.G., Schmitz U., Bose P., Versmold A., Kaufmann P., Frank H.-G. Mechanisms of syncytial fusion: a review. *Placenta*. 2002;23:S107-S113. doi 10.1053/plac.2002.0772

Price E.M., Cotton A.M., Peñaherrera M.S., McFadden D.E., Kobor M.S., Robinson W. Different measures of "genome-wide" DNA methylation exhibit unique properties in placental and somatic tissues. *Epigenetics*. 2012;7(6):652-663. doi 10.4161/epi.20221

Rao S.S.P., Huang S.C., Glenn St Hilaire B., Engreitz J.M., Perez E.M., Kieffer-Kwon K.R., Sanborn A.L., … Schlick T., Bernstein B.E., Casellas R., Lander E.S., Aiden E.L. Cohesin loss eliminates all loop domains. *Cell*. 2017;171(2):305-320.e24. doi 10.1016/j.cell.2017.09.026

Reiss D., Zhang Y., Mager D.L. Widely variable endogenous retroviral methylation levels in human placenta. *Nucleic Acids Res*. 2007;35(14):4743-4754. doi 10.1093/nar/gkm455

Riege K., Kretzmer H., Sahm A., McDade S.S., Hoffmann S., Fischer M. Dissecting the DNA binding landscape and gene regulatory network of p63 and p53. *eLife*. 2020;9:e63266. doi 10.7554/eLife.63266

Robinson W.P., Price E.M. The human placental methylome. *Cold Spring Harb Perspect Med*. 2015;5(5):a023044. doi 10.1101/cshperspect.a023044

Rondinone O., Murgia A., Costanza J., Tabano S., Camanni M., Corsaro L., Fontana L., … Ferrazzi E., Bosari S., Gentilini D., Sirchia S.M., Miozzo M. Extensive placental methylation profiling in normal pregnancies. *Int J Mol Sci*. 2021;22(4):2136. doi 10.3390/ijms22042136

Rudert F., Zimmermann W., Thompson J.A. Intra- and interspecies analyses of the carcinoembryonic antigen (CEA) gene family reveal independent evolution in primates and rodents. *J Mol Evol*. 1989;29(2):126-134. doi 10.1007/BF02100111

Santoni F.A., Guerra J., Luban J. HERV-H RNA is abundant in human embryonic stem cells and a precise marker for pluripotency. *Retrovirology*. 2012;9:111. doi 10.1186/1742-4690-9-111

Schmid D., Bucher P. MER41 repeat sequences contain inducible STAT1 binding sites. *PLoS One*. 2010;5(7):e11425. doi 10.1371/journal.pone.0011425

Sekita Y., Wagatsuma H., Nakamura K., Ono R., Kagami M., Wakisaka N., Hino T., … Ogura A., Ogata T., Yokoyama M., Kaneko-Ishino T., Ishino F. Role of retrotransposon-derived imprinted gene, *Rtl1*, in the feto-maternal interface of mouse placenta. *Nat Genet*. 2008;40(2):243-248. doi 10.1038/ng.2007.51

Shimode S. Acquisition and exaptation of endogenous retroviruses in mammalian placenta. *Biomolecules*. 2023;13(10):1482. doi 10.3390/biom13101482

Simonti C.N., Pavličev M., Capra J.A. Transposable element exaptation into regulatory regions is rare, influenced by evolutionary age, and subject to pleiotropic constraints. *Mol Biol Evol*. 2017;34(11):2856-2869. doi 10.1093/molbev/msx219

Slotkin R.K., Martienssen R. Transposable elements and the epigenetic regulation of the genome. *Nat Rev Genet*. 2007;8(4):272-285. doi 10.1038/nrg2072

Song S.U., Gerasimova T., Kurkulos M., Boeke J.D., Corces V.G. An Env-like protein encoded by a *Drosophila* retroelement: evidence that *gypsy* is an infectious retrovirus. *Genes Dev*. 1994;8(17):2046-2057. doi 10.1101/gad.8.17.2046

Su D., Wang X., Campbell M.R., Song L., Safi A., Crawford G.E., Bell D.A. Interactions of chromatin context, binding site sequence content, and sequence evolution in stress-induced p53 occupancy and transactivation. *PLoS Genet*. 2015;11(1):e1004885. doi 10.1371/journal.pgen.1004885

Sugimoto J., Sugimoto M., Bernstein H., Jinno Y., Schust D. A novel human endogenous retroviral protein inhibits cell-cell fusion. *Sci Rep*. 2013;3:1462. doi 10.1038/srep01462

Sun M., Wolf G., Wang Y., Senft A.D., Ralls S., Jin J., Dunn-Fletcher C.E., Muglia L.J., Macfarlan T.S. Endogenous retroviruses drive lineage-specific regulatory evolution across primate and rodent placentae. *Mol Biol Evol*. 2021;38(11):4992-5004. doi 10.1093/molbev/msab223

Towler C.M., Horne C.H.W., Jandial V., Campbell D.M., MacGillivray I. Plasma levels of pregnancy-specific βl-glycoprotein in complicated pregnancies. *Br J Obstet Gynaecol*. 1977;84(4):258-263. doi 10.1111/j.1471-0528.1977.tb12573.x

van de Lagemaat L.N., Landry J.R., Mager D.L., Medstrand P. Transposable elements in mammals promote regulatory variation and diversification of genes with specialized functions. *Trends Genet*. 2003;19(10):530-536. doi 10.1016/j.tig.2003.08.004

Vasilyev S.A., Tolmacheva E.N., Vasilyeva O.Yu., Markov A.V., Zhigalina D.I., Zatula L.A., Lee V.A., Serdyukova E.S., Sazhenova E.A., Nikitina T.V., Kashevarova A.A., Lebedev I.N. LINE-1 retrotransposon methylation in chorionic villi of first trimester miscarriages with aneuploidy. *J Assist Reprod Genet*. 2021;38(1):139-149. doi 10.1007/s10815-020-02003-1

Vietri M., Schink K.O., Campsteijn C., Wegner C.S., Schultz S.W., Christ L., Thoresen S.B., Brech A., Raiborg C., Stenmark H. Spastin and ESCRT-III coordinate mitotic spindle disassembly and nuclear envelope sealing. *Nature*. 2015;522(7555):231-235. doi 10.1038/nature14408

West R.C., Ezashi T., Schoolcraft W.B., Yuan Y. Beyond fusion: a novel role for ERVW-1 in trophoblast proliferation and type I interferon receptor expression. *Placenta*. 2022;126:150-159. doi 10.1016/j.placenta.2022.06.012

Wicker T., Sabot F., Hua-Van A., Bennetzen J.L., Capy P., Chalhoub B., Flavell A., Leroy P., Morgante M., Panaud O., Paux E., SanMiguel P., Schulman A.H. A unified classification system for eukaryotic transposable elements. *Nat Rev Genet*. 2007;8(12):973-982. doi 10.1038/nrg2165

Xia B., Zhang W., Zhao G., Zhang X., Bai J., Brosh R., Wudzinska A., … Dasen J.S., Maurano M.T., Kim S.Y., Boeke J.D., Yanai I. On the genetic basis of tail-loss evolution in humans and apes. *Nature*. 2024;626(8001):1042-1048. doi 10.1038/s41586-024-07095-8

Yu Y., He J.-H., Hu L.-L., Jiang L.-L., Fang L., Yao G.-D., Wang S.-J., … Shang T., Sato Y., Kawamura K., Hsueh A.J.W., Sun Y.-P. Placensin is a glucogenic hormone secreted by human placenta. *EMBO Rep*. 2020;21(6):e49530. doi 10.15252/embr.201949530

Zimmermann W., Kammerer R. The immune-modulating pregnancy-specific glycoproteins evolve rapidly and their presence correlates with hemochorial placentation in primates. *BMC Genomics*. 2021;22(1):128. doi 10.1186/s12864-021-07413-8

Conflict of interest. The authors declare no conflict of interest.
Received February 18, 2025. Revised July 15, 2025. Accepted July 16, 2025.

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2025;29(5):676-684

doi 10.18699/vjgb-25-74

Rare missense substitutions in the mitochondrial DNA genes in patients with ventricular tachycardia

M.V. Golubenko [1] ✉, N.P. Babushkina [1], V.A. Korepanov [2], N.R. Valiakhmetov [1], T.A. Atabekov [2], K.N. Vitt [2], A.A. Zarubin [1], O.A. Makeeva [1], S.A. Afanasiev [2], R.E. Batalov [2], A.A. Garganeeva [2], M.S. Nazarenko [1], V.P. Puzyrev [1]

[1] Research Institute of Medical Genetics, Tomsk National Research Medical Center of the Russian Academy of Sciences, Tomsk, Russia
[2] Cardiology Research Institute, Tomsk National Research Medical Center of the Russian Academy of Sciences, Tomsk, Russia
✉ maria-golubenko@medgenetics.ru

Abstract. Human mitochondrial DNA (mtDNA) exhibits high population-level polymorphism. While certain pathogenic mtDNA variants are known to cause hereditary mitochondrial syndromes, often presenting with cardiac arrhythmias, life-threatening ventricular tachycardia (VT) itself is a major risk factor for sudden death in cardiovascular diseases. The aim of the work was to study rare ("private") missense substitutions in the mtDNA of patients with documented episodes of ventricular tachycardia in comparison with patients with ischemic heart disease without life-threatening heart arrhythmias and individuals without clinical manifestations of cardiovascular diseases. The sequencing of mtDNA was performed using high-throughput sequencing methods. Specialized algorithms predicting the effect of gene variants were used to assess the effect of missense substitutions. Comparative analysis of the spectrum of the identified amino acid substitutions in the studied groups showed that about 40 % of the individuals in all three groups were carriers of "private" missense variants in mtDNA. However, among such substitutions, the variants classified by the APOGEE2 predictor as "variants of uncertain significance" (VUS) were more common in the group of patients with heart arrhythmias than in the control group, where "private" missense substitutions of the VUS category were not detected ($p = 0.0063$ for Fisher's exact test). In addition, the groups differed in their phred-ranked Combined Annotation Dependent Depletion (CADD) scores, which were lower for individuals in the control group. The results indicate that rare mtDNA variants may contribute to predisposition to cardiovascular disease – in particular, to the risk of developing ventricular tachycardia by some patients.

Key words: mitochondrial DNA; heart arrhythmia; ventricular tachycardia; missense substitutions effects; genetic variant pathogenicity assessment

For citation: Golubenko M.V., Babushkina N.P., Korepanov V.A., Valiakhmetov N.R., Atabekov T.A., Vitt K.N., Zarubin A.A., Makeeva O.A., Afanasiev S.A., Batalov R.E., Garganeeva A.A., Nazarenko M.S., Puzyrev V.P. Rare missense substitutions in the mitochondrial DNA genes in patients with ventricular tachycardia. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov J Genet Breed*. 2025;29(5):676-684. doi 10.18699/vjgb-25-74

Funding. The study was supported by Russian Science Foundation, Grant No. 24-24-00527.

Редкие миссенс-замены в генах митохондриальной ДНК у пациентов с желудочковыми тахикардиями

М.В. Голубенко [1] ✉, Н.П. Бабушкина [1], В.А. Корепанов [2], Н.Р. Валиахметов [1], Т.А. Атабеков [2], К.Н. Витт [2], А.А. Зарубин [1], О.А. Макеева [1], С.А. Афанасьев [2], Р.Е. Баталов [2], А.А. Гарганеева [2], М.С. Назаренко [1], В.П. Пузырёв [1]

[1] Научно-исследовательский институт медицинской генетики Томского национального исследовательского медицинского центра Российской академии наук, Томск, Россия
[2] Научно-исследовательский институт кардиологии Томского национального исследовательского медицинского центра Российской академии наук, Томск, Россия
✉ maria-golubenko@medgenetics.ru

Аннотация. Митохондриальная ДНК (мтДНК) характеризуется высоким полиморфизмом в популяциях. При этом некоторые патогенные варианты мтДНК могут приводить к развитию наследственных митохондриальных синдромов, симптоматика которых включает в том числе нарушения ритма сердца. С другой стороны, жизнеугрожающие аритмии, в виде желудочковой тахикардии, являются фактором риска внезапной смерти у пациентов

© Golubenko M.V., Babushkina N.P., Korepanov V.A., Valiakhmetov N.R., Atabekov T.A., Vitt K.N., Zarubin A.A., Makeeva O.A., Afanasiev S.A., Batalov R.E., Garganeeva A.A., Nazarenko M.S., Puzyrev V.P., 2025

This work is licensed under a Creative Commons Attribution 4.0 License

с сердечно-сосудистыми заболеваниями. Целью работы стало исследование редких («приватных») миссенс-замен в мтДНК пациентов с зарегистрированными эпизодами желудочковой тахикардии в анамнезе в сравнении с пациентами с ишемической болезнью сердца без жизнеугрожающих нарушений ритма и индивидами без клинических проявлений сердечно-сосудистых заболеваний. Определение последовательности мтДНК проводили с помощью методов высокопроизводительного секвенирования, для оценки эффекта миссенс-замен использовали специализированные алгоритмы-предикторы эффекта генных вариантов. Сравнительный анализ спектра выявленных аминокислотных замен в исследованных группах показал, что во всех трех группах около 40 % индивидов были носителями «приватных» миссенс-вариантов в мтДНК, однако среди них в группе пациентов с нарушениями сердечного ритма чаще встречались варианты, классифицируемые предиктором APOGEE2 как «варианты неопределенного значения» (VUS), по сравнению с контрольной группой, в которой «приватных» миссенс-замен категории VUS не обнаружено (p = 0.0063 для точного критерия Фишера). Кроме того, группы различались по значениям phred-ранжированных значений CADD (Combined Annotation Dependent Depletion), которые были ниже для индивидов из контрольной группы. Полученные результаты указывают на то, что редкие варианты мтДНК могут вносить вклад в предрасположенность к сердечно-сосудистым заболеваниям, в частности в риск развития желудочковой тахикардии у некоторых пациентов.

Ключевые слова: митохондриальная ДНК; аритмия; желудочковая тахикардия; эффект миссенс-замен; оценка патогенности генетических вариантов

Introduction

Human mitochondrial DNA (mtDNA) exhibits a high degree of polymorphism, and, consequently, the proteins encoded by mitochondrial genes are similarly polymorphic. These proteins play a critical role in energy metabolism as essential components of the mitochondrial respiratory chain complexes. With the continuous growth of the human population, the burden of rare, so-called "private" genetic variants has increased substantially (Gao, Keinan, 2014), raising the likelihood that newly emerging gene variants – including mtDNA missense substitutions – may persist in the population even if they exert a mildly deleterious effect. While such variants are insufficient to cause severe hereditary disorders, they may contribute to the risk of common polygenic diseases.

The myocardium is one of the most energy-demanding tissues in the body. Most cardiovascular continuum disorders arise from myocardial ischemia, which is characterized by hypoxia, mitochondrial dysfunction, and oxidative stress (Kibel et al., 2020; Severino et al., 2020; Yang et al., 2022). Mitochondrial dysfunction, in turn, can exert an arrhythmogenic effect both through impaired ATP synthesis and via oxidative stress-induced membrane depolarization (Montaigne, Pentiah, 2015; Gambardella et al., 2017; van Opbergen et al., 2019). This is consistent with the frequent occurrence of cardiac arrhythmias in patients with mitochondrial diseases caused by pathogenic mtDNA mutations or nuclear gene defects affecting mitochondrial function (Ng, Turnbull, 2016). Conversely, severe cardiac arrhythmias – particularly paroxysmal ventricular tachycardia – are associated with a high risk of sudden cardiac death (Koplan, Stevenson, 2009; Chao et al., 2017), underscoring the importance of identifying hereditary risk factors for these conditions.

Early research on mtDNA variants associated with cardiovascular disease risk primarily focused on common population variants and their combinations (haplogroups) (Palacín et al., 2011; Hudson et al., 2014; Golubenko et al., 2015, 2021; Kytövuori et al., 2020; Roselló-Díez et al., 2021). Advances in sequencing technologies now enable fast comprehensive analysis of the mitochondrial genome, leading to growing interest in the role of rare mtDNA variants in disease pathogenesis (Govindaraj et al., 2014, 2019; Hagen et al., 2015; Piotrowska-Nowak et al., 2019).

The aim of this work was to study mtDNA rare missense variants in patients with ventricular tachycardia in comparison with patients without ventricular tachycardia and with relatively healthy individuals.

Materials and methods

There were three groups of participants in the study. The "main" group consisted of patients hospitalized in the Department of Surgical Treatment of Complex Heart Rhythm Disorders and Electrical Pacing at the Cardiology Research Institute of Tomsk National Research Medical Center. All patients underwent implantation of a cardioverter-defibrillator (ICD) due to a history of ventricular tachycardia (VT) episodes, as part of primary or secondary prevention of sudden cardiac death (Bockeriaet al., 2017). The group included 127 individuals. Medical histories and diagnostic data were analyzed for all patients. Patients with severe comorbidities (cancer, NYHA class IV heart failure, or chronic kidney disease stages IV–V) were excluded. The majority were male (74.8 %), with a median age of 64.0 years (IQR: 59.0–71.0).

The "comparison" group (n = 53) comprised patients with stable ischemic heart disease and no his-

Table 1. Clinical characteristics of the patients

Parameter	Main group (n = 127)	Comparison group (n = 53)	p
Age and gender			
Age, years, Me (Q_1; Q_3)	64.0 (59.0; 71.0)	67.0 (63.5; 71.5)	0.08420
Males, n (%)	95 (76.7)	23 (43.4)	0.00005
Females, n (%)	32 (23.3)	30 (56.6)	0.00005
Clinical symptoms and comorbidities			
Ischemic heart disease, n (%)	102 (80.3)	53 (100)	< 0.001
History of infarctions, n (%)	72 (56.7)	0	< 0.001
Hypertension, n (%)	118 (92.9)	53 (100)	0.046772
Dyslipidemia, n (%)	91 (71.7)	30 (56.6)	0.049927
Obesity, n (%)	79 (62.2)	32 (60.4)	0.818220
Diabetes mellitus type 2, n (%)	25 (19.7)	6 (11.3)	0.017902
Impaired glucose tolerance, n (%)	10 (7.9)	3 (5.7)	0.601015
Body mass index, kg/m^2, Me (Q_1; Q_3)	28.4 (25.6; 32.1)	29.6 (26.2; 33.8)	0.072343

Note. p – significance level when comparing groups using the Pearson χ^2 test (for frequencies) or the Mann–Whitney U-test (for quantitative characteristics).

tory of myocardial infarction, VT, or indications for ICD implantation. Their median age was 67.0 years (IQR: 63.0–71.0). Clinical characteristics of all patients are provided in Table 1.

In addition to the two groups of patients, a "control" group (n = 58) was formed, which consisted of Tomsk city residents who had no history of cardiovascular symptoms, including absence of heart rhythm disturbances; in addition, these individuals either had no stenosis of the carotid arteries, or the stenosis did not exceed 30 % (estimated by the ultrasound examination). The median age in this sample was 69.0 (62.0; 73.0) years, the ratio of men to women was 40:28 (69 % men).

Informed consent for participation in the study was obtained from all individuals included in the studied groups. The study protocol was approved by the biomedical ethics committees of the Research Institute of Medical Genetics and the Research Institute of Cardiology of the Tomsk National Research Medical Center.

Venous blood samples (6–10 mL, EDTA) were collected, and DNA was isolated using phenol-chloroform extraction.

The complete mitochondrial genome was sequenced via high-throughput sequencing (next-generation sequencing, NGS). Mitochondrial DNA was amplified by long-range PCR with two overlapping fragments: 1) 9,065 bp (primers: 9397-9416 and 1892-1873 of the human mtDNA reference sequence) and 2) 11,170 bp (primers: 15195-15214 and 9796-9777). Overlapping regions spanned 9397-9796 and 15195-1873.

PCR was performed using the BioMaster LR HS-PCR (2x) kit (BioLabMix, Russia). PCR product concentration was quantified via Qubit (Thermo Fisher Scientific, USA) with Spectra Q BR reagents (Raissol, Russia). Equimolar pools of both PCR products (20 ng/μL) were prepared for each sample. DNA libraries were prepared using DNA library preparation kits designed for working with genomic DNA, with double indexing of the libraries. In particular, DNA Prep kits (Illumina, USA) and SG GM Plus kits (Raissol, Russia) were used. The manufacturer's protocols were followed without modifications.

Sequencing was performed either on the MiSeq sequencer (Illumina, USA) using MiSeq reagent v.2 kit, 300 cycles, or on the GenoLab M sequencer (GeneMind, China) using GenoLab M V2.0 FCM reagent kit, 150 cycles.

After the data demultiplexing, fastq nucleotide reads were aligned to the reference human genome sequence (hg38) using DRAGEN 3.9.5 software, DNA pipeline (Illumina, USA). The resulting bam files were analyzed with mtDNA-specific software MtDNA-Server 2 (Weissensteiner et al., 2024). As a result, a list of nucleotide substitutions in comparison with the human mtDNA reference sequence (Andrews et al., 1999) was obtained, and an assessment of the mtDNA haplogroup for the identified haplotype was done according to the generally

accepted human mtDNA tree (van Oven, Kayser, 2008). The mtDNA sequences in the *.fasta or *.txt format were also analyzed in the mtPhyl program (Eltsov, Volodko, 2011), which draws the phylogeny of the analyzed sequences and provides a list of missense variants divided into "haplogroup associated" and "private" substitutions, accompanied with amino acid conservation index for these substitutions.

To assess the effect of missense substitutions in mtDNA genes, we used the APOGEE 2 meta-predictor, which was developed specifically for mitochondrial DNA (Bianco et al., 2023), and in addition, CADD scores (Rentzsch et al., 2021) were analyzed. Data on these and other tools for assessing mtDNA missense substitutions are available online at the MitImpact project address: http://bioinformatics.css-mendel.it/ (Castellana et al., 2015). Statistical analysis was performed in JASP 0.19.3 (JASP Team, 2024). Group comparisons used Pearson's χ^2 test (frequencies) or the Mann–Whitney U-test (quantitative variables).

Results

The mtDNA sequencing results demonstrated high mitochondrial genome diversity in the studied cohorts, with nearly all individuals exhibiting unique mtDNA haplotypes. Only two haplotypes were observed twice, both occurring in the "main" patient group. The frequencies of major mtDNA haplogroups were distributed as follows: haplogroup H occurred at 34 % in the "main" group, 34 % in the "comparison" group, and 41 % in controls; haplogroup J, at 8, 9, and 14 %; haplogroup T, at 12, 9, and 3 %; and haplogroup U, at 30, 34, and 34 % respectively. These frequencies corresponded to the reported Tomsk population data (39 % for H, 7 % for J, 10 % for T, and 25 % for U) (Golubenko et al., 2021). Although trends suggested reduced haplogroup T and elevated haplogroup J frequencies in controls, along with increased haplogroup U frequency both in controls and in "comparison" patients, these differences did not reach statistical significance.

In the "main" patient group, we identified 61 private missense variants and 85 haplogroup-associated missense variants (Table 2). Altogether, 50 individuals (39 % of the group) carried private missense substitutions, including 7 patients with two variants and 2 patients with three variants. The "comparison" group exhibited 28 private missense variants (found in 23 individuals, 43 % of this group, including 5 carriers of two variants) compared to 45 haplogroup-associated variants. The control group showed 35 private missense variants distributed among 23 individuals (40 % of controls), with 8 individuals harboring two variants and 2 individuals carrying three variants.

Elson's neutrality test (Elson et al., 2004) revealed no statistically significant deviations in the ratio of synonymous to non-synonymous substitutions from neutral selection expectations across groups. Similarly, we observed no significant differences in mean amino acid conservation indices between private and haplogroup-associated variants.

The APOGEE2 meta-predictor classifies missense variants using the standard five-tier pathogenicity system (benign, likely benign, variants of uncertain significance (VUS), likely pathogenic, and pathogenic) (McCormick et al., 2020). Variants with APOGEE2 scores of 0.265–0.716 are categorized as VUS, while higher and lower scores indicate likely pathogenic/pathogenic and likely benign/benign variants, respectively (Bianco et al., 2023). No private missense variants in our cohorts met criteria for pathogenic or likely pathogenic. In contrast, the main group contained 11 private VUS variants (18 % of its private variants), while the comparison group had three (10.7 %), and controls showed none (Table 2). This represents a significant accumulation of non-benign

Table 2. Characteristics of mtDNA missense polymorphism in the studied groups

Parameter	Group		
	main (n = 127)	comparison (n = 53)	control (n = 58)
Overall number of private missense variants, *n*	61	28	35
Overall number of non-private (haplogroup-associated) missense variants, *n*	75	40	37
Mean value of conservation index (by MtPhyl estimates) for private missense variants, %	49.54	49.82	52.98
Number of private missense variants classified as VUS (proportion of VUS in all private missense variants), *n* (%)	11 (18.0)	3 (10.7)	0 (0)
Number of individuals with private missense variant, *n* (%)	50 (39.4)	23 (43.4)	23 (39.7)

Note. VUS – variant of uncertain significance.

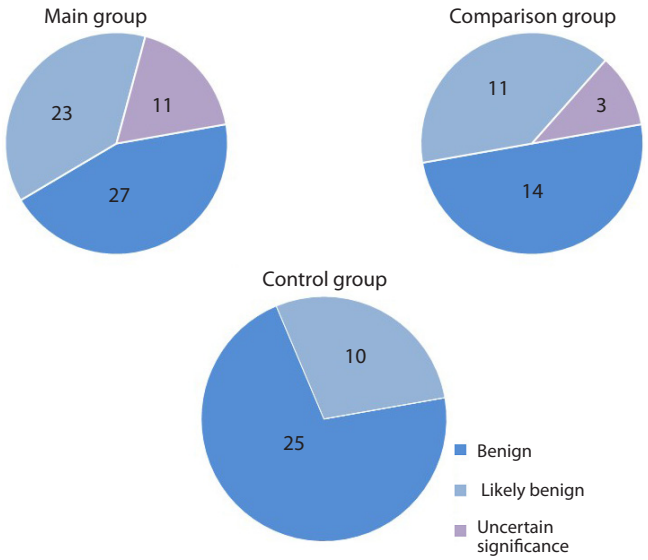


Fig. 1. Attribution of private mtDNA missense variants to the different categories of pathogenicity (the numbers indicate the number of variants in the corresponding category).

private variants in the main group compared to controls (p = 0.0063, Fisher's exact test). Differences between other group pairings were non-significant. The distribution of private variants across pathogenicity categories is illustrated in Figure 1.

The complete list of private missense substitutions classified as VUS is presented in Table 3. Notably, two variants (T3394C and G13708A), though identified as private in our patients, appear in multiple haplogroups on the human mtDNA phylogenetic tree – particularly G13708A, which characterizes West-Eurasian haplogroup J. Another private variant resulted not in an amino acid substitution but in replacement of a stop codon with glutamine (T9205C, *MT-ATP6* Ter227Gln) in the ATP6 gene. While APOGEE2 cannot score such variants, ClinVar database classifies it as VUS (https://www.ncbi.nlm.nih.gov/clinvar/variation/693124/, accessed 24.02.2025). Similarly, a stop-to-lysine variant (A7444G, *MT-COI* Ter514Lys), associated with haplogroup V7 and found in the main group, was previously considered pathogenic due to protein elongation but has been reclassified as "likely benign" after having been reviewed by ClinVar experts (https://www.ncbi.nlm.nih.gov/clinvar/variation/9663/, accessed 24.02.2025).

Of all VUS, 50 % are located in the genes encoding subunits of the first complex of the respiratory chain (NADH dehydrogenase), which is consistent with the total length of these genes, which encompass 65 % of the total length of all protein-coding mtDNA genes. It is interesting, however, that all three VUS identified among the patients of the comparison group were located not in the NADH dehydrogenase genes but in the cytochrome *b* gene (two variants) and cytochrome *c* oxidase gene (one variant). It can also be noted that while haplogroup H is the most common among Europeans (about 40 % of the population), only three VUS, i. e. 21 %, belonged to this haplogroup (H6a1a, H36, H1j8), whereas a signifi-

Table 3. Private mtDNA missense variants classified as VUS (APOGEE2)

No.	mtDNA change	Gene	Amino acid change	APOGEE2 score	Patient group	mtDNA haplogroup
1	T3394C	*MT-ND1*	Y30H	0.5822	Main	J1b1a1
2	C6489A	*MT-CO1*	L196I	0.3289	Main	T
3	G6510A	*MT-CO1*	A203T	0.2836	Comparison	H6a1a
4	C8369T	*MT-ATP8*	P2S	0.2767	Main	U5a2a1b
5	T9205C	*MT-ATP6*	227Q	–	Main	J1a1b1
6	G9738A	*MT-CO3*	A178T	0.3554	Main	R2
7	T10237C	*MT-ND3*	I60T	0.6661	Main	HV
8	G11696A	*MT-ND4*	V313I	0.3383	Main	K1
9	T12075C	*MT-ND4*	M439T	0.3560	Main	U5a1b
10	C13036T	*MT-ND5*	P234S	0.4992	Main	K1b1
11	G13708A	*MT-ND5*	A458T	0.3070	Main	T1a
12	T14291A	*MT-ND6*	E128V	0.4046	Main	H36
13	A14841G	*MT-CYTB*	N32S	0.2743	Comparison	H1j8
14	G15152A	*MT-CYTB*	G136S	0.2924	Comparison	U5a1

cant portion of VUS (36 %) belonged to haplogroup U (U5a2a1b, U5a1b, K1, K1b1, U5a1), and another 36 % belonged to the R2'JT cluster (J1b1a1, T, J1a1v1, R2, T1a). It might be assumed that the appearance of VUS on the background of haplogroup R2'JT may be a risk factor for the development of arrhythmia, including VT, but the total number of VUS in our study is too small to perform any statistical tests, so this issue requires additional research.

The classification of variants into pathogenicity classes is "categorical"; however, it relies on quantitative scales of the effect estimates. One of these estimates is CADD (combined annotation dependent depletion), an integrated metric based on machine learning, which uses more than 60 tools for annotating all possible genetic variants, followed by calculating the probability of their appearance in the genome and ranking all possible variants according to this probability. The logarithm of this score (phred-like ranking) is used to identify the least "probable" variants in the genome, which are therefore the variants with the greatest effect. According to the developers' recommendation, the minimum (threshold) value of the CADD phred score for considering a possibility of functional significance is 10. It means that the variant is among the 10 % most significant of all theoretically possible variants in the genome (Rentzsch et al., 2021).

The plot of CADD phred score values for all "private" missense variants is shown in Figure 2. In all groups, there were missense substitutions with this parameter value greater than 10; however, in the control group, only 26.6 % of "private" missense substitutions were in this zone, and the median value of this parameter was 8.3, while in the patient groups, the median value was 13 (main group) and 13.2 (comparison group). In total, 61.7 % of private substitutions in the main group and 64.3 % in the comparison group had a CADD phred score greater than 10. This differentiation was statistically significant both according to the results of variance analysis (p = 0.014) and according to the nonparametric Kruskal–Wallis criterion (p = 0.011).

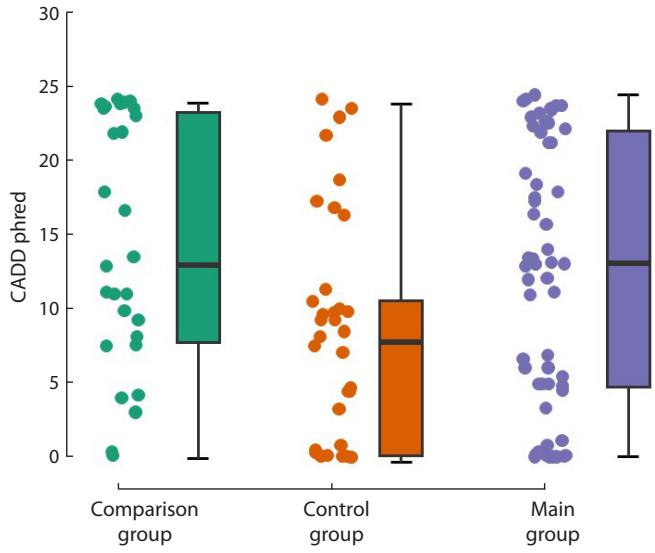


Fig. 2. Plot of CADD phred scores for all private missense substitutions in the studied groups.

Discussion

Unlike evolutionarily "established" combinations of mtDNA variants designated as haplogroups, all newly emerging variants are "private", meaning they are present only in the examined individual and probably in his/her close maternal relatives. When the population size constantly increases, there is an excess of "private" gene variants in the population (Gao, Keinan, 2014). Such newly arising mtDNA variants may influence the phenotype. When the variant has a strong negative effect on the phenotype, it may be eliminated from the population by natural selection; however, if the effect is small, the variant can persist in the population for many generations and even spread due to genetic drift.

Assessing the effect of a missense substitution in a gene on the structure and function of the encoded protein is an important issue. Despite the variety of algorithms and predictors developed for *in silico* effect estimation, the results of these studies do not always correspond to the true effect of specific missense substitutions. This is partly due to the insufficient experimental data on the pathogenicity of various variants, since only a small share of all possible amino acid substitutions has been studied in this regard, so extrapolation of these patterns to the entire data set is not always correct. In addition, epistatic interactions of amino acid residues between amino acid residues within or between protein subunits may contribute to effect variability, where additional amino acid substitutions could compensate for or exacerbate the effect of the analyzed substitution.

APOGEE2 is a meta-predictor that uses for its assessment evolutionary conservation; protein structural characteristics, including tertiary structure data and Gibbs free energy change (ΔΔG); effect estimates obtained from various predictor programs: PolyPhen2, SIFT, Fathmm, PROVEAN, MutationAssessor, EFIN, CADD, PANTHER, PhDSNP, SNAP, and MutationTaster2 (Bianco et al., 2023). This tool has the highest sensitivity (87 %) and specificity (90 %) compared to other predictors, according to the paper.

Distribution of values from the quantitative predictor of functional significance (CADD) showed that private mtDNA missense substitutions in the control group were characterized on average by lower values of this para-

meter, and these differences were statistically significant. Similar to the proportion of VUS, the two groups of patients (the main and comparison groups) did not differ from each other in mean CADD values.

Comparison of the fraction of variants with uncertain significance (VUS) among private missense substitutions showed a higher proportion of VUS in the main patient group versus controls. At the same time, missense substitutions of the VUS category were also registered in the comparison group patients, though their APOGEE2 scores were minimal (Table 3). The ratio of such variants to the total number of private missense substitutions in this sample (3/28, or 10.7 %), while lower than in the main group (11/61, or 18 %), showed no statistically significant difference ($p > 0.05$ by Fisher's exact test). Thus, the high frequency of VUS-category missense substitutions may be associated not with arrhythmia risk specifically but with predisposition to cardiovascular diseases in general. Some previous mtDNA studies have also identified rare and private substitutions, including missense variants, which can be classified as VUS or even as likely pathogenic variants – for example, in patients with hypertrophic cardiomyopathy (Govindaraj et al., 2014; Hagen et al., 2015), dilated cardiomyopathy (Govindaraj et al., 2019), and atherosclerosis (Piotrowska-Nowak et al., 2019).

Notably, two identified variants classified as VUS may be characteristic of certain mtDNA haplogroups. The G13708A substitution is one of the defining variants for haplogroup J, which is known to enhance expressivity of the pathogenic G11778A variant causing Leber's optic atrophy in European populations (Torroni et al., 1997). The T3394C substitution similarly enhances manifestation of pathogenic variant G11778A but in Asian populations (Ji et al., 2019). Both substitutions occur repeatedly in human mtDNA phylogeny (www.phylotree.org). Remarkably, in our study, the patient with private T3394C substitution had mtDNA belonging to haplogroup J (specifically J1b1a1, Table 3), meaning they also carried the G13708A substitution. Thus, this individual had two missense substitutions, each representing an unfavorable "background" promoting manifestation of pathogenic mtDNA variants. In this regard, it is interesting that, according to the published data, similar combinations of mtDNA variants were identified in Parkinson's disease, where variants typically associated with certain haplogroups ("out of place" variants) were more frequent in patients than in controls (Müller-Nedebock et al., 2022). Leber optic atrophy and Parkinson's disease are not cardiovascular diseases, but these examples may present general patterns of mtDNA variants effects manifestation.

Classifying a genetic variant as a VUS does not necessarily indicate a negative effect – it indicates only a higher probability that such a variant somehow influences the phenotype, hence the term "variant of uncertain significance". Nevertheless, the excess of such variants in the group of patients with life-threatening cardiac arrhythmias (ventricular tachycardia) and a high risk of sudden death revealed in our study suggests that, at least in some cases, the risk of sudden death may be increased by rare mtDNA variants with a negative effect, resulting in a decrease in the mitochondrial function. It can be assumed that, under normal conditions, minor deviations from optimal function of mitochondrial protein complexes may be compensated by increased mitochondrial gene expression, mitochondrial biogenesis, or modulation of specific biochemical pathways. Under cellular stress condition, however, this "borderline" mitochondrial dysfunction may become critical for the myocardial pathology development.

Whether such variants represent an arrhythmia-specific risk factor or generally increase cardiovascular disease risk remains an open question. Further studies are required in patient cohorts with diverse cardiovascular pathologies. Assessing genetic background effects on rare missense substitutions (potential epistatic interactions) will require larger samples, as private VUS-category missense variants occur in fewer than 10 % of patients. In addition, it should be noted that our study did not consider heteroplasmy – a situation when only a portion of the mtDNAs have the variant, which can be either inherited or somatically arisen *de novo*. Due to the lack of the possibility of analyzing the DNA of parents (mothers), we could not assess *de novo* variant occurrence. All mtDNA variants described here were homoplasmic.

Conclusion

Comparative analysis of rare (private) missense substitution spectra in mtDNA protein-coding genes among cardiovascular disease patients – particularly those with life-threatening arrhythmias – revealed several missense substitutions that may be classified as VUS, suggesting possible functional impacts on mitochondrial respiratory chain proteins. No such variants were found in the control group of individuals without clinical cardiovascular symptoms.

Groups showed no differences in overall mtDNA missense polymorphism characteristics (the total number of missense substitutions, the proportion of carriers of private missense variants in the group, more than one private missense substitution in one individual, the amino acid conservation index). However, there were statistically significant differences between the main group (with a history of ventricular tachycardia and a high risk of sudden death) and the control group in the proportion of VUS among private missense variants. In

addition, differences between the groups were revealed for the values of the quantitative score characterizing the possibility of the functional significance of variants (CADD score). These results allow us to assume that it is rare missense substitutions of mtDNA that may have functional impact and contribute to the predisposition to the cardiovascular continuum diseases, including the development of ventricular tachycardia in patients.

References

Andrews R.M., Kubacka I., Chinnery P.F., Lightowlers R.N., Turnbull D.M., Howell N. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet*. 1999; 23(2):147. doi 10.1038/13779

Bianco S.D., Parca L., Petrizzelli F., Biagini T., Giovannetti A., Liorni N., Napoli A., … Zhang S., Vescovi A.L., Wallace D.C., Caputo V., Mazza T. APOGEE 2: multi-layer machine-learning model for the interpretable prediction of mitochondrial missense variants. *Nat Commun*. 2023;14(1):5058. doi 10.1038/s41467-023-40797-7

Bockeria L.A., Neminushchiy N.M., Mikhaylichenko S.I., Novichkov S.A., Achkasov E.E. Implantable cardioverter defibrillators in the prevention of sudden cardiac death. *Terapevticheskiy Arkhiv = Therapeutic Archive*. 2017;89(12):103-109. doi 10.17116/terarkh20178912103-109 (in Russian)

Castellana S., Rónai J., Mazza T. MitImpact: an exhaustive collection of pre-computed pathogenicity predictions of human mitochondrial non-synonymous variants. *Hum Mutat.* 2015;36(2):E2413-E2422. doi 10.1002/humu.22720

Chao T.F., Liu C.J., Tuan T.C., Chen S.J., Chen T.J., Lip G.Y.H., Chen S.A. Risk and prediction of sudden cardiac death and ventricular arrhythmias for patients with atrial fibrillation – a nationwide cohort study. *Sci Rep*. 2017;7:46445. doi 10.1038/srep46445

Elson J.L., Turnbull D.M., Howell N. Comparative genomics and the evolution of human mitochondrial DNA: assessing the effects of selection. *Am J Hum Genet*. 2004;74(2):229-238. doi 10.1086/381505

Eltsov N., Volodko N. MtPhyl: Software tool for human mtDNA analysis and phylogeny reconstruction. Version 5.003. [Computer software]. 2011. https://sites.google.com/site/mtphyl/home/

Gambardella J., Sorriento D., Ciccarelli M., Del Giudice C., Fiordelisi A., Napolitano L., Trimarco B., Iaccarino G., Santulli G. Functional role of mitochondria in arrhythmogenesis. *Adv Exp Med Biol*. 2017;982:191-202. doi 10.1007/978-3-319-55330-6_10

Gao F., Keinan A. High burden of private mutations due to explosive human population growth and purifying selection. *BMC Genomics*. 2014;15(Suppl. 4):S3. doi 10.1186/1471-2164-15-S4-S3

Golubenko M.V., Salakhov R.R., Makeeva O.A., Goncharova I.A., Kashtalap V.V., Barbarash O.L., Puzyrev V.P. Association of mitochondrial DNA polymorphism with myocardial infarction and prognostic signs for atherosclerosis. *Mol Biol*. 2015;49(6):867-874. doi 10.1134/S0026893315050088

Golubenko M.V., Shumakova T.V., Makeeva O.A., Tarasenko N.V., Salakhov R.R., Shipulin V.M., Nazarenko M.S. Mitochondrial DNA polymorphism and myocardial ischemia: association of haplogroup H with heart failure. *Sibirskiy Zhurnal Klinicheskoy i Eksperimental'noy Meditsiny = Siberian Journal of Clinical and Experimental Medicine*. 2021;36(4):70-77. doi 10.29001/2073-8552-2021-36-4-70-77 (in Russian)

Govindaraj P., Khan N.A., Rani B., Rani D.S., Selvaraj P., Jyothi V., Bahl A., Narasimhan C., Rakshak D., Premkumar K., Khullar M., Thangaraj K. Mitochondrial DNA variations associated with hypertrophic cardiomyopathy. *Mitochondrion*. 2014;16:65-72. doi 10.1016/j.mito.2013.10.006

Govindaraj P., Rani B., Sundaravadivel P., Vanniarajan A., Indumathi K.P., Khan N.A., Dhandapany P.S., … Rakshak D., Rathinavel A., Premkumar K., Khullar M., Thangaraj K. Mitochondrial genome variations in idiopathic dilated cardiomyopathy. *Mitochondrion*. 2019;48:51-59. doi 10.1016/j.mito.2019.03.003

Hagen C.M., Aidt F.H., Havndrup O., Hedley P.L., Jensen M.K., Kanters J.K., Pham T.T., Bundgaard H., Christiansen M. Private mitochondrial DNA variants in danish patients with hypertrophic cardiomyopathy. *PLoS One*. 2015;10(4):e0124540. doi 10.1371/journal.pone.0124540

Hudson G., Gomez-Duran A., Wilson I.J., Chinnery P.F. Recent mitochondrial DNA mutations increase the risk of developing common late-onset human diseases. *PLoS Genet*. 2014;10(5):e1004369. doi 10.1371/journal.pgen.1004369

JASP Team. Version 0.19.3. [Computer software]. 2024. Available: https://jasp-stats.org/

Ji Y., Zhang J., Yu J., Wang Y., Lu Y., Liang M., Li Q., … Zhu T., Mo J.Q., Huang T., Jiang P., Guan M.X. Contribution of mitochondrial ND1 3394T>C mutation to the phenotypic manifestation of Leber's hereditary optic neuropathy. *Hum Mol Genet*. 2019;28(9):1515-1529. doi 10.1093/hmg/ddy450

Kibel A., Lukinac A.M., Dambic V., Juric I., Selthofer-Relatic K. Oxidative stress in ischemic heart disease. *Oxid Med Cell Longev.* 2020; 2020:6627144. doi 10.1155/2020/6627144

Koplan B.A., Stevenson W.G. Ventricular tachycardia and sudden cardiac death. *Mayo Clin Proc*. 2009;84(3):289-297. doi 10.4065/84.3.289

Kytövuori L., Junttila J., Huikuri H., Keinänen-Kiukaanniemi S., Majamaa K., Martikainen M.H. Mitochondrial DNA variation in sudden cardiac death: a population-based study. *Int J Legal Med*. 2020;134(1):39-44. doi 10.1007/s00414-019-02091-4

McCormick E.M., Lott M.T., Dulik M.C., Shen L., Attimonelli M., Vitale O., Karaa A., … Zhang S., Procaccio V., Wallace D.C., Gai X., Falk M.J. Specifications of the ACMG/AMP standards and guidelines for mitochondrial DNA variant interpretation. *Hum Mutat*. 2020;41(12):2028-2057. doi 10.1002/humu.24107

Montaigne D., Pentiah A.D. Mitochondrial cardiomyopathy and related arrhythmias. *Card Electrophysiol Clin*. 2015;7(2):293-301. doi 10.1016/j.ccep.2015.03.008

Müller-Nedebock A.C., Pfaff A.L., Pienaar I.S., Kõks S., van der Westhuizen F.H., Elson J.L., Bardien S. Mitochondrial DNA variation in Parkinson's disease: Analysis of "out-of-place" population variants as a risk factor. *Front Aging Neurosci*. 2022;14:921412. doi 10.3389/fnagi.2022.921412

Ng Y.S., Turnbull D.M. Mitochondrial disease: genetics and management. *J Neurol*. 2016;263(1):179-191. doi 10.1007/s00415-015-7884-3

Palacín M., Alvarez V., Martín M., Díaz M., Corao A.I., Alonso B., Díaz-Molina B., … Cannata-Andía J., Batalla A., Ruiz-Ortega M., Martínez-Camblor P., Coto E. Mitochondrial DNA and *TFAM* gene variation in early-onset myocardial infarction: evidence for an association to haplogroup H. *Mitochondrion*. 2011;11(1):176-181. doi 10.1016/j.mito.2010.09.004

Piotrowska-Nowak A., Elson J.L., Sobczyk-Kopciol A., Piwonska A., Puch-Walczak A., Drygas W., Ploski R., Bartnik E., Tonska K. New mtDNA association model, MutPred variant load, suggests individuals with multiple mildly deleterious mtDNA variants are more likely to suffer from atherosclerosis. *Front Genet*. 2019;9:702. doi 10.3389/fgene.2018.00702

Rentzsch P., Schubach M., Shendure J., Kircher M. CADD-Splice-improving genome-wide variant effect prediction using deep learning-derived splice scores. *Genome Med*. 2021;13(1):31. doi 10.1186/s13073-021-00835-9

Roselló-Díez E., Hove-Madsen L., Pérez-Grijalba V., Muñoz-Guijosa C., Artigas V., Maria Padró J., Domínguez-Garrido E. Mitochondrial genetic effect on atrial fibrillation: a case-control study. *Mitochondrion*. 2021;56:15-24. doi 10.1016/j.mito.2020.11.007

Severino P., D'Amato A., Pucci M., Infusino F., Birtolo L.I., Mariani M.V., Lavalle C., Maestrini V., Mancone M., Fedele F. Ischemic

heart disease and heart failure: role of coronary ion channels. *Int J Mol Sci*. 2020;21(9):3167. doi 10.3390/ijms21093167

Torroni A., Petrozzi M., D'Urbano L., Sellitto D., Zeviani M., Carrara F., Carducci C., Leuzzi V., Carelli V., Barboni P., De Negri A., Scozzari R. Haplotype and phylogenetic analyses suggest that one European-specific mtDNA background plays a role in the expression of Leber hereditary optic neuropathy by increasing the penetrance of the primary mutations 11778 and 14484. *Am J Hum Genet*. 1997;60(5):1107-1121

van Opbergen C.J.M., den Braven L., Delmar M., van Veen T.A.B. Mitochondrial dysfunction as substrate for arrhythmogenic cardiomyopathy: a search for new disease mechanisms. *Front Physiol*. 2019;10:1496. doi 10.3389/fphys.2019.01496

van Oven M., Kayser M. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum Mutat*. 2008; 30(2):E386-E394. doi 10.1002/humu.20921

Weissensteiner H., Forer L., Kronenberg F., Schönherr S. mtDNA-Server 2: advancing mitochondrial DNA analysis through highly parallelized data processing and interactive analytics. *Nucleic Acids Res*. 2024;52(W1):W102-W107. doi 10.1093/nar/gkae296

Yang L., Wang S., Wu J., Ma L.L., Li Y., Tang H. Editorial: mitochondrial metabolism in ischemic heart disease. *Front Cardiovasc Med*. 2022;9:961580. doi 10.3389/fcvm.2022.961580

Conflict of interest. The authors declare no conflict of interest.
Received March 11, 2025. Revised April 15, 2025. Accepted May 30, 2025.

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2025;29(5):685-692

doi 10.18699/vjgb-25-75

Crouzon syndrome: preimplantation genetic testing for a familial case with a whole and a mosaic variant of the disease

E.V. Soloveva [1], M.M. Skleimova [1], L.I. Minaycheva [1], A.F. Garaeva [1], E.M. Bakulina [2], E.A. Ladygina [2], O.R. Kanbekova [3], G.N. Seitova [1]

[1] Research Institute of Medical Genetics, Tomsk National Research Medical Center of the Russian Academy of Sciences, Tomsk, Russia
[2] Vitromed LLC, Novosibirsk, Russia
[3] Regional Perinatal Center named after I.D. Yevtushenko, Tomsk, Russia
elena.soloveva@medgenetics.ru

Abstract. Crouzon syndrome, which is a hereditary craniosynostosis, can be the result of inheritance from either parent, as well as *de novo* mutations in the *FGFR2* gene. With a confirmed molecular genetic diagnosis, preimplantation genetic testing for monogenic diseases (PGT-M) is available for high-risk families. However, there is currently little information in the literature about using this approach to prevent this condition. The aim of our study was to describe the clinical case of IVF/ICSI with PGT-M for Crouzon syndrome with a successful outcome and confirmatory diagnostics. PGT-M was planned and performed for a married couple (aged 24 and 25), in which the husband had Crouzon syndrome. The husband's father had a milder form of Crouzon syndrome and the pathogenic variant of the *FGFR2* gene was in a mosaic form. During preparation, a testing system was selected for the pathogenic variant NM_000141.5(*FGFR2*):c.1007A>G (p.Asp336Gly) of the *FGFR2* gene, and gene-linked polymorphic microsatellite markers. The STR markers in the husband's father excluded chimerism for the pathogenic variant and indicated mosaicism with the involvement of germ cells. Molecular genetic analysis was performed using a nested PCR, with detection by fragment analysis for STRs and restriction analysis of the pathogenic variant. During the IVF program, superovulation stimulation and embryological procedures were performed according to standard protocols. Fertilization was achieved using the ICSI method, and blastocyst biopsy was done on the sixth day of development. For PGT-M, a direct analysis of pathogenic variants and an indirect analysis of selected informative STRs were used. The thawed embryos were transferred based on the results of preimplantation testing. We selected twelve STRs flanking the *FGFR2* gene, eight informative ones were used during PGT-M. In the IVF program, 15 mature oocytes were obtained, then four blastocysts were biopsied. One of the four embryos inherited a normal paternal chromosome, the other three had the pathogenic variant and the associated risk haplotype. A singleton pregnancy has occurred as a result of embryo transfer recommended after PGT-M. Following the child's birth, molecular diagnostics were performed, confirming the PGT-M result. The presented clinical case provides an effective example of IVF with PGT-M to prevent the birth of affected children in families with hereditary craniosynostosis.

Key words: Crouzon syndrome; PGT-M; preimplantation genetic testing; IVF; mosaicism; *FGFR2* gene

For citation: Soloveva E.V., Skleimova M.M., Minaycheva L.I., Garaeva A.F., Bakulina E.M., Ladygina E.A., Kanbekova O.R., Seitova G.N. Crouzon syndrome: preimplantation genetic testing for a familial case with a whole and a mosaic variant of the disease. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov J Genet Breed*. 2025;29(5):685-692. doi 10.18699/vjgb-25-75

Funding. The work was funded within the state guaranteed order No. 123041700028-8 of the Ministry of Science and Higher Education of the Russian Federation (Research Institute of Medical Genetics, Tomsk National Research Medical Center of the Russian Academy of Sciences).

Синдром Крузона: преимплантационное генетическое тестирование для семейного случая с полным и мозаичным вариантом заболевания

Е.В. Соловьёва [1], М.М. Склеймова [1], Л.И. Минайчева [1], А.Ф. Гараева [1], Е.М. Бакулина [2], Е.А. Ладыгина [2], О.Р. Канбекова [3], Г.Н. Сеитова [1]

[1] Научно-исследовательский институт медицинской генетики Томского национального исследовательского медицинского центра Российской академии наук, Томск, Россия
[2] ООО «Витромед», Новосибирск, Россия
[3] Областной перинатальный центр им. И.Д. Евтушенко, Томск, Россия
elena.soloveva@medgenetics.ru

© Soloveva E.V., Skleimova M.M., Minaycheva L.I., Garaeva A.F., Bakulina E.M., Ladygina E.A., Kanbekova O.R., Seitova G.N., 2025
This work is licensed under a Creative Commons Attribution 4.0 License

Аннотация. Синдром Крузона, относящийся к наследственным краниосиностозам, может быть как результатом наследования от одного из родителей, так и вариантом *de novo* мутаций в гене *FGFR2*. При подтвержденном молекулярно-генетическом диагнозе для семей высокого риска возможно проведение преимплантационного генетического тестирования моногенного заболевания (ПГТ-М). Однако на сегодняшний день в литературе мало сведений о таком подходе к профилактике данного заболевания. Целью нашей работы явилось описание клинического случая ЭКО/ИКСИ с ПГТ-М для синдрома Крузона с успешным исходом и подтверждающей диагностикой. Планирование и проведение ПГТ-М выполнены для супружеской пары (24 и 25 лет), в которой синдром Крузона был у мужа. У отца супруга синдром Крузона был в более легкой форме и патогенный вариант гена *FGFR2* присутствовал в мозаичном варианте. На подготовительном этапе подобрана система тестирования патогенного варианта NM_000141.5(*FGFR2*):c.1007A>G (p.Asp336Gly) гена *FGFR2* и сцепленных с геном полиморфных микросателлитных маркеров. Профиль STR-маркеров у отца супруга исключал химеризм по патогенному варианту и свидетельствовал о мозаицизме, затрагивающем половую линию. Молекулярно-генетический анализ проводили методом гнездовой ПЦР с детекцией фрагментным анализом для STR и рестрикционного анализа для патогенного варианта. В ходе программы ЭКО стимуляцию суперовуляции и эмбриологические процедуры выполняли по стандартным протоколам, оплодотворение проводили методом ИКСИ. Биопсия бластоцист выполнялась на 6-е сутки развития. При ПГТ-М использовали прямой анализ патогенного варианта и косвенный анализ по отобранным на подготовительном этапе информативным STR. Перенос размороженного эмбриона выполнен с учетом результатов преимплантационного тестирования. Нами было отобрано двенадцать STR, фланкирующих ген *FGFR2*, из них восемь информативных были использованы в ходе ПГТ-М. В программе ЭКО получено 15 зрелых ооцитов и в дальнейшем четыре бластоцисты, которые были биопсированы. Один из четырех эмбрионов унаследовал нормальную отцовскую хромосому, в трех остальных выявлялся патогенный вариант и ассоциированный с ним гаплотип риска. В результате переноса в полость матки эмбриона, рекомендованного по результатам ПГТ-М, наступила одноплодная беременность. После рождения ребенка проведена молекулярно-генетическая диагностика, подтвердившая результат ПГТ-М. Представленный нами клинический случай демонстрирует эффективный пример применения ЭКО с ПГТ-М для предотвращения рождения больных детей в семьях с наследственными краниосиностозами.

Ключевые слова: синдром Крузона; ПГТ-М; преимплантационное генетическое тестирование; ЭКО, мозаицизм; ген *FGFR2*

Introduction

Crouzon syndrome belongs to an extensive heterogeneous group of craniosynostosis – birth defects that are characterized by premature fusion of one or more sutures in the cranial vault prior to the completion of growth and development of the brain, which leads to limited growth of the skull, brain, face and central nervous system development. Among craniosynostosis, syndromic craniosynostoses are estimated to comprise 15 % of all cases. To date, there are over 180 craniosynostosis syndromes identified. About 8 % of cases are familial or inherited (Al-Namnam et al., 2019). Crouzon syndrome (OMIM 123500) is the most common syndrome among hereditary craniosynostosis. It is caused by mutations in the fibroblast growth factor receptor 2 (*FGFR2*) gene, which is located on chromosome 10 (10q26.13). The FGFR2 protein is involved in cell signaling, and disruption of the FGFs/FGFR2 signaling pathway leads to abnormal differentiation, proliferation and apoptosis (Al-Namnam et al., 2019; Yapijakis et al., 2023).

The incidence of Crouzon syndrome is approximately 16.5 cases per million live births (1:60,000). We have not found any data on the incidence of the disease in Russia. The disease is inherited as an autosomal dominant trait with incomplete penetrance and variable expressivity. In approximately 70 % of cases the disease is inherited from one parent, while in other cases, it is the result of a *de novo* mutation (Al-Namnam et al., 2019). The syndrome was first described by Louis Edouard Octave Crouzon in 1912.

The disease usually manifests itself in the first three years of life. It can be suspected during the antenatal period using ultrasound examination. It is also often detected at birth due to its classic signs in newborns, which include craniosynostosis, hypoplasia of the middle part of the face, proptosis (exophthalmos), and, in some cases, a beaked-shaped nose. Other common manifestations of the syndrome include coronary craniosynostosis with other cranial sutures fusion, brachycephaly, hypertelorism, prominent frontal tubercles, strabismus, orbital proptosis, mandibular prognathism, and maxillary hypoplasia. These signs either become more pronounced or may regress over time (Al-Namnam et al., 2019). Hearing loss is also common (55 %), and fusion of the C2 and C3 vertebrae occurs in 30 % of cases. Another manifestation may be progressive hydrocephalus (30 %). The patients' mental abilities are usually normal, but in some cases, increased intracranial pressure can lead to intellectual disability. The differential diagnosis of Crouzon syndrome includes Apert syndrome and Pfeiffer syndrome. Historically, these diseases were described separately, but the overlap of spectrum of molecular genetic disorders suggests that these conditions form a continuum (Koltunov, 2011; Al-Namnam et al., 2019).

Treatment of Crouzon syndrome is based on the severity of symptoms. To optimize treatment, a comprehensive assessment by a multidisciplinary team of specialists is needed. The main treatment method is surgery. It allows to perform correction of the face skull and eye sockets in order to optimize cerebral blood flow and prevent the effects of increased

intracranial pressure, blindness and mental retardation. Surgical intervention can be staged or combined, depending on the level of functional impairment in the patients and their age (Koltunov, 2011; Sokolova et al., 2024).

Molecular genetic examination is currently an important part in the diagnosis of Crouzon syndrome. If there is a high risk of transmission of Crouzon syndrome to offspring, preimplantation genetic testing for monogenic disease (PGT-M) may be used (Wenger et al., 1993–2025). To achieve this, the couple needs to resort to IVF (*in vitro* fertilization). PGT-M is currently used worldwide for a wide range of monogenic diseases, however, we found in the literature a description of only one clinical case of PGT-M for Crouzon syndrome (Abou-Sleiman et al., 2002).

In our work, we presented a clinical case of IVF/ICSI with PGT-M with a successful outcome and confirmatory diagnostics.

Material and methods

IVF with PGT-M was performed for a married couple (a 24-year-old woman and a 25-year-old man), in which the husband had the disease status of Crouzon syndrome. The phenotypic manifestations of the man included hypertelorism, a beaked nose, exophthalmos (not very pronounced), hypoplasia of the middle part of the face, a high palate and a protruding chin. His father also had Crouzon syndrome. The father's clinical manifestations were milder, with no exopthalmos, hypoplasia, or hypertelorisms. Familial pathogenic variant of NM_000141.5(FGFR2):c.1007A>G (p.Asp336Gly) was detected in a heterozygous state in the spouse and his father as a result of Sanger sequencing; "14 % mosaicism or chimerism with heterozygosity" for the father was noted (Genoanalytika LLC, Moscow, 2022). The abbreviation for the pathogenic variant, D336G, is used in this article. The variant is described in the ClinVar database (Variation ID: 374815) and is associated with the development of Crouzon syndrome (OMIM 123500). At least three *in silico* pathogenicity prediction programs confirmed the pathogenic effect of the variant on the gene or gene product (AlphaMissense, Revel, Aggregated Prediction). The nucleotide sequence variant was not registered in the control sample of The Genome Aggregation Database v4.1.0 and in the Russian Federal Medical and Biological Agency database of population frequencies of genetic variants for the Russian Federation population (version 1.1.2, database version 59.1, https//gdbpop.nir.cspfmba.ru). This missense variant is located in the gene where missense variants often cause the disease.

The couple underwent detailed medical genetic counseling, including the issues of planned molecular genetic testing system and ART (assisted reproductive technologies) counseling. Voluntary informed consent was obtained from all persons involved in the study for all procedures.

The biological material for the PGT-M setup was venous blood samples from spouses and parents of the husband, as well as from a voluntary unrelated donor (for DNA testing and single lymphocytes) collected into EDTA-containing vacutainer tubes. For postnatal diagnosis, the newborn's dried blood spots on a filter blank were used. DNA isolation was carried out using the DNA-sorb-B kit (Amplisens, Russia) from 100 µl blood. DNA from a dried blood spot sample was isolated using the PREP-MB-DBS DWP kit (DNA-Technology, Russia) at the Auto-Pure 96 robotic station (Hangzhou Allsheng Instruments Co., Ltd, China). Samples of single lymphocytes were used as biological material to validate the system. For this purpose, a suspension of mononuclear cells was isolated from blood by centrifugation in a density gradient of ficoll solution. Single lymphocytes were then taken with a glass biopsy micropipette under a microscope using a polyvinylpyrrolidone solution and each was placed into an individual microtube with a lysing solution. The lysing solution for single lymphocytes and embryo trophectoderm samples contained proteinase K, triton X-100, and twin 20 (Verlinsky, Kuliev, 2004).

For PGT-M, a testing system has been developed that includes analysis of the pathogenic variant D336G of the *FGFR2* gene, as well as polymorphic STR markers (short tandem repeats) linked to the *FGFR2* gene. STRs have been selected within 1 million bp from the pathogenic variant with heterozygosity ranging from 0.8 to 0.9 for dinucleotide repeats, and no less than 0.7 for tri-tetranucleotide repeats. At the PGT-M setup stage, PCR reactions and optimization (gradient PCR if necessary) were performed for all fragments for test DNA, followed by analysis of family samples to identify informative STRs. The developed system was validated on single lymphocytes ($N = 8$).

Standard PCR (2nd PCR) in a volume of 20 µl with fluorescently labeled oligonucleotide primers was used to test DNA samples from family members. Nested PCR was used during single-cell testing, both during PGT-M and single lymphocyte validation. The first round of PCR (1st PCR) was multiplexed in a volume of 50 µl. 1 µl of the 1st PCR product was used as a template for the 2nd PCR. Negative and positive controls were used in all amplification series. The overall scheme of molecular genetic testing, as well as conditions and amplification programs were carried out according to the recommendations of Y. Verlinsky and A. Kuliev (Verlinsky, Kuliev, 2005).

The 2nd PCR products were preliminarily examined in a 7 % polyacrylamide gel to assess the loading of samples on capillary electrophoresis and the quality of negative controls. Fragment analysis of STR markers was performed by capillary electrophoresis on a Nanophore 05 genetic analyzer (Institute of Analytical Instrumentation, Russia). The GeneMarker software (SoftGenetics, USA) was used to evaluate the fragment analysis results. The study of the pathogenic variant involved restriction analysis with HspAI and detection in a 7 % polyacrylamide gel. GelRed dye (Biotinum, USA) was used to color gels.

The couple underwent standard pre-treatment assessment according to national guidelines to plan IVF (Assisted Reproductive Technology…, 2019). Controlled ovarian hyperstimulation was performed using standard short protocol with recombinant FSH (follicle stimulating hormone) and

antagonists. The dose of FSH was selected individually. Ovulation was triggered 0.2 mg of Decapeptil (Ferring GmbH, Germany) when at least three follicles reached 17 mm in size according to ultrasound monitoring. Transvaginal follicle puncture was performed under ultrasound control 36 hours later. WHO standard criteria were used for ejaculate analysis (WHO Laboratory Manual…, 2010).

ICSI (intracytoplasmic sperm injection) was used as a fertilization technique. Embryological procedures and embryo biopsy were carried out taking into account national and international guidelines (ESHRE PGT Consortium and SIG-Embryology Biopsy Working Group et al., 2020; Evaluation of Oocytes…, 2021). A single-step medium SAGE 1-Step™ (ORIGIO, Denmark) was used for embryo culture. Biopsy of good and excellent quality embryos was performed on day 6 using a flip method. The obtained trophectoderm fragments, after washing in phosphate buffer (PBS), were transferred into microtubes containing 5 µl of lysing buffer and frozen at –20 °C. Immediately after the biopsy, blastocysts were vitrified using Kitazato media and carriers (Kitazato Corporation, Dibimed-Biomedical Supply, Spain).

PGT-M was performed using the nested PCR method, according to the scheme developed at the preliminary stage, taking into account the aforementioned conditions and international recommendations (ESHRE PGT-M Working Group et al., 2020).

The preparation for the transfer of the thawed embryo was carried out in a natural cycle, taking into account the results of PGT-M. Kitazato thawing media (Kitazato Corporation, Dibimed-Biomedical Supply, Spain) and a transfer catheter Guardia™ Access ETEmbryo Transfer Catheter (COOK Medical, USA) were used.

Pregnancy diagnosis was performed by standard analysis of human chorionic gonadotropin (HCG) on the 14th day after the embryo transfer, followed by ultrasound examination of the gestation at 7 weeks. No invasive prenatal diagnosis was made.

Postnatal diagnosis was performed using samples of dried newborn bloodstains. Molecular genetic testing of the newborn included all the loci tested during PGT-M.

The study was carried out using the resources of the Biobank of the Population of Northern Eurasia biocollection at the Research Institute of Medical Genetics of the Tomsk National Research Medical Center and the equipment of the Medical Genomics Center for Collective Use at the Tomsk National Research Medical Center.

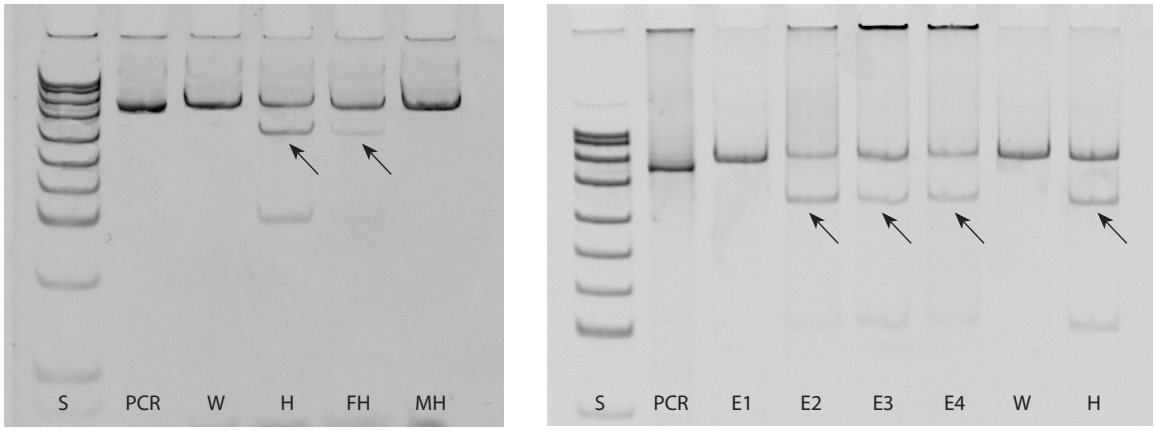
Results

During the PGT-M setup, a test system was developed for a family with a high genetic risk (50 %) of Crouzon syndrome, including the analysis of the pathogenic variant D336G and 12 polymorphic STR markers. The presence of the pathogenic variant was confirmed by restriction analysis in the husband and his father, both for the preliminary examination of the family and for the PGT-M (see the Figure).

We have recorded a difference in the intensity of restriction fragments between the father and the husband. In the husband's father, the fragment corresponding to the allele with the pathogenic variant looked significantly paler than normal. This was consistent with the results of molecular genetic examination provided by the family for PGT-M planning, which also noted mosaicism in the husband's father.

The informative value of STR markers for PGT-M was established through the family analysis based on testing samples from the husband's parents. Table 1 shows the polymorphic STR markers we selected and studied in the family samples.

Nine of the twelve STRs tested were informative regarding the husband's chromosome carrying the pathogenic variant. The husband's father did not have any additional alleles across the entire spectrum of highly polymorphic microsatellite markers studied, which excluded the assumption of chimerism.



The result of testing the pathogenic variant D336G in the FGFR2 gene by restriction analysis with gel detection (inverted image) in family samples from PGT-M setup and embryo samples during PGT-M. W – wife/woman; H – husband/man; FH – husband's father; MH – husband's mother; E1–E4 – the embryo samples; PCR – PCR product without endonuclease digestion; S – the pUC19/MspI size standard.

Table 1. Results of the preliminary stage of PGT-M for a family with Crouzon syndrome

No.	STR/Gene, Repeat (distance from the pathogenic variant, Mb), localization (chr10:/GRCh37/hg19)	Alleles (bp)			
		Husband	Wife	Husband's father	Husband's mother
1	AG (0.98) 122297505–122297597	**255**/257	(239/247)	**255**/245	247/257
2	AT (0.85) 122422706–122422730	**354**/354	374/374	**354**/374	352/354
3	ACAG (0.81) 122466425–122466476	**250**/250	208/208	**250**/208	246/250
4	AAT (0.74)* 122533089–122533125	**282**/273	(291/294)	**282**/291	282/273
5	AC (0.68) 122593191–122593232	**160**/152	(162/160)	**160**/154	160/152
6	AC (0.55) 122725637–122725689	**236**/242	(254/246)	**236**/258	234/242
7	*FGFR2*	**D336G**/N	N/N	**D336G**/N	N/N
8	AC (0.01) 123283616–123283661	**256**/258	(258/260)	**256**/256	246/258
9	AT (0.22) 123493200–123493222	**186**/186	(186/194)	**186**/186	186/186
10	AAT (0.61) 123888576–123888611	**161**/149	(161/149)	**161**/149	152/149
11	AG (0.66) 123935589–123935661	**136**/128	(128/108)	**136**/144	138/128
12	AC (0.76) 124039027–124039061	**246**/244	238/238	**246**/232	248/244
13	ATC (0.95) 124227919–124227959	**90**/96	(96/87)	**90**/90	87/96

Note. Alleles linked to the pathogenic variant are indicated in bold and underlined; uninformative STRs are marked in gray; alleles with coupling that were not established during the preliminary stage are marked in parentheses. * A locus that has not been validated on single cells is marked.

Next, the testing system was validated using samples of single lymphocytes and products obtained from whole-genome amplification of embryos. Acceptable amplification in terms of fragment intensity, absence of non-specific fragments and peak shape was confirmed for all STRs except for one. One STR located at a distance of 0.74 million bp was excluded from the system due to lack of amplification on single cells.

During the IVF treatment cycle, a starting dose of gonadotropins 200 IU FSH + 75 IU LH (luteinizing hormone) was used. The total gonadotropin dose was 1,950 IU FSH + 600 IU LH. No complications occurred during stimulation. 21 oocytes were retrieved from 23 follicles, 15 of them being mature. Fertilization was performed by ICSI in order to reduce contamination of parental DNA during embryo testing. The husband's ejaculate was teratozoospermic; the sperm concentration was 167 million/ml, the percentage of progressively mobile was 49 %. The next day after fertilization, 12 zygotes with two pronuclei were formed. On day 3, 11 embryos were developing. Four embryos have reached the blastocyst stage, and all of them have been successfully biopsied for genetic testing.

Trophectoderm fragments from four embryos were tested for the familial pathogenic variant responsible for Crouzon syndrome and the linked STR markers selected at the preliminary stage. The preimplantation study included only informative STRs validated on single cells.

The results of preimplantation analysis of pathogenic variant of the *FGFR2* gene are shown in the Figure. During PGT-M, the paternal pathogenic variant D336G of the *FGFR2* gene was found in a heterozygous state in three embryo samples. A normal *FGFR2* allele was detected in the sample of the first embryo. The genotype of embryo 1 can be interpreted as homozygous for the normal allele. However, due to the risk of allele dropout (ADO) in PGT-M, the genotype was interpreted only in conjunction with STR results. The full results of preimplantation testing are presented in Table 2.

All presumably homozygous profiles in the embryo samples (Table 2) are indicated as a single allele, as a precautionary measure in interpretation, due to the possible phenomenon of ADO. The results of STR testing showed that the samples from three embryos carrying the pathogenic variant contained paternal alleles linked to the pathogenic variant of the gene. A normal paternal haplotype was detected in embryo 1, along with one of the maternal haplotype, which confirmed the normal homozygous status of the embryo in relation to the pathogenic variant.

Genetic testing for aneuploidy was not carried out, based on the couple's decision following medical genetic counseling.

The transfer of thawed embryo number 1 into the uterine cavity was performed based on the preimplantation testing results. Cryotransfer was performed in a naturally modified ovulation trigger cycle (HCG 6,500 IU).

The result was a singleton pregnancy, confirmed by ultrasound. The pregnancy progressed without complications. Standard prenatal screening did not detect any fetal

Table 2. PGT-M results and confirmatory diagnostic results for a family with Crouzon syndrome

No.	STR/Gene, Repeat (Mb)	Alleles (bp)						
		PGT-M: embryo samples				Control family samples		IVF-PGT-M baby
		E1	E2	E3	E4	Husband	Wife	
1	AG (0.98)	257/247	**255**/239	**255**/247	**255**/247	**255**/257	239/247	257/247
2	AC (0.68)	152/160	**160**/162	*160*	*160*	**160**/152	162/160	152/160
3	AC (0.55)	242/246	**236**/254	**236**/246	**236**/246	**236**/242	254/246	242/246
4	*FGFR2*	*N*	**D336G**/N	**D336G**/N	**D336G**/N	**D336G**/N	N/N	N/N
5	AC (0.01)	258/260	**256**/258	**256**/260	**256**/260	**256**/258	258/260	258/260
6	AAT (0.61)	*149*	***161***	**161**/149	**161**/149	**161**/149	161/149	149/149
7	AG (0.66)	128/108	**136**/128	**136**/108	**136**/108	**136**/128	128/108	128/108
8	AC (0.76)	244/238	**246**/238	**246**/238	**246**/238	**246**/244	238/238	244/238
9	ATC (0.95)	96/87	**90**/96	**90**/87	**90**/87	**90**/96	96/87	96/87

Note. Alleles linked to a pathogenic variant are shown in bold and underlined; loci in italics can be interpreted individually in embryo samples as homozygous or one allele due to the ADO of the second allele.

abnormalities. Invasive prenatal diagnosis, recommended by genetic counseling to check the genotype of the fetus, was not performed due to the patient's concerns about potential complications. The pregnancy was terminated via cesarean section at 39 weeks of gestation. In May 2024, a healthy baby girl weighting 3,480 g was born. Her Apgar score was 7/8 according to the discharge records.

Confirmatory genetic diagnosis was performed using dried newborn blood sample for the familial variant of Crouzon syndrome. The testing was carried out on all loci included in the PGT-M (Table 2). As a result, the child had a genotype homozygous for the normal allele of the *FGFR2* gene, confirmed by the paternal haplotype. The postnatal confirmatory diagnosis results completely matched the results of embryo 1 testing during PGT-M.

Discussion

Our article provides a detailed description of a clinical case of IVF with PGT-M to prevent the birth of a child with Crouzon syndrome in a family where the husband and his father suffered from this disease. The familial pathogenic variant of the *FGFR2* gene in the husband's father was present in a mosaic form and was not a consequence of chimerism, as proven by our data from microsatellite analysis. This most likely caused a milder clinical manifestation of the disease. For embryos, we recorded the standard heterozygous (non-mosaic) pathogenic variant, as well as for the husband (see the Figure).

A case of generative and somatic mosaicism for the same pathogenic variant c.1007A>G(p.Asp336Gly), as in our clinical case, is described in the work of A. Goriely et al. (2010). In a mother without Crouzon syndrome, a heterozygous mosaic variant c.1007A>G(p.Asp336Gly) was detected in approximately 25–30 % of blood and saliva cells. In a child with Crouzon syndrome, this pathogenic variant was present, as it was in our case, in the standard heterozygous form (Goriely et al., 2010).

Cases of transmission of an autosomal dominant disease from a parent who is a mosaic carrier of a pathogenic variant have also been found for other conditions, in particular, for autosomal dominant polycystic kidney disease (Hopp et al., 2020). The prevalence of this phenomenon is currently unclear for various hereditary diseases.

During the IVF program, four blastocysts were obtained, all of which were successfully biopsied and tested for Crouzon syndrome. We used a strategy based on analyzing the pathogenic variant and linked informative STR markers that were selected during the PGT-M setup. We used nested PCR but not genome-wide amplification, because the combination of PGT-M and aneuploidy analysis was not performed for this clinical case. Our approach allowed us to draw definitive conclusions regarding the inheritance of the chromosome responsible for the familial variant of Crouzon syndrome in each embryo. The dominant nature of inheritance of the disease determines a 50 % risk for embryos. In our case, 3 out of 4 (75 %) embryos had the pathogenic variant. One embryo had a normal status, and its transfer led to the achievement of pregnancy and birth of a healthy child.

There is little data on PGT-M for Crouzon syndrome in scientific publications. In 2002, a group from the UK published the results, reporting the first successful PGT for this condition (Abou-Sleiman et al., 2002; Harper et al., 2002). Both articles describe the same clinical case. One article presents the test results in more detail (Abou-Sleiman et al., 2002), while the second one focuses on the outcomes (Harper et al., 2002). A married couple, in which a woman had Crouzon syndrome, underwent pre-implantation testing during two IVF cycles. They had a history of the birth of a girl with

Crouzon syndrome, who later died during corrective surgery. The preimplantation testing system included the analysis of the pathogenic variant using the SSCP (single strand conformational polymorphism) method. Two intragenic SNPs were tested during preparation for the diagnosis, but they were uninformative for the spouses (Abou-Sleiman et al., 2002). In addition to the *FGFR2* gene fragment, unlinked D21S11 locus was included in the preimplantation analysis to control amplification. In just two cycles, 23 day 3 embryos were analyzed. The strategy of taking two blastomeres during biopsy was used. Eight embryos were diagnosed as normal, nine as abnormal and six were inconclusive. Five embryos were transferred during two embryo transfer procedures. The second cycle of PGT resulted in a twin pregnancy, but only one embryo had a heartbeat. Prenatal diagnosis was not performed at the request of the married couple and confirmatory diagnosis was done postnatally. A healthy baby boy was born (Harper et al., 2002).

We have not found any other descriptions of PGT-M for Crouzon syndrome. During our work, we did not encounter any specific difficulties with PGT-M related to the disease, gene, or specific pathogenic variant. Due to the complex molecular mechanisms involved in craniosynostosis development, this group may have attracted less attention from specialists as a method for preventing the disease until now. Our experience shows that PGT-M can successfully be applied in cases where there is a clear monogenic inheritance.

In comparison with the above-described clinical case from the literature, in our case, pregnancy was achieved in the first treatment cycle of IVF with PGT. The differences can be noted in a number of aspects of the entire process related to technology changes in this area. We had fewer embryos for preimplantation testing. In general, when testing trophectoderm samples from 5 or 6-day embryos, as it was in our study, there are fewer samples for PGT-M compared to testing blastomeres on day 3, because of natural selection of embryos from day 3 to 5.

For the genetic analysis we used a more detailed testing system, which included a clear identification of the pathogenic variant by restriction analysis, and at least two informative flanking STRs for the family. To analyze the pathogenic variant properly, a restriction analysis was used based on the natural site of the restriction endonuclease in the case of the pathogenic variant. This approach proved to be more informative than the SSCP method. Currently, a wide range of innovative genome-wide methods have been proposed in the field of PGT-M (De Rycke, Berckmoes, 2020). At the same time, it is important to consider the potential benefits and drawbacks of these methods for each specific case.

The STRs included in our study allowed us to identify all the difficulties of amplification and interpretation of PGT results, including ADO, contamination, and recombination. Even in the cases where amplification of the pathogenic allele fails, haplotypes can allow us to clearly identify embryos carrying a chromosome with the paternal pathogenic variant. According to modern standards, targeted preimplantation analysis without linked informative markers is not recommended (ESHRE PGT-M Working Group et al., 2020).

In the clinical case presented by us, PGT for aneuploidy was not performed. The spouses were of a relatively young age; there were no other indications for this procedure.

Conclusion

Our study presents a clinical case of IVF with PGT-M for a married couple at high risk for Crouzon syndrome, with a successful outcome confirmed by postnatal diagnosis. In comparison with the only case presented in the literature, our results reflect a more modern and, in some aspects, a more reliable approach to IVF with PGT-M for Crouzon syndrome. In the family studied, the pathogenic variant of the *FGFR2* gene was present in a mosaic form in the husband's father, while in the husband and, consequently, in embryos it had the standard heterozygous state. Microsatellite analysis used in our work excluded chimerism in the husband's father. The issues of transmission of dominant diseases in the cases of parental mosaicism require further research. The clinical case presented by us demonstrates an effective example of the use of IVF with PGT-M to prevent the birth of sick children in families with hereditary craniosynostosis.

References

Abou-Sleiman P.M., Apessos A., Harper J.C., Serhal P., Delhanty J.D.A. Pregnancy following preimplantation genetic diagnosis for Crouzon syndrome. *Mol Hum Reprod.* 2002;8(3):304-309. doi 10.1093/ molehr/8.3.304

Al-Namnam N.M., Hariri F., Thong M.K., Rahman Z.A. Crouzon syndrome: genetic and intervention review. *J Oral Biol Craniofac Res.* 2019;9(1):37-39. doi 10.1016/j.jobcr.2018.08.007

Assisted Reproductive Technology and Artificial Insemination. Clinical guidelines (treatment protocol). Ministry of Health of the Russian Federation, 2019 (in Russian)

De Rycke M., Berckmoes V. Preimplantation genetic testing for monogenic disorders. *Genes*. 2020;11(8):871. doi 10.3390/genes 11080871

ESHRE PGT Consortium and SIG-Embryology Biopsy Working Group; Kokkali G., Coticchio G., Bronet F., Celebi C., Cimadomo D., Goossens V., Liss J., Nunes S., Sfontouris I., Vermeulen N., Zakharova E., De Rycke M. ESHRE PGT Consortium and SIG Embryology good practice recommendations for polar body and embryo biopsy for PGT. *Hum Reprod Open.* 2020;2020(3):hoaa020. doi 10.1093/hropen/hoaa020

ESHRE PGT-M Working Group; Carvalho F., Moutou C., Dimitriadou E., Dreesen J., Giménez C., Goossens V., Kakourou G., Vermeulen N., Zuccarello D., De Rycke M. ESHRE PGT Consortium good practice recommendations for the detection of monogenic disorders. *Hum Reprod Open.* 2020;2020(3):hoaa018. doi 10.1093/hropen/ hoaa018

Evaluation of Oocytes and Embryos in the ART Laboratory. Methodological recommendations. Russian Association of Human Reproduction, 2021. https://www.rahr.ru/d_pech_mat_metod/MR_evaluation_of_embryos.pdf (in Russian)

Goriely A., Lord H., Lim J., Johnson D., Lester T., Firth H.V., Wilkie A.O. Germline and somatic mosaicism for *FGFR2* mutation in the mother of a child with Crouzon syndrome: implications for genetic testing in "paternal age-effect" syndromes. *Am J Med Genet A.* 2010;152A(8):2067-2073. doi 10.1002/ajmg.a.33513

Harper J.C., Wells D., Piyamongkol W., Abou-Sleiman P., Apessos A., Ioulianos A., Davis M., Doshi A., Serhal P., Ranieri M., Rodeck C., Delhanty J.D.A. Preimplantation genetic diagnosis for single gene disorders: experience with five single gene disorders. *Prenat Diagn.* 2002;22(6):525-533. doi 10.1002/pd.394

Hopp K., Cornec-Le Gall E., Senum S.R., Te Paske I.B.A.W., Raj S., Lavu S., Baheti S., … Rahbari-Oskoui F.F., Torres V.E.; HALT Progression of Polycystic Kidney Disease Group, the ADPKD Modifier Study; Harris P.C. Detection and characterization of mosaicism in autosomal dominant polycystic kidney disease. *Kidney Int.* 2020; 97(2):370-382. doi 10.1016/j.kint.2019.08.038

Koltunov D.E. Crouzon syndrome: etiology and clinical manifestations. *Clinical Practice in Pediatrics.* 2011;6(5):49-52 (in Russian)

Sokolova M.A., Sarkisyan E.A., Shumilov P.V., Vorona L.D., Levchenko L.A., Ishutina Yu.L., Shabelnikova E.I., Krapivkin A.I. Crouzon syndrome: features of clinical manifestations, management and outcomes in children. *Rossiyskiy Vestnik Perinatologii i Pediatrii = Russian Bulletin of Perinatology and Pediatrics*. 2024;69(1):78-85. doi 10.21508/1027-4065-2024-69-1-78-85 (in Russian)

Verlinsky Y., Kuliev A. An Atlas of Preimplantation Genetic Diagnosis. CRC Press, 2004. doi 10.1201/b14655

Wenger T., Miller D., Evans K. *FGFR* Craniosynostosis syndromes overview. In: GeneReviews® [Internet]. Seattle (WA): University of Washington, Seattle, 1993–2025. Available from: https://www.ncbi.nlm.nih.gov/books/NBK1455/

WHO Laboratory Manual for the Examination and Processing of Human Semen. Geneva: World Health Organization, 2010. Available from: https://iris.who.int/handle/10665/44261

Yapijakis C., Pachis N., Sotiriadou T., Vaila C., Michopoulou V., Vassiliou S. Molecular mechanisms involved in craniosynostosis. *In Vivo*. 2023;37(1):36-46. doi 10.21873/invivo.13052

Conflict of interest. The authors declare no conflict of interest.
Received January 14, 2025. Revised March 7, 2025. Accepted March 26, 2025.

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2025;29(5):693-703

doi 10.18699/vjgb-25-76

Population transcriptomics: a novel tool for studying genetic diversity in human populations under normal and pathological conditions

A.A. Babovskaya, E.A. Trifonova, V.A. Stepanov

Research Institute of Medical Genetics, Tomsk National Research Medical Center of the Russian Academy of Sciences, Tomsk, Russia
anastasia.babovskaya@medgenetics.ru

Abstract. Genetic mechanisms regulating gene expression encompass complex processes such as transcription, translation, epigenetic modifications, and interactions of regulatory elements. These mechanisms play a crucial role in shaping phenotypic diversity in humans. High-throughput technologies, such as expression microarrays and next-generation sequencing (NGS), have enabled precise analysis of transcripts for thousands of genes genome-wide. These methods have enabled researchers to measure gene expression levels in various tissues and cells and to gain deeper insights into previously inaccessible biological processes. Numerous studies show that gene expression varies significantly among individuals. However, there are also notable differences between populations from different continental groups, driven by genetic, epigenetic, environmental factors, and natural selection. Furthermore, disease states represent an important factor influencing gene activity, as they can significantly alter the transcriptomic profiles of individual cells. In this context, comparative population genetic studies help uncover the molecular mechanisms underlying complex phenotypic traits and identify population-specific features of transcriptomic profiles in both health and disease. However, despite significant progress in this field, many aspects remain underexplored. Specifically, the distribution of gene expression variability among populations, the degree of research coverage for specific ethnic groups, the spectrum of biological materials used, and the contribution of population affiliation to observed differences in gene expression during pathological conditions require further investigation. This review presents an overview of contemporary research focused on analyzing variability in expression profiles across different human populations. It summarizes findings from individual studies, outlines the advantages and limitations of the methods employed, highlights key research directions in population transcriptomics, and discusses potential practical applications of the data obtained.

Key words: populations; RNA-seq; next-generation sequencing (NGS); gene expression; transcriptome

For citation: Babovskaya A.A., Trifonova E.A., Stepanov V.A. Population transcriptomics: a novel tool for studying genetic diversity in human populations under normal and pathological conditions. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov J Genet Breed*. 2025;29(5):693-703. doi 10.18699/vjgb-25-76

Funding. This research was funded by a state grant (Basic Scientific Research No. 122020200083-8).

Популяционная транскриптомика: новый инструмент исследования генетического разнообразия популяций человека в норме и при патологии

А.А. Бабовская, Е.А. Трифонова, В.А. Степанов

Научно-исследовательский институт медицинской генетики Томского национального исследовательского медицинского центра Российской академии наук, Томск, Россия
anastasia.babovskaya@medgenetics.ru

Аннотация. Генетические механизмы, регулирующие экспрессию генов, включают в себя сложные процессы: транскрипцию, трансляцию, эпигенетические модификации и взаимодействие регуляторных элементов. Они играют ключевую роль в формировании фенотипического разнообразия человека. Благодаря внедрению высокопроизводительных технологий, таких как экспрессионные микрочипы и массовое параллельное секвенирование (NGS), стало возможным с высокой точностью анализировать транскрипты на уровне тысяч генов по всему геному. Эти методы позволили ученым не только определять уровни экспрессии генов в различных тканях и клетках, но и глубже изучать биологические процессы и феномены, которые ранее оставались недоступными для анализа. Так, многочисленные работы подтвердили, что, несмотря на преобладание индивидуальных различий в уровне экспрессии генов, существуют также значимые вариации между

© Babovskaya A.A., Trifonova E.A., Stepanov V.A., 2025
This work is licensed under a Creative Commons Attribution 4.0 License

популяциями, принадлежащими к разным континентальным группам, обусловленные генетическими, эпигенетическими и средовыми факторами и действием естественного отбора. Кроме того, важным фактором, влияющим на активность генов, являются заболевания, которые могут существенно изменять транскриптом отдельных клеток. В этом контексте сравнительные популяционно-генетические исследования позволяют раскрыть молекулярные механизмы, лежащие в основе сложных фенотипических признаков, и идентифицировать популяционно-специфические особенности транскриптомных профилей как в норме, так и при патологических состояниях. Несмотря на значительный прогресс в этой области, многие аспекты остаются недостаточно изученными. В частности, распределение вариабельности экспрессии генов между популяциями, степень исследованности отдельных этнических групп, спектр используемого биологического материала, а также вклад популяционной принадлежности в наблюдаемые различия экспрессии генов при патологических состояниях требуют дальнейшего изучения. В настоящей статье представлен обзор современных исследований, посвященных анализу вариабельности экспрессионных профилей в различных популяциях человека. Обобщены результаты отдельных экспериментов, описаны преимущества и ограничения используемых методов, выделены основные направления работ в области популяционной транскриптомики, а также обозначены перспективы практического применения полученных данных.
Ключевые слова: популяции; RNA-seq; массовое параллельное секвенирование; экспрессия генов; транскриптом

Introduction

The human phenotype, in both health and disease, is a complex system shaped by genomic, transcriptomic, epigenetic, metabolomic, and environmental factors. From a protein-coding DNA sequence to its functional protein product in the cell, a range of molecular mechanisms is involved, among which processes occurring at the transcriptome level play a pivotal role. The transcriptome, defined as the complete set of transcripts in a cell at a specific moment in time, is central to the mechanisms of genetic regulation of cellular function. Understanding it is essential for identifying the molecular pathways underlying various functional states, including pathological ones. Numerous studies have demonstrated variability in gene expression at the population level (Spielman et al., 2007; Storey et al., 2007; Zhang et al., 2008). This phenomenon is driven by multiple factors, including environmental conditions, dietary habits, epigenetic regulation, and more. Additionally, natural selection likely contributes to the observed differences in gene expression by shaping unique genetic profiles of populations through adaptation processes. Another factor significantly influencing gene expression patterns is the presence of pathological processes.

Extensive research has revealed substantial differences in gene expression in diseases such as diabetes, cancer, cardiovascular disorders, reproductive issues, neurological conditions, and infectious diseases (Wei et al., 2011; Allard et al., 2012; Nédélec et al., 2016; Mitchell et al., 2017). Notably, the prevalence of some of these diseases varies across populations (Kelly et al., 2017). Variability in the transcriptional activity of genes linked to disease pathogenesis, shaped by long-term adaptation and fixed in each population's gene pool, may contribute to these interpopulation differences.

To date, a significant body of data has been accumulated from various experiments, indicating the presence of interpopulation variability in human gene expression. However, the distribution of this variability across populations, the extent of research coverage for specific population groups, the range of biological materials used, and the contribution of population affiliation to the observed differences in gene expression under pathological conditions remain insufficiently explored. This review synthesizes findings from studies investigating patterns of gene expression changes across human populations, outlines the advantages and limitations of the methods employed, and highlights promising directions for future research in this field.

Microarray technologies in studying transcriptomics variability between populations

Nowadays, various technologies have been developed for characterizing and quantifying genome-wide gene expression, including hybridization-based methods (microarrays) and sequencing approaches. Hybridization methods use fluorescently labeled cDNA to bind with high-density commercial microarrays. The widespread use of microarrays for genome-wide gene expression analysis has enabled the identification of multiple levels of gene expression variability within a species: interpopulation, interindividual, and intraindividual (including intertissue and intercellular levels).

One of the earliest studies dedicated to exploring population-level variability in the human transcriptome was conducted using peripheral blood from 52 individuals across three Moroccan Amazigh (Berbers) groups with distinct lifestyles: desert nomads, rural villagers, and urban dwellers (Idaghdour et al., 2010). According to the expression profile analysis obtained via microarrays, the percentage of differentially expressed genes between the studied groups ranged from 16.4 % (desert nomads *vs.* rural villagers) to 29.9 % (rural villagers *vs.* urban dwellers). The authors attributed this primarily to the predominant influence of environmental factors on shaping transcriptome variability.

However, the majority of data confirming interpopulation differences in gene expression have been derived from lymphoblasts (Dixon et al., 2007; Göring et al., 2007) and lymphoblastoid cell lines (LCLs) collected as part of the International HapMap Project (Stranger et al., 2007). These cell lines constitute a biobank of B-lymphocytes gathered from various populations – CEU (Caucasians), CHB (Chinese), JPT (Japanese), YRI (Nigerians), and AA (African Americans) –

and modified with the Epstein–Barr virus to ensure viability (Baust et al., 2017). Initial studies investigating interpopulation variability using microarrays were conducted by research groups led by B. Stranger (Stranger et al., 2007), J. Storey (Storey et al., 2007), and R. Spielman (Spielman et al., 2007). These studies focused on identifying genes with differential expression between populations of European (CEU), East Asian (CHB, JPT), and African (YRI) ancestry.

In the study by R. Spielman and colleagues, a microarray covering more than 4,000 genes was used to compare expression between Caucasians (60 CEU) and Mongoloids (41 CHB and 41 JPT). It was found that over 1,000 genes (approximately 25 %) exhibited differential expression between CEU and the combined CHB/JPT group, while only 27 genes showed differences between Chinese (CHB) and Japanese (JPT) samples (Spielman et al., 2007). Cluster analysis confirmed that samples from Chinese individuals in Los Angeles (CHLA) were more similar in expression profiles to CHB/JPT than to CEU, indicating a characteristic expression pattern associated with Asian ancestry (Spielman et al., 2007). A later study by P. Daca-Roszak and E. Zietkiewicz aimed to identify population-specific genes between Caucasians and Mongoloids. Analysis of B-lymphocyte cell lines revealed 20 genes with interpopulation expression differences. Of the 13 genes selected with the highest fold-change (FC > 2), three (*UTS2*, *UGT2B17*, and *SLC7A7*) confirmed their status as differentially expressed upon validation: *UTS2* showed hyperexpression in Chinese samples, while *UGT2B17* and *SLC7A7* were hyperexpressed in Caucasians (Daca-Roszak, Zietkiewicz, 2019).

Another microarray-based transcriptome study of lymphoblasts found that approximately 17 % of genes exhibited expression differences between Caucasians (CEU) and Negroids (YRI). Many of these genes were associated with immune responses, including cytokines and chemokines (*CCL22*, *CCL5*, *CCR2*, *CXCR3*). Functional analysis revealed an enrichment of inflammatory response categories among genes differentially expressed between CEU and YRI, supporting their role in immune and infectious diseases (Storey et al., 2007).

Population differences in gene expression between Negroids and Caucasians were further analyzed using the Affymetrix GeneChip Human Exon 1.0 ST microarray, which includes over 9,100 transcripts, on an expanded sample of 176 LCLs (87 CEU and 89 YRI). It was determined that 4.2 % of transcripts showed significant expression differences between the groups. Hyperexpression of 156 genes was observed in Caucasian (CEU) samples, while 254 genes were hyperexpressed in African (YRI) samples (Zhang et al., 2008). Subsequent functional analysis identified the involvement of these genes in processes such as ribosome assembly, antimicrobial humoral response, intercellular adhesion, mRNA catabolism, and tRNA processing. Notably, nine genes (*DPYSL2*, *CTTN*, *PLCG1*, *SS18*, *SH2B3*, *CPNE9*, *CMAH*, *CXCR3* and *MRPS7*) had previously been described as differentially expressed between CEU and YRI (Storey et al., 2007).

One of the largest population studies, conducted by H. Fan's group in 2009, encompassed 210 LCLs from four ethnic groups (CEU, CHB, JPT, and YRI) using a high-density Illumina microarray covering over 11,000 transcripts. This study aimed to investigate interindividual and interpopulation differences in gene expression on a genome-wide scale and to determine the proportion of genes contributing to each type of variability. It was found that interindividual variability was a critical component of genetic differences within populations, accounting for nearly half (43 %) of the total variability in gene expression (Fan et al., 2009). These findings align with later data from full-transcriptome studies (RNA-seq) (Hughes et al., 2015), which will be discussed below.

Notably, these studies used lymphoblastoid cell lines from the same populations, but the reported gene expression variability ranged from 8 to 38 % (see the Table). This discrepancy may be attributed to several factors, including technical variability related to cell culture conditions and biological variability caused by epigenetic modifications and adaptation to *in vitro* conditions (Lappalainen et al., 2013). While cell culturing offers advantages, such as minimal biomaterial requirements, high reproducibility, and well-characterized properties, it introduces a significant component of variability into the transcriptomic profile. For instance, freeze–thaw cycles, culture medium composition, and cell density can substantially affect gene expression and transcriptome architecture (Baust et al., 2017). These limitations underscore the need for cautious interpretation of data derived from cell lines.

Another critical factor influencing data variability is the gene expression profiling method. Despite its widespread use, microarray analysis is susceptible to the so-called "batch effect". This effect arises from technical differences between experimental batches, such as the use of different microarray lots, analysis platforms, or variations in experimental conditions (e. g., temperature, humidity, or experiment date) (Fellenberg et al., 2006). To minimize batch effects, bioinformatics correction methods, such as Empirical Bayes methods, are employed to normalize data and reduce the impact of technical artifacts. However, such adjustments may lead to the loss of biologically significant expression differences, limiting the interpretation of results (Johnson et al., 2007).

High-throughput sequencing technologies (NGS) in studying gene expression variability between populations

The advancement of high-throughput sequencing (NGS) technologies has provided the most comprehensive coverage of the genome and transcriptome, enabling researchers to identify key molecular pathways involved in pathological processes with high precision, detect disease biomarkers, and assess dynamic changes in gene expression in response to various stimuli. Unlike microarray profiling, NGS allows for the analysis of not only gene expression levels but also alternative splice variants, rare transcripts, and post-transcriptional modifications, significantly expanding opportunities for understanding gene regulation (Kukurba, Montgomery, 2015). Given the modern research trend toward increasing sample sizes and output data volumes, as well as several limitations imposed by microarray profiling technology (e. g., batch ef-

Transcriptomic studies involving two or more populations compared with each other

No.	Populations	Study object	Genetic differences*	Platform	Reference
	Studies of interpopulation variability in gene expression in groups of healthy individuals				
1	CEU = 60, CHB = 41, JPT = 41	LCL	25 % (939 DEGs between CEU and CHB, 756 DEGs between CEU and JPT, 27 DEGs between CHB and JPT)	Microarrays	Spielman et al., 2007
2	CEU = 8, YRI = 8	LCL	83 % interindividual, 17 % interpopulation		Storey et al., 2007
3	CEU = 87, YRI = 89	LCL	383 DEGs		Zhang et al., 2008
4	YRI = 60, CEU = 60, CHB = 45, JPT = 45	LCL	38 % interindividual, 8–18 % interpopulation		Fan et al., 2009
5	194 individuals (Arabs and Amazighs)	Leukocytes	5.6 to 14 %		Idaghdour et al., 2010
6	CEU = 109, CHB = 80, GIH = 82, JPT = 82, LWK = 82, MEX = 45, MKK = 138, YRI = 108	LCL	472–947 cis-eQTL		Stranger et al., 2012
7	YRI = 30, CEU = 30, CHB = 90, JPT = 45	LCL	205, 192, 193, and 193 cis-eQTL in YRI, CEU, CHB, and JPT, respectively		Yang et al., 2014
8	EA = 211, AA = 112, CA = 78	Monocytes, CD4+ T-lymphocytes	In CD4+ T-cells: 2,352 cis-eQTL in EA, 592 cis-eQTL in CA, 722 cis-eQTL in AA In monocytes: 3,090 cis-eQTL in EA, 1,181 cis-eQTL in CA, 1,318 cis-eQTL in AA		Raj et al., 2014
9	AA = 10, CA = 10, EA = 10	Skin	378 DEGs		Yin et al., 2014
10	CEU = 35/37, CHB = 32/29	LCL/Blood leukocytes	189 DEGs		Daca-Roszak, Zietkiewicz, 2019
11	CEU = 91, FIN = 95, GBR = 94, TSI = 93, YRI = 89	LCL	3 %	RNA-seq	Lappalainen et al., 2013
12	4 San, 7 Mbuti Pygmies, 7 Mozambicans, 7 Pathans, 7 Cambodians, 6 Yakuts, 7 Maya	LCL	25 %		Martin A.R. et al., 2014
13	CEU = 20, CHB = 20	LCL	423 DEGs		Li et al., 2014
14	AA = 24, EA = 148	GTEx	47 % intertissue, 4 % interpopulation		Melé et al., 2015
15	AA = 10, EU = 10, EA = 10, CA = 10	Placental tissue	33.2 % intraindividual, 58.9 % interindividual, 7.8 % interpopulation		Hughes et al., 2015
16	EU = 100, AA = 100	CD14+ monocytes	821 DEGs		Quach et al., 2016
17	Russians = 8, Buryats = 9	Decidual placental cells	4 % interpopulation, 67 % interindividual		Babovskaya et al., 2024
	Studies of interpopulation variability in gene expression in disease				
18	EA = 17, AA = 18	Breast cancer	~490 DEGs	Microarrays	Martin D.N. et al., 2009
19	AA = 21, EA = 17	Blood endothelial cell	31 DEGs		Wei et al., 2011
20	AA = 10, EU = 14	Endometrial tumor tissue	341 DEGs		Allard et al., 2012

Table (end)

No.	Populations	Study object	Genetic differences*	Platform	Reference
21	EU = 52, AA = 11	Prostate tumor and stromal tissue	56 DEGs in tumor tissue, 677 DEGs in stromal tissue	Microarrays	Kinseth et al., 2014
22	AA = 22, EA = 19	Lung tumor tissue	501 DEGs		Mitchell et al., 2017
23	Russians = 21, Yakuts = 23	Placental tissue	~40 %		Trifonova et al., 2022
24	AA = 80, EA = 94	Macrophages	9.3 %	RNA-seq	Nédélec et al., 2016
25	Populations of Southern and West-Central Europe	Blood leukocytes	876 DEGs across all comparison groups		Beretta et al., 2020
26	AA = 48, EA = 23	Blood leukocytes	50 % differences in proinflammatory transcription factor activity		Thames et al., 2019

Note. The table uses standard population designations from the "1000 Genomes Project" (1000 Genomes Project Consortium et al., 2012) and "HapMap Project" (International HapMap Consortium, 2003): CEU – Utah residents of Central European ancestry, CHB – Chinese, GIH – Indians, JPT – Japanese, LWK – Luhya, MEX – Mexicans, MKK – Maasai, YRI – Yoruba, FIN – Finns, GBR – British, TSI – Tuscans. Introduced abbreviations: AA – African Americans, EA – Americans of European ancestry, CA – Americans of Asian ancestry, EU – Caucasians, DEGs – differentially expressed genes.
* The contribution of interpopulation differences to the total variability in gene expression is provided by default.

fects and limited genome coverage), researchers are increasingly turning to NGS methods. These approaches offer greater accuracy, reproducibility, and depth of analysis, making them preferable for studying complex biological processes (Hrdlickova et al., 2017).

In the context of studying gene expression variability across human populations, two key directions can be distinguished. The first focuses on investigating the influence of natural selection and environmental factors on shaping expression profiles. For instance, studies demonstrate that population-specific differences in gene expression are often linked to adaptations to diverse ecological conditions, such as high-altitude hypoxia, ultraviolet radiation levels, or dietary patterns (Hodgson et al., 2014). These studies help elucidate how evolutionary processes shape the transcriptomic landscape across different populations. The second direction examines differences in gene expression associated with diseases, the prevalence or clinical presentation of which varies among populations. For example, research on oncological, autoimmune and infectious diseases shows that population differences in gene expression can influence disease susceptibility, progression and response to therapy (Lappalainen et al., 2013; Quach et al., 2016; Way et al., 2016).

Studies of interpopulation variability in gene expression in groups of healthy individuals

Several studies in the literature have explored expression profiles across different populations using cohorts of conditionally healthy individuals. In one such study, transcriptomic variability was characterized in over 460 lymphoblastoid cell lines (LCLs) derived from individuals of an African population (YRI) and four European subpopulations (CEU, FIN, GBR and TSI) (Lappalainen et al., 2013). Interpopulation differences accounted for only a minor fraction (3 %) of the total gene expression variability. Nevertheless, the number of genes exhibiting statistically significant expression differences between the African and European populations was substantial, ranging from 1,300 to 4,300 genes depending on the European population compared. In contrast, the number of differentially expressed genes identified when comparing European subpopulations among themselves was significantly lower.

In another study, gene expression was examined in 20 LCLs obtained from individuals of European (CEU) and East Asian (CHB) populations (Li et al., 2014). The analysis identified over 400 differentially expressed genes, including 132 genes with elevated expression and 291 genes with reduced expression in the CHB population compared to the CEU one.

A.R. Martin and colleagues investigated transcriptomic profiles in 45 LCLs derived from seven non-European populations (Namibian San, Mbuti Pygmies from the Democratic Republic of Congo, Algerian Mozabites, Pathans from Pakistan, Cambodians from East Asia, Siberian Yakuts, and Mexican Maya). They identified 44 genes with significant differential expression across the studied populations, most of which were associated with immune pathways. The greatest interpopulation variability in expression was observed for the genes *THNSL2*, *DRP2*, *VAV3*, *IQUB*, *BC038731*, *RAVER2*, *SYT2*, *LOC100129055*, *AK126080* and *TTN* (Martin A.R. et al., 2014).

An original approach to minimize the influence of external factors on gene expression patterns was proposed by D. Hughes and colleagues. They studied interpopulation variability in gene expression in placental tissue from individuals of four populations: Americans of European, South Asian, East Asian, and African ancestry (Hughes et al., 2015). The results indicated that approximately 8 % of gene expression variability was attributable to interpopulation differences, while 58.9 % of transcriptome variability was driven by interindividual differences within a single population. The greatest expression variability was recorded in the African and South Asian populations, where over 140 differentially

expressed genes were identified. Genes exhibiting the highest interpopulation variability were predominantly involved in immune response, cellular signaling, and metabolism processes. Despite the strengths of this study, a significant limitation for interpreting the results is the factor of cellular heterogeneity, which cannot be eliminated when using whole tissue. To better understand true variability in gene expression, a study design was proposed in which the transcriptomic landscape was examined in a single cellular subpopulation of decidual cells from individuals of Russian and Buryat populations (Babovskaya et al., 2024). This study was the first to assess intra- and interpopulation variability in genome-wide gene expression at the level of individual placental tissue cells. The findings revealed that interpopulation differences among individuals with physiological pregnancies accounted for 4 %, while interindividual variability contributed 67 % (Nit = 67 %) (Fig. 1). Transcripts involved in regulating apoptotic enzyme activity exhibited the least variability, whereas those participating in renal filtration, blood pressure regulation, TGFβ-mediated processes, and cellular signal transduction showed the greatest variability.

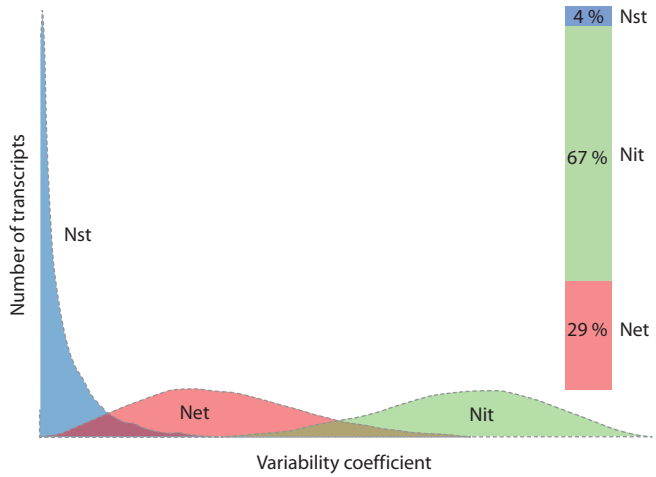


Fig. 1. Interpopulation (Nst), interindividual (Nit), and inter-replicate (Net) differences obtained from the analysis of sums of squares for groups with physiological pregnancy.

RNA-seq data analysis showed that gene expression variability between populations ranges from 4 to 25 %. However, in studies where total gene expression variability is further divided into components characterizing interindividual and intraindividual variance, this range narrows to 3–8 %. Overall, at these levels, gene expression variability among the studied populations does not, on average, reach the level observed at the genome-wide scale, where the proportion of interpopulation genetic differences in the total genetic diversity of human populations is estimated at 10–15 % (Stepanov, 2016). Nevertheless, the population component significantly influences gene expression variability, underscoring its role in shaping transcriptomic diversity.

Studies of interpopulation variability in gene expression in disease contexts

The investigation of polygenic diseases is of significant scientific interest due to their high prevalence across populations. According to epidemiological data, multifactorial diseases (MFDs) account for approximately 90 % of all hereditary pathologies, highlighting their substantial medical and social importance and relevance to modern genetic research (Hoffjan et al., 2016). The risk of MFDs is typically associated with numerous factors, including socioeconomic, demographic, cultural, environmental, and genetic determinants. Our understanding of the genetic determinants of disease risk has expanded considerably with the advent of high-throughput genomic tools, enabling researchers to profile the genome, epigenome, and transcriptome and to broadly analyze the resulting data (Gurdasani et al., 2019; Sirugo et al., 2019). Demographic features, genetic drift, and environmental adaptation over millennia have led to global population differentiation. This genomic diversity opens new opportunities for biomarker discovery, treatment development, and a better understanding of disease risk across populations. Given the accumulated evidence of significant gene expression variability in various pathologies, as well as differences in disease prevalence and clinical manifestations depending on racial background, there is a need to assess the extent to which the population genetic component influences the observed variability.

For example, L. Beretta and coauthors used transcriptomic technologies to study systemic scleroderma in populations of Southern and West-Central Europe, identifying both shared and population-specific pathological pathways in this autoimmune disease (Beretta et al., 2020). The experiment revealed that a key role in the pathogenesis of systemic scleroderma is played by type I interferon activation signaling pathways via TLR receptors, irrespective of population affiliation.

The study by Y. Nédélec and colleagues demonstrated population differentiation in infectious diseases. It was found that 9.3 % of genes expressed in macrophages exhibit differences in regulatory responses to infection linked to population background. Specifically, African ancestry is associated with a stronger inflammatory response and reduced intracellular bacterial growth (Nédélec et al., 2016). However, earlier research indicated that approximately 34 % of genes expressed in macrophages show at least one type of transcriptional divergence related to population affiliation: either differences in gene expression or, less commonly, transcript isoform diversity (Lappalainen et al., 2013). One of the most significant observations by Y. Nédélec and colleagues was the detection of a stronger inflammatory response to infection in individuals of African ancestry. This finding aligns with prior studies showing that Africans more frequently carry alleles associated with heightened proinflammatory responses (Ness et al., 2004) and exhibit elevated levels of circulating C-reactive protein (Kelley-Hedgepeth et al., 2008). Although the precise causal relationship between population affiliation and proinflammatory response remains unestablished, it can be hypothesized that the stronger inflammatory response in individuals of African ancestry explains their enhanced ability to control bacterial growth post-infection (Nédélec et al., 2016).

Cardiovascular diseases have also been a focus of research interest. P. Wei and colleagues examined differential gene expression in endothelial cells between African Americans and Caucasians, exploring the potential contribution of certain genes to observed population differences in the incidence of tumors and cardiovascular diseases. Notably, African Americans (AA) exhibit a 2.4-fold higher incidence and approximately 50 % greater prevalence of hypertension compared to Caucasians (EA). The comparison of AA and EA groups revealed that 31 genes showed differential expression between the two groups at FDR = 0. Four genes exhibited elevated expression in the AA group compared to EA, while 27 genes showed reduced expression in AA (Wei et al., 2011). Among the differentially expressed genes (DEGs), *PSPH* stood out due to its highest differential expression across the studied groups. Interestingly, a homolog of this gene, *PSPHL*, was later identified in another study as the most differentially expressed gene between African Americans and Americans of European ancestry. J. Allard and colleagues investigated genes differentially expressed in endometrial cancer between women of Caucasian and African ancestry. It was reported that, in addition to differences in disease incidence, molecular cancer subtypes also vary between carriers from Caucasian and African populations, suggesting the presence of population-specific features in the expression profiles of tumor tissues. African American women are typically diagnosed with more advanced disease, unfavorable histological types, and higher malignancy grades compared to Caucasians. The study found that gene expression variance between populations accounted for 7.2 % (341 transcripts) at $p < 0.005$ (Allard et al., 2012). Among these, the phosphoserine phosphatase gene *PSPHL* was identified as the most hyperexpressed in both tumor and normal endometrial and ovarian tissues of African American women compared to tissues from women of European ancestry. This suggests that *PSPHL* is not a tumor marker, a conclusion further supported by studies on breast, prostate, and endothelial cells (Wallace et al., 2008; Martin D.N. et al., 2009; Wei et al., 2011).

There is also significant interest in studying population-specific features of oncological diseases, driven by their substantial contribution to global morbidity and mortality. M. Kinseth and colleagues (2014) established population differentiation in the incidence and progression of prostate cancer. Incidence and mortality rates among African Americans (AA) are 1.5 and 2.3 times higher, respectively, than among individuals of Caucasian ancestry. AA also tend to experience more aggressive disease progression and earlier onset. Notably, this study again demonstrated differential expression of the phosphoserine phosphatase genes (*PSPH*) and *CRYBB2* between individuals of European and African ancestry (Wallace et al., 2008). Consistent with prior research, the authors concluded that these genes are not associated with tumor tissue but may serve as markers of racial background (Kinseth et al., 2014).

Another study on population differentiation in gene expression in oncology, conducted by K.A. Mitchell and colleagues, aimed to determine whether racial differences in gene and microRNA expression influence clinically significant differences in lung tumor biology between African Americans (AA) and Americans of European ancestry (EA). They showed that while there are similarities in expression profiles between the populations, differences exist in both protein-coding transcripts and the non-coding genome. The researchers found that the transcriptome of AA tumor cells was enriched in stem cell and invasion pathways, whereas the transcriptomic profile of EA tumor cells showed enrichment in categories related to cell cycle, mitosis, and proliferation processes. The authors noted hypoexpression of the genes *ARL17A*, *LRCC37A3*, and *KANSL1* in lung tissues of AA compared to EA (Mitchell et al., 2017). These genes are located in the 17q21 region, where structural diversity – an inversion polymorphism with duplication – has been previously identified (Steinberg et al., 2012). The presence of a direct (H1) or inverted (H2) haplotype influences differential susceptibility to non-allelic homologous recombination and diseases, including cancer. European populations exhibit a high frequency of duplication events, whereas most West African populations lack those (Steinberg et al., 2012). Other studies on breast, colon, prostate, and endometrial tissues have also identified genes with expression varying by population affiliation, such as *PSPHL*, *CRYBB2*, *AMFR*, and *SOS1* (Martin D.N. et al., 2009; Allard et al., 2012; Field et al., 2012; Mitchell et al., 2017).

Alongside oncological diseases, reproductive health holds high social significance. However, studies of the transcriptome in reproductive system disorders using a population-based approach remain rare. This is due to both limited availability of biological samples and methodological challenges, such as the need to account for hormonal fluctuations and other factors affecting gene expression. Nevertheless, such studies are critical for understanding the etiology of reproductive disorders and developing personalized treatment approaches. Previous research has demonstrated racial and ethnic differences in the incidence of preeclampsia (PE), a severe pregnancy complication associated with high maternal and infant mortality. For instance, studies by Torchin, Ghosh, and Wolf indicate that African American women have a higher risk of developing PE compared to women of European ancestry (Wolf et al., 2004; Ghosh et al., 2014; Torchin et al., 2015). Interpopulation variability in the expression levels of genes and proteins playing key roles in PE pathogenesis may contribute to these differences. For example, A. Whitney and colleagues showed that average levels of prognostic angiogenic (PlGF) and antiangiogenic factors (sEng, sVEGFR1) differ significantly between African Americans and Caucasians living in the United States (Whitney et al., 2003). Notably, both domestic and international studies have reported associations between polymorphic markers of genes encoding angiogenic and antiangiogenic factors and the risk of preeclampsia (Akulenko et al., 2020; Rana et al., 2022). These findings highlight the importance of studying the genetic aspects of preeclampsia while considering population-specific characteristics.

In Russia, one of the first studies aimed at identifying molecular mechanisms underlying severe pregnancy complications, such as preeclampsia, while accounting for the genetic

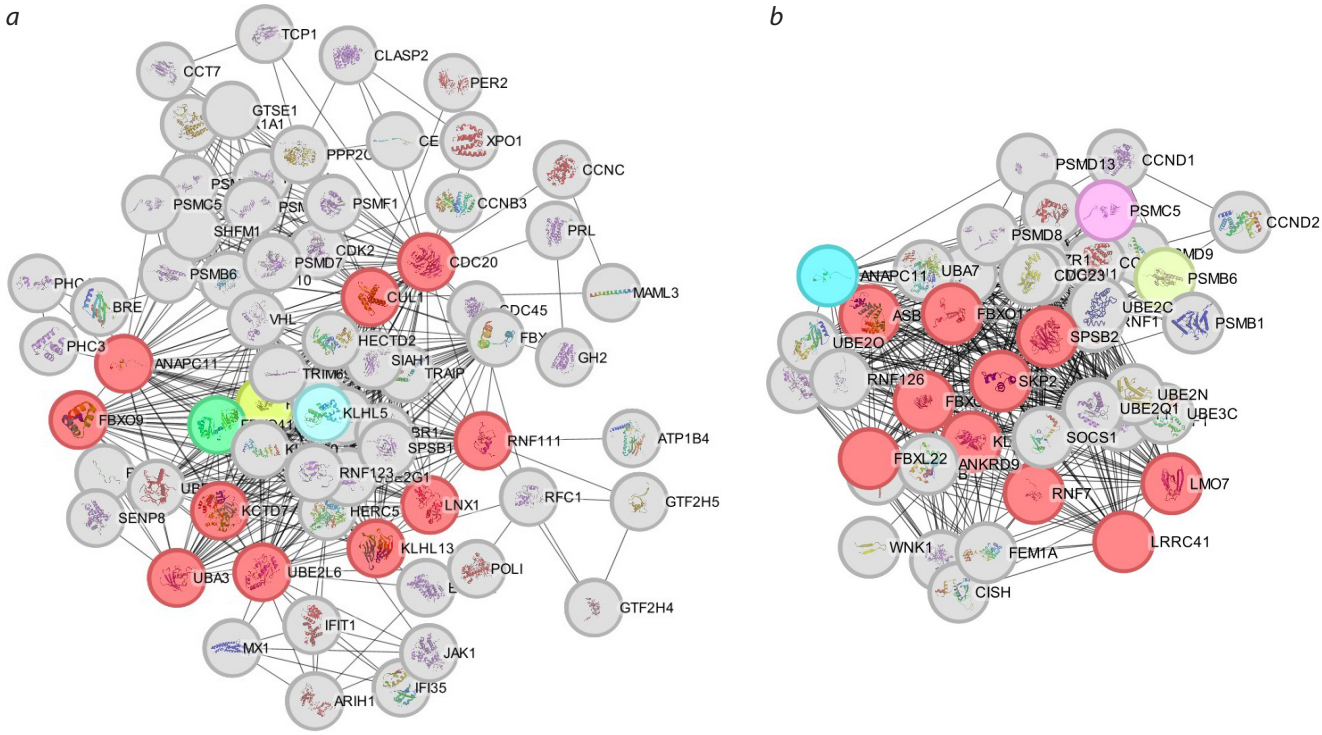


Fig. 2. Top 10 hub genes (red) with their nearest neighboring nodes (gray) in the protein-protein interaction network: *a* – Russians, *b* – Yakuts.

features of Siberian populations, was conducted using high-density microarrays. This study included genome-wide gene expression profiles from 21 women of Russian (European) and 23 women of Yakut (East Asian) ancestry (Trifonova et al., 2022). The study demonstrated a bioinformatics approach to data analysis that not only identified differentially expressed genes between populations but also analyzed clusters of co-expressed genes within protein-protein interaction network models and highlighted functionally active hubs (hub genes) in these networks. Network analysis of co-expression gene blocks yielded 10 clusters containing 7,968 genes associated with PE in Russians and 9 clusters containing 7,966 genes in Yakuts. Construction of protein-protein interaction networks identified the most functionally active genes: *CUL1*, *ANAPC11*, *LNX1*, *CDC20*, *UBE2L6*, *FBXO9*, *KLHL13*, *UBA3*, *KCTD7*, *RNF111* in Russians (Fig. 2*a*), and *KLHL3*, *FB11PSXL*, *ASB2*, *LRRC41*, *LMO7*, *RNF7*, *SKP2*, *FBXO2* in Yakuts (Fig. 2*b*). These genes are predominantly involved in the assembly and regulation of the ubiquitin-ligase complex. In total, 1,701 genes associated with PE were common to both populations. Functional annotation of these genes revealed their involvement in interferon activity and ubiquitin-ligase complex function (Trifonova et al., 2022).

Prospects for applying transcriptome research in forensics

The evidence presented above regarding the existence of interpopulation variability in gene expression across different human populations, along with the identification of specific markers characteristic of certain population groups, suggests that this knowledge could complement standard protocols used in forensic science (Daca-Roszak, Zietkiewicz, 2019).

Transcriptomic data could serve as an additional tool for individual identification, particularly in cases where traditional methods, such as short tandem repeat (STR) analysis, are insufficiently informative due to DNA degradation. Studies demonstrate that gene expression exhibits high individual specificity, driven by both genetic variations and epigenetic modifications (Spielman et al., 2007; Storey et al., 2007; Stranger et al., 2007; Price et al., 2008; Zhang et al., 2008; Fan et al., 2009; Lappalainen et al., 2013). These findings indicate that carefully selected population-specific transcriptomic markers could be utilized in forensics similarly to DNA-based marker methods to infer the population origin of forensic samples, especially when standard protocols are hindered by sample heterogeneity (e. g., mixed samples).

In practice, identifying multiple contributors through DNA marker genotyping in forensic samples is challenging or infeasible when reference DNA profiles are unavailable (Westen et al., 2009). Furthermore, in conjunction with identity determination, population-specific markers could be used to accurately estimate the time of death. Results from using RNA analysis as a complement to the forensic toolkit show that gene expression patterns change post-mortem in a tissue- and individual-specific manner, potentially allowing for the estimation of the time of death (González-Herrera et al., 2013; Ferreira et al., 2018). However, successfully integrating transcriptomic markers into forensic investigations requires further research, standardization of protocols, and the development of specialized databases.

Conclusion

Population transcriptomics studies represent a powerful tool for analyzing the genetic and molecular mechanisms underlying complex phenotypic traits. These studies enable the identification of population-specific features in the transcriptomic profiles of cells and tissues, which can be applied both for fundamental scientific purposes and in practical contexts such as forensic expertise, including biomarker analysis and genetic sample identification. Moreover, population transcriptomics opens new avenues for understanding the molecular mechanisms of diseases across diverse human populations. Recent research has demonstrated that genetic and transcriptional differences between populations may play a key role in the pathogenesis of autoimmune, infectious, oncological, and reproductive system diseases. Further advancements in population transcriptomics could lead to the development of novel approaches for the diagnosis, prevention, and treatment of diseases, tailored to the population-specific characteristics of individuals. Thus, population transcriptomics emerges as a promising field in molecular biology, capable not only of deepening our understanding of the genetic and epigenetic mechanisms underlying phenotypic differences between populations but also of providing a scientific foundation for implementing innovative strategies in medical practice.

References

1000 Genomes Project Consortium; Abecasis G.R., Auton A., Brooks L.D., DePristo M.A., Durbin R.M., Handsaker R.E., Kang H.M., Marth G.T., McVean G.A. An integrated map of genetic variation from 1,092 human genomes. *Nature*. 2012;491(7422):56-65. doi 10.1038/nature11632

Akulenko L.V., Sakvarelidze N.Yu., Mackevich V.A., Tsakhilova S.G. Genotype markers predispositions for preeclampsia. *Problemy Reproduktsii = Russ J Hum Reprod*. 2020;26(6):26-33. doi 10.17116/repro20202606126 (in Russian)

Allard J.E., Chandramouli G.V., Stagliano K., Hood B.L., Litzi T., Shoji Y., Boyd J., Berchuck A., Conrads T.P., Maxwell G.L., Risinger J.I. Analysis of *PSPHL* as a candidate gene influencing the racial disparity in endometrial cancer. *Front Oncol.* 2012;2:65. doi 10.3389/fonc.2012.00065

Babovskaya A.A., Trifonova E.A., Stepanov V.A. Population transcriptomics of preeclampsia. In: Bioinformatics of Genome Regulation and Structure / Systems Biology: Fourteenth International Multiconference: Abstracts, Russia, Novosibirsk, 5–10 August, 2024. Novosibirsk, 2024;762-763. doi 10.18699/bgrs2024-4.2-02

Baust J.M., Buehring G.C., Campbell L., Elmore E., Harbell J.W., Nims R.W., Price P., Reid Y.A., Simione F. Best practices in cell culture: an overview. *In Vitro Cell Dev Biol Anim*. 2017;53(8):669-672. doi 10.1007/s11626-017-0177-7

Beretta L., Barturen G., Vigone B., Bellocchi C., Hunzelmann N., De Langhe E., Cervera R., … Kerick M., Alarcón-Riquelme M.E., Martin J.; PRECISESADS SSc substudy group; PRECISESADS Flow Cytometry study group. Genome-wide whole blood transcriptome profiling in a large European cohort of systemic sclerosis patients. *Ann Rheum Dis*. 2020;79(9):1218-1226. doi 10.1136/annrheumdis-2020-217116

Daca-Roszak P., Zietkiewicz E. Transcriptome variation in human populations and its potential application in forensics. *J Appl Genet*. 2019;60(3-4):319-328. doi 10.1007/s13353-019-00510-1

Dixon A.L., Liang L., Moffatt M.F., Chen W., Heath S., Wong K.C., Taylor J., Burnett E., Gut I., Farrall M., Lathrop G.M., Abecasis G.R., Cookson W.O. A genome-wide association study of global gene expression. *Nat Genet*. 2007;39(10):1202-1207. doi 10.1038/ng2109

Fan H.P.Y., Di Liao C., Fu B.Y., Lam L.C., Tang N.L. Interindividual and interethnic variation in genome-wide gene expression: insights into the biological variation of gene expression and clinical implications. *Clin Chem*. 2009;55(4):774-785. doi 10.1373/clinchem.2008.119107

Fellenberg K., Busold C.H., Witt O., Bauer A., Beckmann B., Hauser N.C., Frohme M., Winter S., Dippon J., Hoheisel J.D. Systematic interpretation of microarray data using experiment annotations. *BMC Genomics*. 2006;7:319. doi 10.1186/1471-2164-7-319

Ferreira P.G., Muñoz-Aguirre M., Reverter F., Sá Godinho C.P., Sousa A., Amadoz A., Sodaei R., … Oliveira C., Dopazo J., Sammeth M., Ardlie K.G., Guigó R. The effects of death and post-mortem cold ischemia on human tissue transcriptomes. *Nat Commun*. 2018;9(1):490. doi 10.1038/s41467-017-02772-x

Field L.A., Love B., Deyarmin B., Hooke J.A., Shriver C.D., Ellsworth R.E. Identification of differentially expressed genes in breast tumors from African American compared with Caucasian women. *Cancer*. 2012;118(5):1334-1344. doi 10.1002/cncr.26405

Ghosh G., Grewal J., Männistö T., Mendola P., Chen Z., Xie Y., Laughon S.K. Racial/ethnic differences in pregnancy-related hypertensive disease in nulliparous women. *Ethn Dis*. 2014;24(3):283-289

González-Herrera L., Valenzuela A., Marchal J.A., Lorente J.A., Villanueva E. Studies on RNA integrity and gene expression in human myocardial tissue, pericardial fluid and blood, and its postmortem stability. *Forensic Sci Int*. 2013;232(1-3):218-228. doi 10.1016/j.forsciint.2013.08.001

Göring H.H.H., Curran J.E., Johnson M.P., Dyer T.D., Charlesworth J., Cole S.A., Jowett J.B., … MacCluer J.W., Kissebah A.H., Collier G.R., Moses E.K., Blangero J. Discovery of expression QTLs using large-scale transcriptional profiling in human lymphocytes. *Nat Genet*. 2007;39(10):1208-1216. doi 10.1038/ng2119

Gurdasani D., Barroso I., Zeggini E., Sandhu M.S. Genomics of disease risk in globally diverse populations. *Nat Rev Genet*. 2019;20(9):520-535. doi 10.1038/s41576-019-0144-0

Hodgson J.A., Mulligan C.J., Al-Meeri A., Raaum R.L. Early back-to-Africa migration into the Horn of Africa. *PLoS Genet*. 2014;10(6):e1004393. doi 10.1371/journal.pgen.1004393

Hoffjan S. Dissecting the genetic background of multifactorial diseases and traits – a major challenge for genetic research. *Mol Cell Probes*. 2016;30(6):345. doi 10.1016/j.mcp.2016.11.003

Hrdlickova R., Toloue M., Tian B. RNA-Seq methods for transcriptome analysis. *Wiley Interdiscip Rev RNA*. 2017;8(1):e1364. doi 10.1002/wrna.1364

Hughes D.A., Kircher M., He Z., Guo S., Fairbrother G.L., Moreno C.S., Khaitovich P., Stoneking M. Evaluating intra- and inter-individual variation in the human placental transcriptome. *Genome Biol*. 2015;16:54. doi 10.1186/s13059-015-0627-z

Idaghdour Y., Czika W., Shianna K.V., Lee S.H., Visscher P.M., Martin H.C., Miclaus K., Jadallah S.J., Goldstein D.B., Wolfinger R.D., Gibson G. Geographical genomics of human leukocyte gene expression variation in southern Morocco. *Nat Genet*. 2010;42(1):62-67. doi 10.1038/ng.495

International HapMap Consortium. The International HapMap Project. *Nature*. 2003;426(6968):789-796. doi 10.1038/nature02168

Johnson W.E., Li C., Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics*. 2007;8(1):118-127. doi 10.1093/biostatistics/kxj037

Kelley-Hedgepeth A., Lloyd-Jones D.M., Colvin A., Matthews K.A., Johnston J., Sowers M.R., Sternfeld B., Pasternak R.C., Chae C.U.; SWAN Investigators. Ethnic differences in C-reactive protein concentrations. *Clin Chem*. 2008;54(6):1027-1037. doi 10.1373/clinchem.2007.098996

Kelly D.E., Hansen M.E.B., Tishkoff S.A. Global variation in gene expression and the value of diverse sampling. *Curr Opin Syst Biol*. 2017;1:102-108. doi 10.1016/j.coisb.2016.12.018

Kinseth M.A., Jia Z., Rahmatpanah F., Sawyers A., Sutton M., Wang-Rodriguez J., Mercola D., McGuire K.L. Expression differences between African American and Caucasian prostate cancer tissue reveals that stroma is the site of aggressive changes. *Int J Cancer*. 2014;134(1):81-91. doi 10.1002/ijc.28326

Kukurba K.R., Montgomery S.B. RNA sequencing and analysis. *Cold Spring Harb Protoc*. 2015;2015(11):951-969. doi 10.1101/pdb.top084970

Lappalainen T., Sammeth M., Friedländer M.R., 't Hoen P.A., Monlong J., Rivas M.A., Gonzàlez-Porta M., … Rosenstiel P., Guigó R., Gut I.G., Estivill X., Dermitzakis E.T. Transcriptome and genome sequencing uncovers functional variation in humans. *Nature*. 2013;501(7468):506-511. doi 10.1038/nature12531

Li J.W., Lai K.P., Ching A.K., Chan T.F. Transcriptome sequencing of Chinese and Caucasian population identifies ethnic-associated differential transcript abundance of heterogeneous nuclear ribonucleoprotein K (*hnRNPK*). *Genomics*. 2014;103(1):56-64. doi 10.1016/j.ygeno.2013.12.005

Martin A.R., Costa H.A., Lappalainen T., Henn B.M., Kidd J.M., Yee M.C., Grubert F., Cann H.M., Snyder M., Montgomery S.B., Bustamante C.D. Transcriptome sequencing from diverse human populations reveals differentiated regulatory architecture. *PLoS Genet*. 2014;10(8):e1004549. doi 10.1371/journal.pgen.1004549

Martin D.N., Boersma B.J., Yi M., Reimers M., Howe T.M., Yfantis H.G., Tsai Y.C., Williams E.H., Lee D.H., Stephens R.M., Weissman A.M., Ambs S. Differences in the tumor microenvironment between African-American and European-American breast cancer patients. *PLoS One*. 2009;4(2):e4531. doi 10.1371/journal.pone.0004531

Melé M., Ferreira P.G., Reverter F., DeLuca D.S., Monlong J., Sammeth M., Young T.R., … Calvo M., Getz G., Dermitzakis E.T., Ardlie K.G., Guigó R. Human genomics. The human transcriptome across tissues and individuals. *Science*. 2015;348(6235):660-665. doi 10.1126/science.aaa0355

Mitchell K.A., Zingone A., Toulabi L., Boeckelman J., Ryan B.M. Comparative transcriptome profiling reveals coding and noncoding RNA differences in NSCLC from African Americans and European Americans. *Clin Cancer Res*. 2017;23(23):7412-7425. doi 10.1158/1078-0432.CCR-17-0527

Nédélec Y., Sanz J., Baharian G., Szpiech Z.A., Pacis A., Dumaine A., Grenier J.C., … Hernandez R.D., Pique-Regi R., Tung J., Yotova V., Barreiro L.B. Genetic ancestry and natural selection drive population differences in immune responses to pathogens. *Cell*. 2016;167(3):657-669.e21. doi 10.1016/j.cell.2016.09.025

Ness R.B., Haggerty C.L., Harger G., Ferrell R. Differential distribution of allelic variants in cytokine genes among African Americans and White Americans. *Am J Epidemiol*. 2004;160(11):1033-1038. doi 10.1093/aje/kwh325

Price A.L., Butler J., Patterson N., Capelli C., Pascali V.L., Scarnicci F., Ruiz-Linares A., … Nemesh J., Arbeitman L., Goldstein D.B., Reich D., Hirschhorn J.N. Discerning the ancestry of European Americans in genetic association studies. *PLoS Genetics*. 2008;4(1):e236. doi 10.1371/journal.pgen.0030236

Quach H., Rotival M., Pothlichet J., Loh Y.E., Dannemann M., Zidane N., Laval G., … Boland A., Deleuze J.F., Kelso J., Albert M.L., Quintana-Murci L. Genetic adaptation and Neandertal admixture shaped the immune system of human populations. *Cell*. 2016;167(3):643-656.e17. doi 10.1016/j.cell.2016.09.024

Raj T., Rothamel K., Mostafavi S., Ye C., Lee M.N., Replogle J.M., Feng T., … Hacohen N., Mathis D., Benoist C., Stranger B.E., De Jager P.L. Polarization of the effects of autoimmune and neurodegenerative risk alleles in leukocytes. *Science*. 2014;344(6183):519-523. doi 10.1126/science.1249547

Rana S., Burke S.D., Karumanchi S.A. Imbalances in circulating angiogenic factors in the pathophysiology of preeclampsia and related disorders. *Am J Obstet Gynecol.* 2022;226(2S):S1019-S1034. doi 10.1016/j.ajog.2020.10.022

Sirugo G., Williams S.M., Tishkoff S.A. The missing diversity in human genetic studies. *Cell*. 2019;177(1):26-31. doi 10.1016/j.cell.2019.04.032

Spielman R.S., Bastone L.A., Burdick J.T., Morley M., Ewens W.J., Cheung V.G. Common genetic variants account for differences in gene expression among ethnic groups. *Nat Genet*. 2007;39(2):226-231. doi 10.1038/ng1955

Steinberg K.M., Antonacci F., Sudmant P.H., Kidd J.M., Campbell C.D., Vives L., Malig M., … Froment A., Donnelly M.P., Kidd K.K., Tishkoff S.A., Eichler E.E. Structural diversity and African origin of the 17q21.31 inversion polymorphism. *Nat Genet*. 2012;44(8):872-880. doi 10.1038/ng.2335

Stepanov V.A. Evolution of genetic diversity and human diseases. *Russ J Genet*. 2016;52(7):746-756. doi 10.1134/S1022795416070103

Storey J.D., Madeoy J., Strout J.L., Wurfel M., Ronald J., Akey J.M. Gene-expression variation within and among human populations. *Am J Hum Genet*. 2007;80(3):502-509. doi 10.1086/512017

Stranger B.E., Nica A.C., Forrest M.S., Dimas A., Bird C.P., Beazley C., Ingle C.E., Dunning M., Flicek P., Koller D., Montgomery S., Tavaré S., Deloukas P., Dermitzakis E.T. Population genomics of human gene expression. *Nat Genet*. 2007;39(10):1217-1224. doi 10.1038/ng2142

Stranger B.E., Montgomery S.B., Dimas A.S., Parts L., Stegle O., Ingle C.E., Sekowska M., Gutierrez-Arcelus M., Nisbett J., Nica A.C., Beazley C., Durbin R. Patterns of cis regulatory variation in diverse human populations. *PLoS Genet*. 2012;8(4):e1002639. doi 10.1371/journal.pgen.1002639

Thames A.D., Irwin M.R., Breen E.C., Cole S.W. Experienced discrimination and racial differences in leukocyte gene expression. *Psychoneuroendocrinology*. 2019;106:277-283. doi 10.1016/j.psyneuen.2019.04.016

Torchin H., Ancel P.Y., Jarreau P.H., Goffinet F. Epidemiology of preterm birth: prevalence, recent trends, short- and long-term outcomes. *J Gynecol Obstet Biol Reprod* (*Paris*). 2015;44(8):723-731. doi 10.1016/j.jgyn.2015.06.010

Trifonova E., Babovskaya A., Zarubin A., Markov A., Stepanov V. Placental tissue co-expression networks across Russians and Yakuts identify key genes and pathways for preeclampsia. *Eur J Hum Genet*. 2022;30:108-109. doi 10.1038/s41431-021-01026-1

Wallace T.A., Prueitt R.L., Yi M., Howe T.M., Gillespie J.W., Yfantis H.G., Stephens R.M., Caporaso N.E., Loffredo C.A., Ambs S. Tumor immunobiological differences in prostate cancer between African-American and European-American men. *Cancer Res*. 2008;68(3):927-936. doi 10.1158/0008-5472.CAN-07-2608

Way G.P., Rudd J., Wang C., Hamidi H., Fridley B.L., Konecny G.E., Goode E.L., Greene C.S., Doherty J.A. Comprehensive cross-population analysis of high-grade serous ovarian cancer supports no more than three subtypes. *G3* (*Bethesda*). 2016;6(12):4097-4103. doi 10.1534/g3.116.033514

Wei P., Milbauer L.C., Enenstein J., Nguyen J., Pan W., Hebbel R.P. Differential endothelial cell gene expression by African Americans versus Caucasian Americans: a possible contribution to health disparity in vascular disease and cancer. *BMC Med*. 2011;9:2. doi 10.1186/1741-7015-9-2

Westen A.A., Matai A.S., Laros J.F., Meiland H.C., Jasper M., de Leeuw W.J., de Knijff P., Sijen T. Tri-allelic SNP markers enable analysis of mixed and degraded DNA samples. *Forensic Sci Int Genet.* 2009;3(4):233-241. doi 10.1016/j.fsigen.2009.02.003

Whitney A.R., Diehn M., Popper S.J., Alizadeh A.A., Boldrick J.C., Relman D.A., Brown P.O. Individuality and variation in gene expression patterns in human blood. *Proc Natl Acad Sci USA.* 2003;100(4):1896-1901. doi 10.1073/pnas.252784499

Wolf M., Hubel C.A., Lam C., Sampson M., Ecker J.L., Ness R.B., Rajakumar A., Daftary A., Shakir A.S., Seely E.W., Roberts J.M., Sukhatme V.P., Karumanchi S.A., Thadhani R. Preeclampsia and future cardiovascular disease: potential role of altered angiogenesis and insulin resistance. *J Clin Endocrinol Metab.* 2004;89(12):6239-6243. doi 10.1210/jc.2004-0548

Yang H.C., Lin C.W., Chen C.W., Chen J.J. Applying genome-wide gene-based expression quantitative trait locus mapping to study population ancestry and pharmacogenetics. *BMC Genomics*. 2014;15:319. doi 10.1186/1471-2164-15-319

Yin L., Coelho S.G., Ebsen D., Smuda C., Mahns A., Miller S.A., Beer J.Z., Kolbe L., Hearing V.J. Epidermal gene expression and ethnic pigmentation variations among individuals of Asian, European and African ancestry. *Exp Dermatol.* 2014;23(10):731-735. doi 10.1111/exd.12518

Zhang W., Duan S., Kistner E.O., Bleibel W.K., Huang R.S., Clark T.A., Chen T.X., Schweitzer A.C., Blume J.E., Cox N.J., Dolan M.E. Evaluation of genetic variation contributing to differences in gene expression between populations. *Am J Hum Genet.* 2008;82(3):631-640. doi 10.1016/j.ajhg.2007.12.015

Conflict of interest. The authors declare no conflict of interest.
Received February 18, 2025. Revised April 3, 2025. Accepted April 4, 2025.

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2025;29(5):704-710

doi 10.18699/vjgb-25-77

Mitochondrial genome polymorphism in the East Slavic population of Northeastern Siberia

B.A. Malyarchuk, G.A. Denisova, A.N. Litvinov

Institute of Biological Problems of the North of the Far Eastern Branch of the Russian Academy of Sciences, Magadan, Russia
malbor@mail.ru

Abstract. Data on mitochondrial DNA (mtDNA) polymorphism at the population level are of significant interest to researchers in the fields of population and ethnic genetics, forensic medicine, and forensic science. In the present study, we have obtained data on the variability of whole mitochondrial genomes in the immigrant East Slavic population of Northeastern Siberia (using the Magadan region as an example). The study yielded novel data concerning mtDNA variability in the Magadan region's inhabitants comprising maternal lineages of Russians ($N = 49$) and Ukrainians ($N = 15$), as well as individuals with a mixture of maternal and paternal ancestries, including Russians on the maternal side and indigenous populations (Koryaks, Evenes, and Itelmens) on the paternal side ($N = 4$). In addition, the mitogenomes of the Russian population from the Novgorod, Kaluga, and Yaroslavl regions ($N = 15$) were sequenced to enhance the power of the phylogeographic analysis. The results of the study demonstrated that the mitochondrial gene pool of the East Slavic immigrant population in the Magadan region is characterized by a high level of diversity. The analysis of genetic differentiation of Russian populations within Russia, as measured by the variability of complete mitochondrial genomes, revealed a low level of interpopulation differences (Fst = 0.15 %, $P = 0.2$). The results of multidimensional scaling of Fst distances indicate that the Russians residing in the Magadan region are genetically similar to the Russian populations inhabiting the southwestern part of the country, specifically the Belgorod and Orel regions. The gene pool of the Russian population in the Magadan region is predominantly characterized by mtDNA haplotypes of West Eurasian (including European) origin. The prevalence of East Asian-derived haplotypes among the Russian population is relatively low, accounting for approximately 4.8 % of the total. However, certain East Asian-specific haplogroups, such as F1b1 and Z1a1a, have demonstrated a prolonged presence in the gene pools of Eastern European populations, as evidenced by phylogeographic analysis. Among the European mtDNA haplotypes of Russians from the Magadan region, Eastern European variants predominate, and they also have a high proportion of mtDNA haplotypes specific to Slavs (19.4 %). Furthermore, rare mtDNA haplotypes have been identified in the mitochondrial gene pools of Russians and Ukrainians residing in the Magadan region. These rare haplotypes are linked to the maternal lines of Empress Alexandra Fedorovna Romanova (haplogroup H1af2) and Prince Dmitry, son of Prince Alexander Nevsky (haplogroup F1b1-a3a2a).

Key words: mitochondrial genome; genetic polymorphism; molecular phylogeography; Eastern Slavs; Northeastern Siberia

For citation: Malyarchuk B.A., Denisova G.A., Litvinov A.N. Mitochondrial genome polymorphism in the East Slavic population of Northeastern Siberia. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov J Genet Breed*. 2025;29(5):704-710. doi 10.18699/vjgb-25-77

Полиморфизм митохондриальных геномов у восточнославянского населения Северо-Востока Сибири

Б.А. Малярчук, Г.А. Денисова, А.Н. Литвинов

Институт биологических проблем Севера Дальневосточного отделения Российской академии наук, Магадан, Россия
malbor@mail.ru

Аннотация. Информация о полиморфизме митохондриальной ДНК (мтДНК) на популяционном уровне вызывает большой интерес со стороны исследователей в области популяционной и этнической генетики, судебной медицины и криминалистики. В настоящей работе получены данные об изменчивости целых митохондриальных геномов у пришлого восточнославянского населения Северо-Востока Сибири (на примере Магаданской области), а также новые сведения об изменчивости мтДНК у жителей Магаданской области – русских ($N = 49$) и украинцев ($N = 15$) – по материнской линии, а также метисного происхождения – русских по материнской линии и коренных жителей (коряки, эвены, ительмены) по отцовской линии ($N = 4$). Кроме этого, для повышения информативности филогеографического анализа секвенированы митогеномы русских Новгородской, Калужской и Ярославской областей ($N = 15$). Результаты проведенного исследования по-

© Malyarchuk B.A., Denisova G.A., Litvinov A.N., 2025
This work is licensed under a Creative Commons Attribution 4.0 License

казали, что митохондриальный генофонд пришлого восточнославянского населения Магаданской области характеризуется высоким уровнем генетического разнообразия. Результаты анализа генетической дифференциации популяций русского населения России по данным об изменчивости целых митохондриальных геномов свидетельствуют о низком уровне межпопуляционных различий (*F*st = 0.15 %, *P* = 0.2). По результатам многомерного шкалирования *F*st-дистанций русские Магаданской области кластеризуются с русскими юго-западной части страны – популяциями Белгородской и Орловской областей. В генофонде русского населения Магаданской области преобладают гаплотипы мтДНК западноевразийского (включая европейское) происхождения. Доля гаплотипов восточноазиатского происхождения у русских невелика (4.8 %), однако часть из них (гаплогруппы F1b1 и Z1a1a), согласно результатам филогеографического анализа, длительное время эволюционировала в генофондах популяций Восточной Европы. Среди европейских по происхождению гаплотипов мтДНК у магаданских русских преобладают восточноевропейские варианты, а также у них высока доля гаплотипов мтДНК, специфичных для славян (19.4 %). В митохондриальных генофондах русских и украинцев Магаданской области обнаружены редкие гаплотипы, родственные материнским линиям императрицы Александры Федоровны Романовой (гаплогруппа H1af2) и князя Дмитрия, сына князя Александра Невского (гаплогруппа F1b1-a3a2a).

Ключевые слова: митохондриальный геном; генетический полиморфизм; молекулярная филогеография; восточные славяне; Северо-Восток Сибири

Introduction

Mitochondrial DNA (mtDNA) is a maternally inherited, non-recombining genetic system that is highly informative for studying the genetic history of populations and reconstructing migrations. The utilization of mtDNA markers to investigate the emergence processes of Siberian immigrant populations of Eastern European (predominantly Russian) origin dates back to the initial population studies of mitochondrial DNA polymorphism in Russia (Lemza, Sokolova, 1992; Malyarchuk et al., 1994; Derenko, Malyarchuk, 1996; Kazakovtseva et al., 1998). Recent studies have demonstrated that the migrant population of Siberia exhibits a high diversity of mitochondrial lineages, predominantly of European origin (Gubina et al., 2014). This phenomenon can be attributed to historical migration processes that commenced in the 16th–17th centuries, coinciding with the exploration of Siberia by Russian pioneers. These processes persisted throughout the capitalist and socialist development of Russian society, contributing to the contemporary genetic landscape of the region.

Meanwhile, investigations of mtDNA polymorphism in the Russian Old Believers of Siberia, who diverged in the mid-17th century, demonstrated that the genetic composition of the Old Believers aligns with that of European populations, including Russians. Nevertheless, the prevalence of the East Asian component in the gene pool of the Russian Old Believers is marginally higher compared to contemporary Russians residing in the Novosibirsk Oblast (Gubina et al., 2014). This observation may be attributed to the increased level of intermarriage between the Old Believers and indigenous Siberians, living in close proximity for several centuries.

Meanwhile, investigations of mtDNA polymorphism in the Russian Old Settlers – descendants of Russian servants and traders who established themselves in the Far North of Eastern Siberia from the mid-17th century onward – revealed a remarkably elevated proportion of East Asian genetic patterns in the gene pools of the Russian Old Settlers (100 % in Pokhodsk and Markovo residents, 67 % in Russkoe-Ust'e residents) (Sukernik et al., 2010; Borisova et al., 2024). This phenomenon can be attributed to the unique historical development of the Russian Old Settlers' populations, characterized by intermarriage between immigrants from Eastern Europe (Cossacks, merchants, and industrialists) and indigenous women, predominantly of Yukaghir origin (Sukernik et al., 2010; Solovyev et al., 2023; Borisova et al., 2024).

The genetic structures of modern indigenous populations in the northeast of Siberia are similarly organized. The gene pools of the Koryaks, Evens, and Chukchi also show asymmetry in the contribution of maternal (by mtDNA) and paternal (by Y chromosome) lineages of European origin, with paternal lineages of this kind sharply prevailing over maternal lineages (Balanovska et al., 2020; Agdzhoyan et al., 2021; Derenko et al., 2023; Solovyev et al., 2023; Borisova et al., 2024; Malyarchuk, Derenko, 2024). The initial settlements established by Russian explorers and traders (Tauysk, Gizhiginsk, Yamsk, Ola, and other settlements) emerged in the 17th–18th centuries within the boundaries of Magadan Oblast. However, it was not until the 1930s that a significant influx of individuals, numbering in the hundreds of thousands, arrived in the region, primarily driven by the economic development of this prosperous Siberian territory. Consequently, the Magadan Oblast began to form the so-called newcomer population, which was predominantly comprised of Russians (81.5 %) and Ukrainians (6.3 %) (according to Rosstat data from 2010). The indigenous aboriginal population constituted approximately 3 % of the total Magadan Oblast population, numbering 133,387 individuals (according to Rosstat data from 2024).

In this paper, we present data on the variability of whole mitochondrial genomes in the immigrant population of the Magadan region. Our objective is to use these data to identify phylogeographic patterns of mitochondrial haplotypes and to study interpopulation relationships of the Russian population of the Russian Federation.

Materials and methods

Mitochondrial genome sequences of the Russian residents of the Magadan region (Magadan city and settlements of the Severo-Evensky district; $N = 49$) were determined. The mitochondrial genomes of Magadan residents were also subjected to phylogeographic analysis, which included 15 Ukrainians (on the maternal side) and four individuals of mixed ancestry,

comprising Russians on the maternal side and indigenous peoples (Koryaks, Evenes, and Itelmens) on the paternal side (Table S1 in the Supplementary Materials)[1].

Furthermore, the mitogenomes of the Russians from the Novgorod, Kaluga, and Yaroslavl oblasts (N = 10, 3, and 2, respectively) were sequenced to enhance the comprehensiveness of phylogeographic analysis (Table S1). The methods and approaches employed for sequencing whole mtDNA molecules were previously described (Derenko, Malyarchuk, 2010). The nucleotide sequences of mitochondrial genomes were deposited in GenBank (www.ncbi.nlm.nih.gov) under accession numbers PQ285752-PQ285800, PQ300111-PQ300129, and PQ283331-PQ283345.

Phylogenetic relationships between mtDNA nucleotide sequences were analyzed using the maximum parsimony method implemented in the mtPhyl v4.015 program (https://isogg.org/wiki/MtPhyl). During the construction of the tree, length polymorphism at sites 16180-16193, 309-315, and 514-524 as well as substitution at position 16519 were not considered. To determine the evolutionary age of monophyletic mtDNA clusters, we employed the molecular clock incorporated in the mtPhyl program based on the correction of the long-term phylogenetic rate of mutations in human mtDNA (one nucleotide substitution in the whole mitogenome over 3,624 years, or 1.665×10^{-8} substitutions per site per year), taking into account the effect of negative selection (Soares et al., 2009).

The phylogenetic classification of mtDNA haplogroups proposed by the developers of the PhyloTree resource (www.phylotree.org) was utilized as a foundation for this study. Furthermore, the regularly updated classification of mtDNA variants provided by the YFull MTree resource (https://www.yfull.com/mtree/) was also considered. The designation of monophyletic mtDNA clusters as ethnospecific was made on the following criteria: at least 75 % of mtDNA haplotypes in the cluster were found to be characteristic of representatives of a certain ethnic group (e. g., Russians) or ethnic community (Slavs). For the phylogeographic analysis of mtDNA, we utilized data on the variability of complete mitogenomes in individuals representing diverse human populations, as reported in GenBank, the Logan DNA Project (http://www.ianlogan.co.uk), and YFull MTree. At the end of 2024, the GenBank database has accumulated a collection of over 61,000 mitochondrial genomes, representing a diverse array of ethnic groups from various geographical regions worldwide (https://www.mitomap.org/foswiki/bin/view/MITOMAP/GBFreqInfo). The ethnic affiliation of the studied samples was determined through the utilization of information derived from the available databases.

In order to undertake a comparative analysis of complete mtDNA variability at the population level, we employed previously published data for the Russian populations of the European part of the country (Malyarchuk et al., 2017). In addition to 49 new mtDNA sequences, the Russian samples from the Magadan region included 13 previously sequenced mitochondrial genomes (Table S1). The genetic diversity parameters of populations were calculated using the DnaSP 5.10.01 software package (Librado, Rozas, 2009). The analysis of molecular variability (AMOVA, Fst-analysis) based on pairwise nucleotide differences between mitogenomes was performed using the Arlequin 3.5.1.2 software package (Excoffier, Lischer, 2010). The location of populations in two-dimensional space was investigated using the multidimensional scaling method of interpopulation Fst-differences, implemented in the STATISTICA10 software package (StatSoft Inc.).

[1] Table S1 and Fig. S1 are available at:
https://vavilov.elpub.ru/jour/manager/files/Suppl_Malyarchuk_Engl_29_5.xlsx

Results and discussion

Interpopulation differences of the Russian population based on whole mitochondrial genome variability data

An analysis of mtDNA variability revealed that the studied Russian samples from the Magadan region (49 new and 13 previously sequenced mitogenomes) do not show significant differences at the population level in the primary parameters of genetic diversity in comparison with other Russian populations of the European part of Russia (Table 1). According to these results, the Magadan Russians occupy an intermediate position among the Russian populations, where the lowest values of the average number of pairwise nucleotide differences (k) are observed among the Russians of the Pskov and Novgorod regions, and the highest, among the Russians of the Vladimir region. Tajima's D parameter, which tests the neutrality of mtDNA evolution within populations, has a significantly negative value in the Magadan region, a finding also observed in other European populations, including Russians (Litvinov et al., 2020). This suggests the influence of negative selection on mtDNA variability.

The analysis of Fst-differences between mitochondrial genomes in Russian populations showed the absence of interpopulation differences (Fst = 0.15 %, P = 0.2). Pairwise comparisons revealed statistically significant differences between the population of the Vladimir region and the northwestern Russian populations of the Pskov and Novgorod regions (Table 2). Conversely, the results of multidimensional scaling of interpopulation Fst-differences indicate that the Russian population of the Magadan region is grouped with the southwestern populations of the Belgorod and Orel regions (Fig. 1). This grouping is significantly different from the Russians of the Vladimir region, as can be seen from the results of multidimensional scaling.

Phylogeographic analysis of mtDNA haplotypes of the East Slavic immigrant population of the Magadan region

The gene pool of Russian population of Magadan region is represented mainly by Western Eurasian mtDNA haplogroups (95.2 %) (Table S1). The most prevalent are haplogroups H (37.1 %), T (16.1 %), U (14.5 %), J (6.5 %), and HV (4.8 %). Other Russian populations in Eastern Europe have similar mitochondrial haplogroup spectra (Morozova et al., 2012; Kushniarevich et al., 2015; Malyarchuk et al., 2017).

The results of the phylogenetic analysis revealed the presence of 62 different mtDNA haplotypes within the cohort of the Russian population residing in the Magadan region (N = 62). This indicates that the sample did not contain identical haplotypes. Furthermore, phylogeographic analysis of

Table 1. Genetic parameters in Russian populations according to the data on variability of whole mitochondrial genomes

Populations	*N*	*n*	*s*	*h*	π	*k*	Tajima's D (*p*-value)
Magadan[1]	62	61	416	0.999 ± 0.003	0.0018 ± 0.0009	30.24	–2.39 (< 0.01)
Belgorod[2]	64	64	437	1.0 ± 0.003	0.0018 ± 0.0009	30.18	–2.39 (< 0.01)
Orel[2]	48	48	311	1.0 ± 0.004	0.0017 ± 0.0009	28.44	–2.18 (< 0.01)
Vladimir[2]	73	71	433	0.999 ± 0.002	0.0019 ± 0.0009	31.38	–2.27 (< 0.01)
Tula[2]	59	59	418	1.0 ± 0.003	0.0018 ± 0.0009	29.38	–2.41 (< 0.01)
Pskov[2]	68	66	368	0.999 ± 0.003	0.0016 ± 0.0008	26.88	–2.29 (< 0.01)
Veliky Novgorod[2]	64	63	404	1.0 ± 0.003	0.0017 ± 0.0009	27.99	–2.41 (< 0.01)
Russians in total[1]	438	419	1,224	0.9997 ± 0.0002	0.0018 ± 0.0009	29.19	–2.59 (< 0.001)

Note. *N* – sample size; *n* – number of haplotypes; *s* – number of polymorphic sites; *h* – haplotypic diversity; π – nucleotide diversity; *k* – average number of pairwise nucleotide differences.
References: [1] present study; [2] Malyarchuk et al., 2017.

Table 2. *Fst*-differences between Russian populations according to the data on the variability of nucleotide sequences of whole mitogenomes

Populations	Belgorod	Orel	Vladimir	Tula	Pskov	Veliky Novgorod
Belgorod	0					
Orel	0	0				
Vladimir	0.0048	0.0025	0			
Tula	0.0007	0	0.0085	0		
Pskov	0	0	0.0115*	0.0043	0	
Veliky Novgorod	0	0	0.0093*	0.0026	0	0
Magadan	0.0005	0	0	0.0039	0.0013	0

Note. *Fst*-values are based on pairwise nucleotide differences between mtDNA haplotypes and are shown under the diagonal.
* Statistically significant differences ($P < 0.05$).

the mtDNA haplotypes identified in the Magadan Russians revealed that these genetic variants are predominantly distributed within the European population (refer to Table S1 and Figure S1 for more details). A total of three distinct haplotypes were identified as belonging to haplogroups that are prevalent in West Asia, including H-7630-11113-12172, R0a1a5, and M5a1b, with samples 10_R, 2_R, and 44_R corresponding to these haplogroups, respectively. However, the origins of two additional haplotypes could not be determined. Approximately a quarter of the Magadan sample (25.8 %) consisted of mtDNA haplotypes common in Eastern European populations, with almost 20 % of the Magadan Russian haplotypes belonging to mtDNA subgroups predominantly distributed among Slavs (haplogroups HV-15617, HV6a1, H1b2g, H13a2b, U4d2b, U5a1a1h, U5a2b1g, U5a2a1o, K1c1e, K2b1, J1c3a1, V7a) (Fig. S1). A proportion of the samples (11.3 %) were found to belong to haplogroups that had been identified primarily within the Russian population (H5a1a, J1c4b1, I1a1c-10454, W1c-10086-12136, R1a1a1, F1b1-a3a2a). The frequency of the Baltic-Finnish component (H1n4, H49, U5b1b1a, Z1a1a) was 8.1 %.

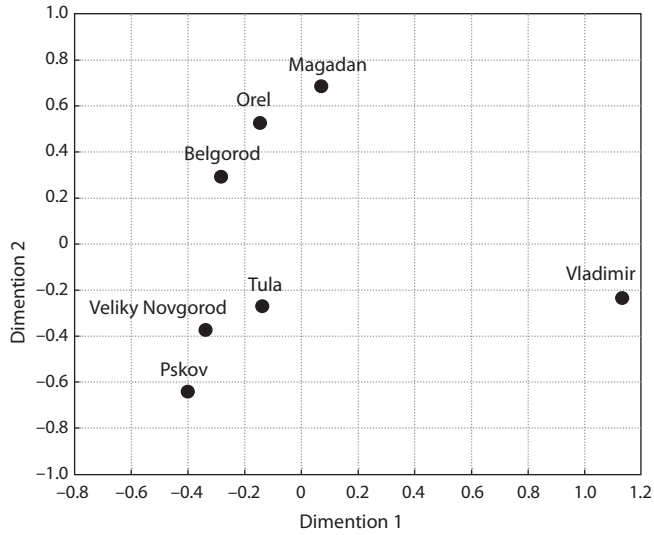


Fig. 1. Results of multidimensional scaling of interpopulation *Fst*-values based on pairwise nucleotide differences between sequences of whole mitochondrial genomes from different Russian populations. The stress value is equal to 0.00003.

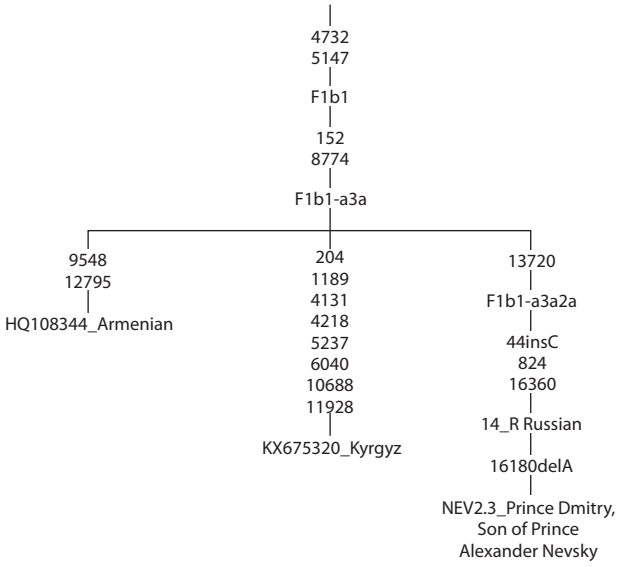


Fig. 2. Phylogenetic tree of mtDNA of haplogroup F1b1-a3a.

Nucleotide positions at which transitions occurred are shown on the branches; 16180delA indicates deletion A at position 16180; the GenBank sample numbers and ethnicity are also indicated.

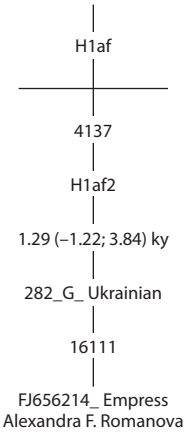


Fig. 3. Phylogenetic tree of mtDNA of haplogroup H1af.

Nucleotide positions at which transitions occurred are shown on the branches, and the evolutionary age of the H1af2 subgroup in thousands of years (ky) is also indicated.

Three mtDNA haplotypes identified in the Russians of the Magadan region belong to haplogroups of East Asian origin: F1b1, Z1a1a, and N9a2a2. However, the phylogeographic analysis demonstrated that haplogroup N9a2a2 is exclusively distributed in East Asia, as it was detected only in the Japanese population (Fig. S1). The presence of the Z1a1a haplotype suggests a potential contribution of Finno-Ugric tribes to the ethnic history of Russians (Lunkina et al., 2004). This mtDNA haplogroup is prevalent in populations from Northeastern Europe, particularly in the Fennoscandian region. It is hypothesized that this haplogroup was introduced into the gene pool of the Sami and Finns from the Volga-Urals region about 3,000 years ago (Tambets et al., 2004; Ingman, Gyllensten, 2007). The F1b1-haplotype identified in the Magadan Russians is noteworthy for its association with the F1b1-a3a2a subgroup, which, according to the YFull MTree classification system, is predominantly present in the modern Russian populations of the Kursk, Belgorod, and Bryansk regions (Table S1). Phylogenetic analysis showed that this haplotype is almost identical (except for deletion at position 16180) to the F1b1-haplotype identified in Prince Dmitry, son of Prince Alexander Nevsky, and possibly of Cuman origin (Zhur et al., 2023) (Fig. 2).

Phylogeographic analysis of 15 mitogenomes of Ukrainian origin also demonstrated that the identified mtDNA haplotypes are predominantly distributed in European populations (Table S1). Three of these haplotypes belong to Slavic-specific mtDNA haplogroups (H6c1, H1u5a2, HV9b1a1), while one more belongs to haplogroup U5b1b1a1d, which is predominantly distributed among Finns. Notably, the Ukrainian sample from the Magadan region exhibited a rare haplotype belonging to haplogroup H1af2, which is known to be associated with the maternal lineage of Empress Alexandra Fedorovna Romanova (Fig. 3).

Despite the substantial size of the GenBank database, which includes more than 61,000 mitogenomes from diverse ethnic backgrounds, it lacks other H1af2 sequences. In contrast, the YFull MTree database contains additional H1af2 sequences (YF129594 and YF109892). However, both databases lack information on ethnic origin and the nucleotide sequences themselves. Our study enabled the detection of an H1af2 haplotype associated with the Russian Tsarina in an individual of Ukrainian origin. The evolutionary age of the H1af2 subgroup was determined to be about 1.3 thousand years, and the YFull MTree database suggests a slightly higher estimate of 3.2 thousand years. Genealogical data indicate that the maternal lineage of the Russian Empress (or Queen Victoria of Great Britain, respectively) is of Western European (especially French) origin.

Phylogeographic analysis of individuals of mixed ancestry revealed that the haplotypes of two individuals (Koryak on the paternal side) belong to haplogroups found predominantly among Finns (K1c1 and U5b1b1a). The remaining two mtDNA haplotypes are associated with haplogroups for which the precise geographical origins remain uncertain; both J2b1a11 and U5a1a1a are common in European populations.

Conclusion

The results of the study demonstrated that the mitochondrial gene pool of the East Slavic immigrant population in Northeastern Siberia (with a focus on the Magadan region) exhibited a high level of diversity. At the same time, the genetic differentiation of Russian populations, as determined by the analysis of variability in whole mitochondrial genomes, was found to be remarkably low. This finding suggests a high degree of similarity among mtDNA haplotypes in diverse Russian populations. The mitochondrial gene pool of the Russian population of the Magadan region does not differ significantly from the gene pools of other Russian populations of the European part of the country. According to the results of multidimensional scaling of *F*st-distances, the Magadan Russians cluster with the Russian populations of the southwestern part of the country, spe-

cifically the Belgorod and Orel regions. This result indicates a distinctive characteristic of the contemporary population of Northeastern Siberia in comparison with the populations of the Russian Old Settlers, who established themselves in the Far North of Eastern Siberia since the middle of the 17th century. The Russian Old Settlers have a markedly high prevalence of East Asian (predominantly Yukaghir) mtDNA haplotypes (Sukernik et al., 2010; Borisova et al., 2024). The analysis of Y chromosome polymorphism and autosomal loci shows that the European variants of polymorphism preserved in these populations (along with the Russian language and culture) are predominantly associated with the population of Northeastern Europe. This finding lends support to the "Pomor" hypothesis of the origin of the Russian Old Settlers of the Arctic coast (Solovyev et al., 2023).

An analysis of the gene pool of the Russian and Ukrainian inhabitants of the Magadan region shows that the mtDNA haplotypes of European origin are predominant. The proportion of East Asian mtDNA haplotypes among Russians is minimal (4.8 %), and further analysis using phylogeographic methods revealed that F1b1 and Z1a1a haplotypes, despite their East Asian origin, have undergone extensive evolution within the gene pools of Eastern European populations. Among mtDNA haplotypes of European origin, Eastern European variants predominate in the Magadan region Russians, accounting for a significant proportion (19.4 %) of mtDNA haplotypes specific to Slavs. Such high frequencies of Slavic-specific mtDNA lineages have been previously reported only among Ukrainians – 23.6 % (Malyarchuk, Derenko, 2023).

It is interesting to note that the findings in the mitochondrial gene pools of modern Russians and Ukrainians have revealed the presence of rare haplotypes that appear to be related to the maternal ancestry of Empress Alexandra Fedorovna Romanova and Prince Dmitry, son of Prince Alexander Nevsky. These findings may offer a promising avenue for further research.

References

Agdzhoyan A.T., Bogunova A.A., Kamenshchikova E.N., Zaporozhchenko V.V., Bogunov Y.V., Balanovsky O.P., Balanovska E.V. The Chukchi of Kamchatka: a genetic portrait based on the wide array of Y-chromosome markers. *Vestnik Moskovskogo Universiteta. Seria XXIII. Antropologia = Moscow University Anthropology Bulletin.* 2021;1:80-92. doi 10.32521/2074-8132.2021.1.080-092 (in Russian)

Balanovska E.V., Bogunov Y.V., Bogunova A.A., Kamenshchikova E.N., Pylev V.Y., Bychkovskaya L.S., Balanovsky O.P., Lavryashina M.B. Demographic portrait of Koryaks from Northern Kamchatka. *Vestnik Moskovskogo Universiteta. Seria XXIII. Antropologia = Moscow University Anthropology Bulletin.* 2020;4:111-122. doi 10.32521/2074-8132.2020.4.111-122 (in Russian)

Borisova T.V., Solovyev A.V., Romanov G.P., Teryutin F.M., Pshennikova V.G., Barashkov N.A., Fedorova S.A. Analysis of the mitochondrial gene pool structure of Russian old-settlers of the Arctic coast of Yakutia from village Russkoye Usty'e. *Russ J Genet.* 2024;60(11):1538-1547. doi 10.1134/S1022795424701096

Derenko M.V., Malyarchuk B.A. Comparative analysis of RFLP of mitochondrial DNA from Eastern Slavic populations in Russia. *Genetika.* 1996;32(6):815-821 (in Russian)

Derenko M.V., Malyarchuk B.A. Molecular Phylogeography of Populations of Northern Eurasia Based on Mitochondrial DNA Variability Data. Magadan: NESC FEB RAS, 2010 (in Russian)

Derenko M., Denisova G., Litvinov A., Dambueva I., Malyarchuk B. Mitogenomics of the Koryaks and Evens of the northern coast of the Sea of Okhotsk. *J Hum Genet.* 2023;68(10):705-712. doi 10.1038/s10038-023-01173-x

Excoffier L., Lischer H.E. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour.* 2010;10(3):564-567. doi 10.1111/j.1755-0998.2010.02847.x

Gubina M.A., Babenko V.N., Damba L.D., Ponomareva M.N., Konovalova N.A., Voevoda M.I. Polymorphism of mitochondrial DNA in Old Believers from Siberia. *Russ J Genet.* 2014;50(6):638-652. doi 10.1134/S1022795414060040

Ingman M., Gyllensten U. A recent genetic link between Sami and the Volga-Ural region of Russia. *Eur J Hum Genet.* 2007;15:115-120. doi 10.1038/sj.ejhg.5201712

Kazakovtseva M.A., Voevoda M.I., Babenko V.N., Osipova L.P. Restriction-site polymorphism of the mitochondrial DNA control region in populations of Russian Old Believers and migrant Slavic populations in Northern Siberia. *Russ J Genet.* 1998;34(4):439-444

Kushniarevich A., Utevska O., Chuhryaeva M., Agdzhoyan A., Dibirova K., Uktveryte I., Möls M., … Balanovska E., Metspalu M., Kivisild T., Villems R., Balanovsky O. Genetic heritage of the Balto-Slavic speaking populations: a synthesis of autosomal, mitochondrial and Y-chromosomal data. *PLoS One.* 2015;10(9):e0135820. doi 10.1371/journal.pone.0135820

Lemza S.V., Sokolova O.V. Restriction polymorphism of mitochondrial DNA among the Russian population of Western Siberia. *Genetika = Genetics* (*Moscow*). 1992;28(5):136-140 (in Russian)

Librado P., Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics.* 2009;25(11):1451-1452. doi 10.1093/bioinformatics/btp187

Litvinov A.N., Malyarchuk B.A., Derenko M.V. The nature of the molecular evolution of the mitochondrial genomes of the Russian population of East Europe. *Vestnik Severo-Vostochnogo Nauchnogo Centra DVO RAN = The Bulletin of the North-East Scientific Center FEB RAS.* 2020;2:107-113. doi 10.34078/1814-0998-2020-2-107-113 (in Russian)

Lunkina A.V., Denisova G.A., Derenko M.V., Malyarchuk B.A. Mitochondrial DNA variation in two Russian populations from Novgorod oblast. *Russ J Genet.* 2004;40(7):795-799. doi 10.1023/B:RUGE.0000036530.09850.70

Malyarchuk B.A., Derenko M.V. Mitochondrial gene pool of Ukrainians in the context of variability of whole mitogenomes in Slavic peoples. *Russ J Genet.* 2023;59(1):88-96. doi 10.1134/S1022795423010088

Malyarchuk B.A., Derenko M.V. Genetic history of the Koryaks and Evens of the Magadan region based on Y-chromosome polymorphism data. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov J Genet Breed.* 2024;28(1):90-97. doi 10.18699/vjgb-24-11

Malyarchuk B.A., Lapinsky A.G., Balmysheva N.P., Butorina O.T., Solovenchuk L.L. RFLP of mitochondrial DNA in the population of Magadan city. *Genetika = Genetics* (*Moscow*). 1994;30(1):112-114 (in Russian)

Malyarchuk B., Litvinov A., Derenko M., Skonieczna K., Grzybowski T., Grosheva A., Shneider Y., Rychkov S., Zhukova O. Mitogenomic diversity in Russians and Poles. *Forensic Sci Int Genet.* 2017;30:51-56. doi 10.1016/j.fsigen.2017.06.003

Morozova I., Evsyukov A., Kon'kov A., Grosheva A., Zhukova O., Rychkov S. Russian ethnic history inferred from mitochondrial DNA diversity. *Am J Phys Anthropol.* 2012;147(3):341-351. doi 10.1002/ajpa.21649

Soares P., Ermini L., Thomson N., Mormina M., Rito T., Röhl A., Salas A., Oppenheimer S., Macaulay V., Richards M.B. Correcting for purifying selection: an improved human mitochondrial molecular clock. *Am J Hum Genet.* 2009;84(6):740-759. doi 10.1016/j.ajhg.2009.05.001

Solovyev A.V., Borisova T.V., Romanov G.P., Teryutin F.M., Pshennikova V.G., Nikitina S.E., Alekseev A.N., Barashkov N.A., Fedorova S.A. Genetic history of Russian Old-Settlers of the Arctic coast of Yakutia from the settlement of Russkoye Ust'ye inferred from Y chromosome data and genome-wide analysis. *Russ J Genet.* 2023; 59(9):949-955. doi 10.1134/S1022795423090119

Sukernik R.I., Volod'ko N.V., Mazunin I.O., Eltsov N.P., Starikovskaya E.B. The genetic history of Russian old settlers of polar northeastern Siberia. *Russ J Genet.* 2010;46(11):1386-1394. doi 10.1134/S1022795410110153

Tambets K., Rootsi S., Kivisild T., Help H., Serk P., Loogväli E.L., Tolk H.V., … Ferák V., Füredi S., Komel R., Beckman L., Villems R. The western and eastern roots of the Saami – the story of genetic "outliers" told by mitochondrial DNA and Y chromosomes. *Am J Hum Genet.* 2004;74(4):661-682. doi 10.1086/383203

Zhur K.V., Sharko F.S., Sedov V.V., Dobrovolskaya M.V., Volkov V.G., Maksimov N.G., Seslavine A.N., Makarov N.A., Prokhortchouk E.B. The Rurikids: the first experience of reconstructing the genetic portrait of the ruling family of Medieval Rus' based on paleogenomic data. *Acta Naturae.* 2023;15(3):50-65. doi 10.32607/actanaturae.23425

Conflict of interest. The authors declare no conflict of interest.
Received January 15, 2025. Revised February 24, 2025. Accepted February 25, 2025.

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2025;29(5):711-721

doi 10.18699/vjgb-25-78

The Forest and Tundra Nenets: differences in Y-chromosome haplogroups

V.N. Kharkov [1], L.V. Valikhova[1], D.S. Adamov[2], A.A. Zarubin [1], I.Yu. Khitrinskaya [1], V.A. Stepanov [1]

[1] Research Institute of Medical Genetics, Tomsk National Research Medical Center of the Russian Academy of Sciences, Tomsk, Russia
[2] Bochkov Research Centre for Medical Genetics, Moscow, Russia

vladimir.kharkov@medgenetics.ru

Abstract. The Forest and Tundra Nenets in different areas of the Yamalo-Nenets Autonomous Okrug were studied using Y-chromosome markers. The results of analyzing the genetic structure of Nenets clans using 44 STR markers of the Y chromosome are presented, taking into account their presence in subethnoses (Tundra and Forest Nenets), as well as to the Kharyuchi ("true Nenets") and Vanuito ("foreigners") phratries. The number of the Nenets ($N = 606$) includes the Tundra ($N = 536$) and Forest ($N = 70$) Nenets. Sublineage N1a2b1b1a~-B170 is specific for the clans in the Kharyuchi phratry, and sublineage N1a2b1b1b-B172, for the clans in the Vanuito phratry. Most Forest Nenets clans have haplogroup N1a2b1-B478. All males of the Pyak clan, which is prevalent in the Forest Nenets, have a specific haplogroup, N1a1a1a1a2a1c1~. The results of the study suggest that the Nenets clan associations typically have a common ancestor in the male line and are characterized by a recent founder effect. Each Nenets clan has its own specific cluster of haplotypes, equidistant from each other. The structure of Y-chromosome haplotypes and haplogroups in the Nenets gene pool includes the Nenets heritage from the Khanty and Enets. Many samples from these sample sets were shown to have rare haplotypes that were absent from the baseline data and to differ significantly from the other haplotypes found in the populations. They belong to various rare branches of the Y-chromosome haplogroups found only in these sample sets. Some samples form haplotype variants that have not been described previously and allow us to characterize the phylogeny of these lineages in more detail. The Forest and Tundra Nenets differ greatly in the composition of haplogroups, which is fully consistent with ethnological and linguistic data on the origin of these populations. The predominant haplogroups are N1a1a1a1a2a1c1~-Y13850, Y13852, Y28540 CTS9108 (xY24219, Y24375) and N1a2b1-B478, Z35080, Z35081, Z35082, Z35083, Z35084 (xB169) in the Forest Nenets, and N1a2b1b1a~-B170 (xZ35104), N1a1a1a1a2a1c~-Y13850, Y13852, Y13138, PH3340 (xY24219, Y24365) and N1a2b1b1b-B172, Z35108 in the Tundra Nenets.

Key words: gene pool; human populations; genetic diversity; Y chromosome; Nenets

For citation: Kharkov V.N., Valikhova L.V., Adamov D.S., Zarubin A.A., Khitrinskaya I.Yu., Stepanov V.A. The Forest and Tundra Nenets: differences in Y-chromosome haplogroups. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov J Genet Breed*. 2025;29(5):711-721. doi 10.18699/vjgb-25-78

Funding. The study was supported by the grant of the Russian Science Foundation No. 22-64-00060, https://rscf.ru/project/22-64-00060/

Особенности генофондов лесных и тундровых ненцев по гаплогруппам Y-хромосомы

В.Н. Харьков [1], Л.В. Валихова[1], Д.С. Адамов[2], А.А. Зарубин [1], И.Ю. Хитринская [1], В.А. Степанов [1]

[1] Научно-исследовательский институт медицинской генетики Томского национального исследовательского медицинского центра Российской академии наук, Томск, Россия
[2] Медико-генетический научный центр им. академика Н.П. Бочкова, Москва, Россия

vladimir.kharkov@medgenetics.ru

Аннотация. Исследованы лесные и тундровые ненцы из различных районов Ямало-Ненецкого автономного округа по маркерам Y-хромосомы. Представлены результаты генетической структуры ненецких родов по 44 STR-маркерам Y-хромосомы с учетом их принадлежности к субэтносам (тундровые и лесные ненцы), а также к фратриям Харючи («настоящие ненцы») и Вануйто («иноплеменники»). Количество ненцев $N = 606$ делится на тундровых ($N = 536$) и лесных ($N = 70$). Для родов, принадлежащих к фратрии Харючи, специфична сублиния N1a2b1b1a~-B170, а для родов фратрии Вануйто – N1a2b1b1b-B172. У большинства родов, относящихся к субэтносу лесных ненцев, присутствует гаплогруппа N1a2b1-B478. Все мужчины рода Пяк, доминирующего у лесных ненцев, принадлежат к специфичной гаплогруппе N1a1a1a1a2a1c1~. Результаты работы

© Kharkov V.N., Valikhova L.V., Adamov D.S., Zarubin A.A., Khitrinskaya I.Yu., Stepanov V.A., 2025
This work is licensed under a Creative Commons Attribution 4.0 License

свидетельствуют о том, что родовые объединения ненцев, как правило, имеют общего предка по мужской линии, для них характерен недавний эффект основателя. Каждый ненецкий род имеет свой специфичный кластер гаплотипов, равноудаленных друг от друга. Структура гаплотипов и гаплогрупп Y-хромосомы в составе ненецкого генофонда демонстрирует наследие ненцев от хантов и энцев. Для многих образцов из этих выборок показаны индивидуальные редкие гаплотипы, которые отсутствуют в базах данных и значительно отличаются от других гаплотипов, обнаруженных в этих популяциях. Они относятся к различным редким ветвям гаплогрупп Y-хромосомы, обнаруженным только в этих выборках. Часть образцов формирует отдельные варианты гаплотипов, которые не были описаны ранее, и позволяет более подробно охарактеризовать филогению этих линий. Лесные и тундровые ненцы сильно различаются по составу гаплогрупп, что полностью соответствует данным этнологов и лингвистов о происхождении этих популяций. У лесных ненцев преобладают гаплогруппа N1a1a1a1a2a1c1~-Y13850, Y13852, Y28540 CTS9108 (xY24219,Y24375) и с небольшой частотой гаплогруппа N1a2b1-B478, Z35080, Z35081, Z35082, Z35083, Z35084 (xB169). У тундровых ненцев доминируют три гаплогруппы, N1a2b1b1a~-B170 (xZ35104), N1a1a1a1a2a1c~-Y13850, Y13852, Y13138, PH3340 (xY24219,Y24365) и N1a2b1b1b-B172, Z35108.

Ключевые слова: генофонд; популяции человека; генетическое разнообразие; Y-хромосома; ненцы

Introduction

The Nenets are an indigenous people of the northern territories of Western Siberia and the Eastern European part of the Urals. According to the All-Russian Census of 2021, there were 49,646 of them. They are divided into European and Siberian. European Nenets live in the Nenets Autonomous Okrug of the Arkhangelsk Region, Siberian Nenets live in the Yamalo-Nenets Autonomous Okrug of the Tyumen Region and in the Dolgano-Nenets Taimyr Municipal District of the Krasnoyarsk Territory. A very small number of Nenets live in the Khanty-Mansi Autonomous Okrug, the Murmansk Region and the Komi Republic. Together with the Enets, Nganasan and Selkup languages, the Nenets language belongs to the Samoyedic group of the Uralic language family. According to anthropological characteristics, the Nenets belong to the Mongoloids.

The Nenets are divided into Forest Nenets (living in the area of the Taz and Pur rivers of the Purovsky district of the Yamal-Nenets Autonomous Okrug) and Tundra Nenets, inhabiting the northern Priobye of the Yamal, Tazovsky and Nadym districts. The number of the Forest Nenets is very small and amounts to approximately 1,500 people. According to available data, the Forest Nenets have retained the archaic features of the Nenets community. The phenotype of the Forest Nenets has a more pronounced Mongoloid component (a less developed beard, the presence of epicanthus, a flatter face, a lower bridge of the nose), but at the same time these features are combined with Caucasian features: light eyes, a raised tip of the nose and the base of the nose. The narrow nose brings the Forest Nenets closer to the Tungus-Manchu peoples of Siberia and the Yukaghirs, typologically similar to each other; based on these data, some anthropologists suggest that it was the aboriginal tribes related to the Yukaghirs that were the predecessors of the Samoyeds in the modern territory inhabited by the Nenets (Alekseeva et al., 1972).

The Samoyedic languages are divided into two groups: Northern Samoyedic (Nenets, Enets, Nganasan) and Southern Samoyedic (Selkup and the languages of the Sayan Highlands) (Maitinskaya, 1966). Within the Nenets language, there are two dialects: the Tundra dialect, spoken by 95 % of the Nenets, and the Forest dialect. The Tundra dialect, in turn, is subdivided into three subdialects: western, eastern, and Bolshezemelsky, which formed the basis of the literary Nenets language (Tereshchenko, 1956; Khomich, 1976). The Forest dialect is used by the Nenets inhabiting the taiga zone. Understanding between Nenets who speak the Tundra and Forest dialects is very scarce, since the two dialects have significant differences in phonetics. Some phonetic features of the Tundra dialect may be related to the preservation of these sounds from the Uralic proto-language or they appeared secondarily, thanks to the language of the aborigines living in the tundra territories. The second option is supported by the fact that similar sounds are also found in the languages of the Chukchi, Koryaks, and Eskimos (UNESCO, 2010).

Such a significant difference in the two Nenets dialects in phonetic, lexical and morphological features can be explained by a closer connection of the Forest Nenets with representatives of the Enets, Nganasan, and also some features of the Forest Nenets dialect bring it closer to the Khanty language. This once again confirms that the Forest Nenets are an earlier autochthonous group of the Nenets.

The clan structure of the Nenets has been well studied by anthropologists and ethnographers (Volzhanina, 2017). Marriages among the Nenets have always been strictly exogamous. Among the Tundra Nenets, clans were united into two phratries. The first of them is Kharyuchi ("real Nenets"), which included all the clans of Samoyed origin. The second phratry is Vanuito ("foreigners"), consisting of clans that go back to the indigenous population of these territories, as well as clans of Enets and Khanty origin included in it (Khomich, 1976).

Differences in the composition and frequencies of Y-chromosome haplogroups between these two phratries are shown in our previous article (Kharkov et al., 2021). It revealed the features of haplogroup frequencies between the two Nenets phratries and the clans of the Tundra Nenets, who originated from the Samoyeds, Enets and Khanty.

The objectives of this article are to increase the samples of Tundra Nenets and compare them with new samples of Forest Nenets by Y haplogroups.

The study of the structure of human population gene pools is one of the key areas of modern genetics. In recent years, there has been a real breakthrough in population genetics associated with the widespread introduction of sequencing methods, both for whole-genome genotyping of samples and for searching for new informative SNP markers in various Y-chromosome haplogroups. This makes it possible to analyze

the differences in the gene pools of the Tundra and Forest Nenets of the Yamal-Nenets Autonomous Area not only at the ethnic and subethnic levels, but also at the clan level. If two people belong to different haplogroups, there can be no relationship between them in the male line. Genetic relationship is determined by the similarity of other indicators – haplotypes with an increase in the number of YSTR for a more detailed specification of the differences between them.

Until recently, the main problem was the lack of informative SNP markers for detailed analysis of the phylogenetic structure and origin of various haplogroups. New basic and terminal SNPs of the Y chromosome are extracted from the data of complete genomes for a more detailed analysis of the division of specific lines of these haplogroups. The number of SNPs discovered in recent years has already reached many thousands. Many SNPs have been confirmed on a limited set of samples and data on the frequencies of the sublineages they determine in real ethnic groups are absent or are very approximate, due to the non-representativeness of the samples studied.

Detailed analysis of haplogroups based on Y-chromosome SNP and STR genotyping is one of the most effective methods for studying the genetic diversity of human populations. It allows for a more accurate reconstruction of the origin of individual sublineages within haplogroups, calculation of their age and founder effects, as well as description of the demographic growth of populations and the phylogeny of specific variants of all Y-chromosome haplogroups. This method provides a higher level of geographic differentiation among Y-chromosome variants compared to mitochondrial DNA (mtDNA) and autosomes. These data can be used to study migration events and the history of ethnic groups (Underhill et al., 2000; Adamov, Fedorova, 2024; Adamov et al., 2024). Y-chromosome DNA markers have shown the highest level of genetic differentiation between populations compared to any other genetic systems.

The aim of this study is a comprehensive analysis of haplogroup frequencies and differences in Y-chromosome haplotypes in the Forest and Tundra Nenets populations. To address the differences between them, the structure of Asian haplogroups N1a1 and N1a2 was determined and YSTR data were obtained to clarify the age and relationships between the different branches of these haplogroups.

Materials and methods

The study material consisted of DNA samples of men from various populations of Tundra (N = 536) and Forest Nenets (N = 70). Over the past few years, we have increased the number of samples of Forest and Tundra Nenets, which included additional small clans that were not included in the population sample published in the previous article (Kharkov et al., 2021). Increasing the number of samples for various population samples of the indigenous population of Russia makes a significant contribution to the study of the specific features of their gene pools according to Y-chromosome haplogroups. Tundra Nenets were collected in the villages of Tazovsky, Antipayuta, Gyda, Samburg, Aksarka, Beloyarsk, Yar-Sale, Syunai-Sale, Nadym. Forest Nenets were collected in the villages of Tarko-Sale and Kharampur. The Kharyuchi phratry includes the Samoyedic clans Ader, Anagurichi, Vora, Vylko, Evai, Lapsui, Nenyan, Okotetto, Susoi, Serotetto, Tadibe, Taleev, Togoi, Tesida, Khudi, Heno, Yadne, Yando, Yaptunai. The Vanuito phratry includes the Samoyedic clans Vanuito, Vengo, Puiko, Yar, Yaptik, Yaungad; the Enets clans Maryik, Okovai, Ter; the Khanty clans Vekho, Nerkagi, Purungui, Salinder, Tibichi. The European Nenets clans include Laptander, Pyryrko, Syadai, Taiberi, which are not part of the Kharyuchi and Vanuito phratries. The Forest Nenets include the main clan Pyak, the Samoyedic clan Vello, Segoi; Enets clans Ayvasedo, Nyach, Ter (Kvashnin, 2011).

The material was obtained during joint scientific and practical medical expeditions from 2019 to 2024 and deposited in the bioresource collection “Biobank of the Population of Northern Eurasia”. Primary biological material (venous blood) was collected from donors in compliance with the written informed consent procedure for the study. A questionnaire was compiled for each donor with his pedigree, ethnicity and places of birth of ancestors. The study included only DNA samples from male donors who, according to the questionnaire, denied the fact of miscegenation on the paternal line with representatives of other ethnic groups in at least three generations.

To study the composition and structure of Y-chromosome haplogroups, two systems of genetic markers were included in the study: diallelic loci represented by SNPs and polyallelic highly variable microsatellites (YSTRs). Using 357 SNP markers, the belonging of men to different haplogroups was determined. Some of them form the main base lines of haplogroups, while the remaining terminal SNPs are present in specific sublines in different related clans.

Genotyping of terminal SNP markers was performed using polymerase chain reaction and subsequent analysis of DNA fragments using RFLP (restriction fragment length polymorphism) analysis. For specific terminal SNPs for individual sublines, genotyping of a small number of samples was carried out based on their YSTR haplotypes and the results of NGS sequencing of the Y chromosome. The designation of haplogroups was given with reference to the ISOGG 2019 Y-DNA Haplogroup Tree.

Analysis of STR haplotypes within haplogroups was performed using 44 STR markers of the non-recombining part of the Y chromosome (DYS19, 385a, 385b, 388, 389I, 389II, 390, 391, 392, 393, 426, 434, 435, 436, 437, 438, 439, 442, 444, 445, 448, 449, 456, 458, 460, 461, 481, 504, 505, 518, 525, 531, 533, 537, 552, 570, 576, 635, 643, YCAIIa, YCAIIb, GATA H4.1, Y-GATA-A10, GGAAT1B07). STR markers were genotyped using capillary electrophoresis on ABI Prism 3730 and Nanofor-05 devices.

The experimental studies were carried out at the Center for Collective Use of Research Equipment “Medical Genomics” (Research Institute of Medical Genetics, Tomsk National Research Medical Center). Construction of median networks of Y-chromosome haplotypes was performed using the Network v.10.2.0.0 program (Fluxus Technology Ltd; www.fluxus-engineering.com) using the Bandelt median network method (Bandelt et al., 1999). The generation age of the observed haplotype diversity in haplogroups was estimated using the ASD method (Zhivotovsky et al., 2004) based on the root-mean-square differences in the number of repeats between all markers. When calculating the age of births for individual haplogroups, single samples that significantly stood out from

the general cluster of haplotypes were excluded. Calculations were performed for birth groups of at least five samples. The generation age was taken to be 30 years, the mutation rate was 0.0033 per locus per generation (Balanovsky, 2017).

The selection of derivative YSNP variants for haplogroup age estimation was performed based on the coordinate of the hg38 reference sequence, which falls within the combBED regions that roughly correspond to X-degenerated euchromatin sequences. The combBED sequence consists of 857 Y-chromosome regions with a total length of 8.47 Mb (Adamov et al., 2015). The age estimation error is calculated based on the assumption of the Poisson nature of the SNP mutation process (Poznik et al., 2013). These SNP positions were extracted from whole-genome sequencing data for 54 Nenets male samples.

Results and discussion

After genotyping of Y-chromosome SNP markers and YSTR markers, a strong difference was shown between the Forest and Tundra Nenets in the composition of haplogroups.

The clans of the Forest Nenets belong to two main haplogroups, N1a1a1a1a2a1c1~-Y13850, Y13852, Y28540 CTS9108 (xY24219,Y24375) and N1a2b1-B478, Z35080, Z35081, Z35082, Z35083, Z35084 (xB169), which are dominant in frequency, and only three samples belonging to haplogroup N1a2b1b1-B170 belong to the Segoi clan. The set of haplogroups is more diverse among the Tundra Nenets. The most frequent haplogroups are N1a2b1b1a~-B170 (xZ35104), N1a1a1a1a2a1c~-Y13850, Y13852, Y13138, PH3340 (xY24219,Y24365) and N1a2b1b1b-B172, Z35108, the remaining haplogroups with lower frequencies are listed in Table 1. Forest and Tundra Nenets do not coincide in the frequencies of Y-chromosome haplogroups and their clans differ greatly in the haplotypes of various sublineages.

Y-chromosome haplogroup N1a2b-P43 (previously designated as N1b and N2), dominant in the Tundra Nenets (indicate the total 69.8 %), is found with uneven distribution among East Asian, Siberian and Eastern European populations. Its highest proportion is present in the populations of Western Siberia among the Nganasans (92 %), Enets (78 %) and Nenets (57 %) (Karafet et al., 2002; Tambets et al., 2004; Derenko et al., 2007; Rootsi et al., 2007; Mirabal et al., 2009; Ilumäe et al., 2016; Kharkov et al., 2021).

Two branches of the Y-chromosome haplogroup N1a1a1a1a2-Z1936 are present in the clans of the Tundra Nenets (Lapthander, Nerkagi, Salinder, Tibichi and Yar) and a parallel branch of this haplogroup is present in the Forest Nenets clan Pyak.

Haplogroup Q1b1a3b1-BZ99 was found only in three men from the Anagurichi clan of the Yamal Nenets. It has a southern Siberian origin, but a very low frequency in most Siberian populations, so it is not yet possible to compare its heritage in the Nenets with the aboriginal or alien Samoyedic population. In the whole-genome data obtained using the Admixture method, there are almost no data on the composition of Y-haplogroups of the ancient aboriginal population of these territories called Sikhirtya. According to ethnographers, the alien Samoyeds took their daughters as wives (Peoples of West Siberia, 2005). The number of men of this aboriginal population has greatly decreased due to clashes with the ancestors of the Nenets and the spread of various infectious diseases over the past several hundred years. They were hunters similar to the Yukaghirs, and their numbers did not increase from generation to generation, because they did not master reindeer herding (Khomich, 1970).

The population samples of the Gydan and Yamal Nenets completely coincide in the composition of haplogroups, but differ significantly in their frequency (Table 2). Among the Taz Nenets, the maximum frequency is characteristic of N1a2b1b1-B170, which corresponds to the proportion of men belonging to the Kharyuchi phratry. Among the Yamal Nenets, more than a third in frequency is occupied bY haplogroup N1a1a1a1a2-Z1936, to which belong almost all representatives of the Laptander clan and part of the Yar clan, who are descendants of the European Tysiya Nenets (Kvashnin, 2011).

Haplogroup N1a2b1b1-B170. This haplogroup is the most frequent among the Tundra Nenets. The Nenets are characterized by its division into specific haplotype variants, which almost completely coincide with their division into clans. It completely dominates in the Kharyuchi phratry, which includes the clans Ader, Evai (Yavai), Lapsui, Nenyan,

Table 1. Frequency of occurrence of Y-chromosome haplogroups in Tundra and Forest Nenets

Haplogroups	Nenets,	
	Tundra	Forest
N1a2b1-B478, Z35080, Z35081, Z35082, Z35083, Z35084 (xB169)	6.8 % (36)	18.6 % (13)
N1a2b1b1a~-B170 (xZ35104)	45.8 % (243)	4.2 % (3)
N1a2b1b1b-B172, Z35108	14.1 % (75)	–
N1a2b2a1-VL97, Y3185 (xZ35049,Z35070)	3.1 % (16)	–
N1a1a1a1a2-Z1936 (xL1034, CTS9925, PF967.2)	10.7 % (57)	–
N1a1a1a1a2a1c~-Y13850, Y13852, Y13138, PH3340 (xY24219,Y24365)	18.8 % (100)	–
N1a1a1a1a2a1c1~-Y13850, Y13852, Y28540 CTS9108 (xY24219,Y24375)	–	77.2 % (54)
Q1b1a3b1-BZ99 (xB30)	0.7 % (4)	–

Table 2. Distribution of Y-chromosome haplogroups among Tundra and Forest Nenets by clans

Clan	Haplogroups								
	N1a2b1-B478, Z35080, Z35081, Z35082, Z35083, Z35084 (xB169)	N1a2b1b1a~ -B170 (xZ35104)	N1a2b1b1b-B172, Z35108	N1a2b2a1-VL97, Y3185 (xZ35049, Z35070)	N1a1a1a1a2-Z1936 (xL1034, CTS9925, PF967.2)	N1a1a1a1a2a1c~ -Y13850, Y13852, Y13138, PH3340 (xY24219, Y24365)	N1a1a1a1a2a1c1~ -Y13850, Y13852, Y28540 CTS9108 (xY24219, Y24375)	Q1b1a3b1-BZ99 (xB30)	Total number
Ader	–	6 ~440 years (SD = 172)	–	–	1	–	–	–	7
Agichev	–	–	–	–	–	–	1	–	1
Aivasedo	10 ~520 years (SD = 159)	1	–	–	–	–	1	–	12
Anagurichi	–	–	–	–	–	–	–	3	3
Vanuito	–	–	58 ~110 years (SD = 19)	–	–	–	–	–	58
Vora	3	–	–	–	–	2	–	–	5
Vylkoo	4	–	–	–	–	–	–	–	4
Vello	–	6 ~460 years (SD = 164)	2	–	–	–	–	–	8
Vengo	–	1	1	5	–	–	–	–	7
Veho	–	–	–	1	–	–	–	–	1
Evay Yavay	–	7 ~420 years (SD = 135)	–	–	–	–	–	–	7
Lapsui	–	55 ~310 years (SD = 140)	–	–	–	2	–	–	57
Laptander	–	–	–	–	31 ~110 yearsт (SD = 41)	–	–	–	31
Ledkov	–	–	–	–	1	–	–	–	1
Marik	6 ~230 years (SD = 115)	–	–	–	–	–	–	–	6
Nenyang	–	4	–	–	–	–	–	–	4
Nerkagi	–	–	–	–	–	10 ~280 years (SD = 117)	–	–	10
Nyach	8 ~400 years (SD = 210)	1	–	–	–	1	–	–	10
Okovay	2	–	–	–	–	–	–	–	2
Okotetto	–	4	1	–	–	–	–	–	5
Puyko	–	–	4	–	–	–	–	–	4
Purungui	–	–	6 ~140 years (SD = 74)	–	–	–	–	–	5
Pyryrko	1	–	–	–	–	–	–	–	1
Pyak	–	–	–	–	–	–	52 ~560 years (SD = 183)	–	52

Table 2 (end)

Clan	Haplogroups								
	N1a2b1-B478, Z35080, Z35081, Z35082, Z35083, Z35084 (xB169)	N1a2b1b1a~ -B170 (xZ35104)	N1a2b1b1b-B172, Z35108	N1a2b2a1-VL97, Y3185 (xZ35049, Z35070)	N1a1a1a1a2-Z1936 (xL1034, CTS9925, PF967.2)	N1a1a1a1a2a1c~ -Y13850, Y13852, Y13138, PH3340 (xY24219, Y24365)	N1a1a1a1a2a1c1~ -Y13850, Y13852, Y28540 CTS9108 (xY24219, Y24375)	Q1b1a3b1-BZ99 (xB30)	Total number
Salinder	–	–	–	–	–	56 ~600 years (SD = 230)	–	–	56
Segoy	–	6 ~450 years (SD = 189)	–	–	–	–	–	–	6
Susoi	–	7 ~80 years (SD = 45)	–	–	–	–	–	–	7
Serotetto	–	6 ~110 years (SD = 84)	–	–	–	–	–	–	6
Syadai	2	–	–	–	–	–	–	–	2
Tadibe	–	2	–	–	–	–	–	–	2
Taiberi	–	–	–	1	4	2	–	–	7
Taleev	2	–	–	–	2	–	–	–	4
Tyor	5	3	–	–	–	–	–	–	8
Tibichi	–	3	–	–	–	14 ~160 years (SD = 71)	–	–	17
Togoy	–	4	–	–	–	3	–	–	7
Togolya	–	–	–	–	1	–	–	–	1
Tedido	–	–	–	–	–	1	–	–	1
Tesida	–	21 ~70 years (SD = 41)	–	–	–	–	–	–	21
Hardyu	1	–	–	–	–	–	–	–	1
Khudi	1	8 ~90 years (SD = 48)	–	–	–	–	–	–	9
Heno	1	–	–	–	–	1	–	–	2
Yadne	–	53 ~160 yearsт (SD = 85)	–	1	–	2	–	–	56
Yando	–-	19 ~180 years (SD = 86)	–	–	–	–	–	–	19
Yaptik	3	1	–	–	–	–	–	–	4
Yaptunay	–	25 ~230 years (SD = 92)	–	–	–	–	–	–	25
Yar	–	11 ~160 years (SD = 94)	2	13 ~520 years (SD = 111)	17 ~581 years (SD = 240)	6 ~33 years (SD = 35)	–	–	49
Yaungat	–	–	1	–	–	–	–	–	1

Segoi, Susoi, Serotetto, Tesida, Khudi, Yadne, Yando, Yaptunai (Table 2). Its age, estimated by us based on YSTR markers, is ~500 years (SD = 111). The age of this lineage, according to YSNP sequencing data from whole-genome data of Nenets samples, is ~1,030 years; in earlier studies, it was determined as ~1,650 years (Ilumäe et al., 2016). This haplogroup was also found in several men of the Vello clan, belonging to the Forest Nenets, which is not formally part of the Kharyuchi phratry. Another clan that is not part of this phratry, but belongs to this sublineage is Yar (part of the Vanuito phratry).

The median network N1a2b1b1a-B170, which includes men of the Kharyuchi phratry, has a typical star-shaped structure, with a clearly distinguished central haplotype, from which all the others originate. The inclusion of almost all clans of the phratry in this haplogroup confirms the theory of the community and unity of origin of the Tundra Nenets from a common male ancestor along the paternal lines. The age of this haplogroup, estimated by SNP markers (~1,030 years) does not quite coincide with the approximate time of the main migration of the Samoyedic tribes from the territory of the Kulai culture to the far North (Peoples of West Siberia, 2005). Their migration began in the 3rd-2nd centuries BC, and the common ancestor of all men of the Kharyuchi phratry along the male line was formed about a thousand years ago.

The age calculation of the Nenets clans that belong to this haplogroup according to YSTR is given in Table 2. The time of the common ancestor for each individual clan approximately coincides with the data of ethnographers. The Ader clan (~440 years) is one of the most ancient clans among the Nenets of the Kharyuchi phratry, and the ancestors of this clan were among the first to develop the territory to the east of Yamal. It was first recorded in the yasak book of 1695 (Kvashnin, 2001). The Lapsui clan (~310 years) separated from the third Kharyuchi wave in the mid-19th century, and the Susoi clan (~80 years), in turn, separated from the Okotetto clan in the 1920s (Kvashnin, 2001). The age of the second largest Yadne clan is ~160 years. The Figures 1 and 2 show the median networks of the clans with the largest number of men: Lapsui $N = 55$ (Fig. 1) and Yadne $N = 53$ (Fig. 2).

Haplogroup N1a2b1b1b-B172, Z35108. Another parallel line, N1a2b1-B172, is characteristic of the Vanuito phratry of the Tundra Nenets. Its carriers belong to the Vanuito, Vello, Purunguy, Puiko, Yar and Yaungat clans. The formation time of this more ancient Nenets branch according to YSNP is ~1,420 years. The age of these clans is also indicated in Table 2. The total age of the Purunguy, Puiko, Vello and Yar clans is ~240 years (SD = 71). Just like for the Kharyuchi phratry, the age of these clans coincides with the time of their origin. These clans were first recorded in the materials of the Surgut Church in 1880, and the Vello clan, in 1860 (Kvashnin, 2001). This lineage, present in the Nenets, was inherited by them from the Khanty ancestors in the male line. Almost each of these Nenets clans is characterized by the complete dominance of a specific subbranch N1a2b1b1b-B172, Z35108 with terminal SNPs and with a specific haplotype spectrum emphasizing the recent founder effect. In the north, the Khanty came into contact with the Nenets, some of them were assimilated by them, which is confirmed by ethnographic data, as well as our study of the clan structure of the Gydan Nenets using Y-chromosome markers (Kharkov et al., 2021).

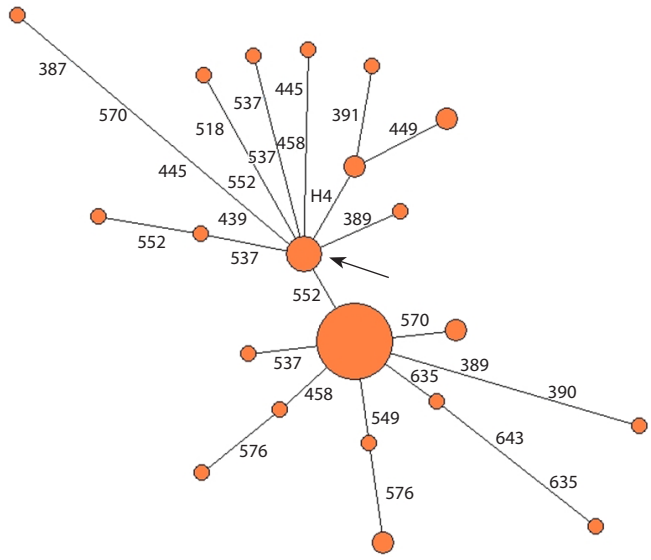


Fig. 1. Median network of YSTR haplotypes of haplogroup N1a2b1b1a-B170 in the Lapsui clan.

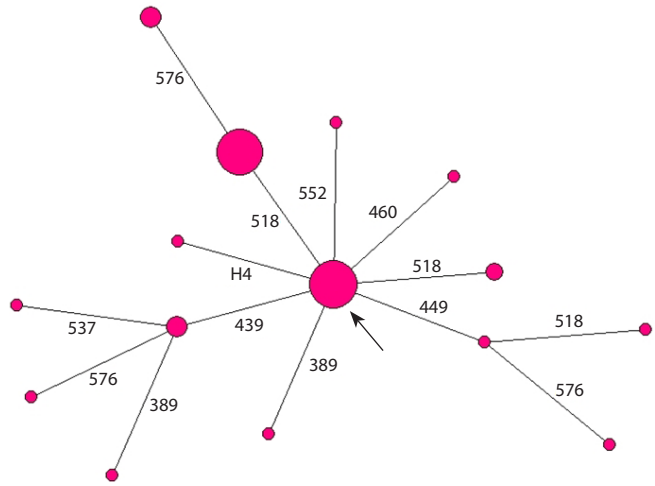


Fig. 2. Median network of YSTR haplotypes of haplogroup N1a2b1b1a-B170 in the Yadne clan.

The median network of this haplogroup (Fig. 3) has a clearly visible predominant haplotype, but the network itself consists of two clusters and, unlike the median network of the Kharyuchi phratry, is more fragmented. This difference between the two phratries is easily explained by the origin of these phratries. Vanuito developed on a local aboriginal substrate and also includes clans of Enets and Khanty origin, unlike Kharyuchi, who are carriers of the South Siberian Samoyed component (Prokofiev, 1940). The high frequency of this main haplotype of the median network in many men proves the strong demographic growth of these clans over the past two hundred years with a several-fold increase in the number of sons from generation to generation.

Haplogroup N1a2b1-B478, Z35080, Z35081, Z35082, Z35083, Z35084 (xB169). This line is specific to the Enets clans that became part of the Nenets. It is included in the Tun-

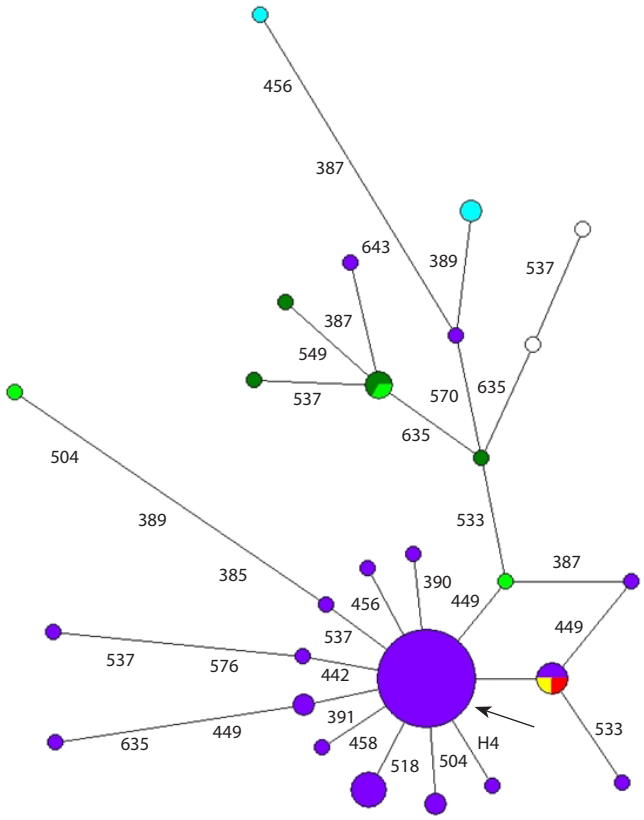


Fig. 3. Median network of YSTR haplotypes of haplogroup N1a2b1b1a-B170 in the Vanuito, Purungui, Yar, Puiko, Wello, Okotetto and Yaungat clans.

The genus Vanuito is shown in lilac, the genus Purungui in dark green, the genus Yar in light green, the genus Puiko in blue, the genus Wello in white, the genus Okotetto in red, and the genus Yaungat in yellow.

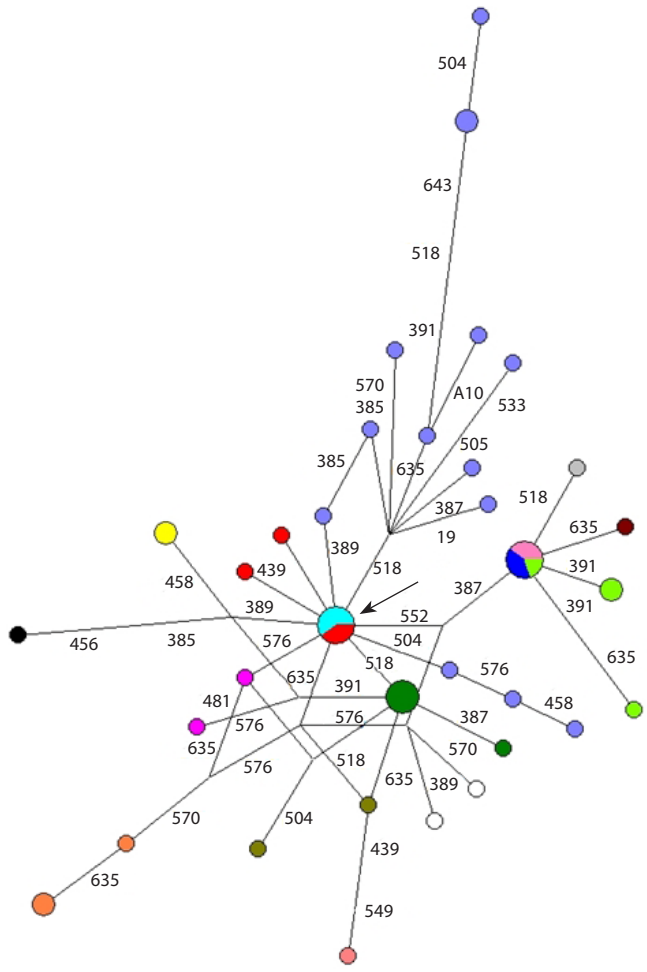


Fig. 4. Median network of YSTR haplotypes of haplogroup N1a2b1-B478.

The Aivasedo genus is shown in light blue, the Nyach genus in dark green, the Vylko genus in light green, the Syadai genus in blue, the Taleev genus in pink, the Maryik genus in light blue, the Yaptik genus in red, the Okovai genus in yellow, the Khudi genus in brown, the Hardyu genus in black, the Pyryrko genus in grey, and the Heno genus in white.

dra and Forest Nenets with a frequency of 6.8 % and 18.6 %, respectively (Table 1). In particular, N1a2b1-B478 was found in the Maryik, Nyach, Okovai and Ter clans – Enets clans that became part of the Tundra Nenets. This haplogroup is also present in the Forest Nenets clans Ayvasedo, Nyach and Ter, which are a branch of the Enets clan Muggadi and belong to the Vanuito phratry. The total age of haplogroup N1a2b1-B478 according to YSTR was ~680 years (SD = 202), according to YSNP data of their complete genomes, it is much older than other haplogroups of the Nenets – ~3,430 years.

The YSTR data apparently determine the age of their separation from a common ancestor in the male line. The SNP data determine the time of occurrence of these mutations several thousand years before the demographic growth of the Samoyedic peoples in the tundra territory. In the median network constructed for these clans (Fig. 4), one can notice the absence of a dominant haplotype, as was typical for the Kharyuchi and Vanuito phratry. The demographic growth of the clans belonging to this haplogroup from generation to generation is much smaller, compared to other clans of the Kharyuchi and Vanuito phratry. Most likely, this was due to military clashes between the Enets and Nenets, the small number of descendants of the Enets clans, and their mortality from infectious diseases.

The division of this median network into small clusters may indicate heterogeneity of the origin of the Forest Nenets group. For example, the Aivasedo clan was formed by including components of the Tundra Nenets (the Chor clan) during their interaction with the Forest Enets (Vasiliev, 1973). The origin of the Nyach clan is still debated among ethnographers (Verbov, 1939; Dolgikh, 1960). According to the first version, this clan is one of the fundamental clans of the Forest Nenets of unknown origin, according to another version, it is a clan not of the Forest, but of the Tundra Nenets, which is known in the tundra under the name Nyats. All these Enets clans were assimilated by the Nenets and became part of the Vanuito phratry and the Forest Nenets. They are descendants of the local indigenous population of these territories, belonging to the Samoyedic group of languages, who mastered these areas long before the arrival of the Kharyuchi phratry, who are descendants of the later second wave of Samoyeds from Southern Siberia (Vasiliev 1975).

Haplogroup N1a1a1a1a2a1c1~-Y13850, Y13852, Y28540 CTS9108 (xY24219,Y24375). This haplogroup is completely dominant among the Forest Nenets, to which all men of the Pyak clan belong (Fig. 5). According to YSNP data, in comparison with the calculation of a common ancestor ac-

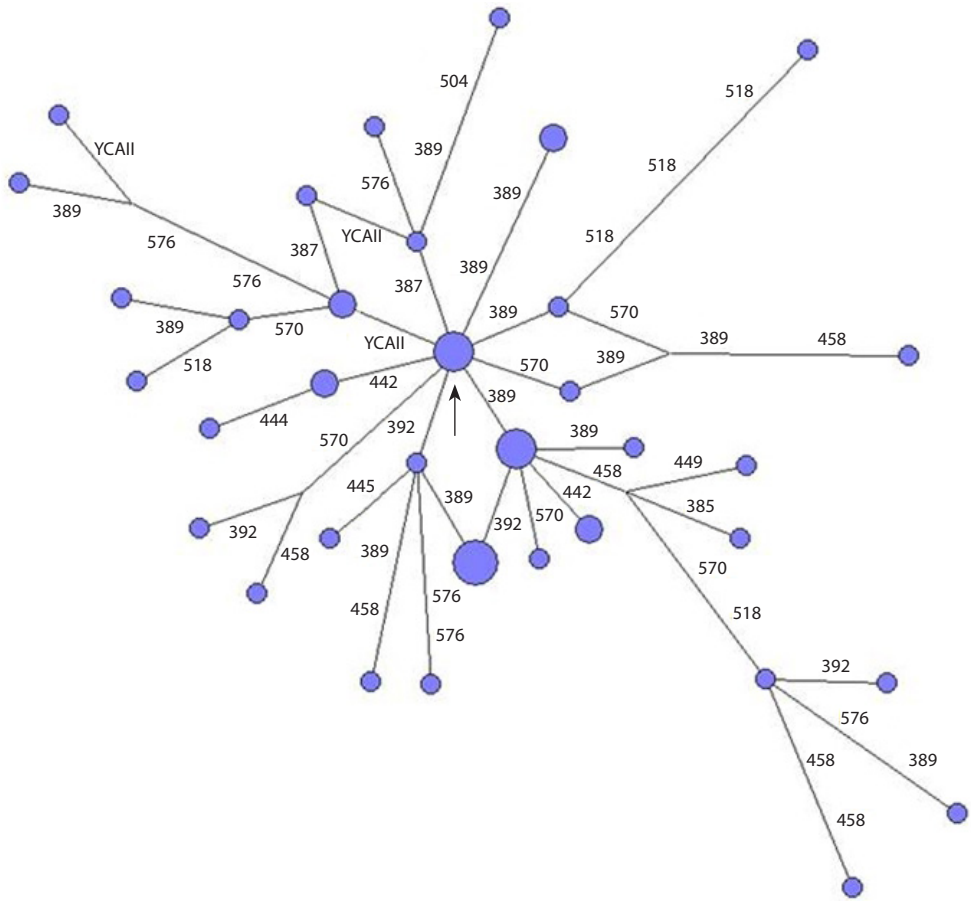


Fig. 5. Median network of YSTR haplotypes of haplogroup N1a1a1a1a2a1c1~-Y13850 of the Pyak clan.

cording to YSTR (~560 years), its age turned out to be much more ancient – ~2,680 years. There is no clearly distinguishable dominant haplotype in the median network. The Pyak clan is one of the oldest clans, first mentioned in 1627, which consists of many divisions (Vasiliev, 1973). This haplogroup was not found among the Tundra Nenets.

Haplogroup N1a1a1a1a2a1c~-Y13850, Y13852, Y13138, PH3340 (xY24219,Y24365). This haplogroup is sister to the haplogroup of the Pyak clan. Its age according to YSTR is ~1,405 years (SD = 374). It is found in the clans of the Tundra Nenets Salinder, Tibichi, Nerkagi, Taiberi, Vora, Yar, Lar, Togoy, Lapsui.

Haplogroup N1a2b2a1-VL97, Y3185 (xZ35049,Z35070). This is a clade of the European branch of N1a2b, which was possibly inherited by the Forest Nenets from the Khanty or Komi. A description of this line is given in our article on the Khanty (Kharkov, 2023). It is present in the Yar and Vengo clans with a low frequency. The total age of this haplogroup in the Nenets was ~1,110 years (SD = 185).

Each of the above-mentioned clans is characterized by the presence of a dominant haplotype, as well as a pronounced genetic proximity to individuals belonging to it, which indicates a founder effect for each individual clan and a biological relationship of men of each clan along the male line. All samples belonging to different sublineages have a pronounced founder effect. They are characterized by a star-shaped phylogeny of median networks, forming separate groups by clans. Their male ancestors separated from the southern Siberian populations quite a long time ago and went north into the tundra territory.

The results of genetic analysis prove that all Nenets clans are, first of all, a union of relatives on the paternal line. These data indicate a close relationship between all clans belonging to one of the phratries. With isolated exceptions, each phratry has its own specific cluster of haplotypes, equidistant from each other. Most clans of Forest and Tundra Nenets are characterized by the complete dominance of one haplogroup with a specific spectrum of haplotypes, emphasizing the recent founder effect. Two Nenets phratries differ significantly in the genetic structure of clans by Y-chromosome markers, which confirms their formation on the basis of different ancestral components from the autochthonous and migrant population. All samples belonging to different sublineages show a pronounced founder effect.

The fundamental differences between the Forest and Tundra Nenets in Y-chromosome haplogroups are shown. Three ethnic groups took part in the ethnogenesis of the Nenets: the Samoyedic group, the local aboriginal group, and clans of Khanty origin. As a result of their interaction, according to B.O. Dolgikh (1960), two phratries of the Nenets were formed. One goes back to the Samoyeds (Kharyuchi), and the other to the aborigines (Vanuito). Our results on the

distribution of various Y-chromosome haplogroups in the Nenets clans are completely consistent with these data. The clan structure of the Forest and Tundra Nenets within their subpopulations was revealed.

Conclusion

The Nenets clans of Khanty and Enets origin are completely different from each other and from the Samoyedic Kharyuchi phratry by Y-chromosome haplogroup composition. The age of YSTR haplotypes for Y-chromosome haplogroups coincides with the beginning of the demographic growth of mixed populations. The data obtained in this work supplement the information on the differences between Forest and Tundra Nenets. They are in good agreement with the accumulated body of knowledge from other disciplines studying Siberian populations: linguistics, areology, anthropology. This makes it possible to describe in more detail the history of the formation of the gene pool of Forest and Tundra Nenets. In the course of the work, an expanded set of YSTR markers was selected and used, which made it possible to move to a fundamentally new level of detail in the molecular phylogenetic structure of haplogroups and differentiation of samples.

References

Adamov D.S., Fedorova S.A. Y-chromosomal STRs mutation rates in the Yakut population. *Russ J Genet*. 2024;60(4):533-542. doi 10.1134/S1022795424040021

Adamov D.S., Alekseev A.N., Fedorova S.A. Dating the time to the most recent common ancestor of the Sakha (Yakuts) with Y chromosomal haplogroup N3A2-M1982: new ethnogenetic reconstructions. *Yakutskiy Meditcinskiy Zhurnal = Yakut Med J*. 2024;2:88-93. doi 10.25789/YMJ.2024.86.21 (in Russian)

Adamov D., Guryanov V., Karzhavin S., Tagankin V., Urasin V. Defining a new rate constant for Y-chromosome SNPs based on full sequencing data. *Russ J Genet Geneal*. 2015;7(1):68-89

Alekseeva T.I., Volkov-Dubrovin V.P., Golubchikova Z.A., Pavlovsky O.M., Smirnova N.S., Spitsyn V.A. Anthropological study of the Forest Nenets (morphology, physiology and population genetics). Part II. *Voprosy Antropologii = Herald of Anthropology*. 1972;42: 31-54 (in Russian)

Balanovsky O. Toward a consensus on SNP and STR mutation rates on the human Y-chromosome. *Hum Genet.* 2017;136(5):575-590. doi 10.1007/s00439-017-1805-8

Bandelt H.J., Forster P., Röhl A. Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol.* 1999;16(1):37-48. doi 10.1093/oxfordjournals.molbev.a026036

Derenko M., Malyarchuk B., Denisova G., Wozniak M., Grzybowski T., Dambueva I., Zakharov I. Y-chromosome haplogroup N dispersals from south Siberia to Europe. *J Hum Genet.* 2007;52(9): 763-770. doi 10.1007/s10038-007-0179-5

Dolgikh B.O. The Tomsk Uezd. In: Dolgikh B.O. The Clanish and Tribal Composition of the Population of Siberia in the 17th Century. Moscow: USSR Academy of Sciences Publ, 1960;94-104 (in Russian)

Ilumäe A.-M., Reidla M., Chukhryaeva M., Järve M., Post H., Karmin M., Saag L., … Balanovsky O., Kivisild T., Underhill P.A., Villems R., Rootsi S. Human Y chromosome haplogroup N: a non-trivial time-resolved phylogeography that cuts across language families. *Am J Hum Genet.* 2016;99(1):163-173. doi 10.1016/j.ajhg.2016.05.025

Karafet T.M., Osipova L.P., Gubina M.A., Posukh O.L., Stephen L., Zegura S.L., Hammer M.F. High levels of Y-chromosome differentiation among native Siberian populations and the genetic signature of a boreal hunter-gatherer way of life. *Hum Biol*. 2002;74(6):761-789. doi 10.1353/hub.2003.0006

Kharkov V.N., Valikhova L.V., Yakovleva E.L., Serebrova V.N., Kolesnikov N.A., Petelina T.I., Khitrinskaya I.Yu., Stepanov V.A. Reconstruction of the origin of the Gydan Nenets based on genetic analysis of their tribal structure using a new set of YSTR markers. *Russ J Genet*. 2021;57(12):1414-1423. doi 10.1134/S102279 5421120061

Kharkov V.N., Kolesnikov N.A., Valikhova L.V., Zarubin A.A., Svarovskaya M.G., Marusin A.V., Khitrinskaya I.Yu., Stepanov V.A. Relationship of the gene pool of the Khants with the peoples of Western Siberia, Cis-Urals and the Altai-Sayan Region according to the data on the polymorphism of autosomic locus and the Y-chromosome. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov J Genet Breed*. 2023;27(1):46-54. doi 10.18699/VJGB-23-07

Khomich L.V. Nenets legends about Sihirtya. In: Folklore and Ethnography. Leningrad, 1970;59-69 (in Russian)

Khomich L.V. Issues of the Ethnogenesis and Ethnic History of the Nenets. Leningrad: Nauka Publ., 1976 (in Russian]

Kvashnin Yu.N. Basic elements of the clan structure of the Gydan Nenets. *Vestnik Arkheologii, Antropologii i Etnografii = Herald of Archeology, Anthropology, and Ethnology*. 2001;3:199-204 (in Russian)

Kvashnin Yu.N. Tribal names in Nenets toponomy. *Anthropologicheskij Forum = Forum for Anthropology and Culture.* 2011;S14: 23-66 (in Russian)

Maitinskaya K.E. Finno-Ugric languages. Introduction. In: Languages of the Peoples of the USSR. Vol. 3. Moscow: Nauka Publ., 1966;9-25 (in Russian)

Mirabal S., Regueiro M., Cadenas A.M., Cavalli-Sforza L.L., Underhill P.A., Verbenko D.A., Limborska S.A., Herrera R.J. Y-chromosome distribution within geo-linguistic landscape of northwestern Russia. *Eur J Hum Genet.* 2009;17:1260-1273. doi 10.1038/ejhg.2009.6

Peoples of West Siberia: Khanty, Mansi, Selkups, Nenets, Enets, Nganasans, Kets. Moscow: Nauka Publ., 2005 (in Russian)

Poznik D., Henn B., Yee M.-C., Sliwerska E., Euskirchen G., Lin A., Snyder M., Quintana-Murci L., Kidd J.M., Underhill P.A., Bustamante C.D. Sequencing Y chromosomes resolves discrepancy in time to common ancestor of males versus females. *Science.* 2013; 341(6145):562-565. doi 10.1126/science.1237619

Prokofiev G.N. Ethnogenesis of the peoples of the Ob-Yenisei region (Nenets, Nganasans, Enets, Selkups, Kets, Khanty, and Mansi). *Sovetskaya Etnografiya = Soviet Ethnology.* 1940;3:67-76 (in Russian)

Rootsi S., Zhivotovsky L.A., Baldovic M., Kayser M., Kutuev I.A., Khusainova R., Bermisheva M.A., … Ferak V., Parik J., Kivisild T., Underhill P.A., Villems R. A counter-clockwise northern route of the Y-chromosome haplogroup N from Southeast Asia towards Europe. *Eur J Hum Genet*. 2007;15(2):204-211. doi 10.1038/sj.ejhg.5201748

Tambets K., Rootsi S., Kivisild T., Help H., Serk P., Loogväli E.-L., Tolk H.-V., … Ferák V., Füredi S., Komel R., Beckman L., Villems R. The western and eastern roots of the Saami – the story of genetic "outliers" told by mitochondrial DNA and Y chromosomes. *Am J Hum Genet*. 2004;74(4):661-682. doi 10.1086/383203

Tereshchenko N.M. Materials and Research on the Nenets Language. Moscow, 1956 (in Russian)

UNESCO. Atlas of the World's Languages in Danger. 2010. Available: https://unesdoc.unesco.org/ark:/48223/pf0000187026

Underhill P.A., Shen P., Lin A.A., Passarino G., Yang W.H., Kauffman E., Bonné-Tamir B., … Piazza A., Davis R.W., Feldman M.W., Cavalli-Sforza L.L., Oefneret P.J. Y chromosome sequence variation and the history of human populations. *Nat Genet*. 2000;26(3):358-361. doi 10.1038/81685

Vasiliev V.I. On the genetic nature of the ethnic components of the Forest Nenets. *Sovetskaya Etnografiya = Soviet Ethnology.* 1973;4: 106-112 (in Russian)

Vasiliev V.I. The formation of the Yenisei Nenets: on the ethnic nature of ethnographic groups in modern peoples of the North). In: Ethnogenesis and Ethnic History of the Peoples of the North. Moscow: Nauka Publ., 1975;111-147 (in Russian)

Verbov G.D. Forest Nenets. *Sovetskaya Etnografiya = Soviet Ethnology.* 1939;2:43-65 (in Russian)

Volzhanina E.A. Sociodemographic aspects of the Yamal Nenets people according to all-Russia censuses of the 21st century. *Vestnik Arkheologii, Antropologii i Etnografii = Herald of Archeology, Anthropology, and Ethnology*. 2017;3(38):120-130. doi 10.20874/2071-0437-2017-38-3-120-130 (in Russian)

Zhivotovsky L.A., Underhill P.A., Cinnioglu C., Kayser M., Morar B., Kivisild T., Scozzari R., … Yong K.K., Gresham D., Tournev I., Feldman M.W., Kalaydjieva L. On the effective mutation rate at Y-chromosome STRs with application to human population divergence time. *Am J Hum Genet.* 2004;74(1):50-61. doi 10.1086/380911

Conflict of interest. The authors declare no conflict of interest.
Received February 14, 2025. Revised June 11, 2025. Accepted June 12, 2025.

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2025;29(5):722-731

doi 10.18699/vjgb-25-79

Mitogenomic analysis of a representative of the Chernyakhov culture in the Middle Dniester and their genetic relationship with the Slavs in the context of paleoanthropological data

E.V. Rozhdestvenskikh [1], T.V. Andreeva [1, 2, 3] ✉, A.B. Malyarchuk [2, 3], I.Yu. Adrianova [3], D.S. Khodyreva [4], A.A. Evteev [4], A.P. Buzhilova [4], E.I. Rogaev [1, 5] ✉

[1] Research Center for Genetics and Life Sciences, Sirius University of Science and Technology, Sirius Federal Territory, Krasnodar region, Russia
[2] Centre of Genetics and Genetic Technologies, Faculty of Biology, Lomonosov Moscow State University, Moscow, Russia
[3] Vavilov Institute of General Genetics of the Russian Academy of Sciences, Moscow, Russia
[4] Research Institute and Museum of Anthropology, Lomonosov Moscow State University, Moscow, Russia
[5] Department of Psychiatry, UMass Chan Medical School, Shrewsbury, MA, USA
✉ andreeva@rogaevlab.ru; evivrecc@gmail.com

Abstract. Occupying a fairly extensive territory within the East European Plain, representatives of the Chernyakhov culture interacted with many synchronous tribes of other cultures inhabiting neighbouring regions. The question of a possible Proto-Slavic component in the population of the Chernyakhov culture is a subject of many years of discussion, but there is still no evidence for the genetic contribution of representatives of this culture to the gene pool of the Slavs in the subsequent historical period. In this study, we present the results of the craniological and genetic analysis of an individual from the Krynichki burial ground, presumably belonging to the Slavic part of the population of the Chernyakhov culture. A craniometric comparative analysis was conducted for several series of skulls of the East Slavs and representatives of the Chernyakhov culture. The comparison of intragroup variability in the groups of the two cultures showed marked differences between them in the first three principal components. At the same time, the East Slavic and Chernyakhov cultures have similar levels of craniological variability. Differences between female specimens are not so pronounced as those of males'. Based on the analysis of whole-genome sequencing data, the individual from the Krynichki was identified as being a female. The complete sequence of mitochondrial DNA, which belongs to the haplogroup H5a1a1, was reconstructed. For this mitochondrial lineage, a phylogenetic relationship was revealed with eight specimens from publicly available genomic databases, five of which belong to representatives of the present-day West and East Slavic populations. Furthermore, we revealed a mitochondrial sequence identical to that from our previous research on an individual from a medieval burial site located in the modern Vologda region, which is thought to have Slavic ancestry. The complete match between the medieval individual's mtDNA sequence and that of a representative of the Chernyakhov culture points to their likely maternal ancestry. Thus, a possible continuity between representatives of the Chernyakhov culture (3rd century AD) and the population of Ancient Rus' (the second half of the 12th–early 13th centuries AD) has for the first time been shown, as genomic data suggest.

Key words: Chernyakhov culture; Slavs; ancient DNA; mitochondrial DNA; H5a1a1; craniology; phylogeographic analysis

For citation: Rozhdestvenskikh E.V., Andreeva T.V., Malyarchuk A.B., Adrianova I.Yu., Khodyreva D.S., Evteev A.A., Buzhilova A.P., Rogaev E.I. Mitogenomic analysis of a representative of the Chernyakhov culture in the Middle Dniester and their genetic relationship with the Slavs in the context of paleoanthropological data. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov J Genet Breed*. 2025;29(5):722-731. doi 10.18699/vjgb-25-79

Funding. This work was supported by the grant of the state programme of the "Sirius" Federal Territory "Scientific and technological development of the "Sirius" Federal Territory", Agreement No. 18–03 date 10.09.2024 (TVA).

Митогеномный анализ представителя черняховской культуры в Среднем Поднестровье и его генетическая связь со славянами в контексте палеоантропологических данных

Е.В. Рождественских [1], Т.В. Андреева [1, 2, 3] ✉, А.Б. Малярчук [2, 3], И.Ю. Адрианова [3], Д.С. Ходырева [4], А.А. Евтеев [4], А.П. Бужилова [4], Е.И. Рогаев [1, 5] ✉

[1] Научный центр генетики и наук о жизни, Научно-технологический университет «Сириус», федеральная территория «Сириус», Краснодарский край, Россия
[2] Центр генетики и генетических технологий, биологический факультет, Московский государственный университет им. М.В. Ломоносова, Москва, Россия

© Rozhdestvenskikh E.V., Andreeva T.V., Malyarchuk A.B., Adrianova I.Yu., Khodyreva D.S., Evteev A.A., Buzhilova A.P., Rogaev E.I., 2025
This work is licensed under a Creative Commons Attribution 4.0 License

[3] Институт общей генетики им. Н.И. Вавилова Российской академии наук, Москва, Россия
[4] Научно-исследовательский институт и Музей антропологии им. Д.Н. Анучина Московского государственного университета им. М.В. Ломоносова, Москва, Россия
[5] Медицинская школа Чан Массачусетского университета, департамент психиатрии, Шрусбери, США
andreeva@rogaevlab.ru; evivrecc@gmail.com

Аннотация. Занимая довольно обширную территорию в пределах Русской равнины, представители черняховской культуры взаимодействовали со многими синхронными племенами, принадлежащими к другим культурным группам и населявшими смежные регионы. Предметом многолетних дискуссий является вопрос о возможном присутствии протославянского компонента в структуре населения черняховской культуры, однако до сих пор отсутствует доказательная база генетического вклада представителей данной культуры в генофонд славян в последующем историческом периоде. В статье представлены результаты краниологического и генетического анализа индивида из могильника Кринички, предположительно, относящегося к славянской части населения черняховской культуры. Проведен краниометрический сравнительный анализ нескольких серий черепов восточных славян и представителей черняховской культуры. Сопоставление внутригрупповой изменчивости в группах двух культур показало заметные различия между ними по трем первым главным компонентам. В то же время диапазон краниологической изменчивости восточных славян и носителей черняховской культуры сопоставим. Различия в женских выборках не так резко выражены, как в мужских. На основе анализа данных полногеномного секвенирования определен женский пол исследуемого индивида. Реконструирована полная последовательность митохондриальной ДНК, которая относится к гаплогруппе H5a1a1. Для этой митохондриальной линии выявлена филогенетическая связь с восемью образцами, найденными в открытых базах данных геномных последовательностей, пять из которых относятся к представителям современных западно- и восточнославянских популяций. Более того, обнаружена идентичная митохондриальная последовательность, принадлежащая исследованному нами ранее индивиду из средневекового могильника на территории современной Вологодской области, для которого рассматривается славянское происхождение. Полное совпадение последовательности мтДНК представителя черняховской культуры и этого средневекового индивида предполагает их вероятное родство по материнской линии. Таким образом, впервые на основе геномных данных показана возможная преемственность между некоторыми представителями черняховской культуры (III в. н.э.) и населения Древней Руси (вторая половина XII–начало XIII вв.).

Ключевые слова: черняховская культура; славяне; древняя ДНК; митохондриальная ДНК; H5a1a1; краниология; филогеографический анализ

Introduction

A new cultural entity, known as the Chernyakhov culture, arose at the end of the Roman era in the territories of the Northern Black Sea Region and the upper reaches of the Dniester-Western Bug rivers (Magomedov, 2001). The multi-ethnic origin of this group is currently the prevailing theory among experts (Sedov, 1979; Magomedov, 2001; Zinkovskaya, Kolesnikova, 2020). Nonetheless, it remains ambiguous which "barbarian" tribes formed the basis of the Chernyakhov culture, and which ones exerted influence subsequent to its final formation (Shchukin, 2005). Archaeologists examine the existence of a Slavic or Proto-Slavic element in its composition, the carriers of which thereafter developed the areas previously occupied by the representatives of the Chernyakhov culture (Sedov, 1979; Shchukin, 1997; Terpilovsky, 2000). The presence of Early Slavic cultural components can be seen in the Chernyakhov settlements in the Middle Dniester region (Lyapushkin, 1968; Rickman, 1975; Vinokur, 2002); however, pinpointing the purported Proto-Slavic component is a challenging objective when relying solely on archaeological and anthropological approaches (Terpilovsky, 2000; Magomedov, 2001).

Krynichki is a burial site located in the Middle Dniester, specifically in the Balti district of the Odessa region. As an archaeological site, it has been known since the end of the 19th century, but it began to be considered as a site of the Chernyakhov culture in the 20th century (Gamchenko, 1911; Symonovich, 1960). The existence of the Chernyakhov culture in the 3rd-4th centuries AD was verified by the artefacts found during the 1957–1958 excavations at this location carried out by the South Russian Expedition of the IIMK of the USSR Academy of Sciences. During the works on the Labushna Posad gully, a single-layer settlement and burial ground dated to the 3rd-4th centuries AD were discovered (Symonovich, 1960). There were only indications of inhumation in the burial. This characteristic distinctly differentiates this burial from the sites of the Chernyakhov culture, which are characterised by varying proportions of both inhumation and cremation rites (Nikitina, 1985). Furthermore, the grave was positioned separately from the others, which cannot be explained by family ties or belonging to a particular social group (Symonovich, 1960). In the burial site, only three of the five found skeleton remains burial site were quite well preserved, allowing for their anthropological and genetic investigations.

A young girl individual from Burial 4 was studied in this research. The skeleton was found in an oval-shaped grave, lying on its back in an extended position, with the head orientated northeast. The burial contained two garter-style bronze crossbow fibulae, glass beads, a multi-part bone comb secured with bronze nails, and a cylindrical clay spindle adorned with a circular pattern on its sides. Also in the burial at the left elbow was a bronze staff-like pin, which apparently served to attach the ribbon to the braid (Nikitina, 2008). The aforementioned inventory is typical of female graves. It is significant that the habit of embellishing the braid with a ribbon was prevalent

among East Slavic women up to the 19th-20th centuries (Chistov, 1987). This trait may be thought as indirect indication of the buried female's relationship to the Slavic community.

Complex genetic research of the Chernyakhov culture representatives, together with anthropological data, can aid in estimating their potential genetic contribution to the formation of Slavic populations. Previously, mitochondrial genome analysis has proved its effectiveness to obtain information about historical processes, particularly migratory occurrences (Andreeva et al., 2024), as well as to assess the kinship between individuals (Andreeva et al., 2023a) and the probable origins of the studied people (Andreeva et al., 2023b).

The purpose of this study was to conduct a comprehensive analysis of the presence of the Slavic component in Chernyakhov culture representatives using paleoanthropological and genetic data.

Materials and methods

The skulls of 153 Chernyakhov culture representatives from the funds of the Research Institute and Museum of Anthropology of Lomonosov Moscow State University were measured and analysed; for comparative analysis by statistics methods, several craniological series of East Slavs (229 skulls) were also studied (Table 1). A part of the studied materials included in the analysis overlapped geographically. This is because the Slavic tribes inhabited far more territory than the Chernyakhov groups.

Eight measurements of the facial section of the skull (Alekseev, Debets, 1964) were used in the analysis: mean facial width (46 Mar.), upper facial height (48 Mar.), orbital width (51 Mar.), orbital height (52 Mar.), nasal width (54 Mar.), subspinale height above the zygomaxillary chord, simotic width (SC Biom.) and height (SS Biom.). By World PCA software and the set of additional analytical techniques based on it, craniometric data were statistically analysed using the principal component (PC) approach (for a thorough description, see (Evteev et al., 2021)).

A fragment of the petrous part of temporal bone of museum specimen No. 10917 (Fig. 1*a*) was used for genetic research. It belongs to a juvenile individual from grave 4 of the archaeological site of Krynichki (Fig. 1*b*). This burial site dates back to 230–270 AD (Nikitina, 2008).

Ancient DNA was extracted from the cochlea specimen weighing 0.236 grams. The DNA extraction was carried out in sterile rooms of the Sirius University of Science and Technology in accordance with the previously published methodology (Andreeva et al., 2022). The sample of DNA was assigned the number AB93.

The DNA quality was evaluated using an Agilent Bioanalyzer 2100 using a High Sensitivity DNA chip kit ("Agilent"). The ancient DNA was used to create a genomic library (Gansauge et al., 2017), which was then sequenced on the Illumina HiSeq 2500 platform using single-end reads.

The bioinformatics analysis of the sequencing data involved multiple stages. AdapterRemoval v2 (Schubert et al., 2016) was used to remove adapter sequences. The nucleotide sequences with a length greater than 30 nucleotides were aligned to the reference genomes rCRS/NC_012920.1 (mitochondrial genome) (Anderson et al., 1981) and hg19/GRCh37 (human reference genome) using the BWA tool (Li, Durbin, 2009). MapDamage v.2.2.1 (Jónsson et al., 2013) was used to validate the authenticity of ancient DNA by analysing the frequency of C→T substitutions at the ends of reads. Contamination was evaluated using the Schmutzi software (Renaud et al., 2015). To identify the genetic sex of the sample, the ratios of the average number of reads covering sex chromosomes (X and Y separately) to the number of reads covering autosomal ones were calculated.

Reads with a mapping quality score greater than 20 were used to reconstruct the mitochondrial sequence (mtDNA). The genetic variations were determined by the BCFtools program (Danecek et al., 2021). Additionally, quality filtering was applied to the genotypes (QUAL > 30), as well as normalisation of the identified insertions and deletions. The nucleotide variations found in mitochondrial DNA were verified using the IGV (integrative genomics viewer) tool (Robinson et al., 2011).

Mitochondrial haplogroup determination of AB93 was performed using Haplogrep3 (Schönherr et al., 2023), based on Phylotree build 17 (Van Oven, 2015). The Yfull MTree 1.02 database (https://www.yfull.com/mtree/) was used for validation.

The search for the mtDNA sequences of both present-day and ancient individuals that had the highest similarity to the AB93 sample was conducted using public databases such as NCBI (https://www.ncbi.nlm.nih.gov/nuccore), Allen Ancient DNA Resource (AADR) (Mallick et al., 2024), Yfull MTree 1.02 (https://www.yfull.com/mtree/), and AmtDB (Ehler et al., 2019). The BLAST service (https://blast.ncbi.nlm.nih.gov/Blast.cgi) was used to search and select sequences in the NCBI database. The search conditions were set to 100 % Query coverage and a Percent Identity of at least 99.98 %.

The phylogeographic analysis was carried out using the maximum likelihood method in the mtPhyl program (Eltsov, Volodko, 2016). The major clade was constructed by grouping samples with the fewest number of nucleotide substitutions, taking into consideration the whole mitochondrial sequence. For samples from the ancestral haplogroup poly-C tracts, tandem repeat sections 522–524 and 573–576, and nucleotide position 16519 (a so-called "hot spot") were excluded from the analysis.

Results

At the first stage, several craniological series of East Slavic and Chernyakhov culture skulls were examined. On the basis of Chernyakhov male skulls, it was found that the first two principal components (PC1 and PC2) account for 47.5 % of the variability, while the first four ones account for 78.5 %. Therefore, the first two components show notable variance (Fig. 2). Individuals with high PC1 values have increased face height and width and facial profile and, to a lesser extent, increased eye socket height and nose width. Individuals with high PC2 values are characterised by a decrease in facial width combined with an increase in the size of the back of the nose. In general, the analysed groupings do not form clearly defined clusters. Nevertheless, the calculation of intragroup average pairwise Euclidean distances by PC1–4 values (SPER)

Table 1. Craniological series used in anthropological analysis

No.	Series title	Geographic location of the burial site	Number of skulls	
			male	female
	Craniological samples of representatives of the Chernyakhov culture (153 ind.)			
1	Baev	Volyn region		2
2	Budesti	Dubossari region	14	9
3	Viktorovka	Nikolaev region	1	1
4	Gavrilovka	Kherson region	6	8
5	Gorodok	Rivne region		1
6	Danilova balka	Kirovograd region	7	1
7	Derevyannaya	Kiev region	1	1
8	Zhuravka	Cherkassy region	32	27
9	Koblevo	Mykolaiv region	7	2
10	Krivchany	Khmelnitsky region		1
11	Krynichki	Odessa region	2	
12	Malaeshty	Dniester	2	4
13	Mikhailovka	Kherson region	3	
14	Ranzhevoe	Odessa region	1	1
15	Ridkoduby	Khmelnitsky region	1	
16	Romashki	Kiev region	2	
17	Ruzhichanka	Khmelnitsky region	1	1
18	Sabodash	Cherkasy region		1
19	Furmanovka	Odessa region	1	
20	Chernyakhov	Kiev region	8	2
21	Chistilov	Ternopil region	2	
		Total	91	62
	Craniological samples of East Slavs (229 ind.)			
1	Novgorod barrows (Kotorsk III, Udrai II, Udrai IV, Retenskoe ozero, Logoveshche, Konezer'e, Slavenka, Ozertitsy and Khreple)	Novgorod region, according to (Sankina, 2012)	48	
2	Kostroma group of barrows (Ples, Gorodok, Novoselki)	Kostroma region, according to (Alexeyeva, 1966)	12	
3	Groups of barrows in the basin of the middle reach of the Moscow River (Il'inskoe, Kosino, Spas-Tushino, Lepeshki)	Moscow region, according to (Alexeyeva, 1966)	16	10
4	Groups of barrows in the area of the Smolensk reach of the Dnieper River (Volochek, Seltso, Varnavino, Selishche, Staraya Rudnya, Ivanovichi)	Smolensk region, according to (Alexeyeva, 1966)	27	28
5	A series of the Vladimir-Ryazan-Nizhegorod town groups of barrows (Ziminki, Murom, Popovskaya, Gorodishche, Terekhovo)	Vladimir, Ryazan, Nizhny Novgorod regions, according to (Alexeyeva, 1966)	22	
6	Chernigov group of barrows (Gushchino, Bakhmach, Stol'noe, Shestovitsy)	Chernigov region, according to (Alexeyeva, 1966)	13	
7	Pereyaslavl group of barrows (Pereyaslav-Khmel'nitskii, Medvezh'e, Lipovoe)	Poltava region, according to (Alexeyeva, 1966)	32	
8	Groups of barrows in the middle reach of the Desna River (Gochevo, Aleksandrovka, Krasnoe, Golubitsa, Setnyi khutor)	Kursk and Chernigov regions, according to (Alexeyeva, 1966)		21
		Total	170	59

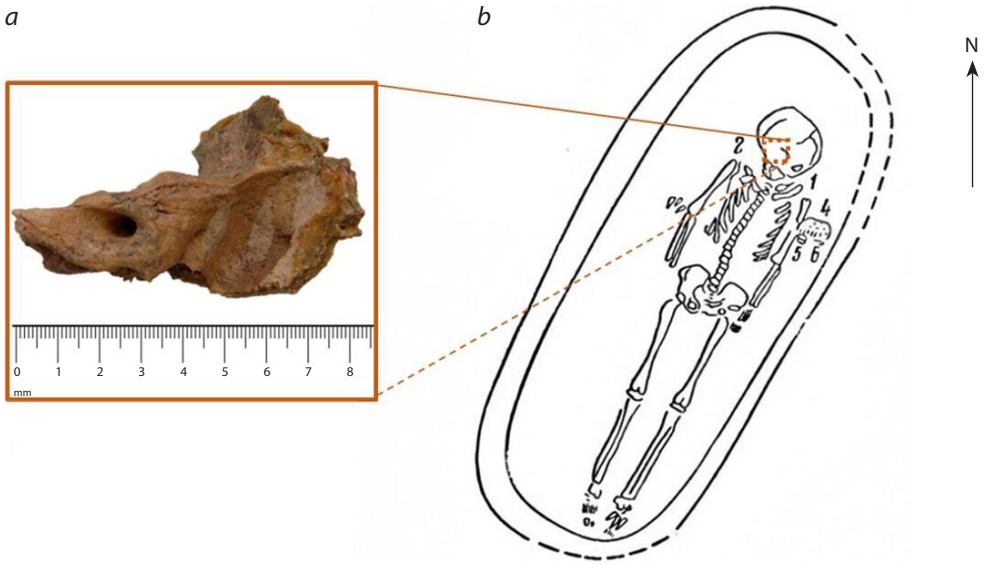


Fig. 1. Anthropological material used for genetic analysis.

a – the petrous part of right temporal bone of female individual (museum No. 10917); *b* – scheme of burial 4 from the burial site in the settlement of Krynichki (Symonovich, 1960).

showed that the samples from Budesti and Chernyakhov (left bank of the Dniester River, present-day Dubossari district, and present-day Kiev region, Obukhov district, respectively) are the most homogeneous (SPER ≤ 2.7), while the series from Gavrilovka (present-day Kherson region) is significantly more heterogeneous (SPER ≤ 4.5).

In the study of female skulls, the first two principal components explain 46 % of the overall variability, whereas the first four explain 73.7 %. The morphological meaning of PC1 is completely consistent with that of the male skulls: individuals with high values of this component have increased facial size, eye socket and nose width, as well as enhanced horizontal profiling. PC2 has a similar connotation in general, describing a rise in the size of the nasal bones. Thus, the basic tendencies of morphological diversity of the facial region of the skull in the male and female sections of the combined series of the Chernyakhov culture are similar (Fig. 2).

Additionally, principal component analysis was applied to the representatives of the East Slavs and the Chernyakhov culture. This analysis revealed substantial discrepancies between them for the first three PCs (Fig. 3). According to the Kaiser criterion (Deryabin, 2008), the first three principal components can be considered significant, as well as conditionally the fourth (eigenvalue 0.97). At the same time, East Slavs and individuals of the Chernyakhov culture have a similar range of overall craniological diversity (intragroup SPER – 3.18 and 3.07, respectively). Some local male series of skulls (Budeshti and Chernyakhov) overlap with East Slavic groupings (Smo-

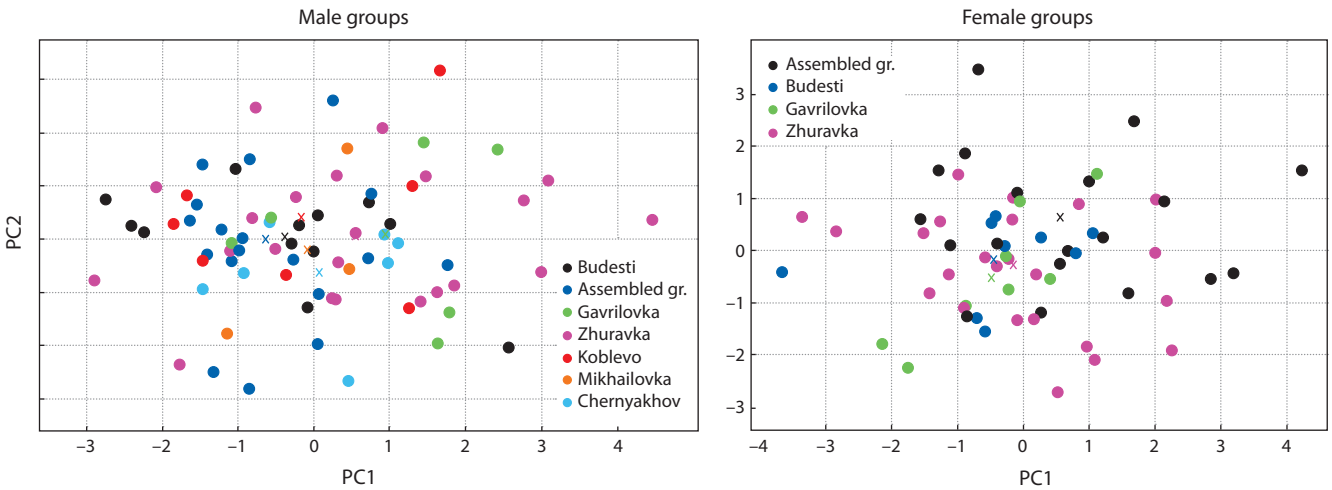


Fig. 2. Location of the studied groups of the Chernyakhov culture in the space of PC1 and PC2.

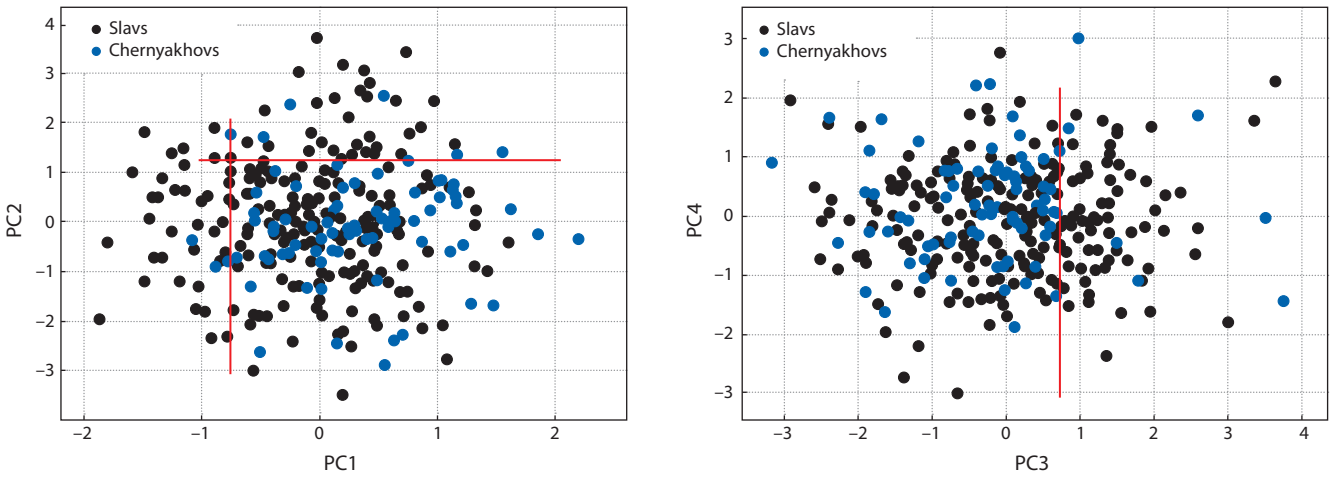


Fig. 3. Location of the studied male skulls in the space of four PCs (East Slavs and representatives of the Chernyakhov culture together).

The red lines in the figure restrict the cluster of East Slavs and Chernyakhov culture individuals with similar PC scores from the cluster of East Slavs with very different scores.

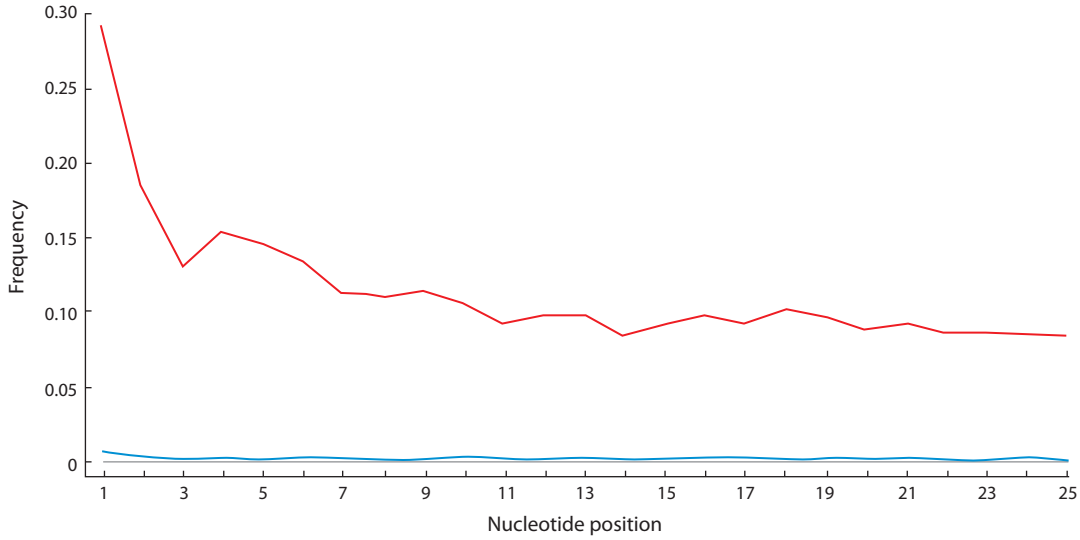


Fig. 4. Nucleotide substitution profile obtained using the mapDamage2 programme (Jónsson et al., 2013) for reads mapped to the mitochondrial reference genome.

The red line indicates specific for ancient DNA C→T substitutions among the first 25 nucleotides from the 5′-end of the DNA fragments.

lensk, Kostroma, and Vladimir-Ryazan-Nizhegorod regions). The disparities in the generalised female series of skulls are less pronounced than those observed in the male series.

The second stage was genetic analysis of AB93. The whole mitochondrial DNA sequence of the Chernyakhov culture representative was reconstructed. This individual presumably has a Proto-Slavic component based on the archaeological context.

According to the results of sequencing and primary bioinformatic analysis, about 115.8 million short reads were obtained, of which 51.5 % were mapped to the human reference genome. The higher frequency of C→T substitutions observed throughout the entire length of the fragments (Fig. 4) confirms that the AB93 sample belongs to ancient DNA. This specific DNA deamination deals with postmortem alterations and marks DNA extracted from archaeological and anthropological samples well. The ratio of the average number of reads covering sex chromosomes (X and Y separately) to the number of reads covering autosomal ones revealed that the studied sample belonged to a female, which is consistent with the archaeological context.

Mapping reads to the reference human mitochondrial genome (NC_012920.1) resulted in the reconstruction of the whole mitochondrial DNA sequence. The average coverage was x48,32, which allowed us to identify the mitochondrial haplogroup of AB93 and to conduct phylogeographic analysis. The analysis of the AB93 mtDNA sequence revealed variations associated with haplogroup H5a1a1 (Table 2).

Table 2. Variants of single nucleotide substitutions, forming mitochondrial haplogroups and detected in the mitochondrial sequence of AB93

Coordinate of the nucleotide position on the rCRS mitochondrial reference genome	Reference variant → variant detected in AB93	Haplogroup formed by the detected substitutions
456	C→T	H5'36
16304	T→C	H5
4336	T→C	H5a
15833	C→T	H5a1
721	T→C	H5a1a
16483	G→A	H5a1a1

Table 3. Ancient samples with mitochondrial DNA belonging to haplogroup H5a1a

Sample name	Dating	Cultural background	Region, burial site	Haplogroup*	Reference
AB93	230–270 AD	Chernyakhov culture	Odessa region, Krynichki burial ground	H5a1a	This study
MH605032 (BM5)	Early Bronze Age	Thracian culture	Bulgaria, Bereketska grave	H5a1a	Modi et al., 2019
DB37	Second half of 12th – beginning of 13th centuries AD	Russian North	Vologda region, Minino II	H5a1a	Rozhdestvenskikh et al., 2024
PCA0044	100–300 AD	Wielbark culture	Poland, Kowalewko	H5a1a	Stolarek et al., 2023
PCA0403	11th–12th centuries AD	–	Poland, Lubusz Voivodeship, Santok	H5a1a	Stolarek et al., 2023

* Haplogroup is indicated according to the PhyloTree consensus classification; samples not included in the phylogeographic analysis are marked in grey.

In addition to haplogroup-forming variations, the mtDNA sequence of AB93 had an A-to-G substitution at position 93 located in hypervariable region 2 (HVR2). This variant does not define a haplogroup and is a private variant of the AB93 sample.

Mitochondrial sequences of individuals from public databases belonging to haplogroups H5a1a and H5a1a1 were used for phylogeographic analysis. Additionally, mtDNA sequences diverging by no more than three haplogroup-forming substitutions were included in the analysis. Thus, the sequences of 38 samples, including AB93, with geographic, ethnic, or cultural affiliation were selected for phylogeographic analysis. Of these, three sequences belonged to ancient individuals (Table 3) and 35 were from present-day individuals.

Figure 5 shows a fragment of the reconstructed phylogeographic tree of the mitochondrial lineage H5a1a. Five out of the eight samples belong to representatives of present-day Slavic populations and one is a previously studied individual from a medieval burial site in the modern Vologda region. This individual was dated to the latter half of the 12th to early 13th centuries AD and was identified as having Slavic origins (Rozhdestvenskikh et al., 2024). The mitochondrial DNA sequences of a Chernyakhov culture representative (AB93) and a medieval individual (DB37) from the Minino II burial site were found to be identical.

Discussion

The results of the anthropological analysis demonstrate that the morphological variability of the Chernyakhov culture's individuals has a generally uniform character in the space of principal components, with no pronounced clusters. At the same time, the overall range of craniological variability is significant and comparable to that in generalised craniological series for the entire territory of the settlement of the medieval East Slavs. This result indicates that the craniological series of Chernyakhov representatives have heterogeneity, potentially attributable to genetic factors.

The mitochondrial sequence of the studied individual of the Chernyakhov culture belongs to haplogroup H5a1a1. This lineage is part of the ancestral clade H5a1a, which is prevalent in present-day Slavic groups in Central and Eastern Europe (Mielnik-Sikorska et al., 2013; Malyarchuk et al., 2017). It was found in East Slavic populations (including Russians, Ukrainians, and Belarusians) as well as among West Slavs (Poles, Czechs, Slovaks, and Kashubians) (Mielnik-Sikorska et al., 2013; Malyarchuk et al., 2017). Haplogroup H5a1a has

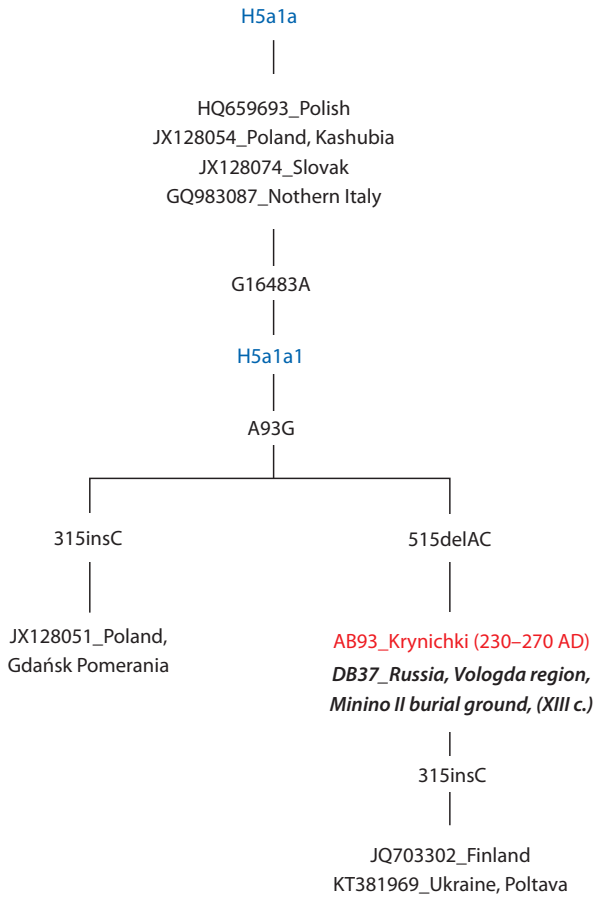


Fig. 5. Fragment of the phylogeographic tree constructed using the mtPhyl program for mtDNA samples of the mitochondrial branch of H5a1a.

The studied sample is shown in red. The ancient individual highlighted in bold italics. Present-day samples have an identifying number as well as geographic/ethnic information. Haplogroups of the H5a1 lineage are indicated in blue. For transitions, the position number and the substitution variant are indicated. ins – insertion and its location in the genome; del – deletion and its location in the genome.

the highest frequency (13.29 % of the entire sample) among present-day Poles (https://www.familytreedna.com/).

All ancient individuals whose mitochondrial sequences belong to haplogroup H5a1a have been found in the territory of Eastern Europe. The oldest bearer of this haplogroup dating back to the Early Bronze Age was found in present-day Bulgaria, in the eastern part of the Balkan Peninsula (Thrace) (Modi et al., 2019). In Poland, haplogroup H5a1a was found in a representative of the Wielbarka culture (100–300 AD) (Stolarek et al., 2023), as well as in a medieval individual from the Santok necropolis in western Poland. This necropolis is assigned to the local Pomeranian people and dates back to the 11th–12th centuries AD (Stolarek et al., 2023). The hypothesis has been proposed that the Wielbark civilisation was based on the indigenous population of the Vistula and Western Bug river basins, as well as Gothic tribes who migrated from southern Scandinavia (Stolarek et al., 2023). Furthermore, the local tribes are thought to represent Eastern Europe's Proto-Slavic population (Grzesik, 2017). Significantly, an analysis of the hypervariable regions of mitochondrial sequences demonstrated the continuity of mitochondrial lineages in present-day Poland, dating back at least to the Roman period. It was also found that the local maternal lineages are part of the mitochondrial branch H5a1 (Juras et al., 2014).

We revealed that the mitochondrial DNA sequence of the representative of the Chernyakhov culture from the Middle Dniester exactly matches that of a medieval young man from the Russian North (present-day Vologda region) (Rozhdestvenskikh et al., 2024). It is important to note that, despite the fact that this individual from the Russian North was buried in a region primarily inhabited by local Finno-Ugric tribes, his burial customs followed Christian practices (Archaeology…, 2007). Also, this period is characterised by active interaction between Slavic and Finno-Ugric groups in the early stages of the Ancient Rus' state formation. Most likely, this young man was a non-local Slavic representative (Rozhdestvenskikh et al., 2024). The identity of the mitochondrial sequences suggests a probable maternal relationship between these two individuals – a teenage girl from a burial in the south of Rus' and a young man from a northern burial ground. This allows us to identify potential migration routes of the ancient population within the East European Plain.

Conclusion

As a result, our craniological analysis shows that East Slavs and Chernyakhov culture representatives share a similar range of craniological diversity. Genetic analysis also reveals that the mitochondrial lineage identified in the individual of the Chernyakhov culture is characteristic of Slavic groups, both present-day and ancient, who inhabited territories associated with the probable origin and settlement of the Slavs. In this regard, we can assume the presence of a genetic connection of the maternal lineage between representatives of the Chernyakhov culture and the ancient population of Eastern Europe, on the basis of which the Slavic community was formed. However, it should be noted that the data on the genetic connection between representatives of the Chernyakhov culture and the Slavs that we obtained for the first time on the basis of mitochondrial genomes require further confirmation using additional genetic markers and anthropological material.

References

Alekseev V.P., Debets G.F. Craniometry. Anthropological Research Methods. Moscow: Nauka Publ., 1964 (in Russian)

Alexeyeva T. East Slavic skulls from tribal barrows. *Materiały i Prace Antropologiczne = Anthropological Materials and Studies*. Kraków, 1966;3-142 (in Polish)

Anderson S., Bankier A.T., Barrell B.G., de Bruijn M.H., Coulson A.R., Drouin J., Eperon I.C., Nierlich D.P., Roe B.A., Sanger F., Schreier P.H., Smith A.J., Staden R., Young I.G. Sequence and organization of the human mitochondrial genome. *Nature*. 1981;290(5806):457-465. doi 10.1038/290457a0

Andreeva T.V., Manakhov A.D., Gusev F.E., Patrikeev A.D., Golovanova L.V., Doronichev V.B., Shirobokov I.G., Rogaev E.I. Genomic analysis of a novel Neanderthal from Mezmaiskaya Cave provides insights into the genetic relationships of Middle Palaeolithic population. *Sci Rep*. 2022;12(1):13016. doi 10.1038/s41598-022-16164-9

Andreeva T.V., Dobrovolskaya M.V., Sedov V.V., Vdovichenko M.V., Reshetova I.K., Soshkina A.D., Dudko N.A., Bydanov A.S., Adrianova I.Yu., Buzhilova A.P., Rogaev E.I. People from the stone sarcophagus No. 11 at the St. Georges (Yuriev) Monastery: a genetic history based of mitochondrial genomes. *Kratkie Soobshcheniya Instituta Archeologii = Brief Communications of the Institute of Archaeology.* 2023a;270:418-437. doi 10.25681/IARAS.0130-2620.270.418-437 (in Russian)

Andreeva T.V., Malyarchuk A.B., Rodinkova V.E., Soshkina A.D., Rozhdestvenskikh E.V., Dobrovolskaya M.V., Rogaev E.I. An individual of the Volyntsevo period from Kurilovka: the first archeogenetic data. *Russijskaja Archeologija = Russian Archaeology.* 2023b;(3):57-71. doi 10.31857/S0869606323030030 (in Russian)

Andreeva T.V., Zhilin M.G., Malyarchuk A.B., Engovatova A.V., Soshkina A.D., Dobrovolskaya M.V., Buzhilova A.P., Rogaev E.I. Archaeogenomics of humans from the layer of the Upper Volga Culture revealed their greatest genetic similarity with Eastern European hunter-gatherers and ancient representatives of Mesolithic/Neolithic Europe. *Vestnik Arheologii, Antropologii i Etnografii.* 2024;64:113-125. doi 10.20874/2071-0437-2024-64-1-10 (in Russian)

Archaeology of the Northern Russian Village of the 10–13th Centuries: Medieval Settlements and Burial Grounds on Lake Kubenskoye, vol. I: Settlements and Burial Grounds. Moscow: Nauka Publ., 2007. Available at: https://archaeolog.ru/media/series/arch-northrus-village/Arch_northrus_village-1.pdf (in Russian) (in Russian)

BLAST – Basic Local Alignment Search Tool. Available: https://blast.ncbi.nlm.nih.gov/Blast.cgi

Chistov K.V. Ethnography of Eastern Slavs. Moscow: Nauka Publ., 1987 (in Russian)

Danecek P., Bonfield J.K., Liddle J., Marshall J., Ohan V., Pollard M.O., Whitwham A., Keane T., McCarthy S.A., Davies R.M., Li H. Twelve years of SAMtools and BCFtools. *GigaScience.* 2021;10(2):giab 008. doi 10.1093/gigascience/giab008

Deryabin V.E. Lectures on Multidimensional Biometrics for Anthropologists. Moscow: Moscow State University Publ., 2008 (in Russian)

Ehler E., Novotný J., Juras A., Chylenski M., Moravcík O., Paces J. AmtDB: a database of ancient human mitochondrial genomes. *Nucleic Acids Res*. 2019;47(D1):D29-D32. doi 10.1093/nar/gky843

Eltsov N., Volodko N. MtPhul: Software tool for human myDNA analysis and phylogeny reconstruction. 2016. Available online: http://eltsov.org (Available from: https://sites.google.com/site/mtphyl/home)

Evteev A.A., Staroverov N.E., Potrakhov N.N. Analytical techniques and software for the study of intragroup metric variation using principal component analysis. *Bulletins et Mémoires de la Société d'Anthropologie de Paris*. 2021;33(1):7539. doi 10.4000/bmsap.7539

FamilyTreeDNA. Available: https://www.familytreedna.com/

Gamchenko S.S. Archeological Research on the Tripoli Culture in Podillya in 1909. St. Petersburg, 1911

Gansauge M.T., Gerber T., Glocke I., Korlevic P., Lippik L., Nagel S., Reihl L.M., Schmidt A., Meyer M. Single-stranded DNA library preparation from highly degraded DNA using *T4* DNA ligase. *Nucleic Acids Res*. 2017;45(10):e79. doi 10.1093/nar/gkx033

Grzesik R. Ethnogenesis of the Slavs in the Polish historical reflection of the 20th–21st century. *Studia Slavica et Balcanica Petropolitana.* 2017;(2):107-121. doi 10.21638/11701/spbu19.2017.209 (in Russian)

Jónsson H., Ginolhac A., Schubert M., Johnson P.L., Orlando L. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics*. 2013;29(13):1682-1684. doi 10.1093/bioinformatics/btt193

Juras A., Dabert M., Kushniarevich A., Malmström H., Raghavan M., Kosicki J.Z., Metspalu E., Willerslev E., Piontek J. Ancient DNA reveals matrilineal continuity in present-day Poland over the last two millennia. *PLoS One*. 2014;9(10):110839. doi 10.1371/journal.pone.0110839

Li H., Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics.* 2009;25(14):1754-1760. doi 10.1093/bioinformatics/btp324

Lyapushkin I.I. Slavs of Eastern Europe on the eve of the formation of the Old Russian state (8th–the first half of the 9th century). Leningrad: Nauka Publ., 1968 (in Russian)

Magomedov B.V. Chernyakhov Culture: The Issue of Ethnicity. Lublin, 2001. Available at: http://www.archaeology.ru/Download/Magomedov/Magomedov_2001_Czerniachov_ethnos.pdf (in Russian)

Mallick S., Micco A., Mah M., Ringbauer H., Lazaridis I., Olalde I., Patterson N., Reich D. The Allen Ancient DNA Resource (AADR) a curated compendium of ancient human genomes. *Sci Data*. 2024;11(1):182. doi 10.1038/s41597-024-03031-7

Malyarchuk B., Litvinov A., Derenko M., Skonieczna K., Grzybowski T., Grosheva A., Shneider Y., Rychkov S., Zhukova O. Mitogenomic diversity in Russians and Poles. *Forensic Sci Int Genet*. 2017;30:51-56. doi 10.1016/j.fsigen.2017.06.003

Mielnik-Sikorska M., Daca P., Malyarchuk B., Derenko M., Skonieczna K., Perkova M., Dobosz T., Grzybowski T. The history of Slavs inferred from complete mitochondrial genome sequences. *PLoS One*. 2013;8(1):e54360. doi 10.1371/journal.pone.0054360

Modi A., Nesheva D., Sarno S., Vai S., Karachanak-Yankova S., Luiselli D., Pilli E., … Antonova O., Hadjidekova S., Galabov A., Toncheva D., Caramelli D. Ancient human mitochondrial genomes from Bronze Age Bulgaria: new insights into the genetic history of Thracians. *Sci Rep*. 2019;9(1):5412. doi 10.1038/s41598-019-41945-0

Nikitina G.F. Systematics of Burial Rituals among Chernyakhov Culture Tribes. Moscow: Nauka Publ., 1985 (in Russian)

Nikitina G.F. Chernyakhov Culture in Transnistria according to the Analysis of Archaeological Data. Moscow: Taus Publ., 2008. Available at: https://archaeolog.ru/media/books_2008/Nikitina_2008.pdf (in Russian)

Renaud G., Slon V., Duggan A.T., Kelso J. Schmutzi: estimation of contamination and endogenous mitochondrial consensus calling for ancient DNA. *Genome Biol*. 2015;16:224. doi 10.1186/s13059-015-0776-0

Rickman E.A. The Ethnic History of Transnistria and the Adjacent Danube Region in the First Centuries of Common Era. Moscow: Nauka Publ., 1975 (in Russian)

Robinson H.J., Thorvaldsdóttir H., Winckler W., Guttman M., Lander E.C., Getz G., Mesirov J.P. Integrative genomics viewer. *Nat Biotechnol*. 2011;29(1):24-26. doi 10.1038/nbt.1754

Rozhdestvenskikh E.V., Manakhov A.D., Andreeva T.V., Kunizheva S.S., Malyarchuk A.B., Bydanov A.S., Savinetsky A.B., Krylovich O.A., Dobrovolskaya M.V., Buzhilova A.P., Makarov N.A., Rogaev E.I. A youth with a polecat from the medieval Minino cemetery at the Kubenskoe Lake: archaeogenetic identification. *Kratkiye Soobshcheniya Instituta Arkheologii = Brief Communications of the Institute of Archaeology*. 2024;274:348-368. doi 10.25681/IARAS.0130-2620.274.348-368 (in Russian)

Sankina S.L. The Ethnic History of the Medieval Population of the Novgorod Land According to Anthropological Data. Saarbrucken: LAP LAMBERT Academic Publishing, 2012 (in Russian)

Schönherr S., Weissensteiner H., Kronenberg F., Forer L. Haplogrep 3 – an interactive haplogroup classification and analysis platform. *Nucleic Acids Res*. 2023;51(W1):W263-W268. doi 10.1093/nar/gkad284

Schubert M., Lindgreen S., Orlando L. AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Res Notes*. 2016;12(9):88. doi 10.1186/s13104-016-1900-2

Sedov V.V. The Origin and Early History of the Slavs. Moscow: Nauka Publ., 1979 (in Russian)

Shchukin M.B. Stratum: Structures and Catastrophes: The Birth of Slavs. St. Petersburg: Nestor Publ., 1997;110-147 (in Russian)

Shchukin M.B. The Gothic Way: Goths, Rome, and the Chernyakhov Culture). Saint-Petersburg: St. Petersburg State University Publ., 2005 (in Russian)

Stolarek I., Zenczak M., Handschuh L., Juras A., Marcinkowska-Swojak M., Spinek A., Dębski A., Matla M., Kóčka-Krenz H., Piontek J.; Polish Archaeogenomics Consortium Team; Figlerowicz M. Genetic history of East-Central Europe in the first millennium CE. *Genome Biol*. 2023;24(1):173. doi 10.1186/s13059-023-03013-9

Symonovich E.A. Artifacts of the Chernyakhov culture in the village of Krinichki. Chernyakhov culture. *Materialy i Issledovaniya po Arkheologii SSSR = Materials and Studies on USSR Archeology*. 1960;82:239-252 (in Russian)

Terpilovsky R.V. The Kiev and Chernyakhovskaya cultures: contacts. *Stratum plus. Archaeology and Cultural Anthropology*. 2000;(4): 303-311 (in Russian)

Van Oven M. PhyloTree Build 17: growing the human mitochondrial DNA tree. *Forensic Sci Int Genet Suppl Ser*. 2015;5:e392-e394. doi 10.1016/j.fsigss.2015.09.155

Vinokur I.S. Slavic-Thracian contacts in the Middle and Upper Transnistria. In: Ancient Communities of Farmers and Herders in the Northern Black Sea Region (5th millennium BC–5th century AD). Tiraspol, 2002;400-404 (in Russian)

YFull MTree 1.02 (under construction). Available: https://www.yfull.com/mtree/

Zinkovskaya I.V., Kolesnikova A.Y. Problems of the Chernyakhov culture in modern Ukrainian archaeology (2000–2019). *Vestnik VGU. Seriya: Istoriya. Politologiya. Sotsiologiya = Proceedings of Voronezh State University. Series: History. Political Science. Sociology*. 2020;(1):35-40 (in Russian)

Conflict of interest. The authors declare no conflict of interest.
Received November 19, 2024. Revised December 26, 2024. Accepted December 28, 2024.

Прием статей через электронную редакцию на сайте http://vavilov.elpub.ru/index.php/jour
Предварительно нужно зарегистрироваться как автору, затем в правом верхнем углу страницы выбрать «Отправить рукопись». После завершения загрузки материалов обязательно выбрать опцию «Отправить письмо», в этом случае редакция автоматически будет уведомлена о получении новой рукописи.

«Вавиловский журнал генетики и селекции (Vavilov Journal of Genetics and Breeding)» до 2011 г. выходил под названием «Информационный вестник ВОГиС»/ "The Herald of Vavilov Society for Geneticists and Breeding Scientists".

Сетевое издание «Вавиловский журнал генетики и селекции (Vavilov Journal of Genetics and Breeding)» – реестровая запись СМИ Эл № ФС77-85772, зарегистрировано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций 14 августа 2023 г.

Издание включено ВАК Минобрнауки России в Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук, Russian Science Citation Index, Российский индекс научного цитирования, ВИНИТИ, Web of Science CC, Scopus, PubMed Central, DOAJ, ROAD, Ulrich's Periodicals Directory, Google Scholar.

Открытый доступ к полным текстам:
русскоязычная версия – на сайте https://vavilovj-icg.ru/
и платформе Научной электронной библиотеки, elibrary.ru/title_about.asp?id=32440

англоязычная версия – на сайте vavilov.elpub.ru/index.php/jour
и платформе PubMed Central, https://www.ncbi.nlm.nih.gov/pmc/journals/3805/

При перепечатке материалов ссылка обязательна.

✉ email: vavilov_journal@bionet.nsc.ru

Издатель: Федеральное государственное бюджетное научное учреждение
«Федеральный исследовательский центр Институт цитологии и генетики
Сибирского отделения Российской академии наук»,
проспект Академика Лаврентьева, 10, Новосибирск, 630090.
Адрес редакции: проспект Академика Лаврентьева, 10, Новосибирск, 630090.
Секретарь по организационным вопросам С.В. Зубова. Тел.: (383)3634977.
Издание подготовлено информационно-издательским отделом ИЦиГ СО РАН. Тел.: (383)3634963*5218.
Начальник отдела: Т.Ф. Чалкова. Редакторы: В.Д. Ахметова, И.Ю. Ануфриева. Дизайн: А.В. Харкевич.
Компьютерная графика и верстка: Т.Б. Коняхина, О.Н. Савватеева.
Дата выхода в свет 28.08.2025. Формат 60 × 84 $^{1}/_{8}$. Уч.-изд. л. 15.1.