2024 • 28 • 7

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-24-75

2024 • 28 • 7

Online ISSN 2500-3259

Online edition

# VAVILOVSKII ZHURNAL GENETIKI I SELEKTSII

## VAVILOV JOURNAL OF GENETICS AND BREEDING

*Founded in 1997*
*Published 8 times annually*
DOI 10.18699/vjgb-24-75

# СОДЕРЖАНИЕ • 2024 • 28 • 7

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

# CONTENTS • 2024 • 28 • 7

## Molecular and cell biology



## Plant genetics and breeding



## Animal genetics



## Medical genetics

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2024;28(7):679-687

DOI 10.18699/vjgb-24-76

# Generation and characterization of two induced pluripotent stem cell lines (ICGi052-A and ICGi052-B) from a patient with frontotemporal dementia with parkinsonism-17 associated with the pathological variant c.2013T>G in the *MAPT* gene

E.V. Grigor'eva [1, 2], A.A. Malakhova [1, 2], E.S. Yarkova[1, 3], J.M. Minina [1], Y.V. Vyatkin[4], J.A. Nadtochy[1, 3], E.A. Khabarova[1, 5], J.A. Rzaev[5], S.P. Medvedev [1, 2], S.M. Zakian [1, 2]

[1] Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
[2] Institute of Chemical Biology and Fundamental Medicine of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
[3] Novosibirsk State University, Novosibirsk, Russia
[4] NOVEL Ltd., Novosibirsk, Russia
[5] Federal Neurosurgical Center of the Ministry of Health of the Russian Federation, Novosibirsk, Russia
evlena@bionet.nsc.ru

**Abstract.** Frontotemporal dementia with parkinsonism-17 is a neurodegenerative disease characterised by pathological aggregation of the tau protein with the formation of neurofibrillary tangles and subsequent neuronal death. The inherited form of frontotemporal dementia can be caused by mutations in several genes, including the *MAPT* gene on chromosome 17, which encodes the tau protein. As there are currently no medically approved treatments for frontotemporal dementia, there is an urgent need for research using *in vitro* cell models to understand the molecular genetic mechanisms that lead to the development of the disease, to identify targets for therapeutic intervention and to test potential drugs to prevent neuronal death. Analysis of exome sequencing data from a 46-year-old patient with a clinical diagnosis of Parkinson's disease revealed the presence of the pathological variant c.2013T>G (rs63750756) in the *MAPT* gene, which is associated with frontotemporal dementia with parkinsonism-17. By reprogramming the patient's peripheral blood mononuclear cells, we obtained induced pluripotent stem cells (iPSCs). Two iPSC lines were characterised in detail. Reprogramming was performed by transfection with non-integrating episomal vectors expressing the OCT4, SOX2, KLF4, LIN28, L-MYC and mp53DD proteins. The iPSC lines ICGi052-A and ICGi052-B proliferate stably, form colonies with a morphology characteristic of human pluripotent cells, have a normal diploid karyotype (46,XX), express endogenous alkaline phosphatase and pluripotency markers (OCT4, NANOG, SSEA-4 and TRA-1-60) and are able to differentiate into derivatives of three germ layers: ento-, ecto- and mesoderm. The iPSC lines obtained and characterised in detail in this work represent a unique tool for studying the molecular genetic mechanisms of the pathogenesis of frontotemporal dementia with parkinsonism-17, as well as for testing potential drugs *in vitro*.

**Key words:** frontotemporal dementia with parkinsonism-17; induced pluripotent stem cells; *MAPT* gene.

**For citation:** Grigor'eva E.V., Malakhova A.A., Yarkova E.S., Minina J.M., Vyatkin Y.V., Nadtochy J.A., Khabarova E.A., Rzaev J.A., Medvedev S.P., Zakian S.M. Generation and characterization of two induced pluripotent stem cell lines (ICGi052-A and ICGi052-B) from a patient with frontotemporal dementia with parkinsonism-17 associated with the pathological variant c.2013T>G in the *MAPT* gene. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2024;28(7):679-687. DOI 10.18699/vjgb-24-76

**Funding.** The study was supported by the Ministry of Science and Higher Education of the Russian Federation, Agreement No. 075-15-2021-1063/10.

**Acknowledgements.** The analysis of preparations with immunofluorescence staining and differential chromosome staining was performed using the resources of the Common Use Center for Microscopy of Biological Objects of ICiG SB RAS (https://ckp.icgen.ru/ckpmabo), supported by the budget project of the Institute of Cytology and Genetics SB RAS (FWNR-2022-0015).

# Создание и характеристика двух линий индуцированных плюрипотентных стволовых клеток (ICGi052-A и ICGi052-B) от пациента с лобно-височной деменцией с паркинсонизмом-17, ассоциированной с патологическим вариантом c.2013T>G в гене *MAPT*

Е.В. Григорьева [1, 2], А.А. Малахова [1, 2], Е.С. Яркова[1, 3], Ю.М. Минина [1], Ю.В. Вяткин[4], Ю.А. Надточий[1, 3], Е.А. Хабарова[1, 5], Дж.А. Рзаев[5], С.П. Медведев [1, 2], С.М. Закиян [1, 2]

[1] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
[2] Институт химической биологии и фундаментальной медицины Сибирского отделения Российской академии наук, Новосибирск, Россия
[3] Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия
[4] ООО «Новые Программные Системы», Новосибирск, Россия
[5] Федеральный центр нейрохирургии Министерства здравоохранения Российской Федерации, Новосибирск, Россия
evlena@bionet.nsc.ru

**Аннотация.** Лобно-височная деменция с паркинсонизмом-17 – нейродегенеративное заболевание, характеризующееся патологической агрегацией белка тау с образованием нейрофибриллярных клубков и дальнейшей гибелью нейронов. Наследственная форма лобно-височной деменции может быть вызвана мутациями в различных генах, одним из которых является ген *MAPT* на хромосоме 17, кодирующий тау-белок. Поскольку на данный момент отсутствуют утвержденные медицинским сообществом способы борьбы с лобно-височной деменцией, исследование на клеточных моделях *in vitro* молекулярно-генетических механизмов, приводящих к развитию заболевания, поиск мишеней для терапевтического воздействия и возможность тестирования потенциальных лекарственных препаратов для предотвращения гибели нейронов являются актуальной задачей. Анализ данных секвенирования экзома 46-летней пациентки с клиническим диагнозом болезнь Паркинсона показал наличие патологического варианта c.2013T>G (rs63750756) в гене *MAPT*, который ассоциирован с лобно-височной деменцией с паркинсонизмом-17. При помощи репрограммирования мононуклеарных клеток периферической крови пациентки нами были получены десять линий индуцированных плюрипотентных стволовых клеток (ИПСК), из которых детально охарактеризованы две. Репрограммирование проводили с помощью трансфекции неинтегрирующимися эписомными векторами, которые экспрессируют белки OCT4, SOX2, KLF4, LIN28, L-MYC и mp53DD. Линии ИПСК ICGi052-A и ICGi052-B стабильно пролиферируют, образуют колонии с характерной для плюрипотентных клеток человека морфологией, имеют нормальный диплоидный кариотип (46,XX), экспрессируют эндогенную щелочную фосфатазу и маркеры плюрипотентности (OCT4, NANOG, SSEA-4 и TRA-1-60) и способны дифференцироваться в производные трех зародышевых листков: энто-, экто- и мезодерму. Благодаря тому, что ИПСК можно направленно дифференцировать в широкий спектр типов клеток, полученные в данной работе и детально охарактеризованные линии ИПСК являются уникальным инструментом для изучения молекулярно-генетических механизмов патогенеза лобно-височной деменции с паркинсонизмом-17, а также тестирования потенциальных лекарственных препаратов *in vitro*.

**Ключевые слова:** лобно-височная деменция с паркинсонизмом-17; индуцированные плюрипотентные стволовые клетки; ген *MAPT*.

## Introduction

Frontotemporal dementia with parkinsonism linked to chromosome 17 was first described by T. Lynch and colleagues in 1994 (Lynch et al., 1994). This inherited early-onset dementia syndrome also includes parkinsonism, a set of symptoms that includes both non-motor manifestations (such as cognitive impairment, depression, bipolar disorder, sleep disturbances) and motor disorders such as muscle rigidity, resting tremor, bradykinesia. Early mental disorders and personality changes often precede severe dementia.

The *MAPT* gene encodes the tau protein, which regulates the assembly and stabilisation of microtubules involved in central nervous system signalling and axonal transport (Esmaeli-Azad et al., 1994). Six isoforms of the tau protein are expressed in the human brain due to alternative splicing of exons 2, 3 and 10 of the mRNA of the *MAPT* gene. Alternative splicing of exon 10 results in a tau protein with three (3R) or four (4R) microtubule-binding (MT-binding) domains. The 4R tau isoforms have an increased affinity for microtubules compared to the 3R isoforms. In childhood, the 3R form predominates in the brain, whereas in the adult brain, the 3R and 4R forms are present in a 1:1 ratio (D'Souza, Schellenberg, 2006).

The *MAPT:c.2013T>G* genetic variant (rs63750756, p.N279K) is a missense mutation in exon 10 resulting in the substitution of asparagine for lysine at position 279 of the *MAPT* protein (Hasegawa et al., 1999). As a result of this substitution, the polypurine-positive cis-element present in exon 10 is enhanced, leading to an increase in the frequency of inclusion of exon 10 in the transcript during splicing (Ritter et al., 2018). This results in a change in the ratio of 3R and 4R forms, leading to microtubule destabilisation, disruption of vesicle intracellular trafficking (Wren et al., 2015), formation of filamentous inclusions (Ghetti et al., 2015) and subsequent neuronal death. In addition, neurons with this genetic variant are characterised by increased spontaneous calcium fluctuations, mitochondrial dysfunction and increased production of reactive oxygen species, which also leads to cell death (Korn et al., 2023).

These processes lead to the development of frontotemporal dementia with parkinsonism-17. The exact mechanism of disease development in this genetic variant is not yet known, making it impossible to develop a treatment regimen.

Several models are used to study the development of frontotemporal dementia. The first are animal models, mainly in rodents: laboratory mice and rats. Rodents are actively used to create models of neurodegenerative diseases (Britti et al., 2020; Esteras et al., 2021). There are also transgenic lines of laboratory animals carrying genetic variants that cause hereditary frontotemporal dementia (Dawson et al., 2007). However, these model systems have some limitations due to differences in the signs of ageing between mice and humans and even differences between mouse and human tau protein

(Iovino et al., 2015; Hernández et al., 2020). Therefore, in addition to animal models, models based on induced pluripotent stem cells (iPSCs) are now the most promising.

iPSCs are cells obtained by reprogramming somatic cells and are characterised by the ability to differentiate into derivatives of three germ layers like embryonic stem cells. iPSCs have all the properties of embryonic stem cells and are at the same time autologous with respect to the somatic cell donor (Valetdinova et al., 2021). iPSCs have great potential for personalised medicine, as they can be derived at any time in a patient's life and will be patient-specific. iPSCs make it possible to circumvent immunogenic and ethical issues, so the demand for them is growing. In principle, iPSCs can be generated by reprogramming any mature cell type isolated from the body. Fibroblasts, mononuclear blood cells, keratinocytes and melanocytes are widely used.

Because iPSCs can be differentiated into any cell type, their range of applications is very broad (Liu et al., 2020). Differentiated iPSC derivatives, such as neuron-like cells, can be used to study the molecular genetic mechanisms of the development of neurodegenerative diseases, including during the early stages of cell differentiation, to search for target molecules, and to test various chemical compounds as potential drugs (Grigor'eva et al., 2023).

In this work, we have obtained, characterised and certified two lines of iPSCs from a patient with a pathological genetic variant of *MAPT:c.2013T>G* (rs63750756, p.N279K), which leads to the development of frontotemporal dementia with parkinsonism-17. The resulting iPSC lines provide a cellular model to study the mechanisms of development of this disease and can also be used for drug screening.

## Materials and methods

**Compliance with ethical standards.** The study was approved by the Research Ethics Committee of the Federal Neurosurgical Center (Novosibirsk, Russia), Protocol No. 1 dated 14 March 2017. Peripheral blood samples of the patient were provided by the Federal Neurosurgical Center (Novosibirsk, Russia). The patient signed a voluntary informed consent and an information sheet.

**Isolation of PBMCs using a ficoll gradient.** Peripheral blood was collected in two 9 ml vacuum tubes containing K3EDTA (VACUETTE, Greiner Bio). To isolate PBMCs, 3–4 ml of blood was layered on 3–4 ml of ficoll solution (Histopaque-1077, Sigma-Aldrich or Biolot, density 1.077) and centrifuged at 400 g for 35–40 min on a centrifuge with slow rotor acceleration and deceleration (SL 16 Cenrtifuge, Thermo Fisher Scientific). A whitish interphase layer containing PBMCs was then carefully collected, transferred to a 15 ml tube and washed twice in the maximum volume of PBS by centrifugation at 300 g for 15–20 min. PBMCs were frozen in medium containing 90 % KnockOut Serum substitute (Thermo Fisher Scientific) and 10 % DMSO (Sigma-Aldrich), 5–10 million cells per cryovial.

**Reprogramming of patient-specific PBMCs and conditions for culturing iPSCs.** PBMC transfection was performed on a Neon Transfection System (Thermo Fisher Scientific) device using episomal vectors encoding OCT4, SOX2, KLF4, L-MYC, LIN28, mp53DD and EBNA1 (Addgene ID No. 41813-14, No. 41855-57), as previously described (Grigor'eva et al., 2023). 0.5 micrograms of each vector was used for transfection of $1 \times 10^6$ PBMCs, using the programme 1650 V, 10 ms, 3 pulses. Cells were plated on a feeder layer of mitomycin C (Sigma-Aldrich)-treated embryonic fibroblasts. Single cell colonies were manually plated into 1 $cm^2$ wells with a pre-plated feeder and cultured in iPSC medium containing DMEM/F12, 15 % KnockOut Serum Replacement, 1 % GlutaMAX-I, 0.1 mM NEAA, 1 % penicillin-streptomycin (all Thermo Fisher Scientific), 0.1 mM 2-mercaptoethanol (Sigma-Aldrich) and 10 ng/ml bFGF (SCI Store). Cell colonies were enzymatically transplanted every 4–5 days using TrypLE Express (Thermo Fisher Scientific) at a ratio of 1:8–1:10 with the addition of 2 μg/ml ROCK inhibitor Thiazovivin (Sigma-Aldrich) for one day. All cells were cultured in a $CO_2$ incubator at 37 °C in a humid atmosphere with 5 % $CO_2$.

**Detection of endogenous alkaline phosphatase.** Cells were fixed by air drying and stained with SIGMAFAST BCIP/NBT reagent (Sigma-Aldrich) according to the manufacturer's protocol for 10–15 minutes in the dark at room temperature (RT). The preparations were washed with PBS and visualised using a Nikon Eclipse Ti-E microscope (Nikon).

**Immunofluorescence staining.** For immunofluorescence analysis, cells were placed on 8-well chambered coverslips (Thermo Fisher Scientific), fixed in 4 % paraformaldehyde (PFA, Sigma-Aldrich) for 10 minutes at RT, permeabilized with 0.5 % Triton-X 100 (Sigma-Aldrich) for 30 minutes at RT, and incubated with 1 % BSA (VWR) for 30 minutes at RT. Cells were incubated with primary antibodies overnight at 4 °C, secondary antibodies were added for 1.5 hours at RT (Table 1). The nucleus was counterstained with DAPI. Images were captured using a Nikon Eclipse Ti-E microscope and NIS Elements Advanced Research software version 4.30.

**Spontaneous differentiation of iPSCs in embryoid bodies.** In order to determine the potential of the cells to produce three germ layers, spontaneous differentiation of iPSCs through embryoid body formation was performed as previously described (Grigor'eva et al., 2023). Briefly, cells were detached with 0.15 % collagenase type IV (Thermo Fisher Scientific) and plated on 1 % agarose-coated dishes in iPSC culture medium without the addition of bFGF. After 9–14 days, the embryoid bodies were transferred to 8-well Chambered Coverglass Matrigel-coated plates (Thermo Fisher Scientific) and cultured for a further 7–9 days. They were then fixed in 4 % PFA and subjected to immunofluorescence staining. The list of antibodies is shown in Table 1.

**Karyotyping of iPSC lines.** Cells were expanded to a monolayer and plated on four wells of a 12-well tablet coated with Matrigel (Corning) extracellular matrix proteins and cultured for 48–72 hours, depending on the rate of cell proliferation. 2.5 hours before fixation, the medium was changed to fresh medium, 3 μg/ml ethidium bromide and 50 ng/ml colcemide were added and the cells were left in a $CO_2$ incubator at 37 °C. The cells in the wells were then disaggregated with 300 μl TrypLE Express and 3 ml of a hypotonic solution of 0.28 % KCl were added for 20 minutes at 37 °C. After incubation, 2 drops of Carnois fixative (3 parts methanol, 1 part glacial acetic acid) were added, the cell suspension was carefully transferred to centrifuge tubes and centrifuged at

**Table 1.** Antibodies used in the work

| Antibody | Dilution | Manufacturer, cat. No. | RRID |
|---|---|---|---|
| Markers of pluripotency | | | |
| Rabbit IgG anti-OCT4 | 1:200 | Abcam, ab18976 | RRID:AB_444714 |
| Rabbit IgG anti-NANOG | | Abcam, ab62734 | RRID:AB_956161 |
| Mouse IgG3 anti-SSEA-4 | | Abcam, ab16287 | RRID:AB_778073 |
| Mouse IgM anti-TRA-1-60 | | Abcam, ab16288 | RRID:AB_778563 |
| Markers of differentiated derivatives | | | |
| Mouse IgG2a anti-αSMA | 1:200 | Dako, M0851 | RRID:AB_2223500 |
| Mouse IgG1 anti-CD29 (Integrin beta 1) (TS2/16) | 1:100 | Thermo Fisher Scientific, 14-0299-82 | RRID:AB_1210468 |
| Chicken IgY anti-MAP2 | 1:1000 | Abcam, ab5392 | RRID:AB_2138153 |
| Mouse IgG2a anti-Tubulin β 3 (TUBB3)/ Clone: TUJ1 | | BioLegend, 801201 | RRID:AB_2313773 |
| Mouse IgG1 anti-Cytokeratin 18 (KRT18) | 1:250 | Abcam, ab668 | RRID:AB_305647 |
| Mouse IgG1 anti- HNF3β/FOXA2 | 1:50 | Santa Cruz Biotechnology, sc-374376 | RRID:AB_10989742 |
| Secondary antibodies | | | |
| Goat anti-Mouse IgG1 Alexa Fluor 568 | 1:400 | Thermo Fisher Scientific, A21124 | RRID:AB_2535766 |
| Goat anti-Mouse IgG1 Alexa Fluor 488 | | Thermo Fisher Scientific, A21121 | RRID:AB_2535764 |
| Goat anti-Chicken IgY H&L, Alexa Fluor 488 | | Abcam, ab150173 | RRID:AB_2827653 |
| Goat anti-Mouse IgG2a Cross-Adsorbed Secondary Antibody, Alexa Fluor 568 | | Thermo Fisher Scientific, A21134 | RRID:AB_2535773 |
| Goat anti-Mouse IgG2a Alexa Fluor 488 | | Thermo Fisher Scientific, A21131 | RRID:AB_2535771 |
| Goat anti-Mouse IgG3 Cross-Adsorbed Secondary Antibody, Alexa Fluor 488 | | Thermo Fisher Scientific, A21151 | RRID:AB_2535784 |
| Goat anti-Mouse IgM Heavy Chain Cross-Adsorbed Secondary Antibody, Alexa Fluor 568 | | Thermo Fisher Scientific, A21043 | RRID:AB_2535712 |
| Goat anti-Rabbit IgG (H + L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 488 | | Thermo Fisher Scientific, A11008 | RRID:AB_143165 |
| Goat anti-Rabbit IgG (H + L) Alexa Fluor 568 | | Thermo Fisher Scientific, A11011 | RRID:AB_143157 |

1,300 rpm for 7 minutes. The cells were fixed by adding 1.5 ml fresh Carnois fixative to the cell pellet for 15 minutes on ice. The cells were then centrifuged for 5 minutes at 1,300 rpm, the Carnois fixative was changed twice and 70–80 µl of the cell suspension was pipetted onto cooled wet slides from a height of 10–20 cm. The preparations were allowed to dry at room temperature.

For differential chromosome staining, the preparations were stained with DAPI solution (200 ng/ml, in 2xSSC) for 5 minutes. The preparations were then rinsed in a 2xSSC buffer and in water. After drying the preparations in air, 7–10 µl of antifade (Vector) was applied under a cover glass.

Chromosome analysis was performed under an Axioplan 2 microscope (Zeiss) equipped with a CV-M 300 CCD camera (JAI Corp.) at the Common Use Center of Microscopy of Biological Objects (https://ckp.icgen.ru/ckpmabo) at the Institute of Cytology and Genetics SB RAS. ISIS 5 software (MetaSystems Group) was used to process the metaphases.

**Isolation of genomic DNA and RNA.** Genomic DNA was isolated from PBMCs and iPSCs using the DNeasy Blood & Tissue Kit (Qiagen) or by extraction using QuickExtract DNA Extraction Solution (Lucigen). RNA was isolated using Trizol (Thermo Fisher Scientific) according to the manufacturer's protocol.

**Identification of a pathological genetic variant in patient-specific PBMCs and obtained iPSCs.** Clinical exome sequencing was performed at Genoanalytica LLC, Moscow (https://www.genoanalytica.ru), using a DNA sample from patient-specific PBMCs. Genomic DNA was ultrasonically fragmented to an average size of 300 nucleotides using the Covaris S2 instrument. After concentration measurement, 800 ng of DNA was used to prepare libraries using the NEBNext® Ultra™ II DNA Library Prep Kit for Illumina (New England Biolabs) according to the manufacturer's instructions. The resulting library was then hybridised with probes corresponding to protein-coding parts of the human

**Table 2.** Sequences of primers used in the work

| Application | Gene/locus | Product size, bp | Forward/Reverse primer (5′-3′) |
|---|---|---|---|
| Detection of episomal vectors | *oriP* | 544 | TTCCACGAGGGTAGTGAACC/<br>TCGGGGGTGTTAGAGACAAC |
| Reference gene (RT-qPCR) | *Beta-2-microglobulin* | 90 | TAGCTGTGCTCGCGCTACT/<br>TCTCTGCTGGATGACGTGAG |
| Markers of pluripotency (RT-qPCR) | *NANOG* | 116 | TTTGTGGGCCTGAAGAAAACT/<br>AGGGCTGTCCTGAATAAGCAG |
| | *OCT4* | 94 | CTTCTGCTTCAGGAGCTTGG/<br>GAAGGAGAAGCTGGAGCAAA |
| | *SOX2* | 100 | GCTTAGCCTCGTCGATGAAC/<br>AACCCCAAGATGCACAACTC |
| Mycoplasma detection | *Ribosomal 16S RNA gene* | 280 | GGGAGCAAACAGGATTAGATACCCT/<br>TGCACCATCTGTCACTCTGTTAACCTC |
| Confirmation of the mutation | *MAPT:c.2013T>G* | 427 | TCGTAAAGCCCGCTGGAAAT/<br>GTGTACGCACTCACACCACT |

genome using the Sure Select AllExome V7 Kit (Agilent) according to the instructions. After hybridisation, the library was sequenced using paired-end reads of 150 nucleotides on the HiSeq 2500 (Illumina). The raw sequencing data of the PD57 patient exome are available in the SRA database (project PRJNA563295, sample SAMN42050731, https://www.ncbi.nlm.nih.gov/biosample/42050731).

To confirm the presence of the single nucleotide polymorphism, the PCR products were Sanger sequenced using the primers listed in Table 2. The PCR was performed on a T100 thermal cycler (Bio-Rad) using HS-Taq PCR-Color (2×) BioMaster (Biolabmix) and the following programme: 95 °C – 3 min; 35 cycles: 95 °C – 30 s; 68 °C – 30 s; 72 °C – 30 s; and 72 °C – 5 min. The PCR product was purified by electrophoresis in agarose gel followed by separation using the Cleanup Mini Kit (Eurogene) according to the manufacturer's protocol. Sanger sequencing reactions were performed using the Big Dye Terminator V. 3.1. Cycle Sequencing Kit (Applied Biosystems) and analysed on the ABI 3130XL Genetic Analyser at the Genomics Core Facility SB RAS (http://www.niboch.nsc.ru/doku.php/corefacility ).

**Quantitative RT-PCR for pluripotency markers.** Reverse transcription of RNA was performed using M-MuLV Revertase (Biolabmix). Quantitative PCR was performed on a LightCycler 480 Real-Time PCR System (Roche) using the BioMaster HS-qPCR SYBR Blue 2× Kit (Biolabmix) according to the following programme: 95 °C – 5 min; 40 cycles: 95 °C – 10 s, 60 °C – 1 min. Primer sequences for pluripotency genes are shown in Table 2. CT values were normalised to beta-2-microglobulin using the ΔΔCT method.

**PCR analysis for the detection of episomes and mycoplasma.** Detection of mycoplasma contamination and presence of episome sequences in cells was performed by PCR (95 °C – 5 min; 35 cycles: 95 °C – 15 s, 60 °C – 15 s, 72 °C – 20 s) on a T100 Thermal Cycler (Bio-Rad) (Choppa et al., 1998; Okita et al., 2013). Primer sequences are listed in Table 2.

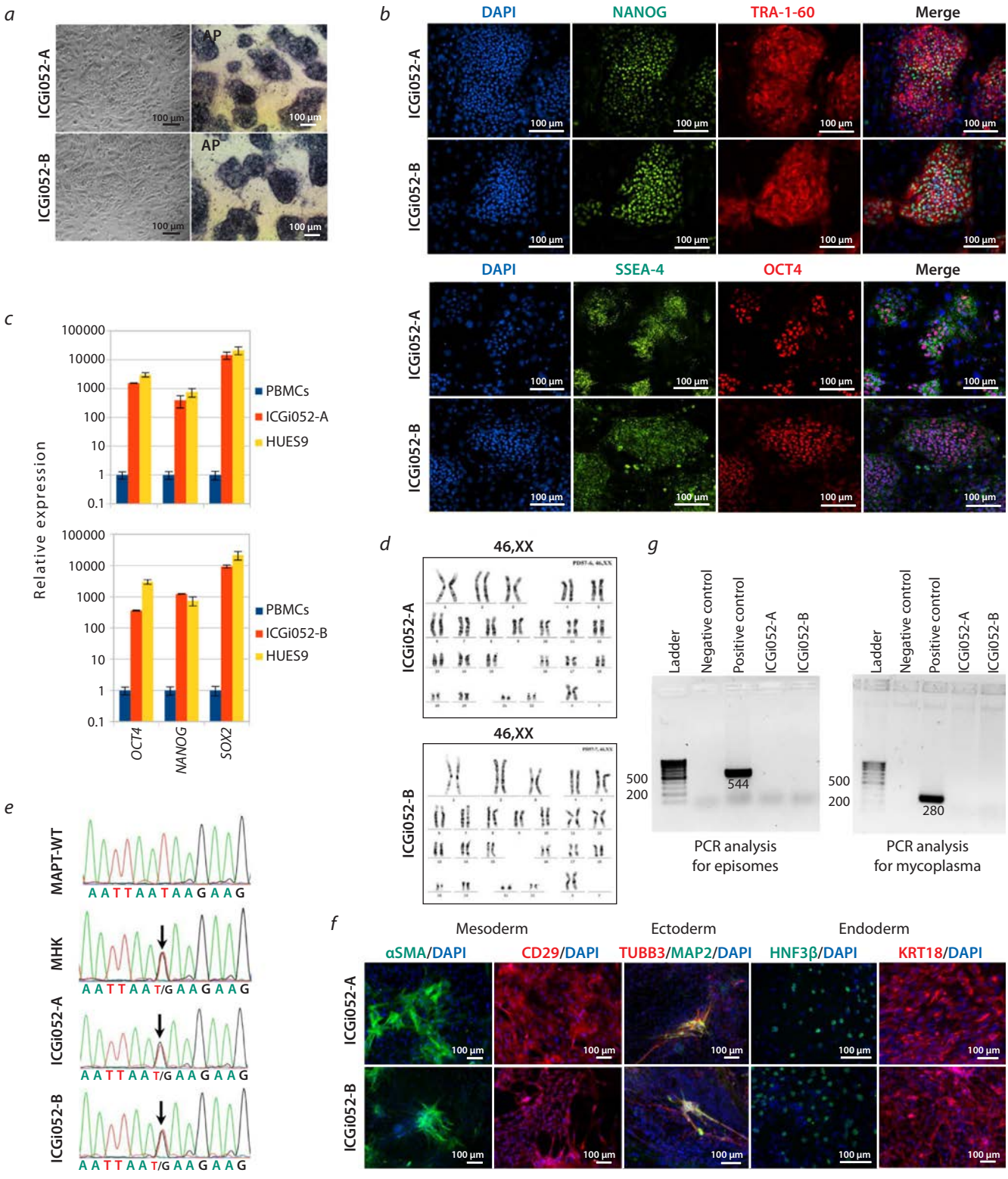
**STR analysis.** Genotyping of the studied DNA samples was performed at Genoanalytica LLC by polymerase chain reaction using a set of PCR reagents for direct amplification COrDIS EXPERT 26 (Russia) according to the manufacturer's protocol, followed by separation of the amplification products on a capillary electrophoresis device 3130 Genetic Analyzer (HITACHI, Applied Biosystems Group of The Applera Corporation, Japan, registration certificate No. FSZ 2004/1586).

## Results and discussion

Analysing the results of exome sequencing of patients from the Federal Neurosurgical Centre (Novosibirsk) with a clinical diagnosis of Parkinson's disease, a pathogenic genetic variant (c.2013T>G, rs63750756) in the *MAPT* gene associated with frontotemporal dementia with parkinsonism-17 was found in a 46-year-old patient. The first signs of parkinsonism in the patient were detected at the age of 44 years. In the family history, all female relatives had signs of the disease.

Mononuclear cells were isolated from the patient's peripheral blood and reprogrammed by transfection with non-integrating episomes expressing OCT4, SOX2, KLF4, LIN28, L-MYC and mp53DD (Okita et al., 2013). This resulted in ten lines being obtained. In the initial stages of the analysis, two lines were identified that met all the requirements for pluripotent stem cells. These lines (ICGi052-A/PD57-6 and ICGi052-B/PD57-7) have been characterized and registered in the Human Pluripotent Stem Cell Registry (hPSCreg, https://hpscreg.eu). Full information on these lines is available in hPSCreg via the links https://hpscreg.eu/cell-line/ICGi052-A and https://hpscreg.eu/cell-line/ICGi052-B.

Both lines grow in dense monolayer colonies of cells with a high nuclear-cytoplasmic ratio and express an early marker of pluripotent cells, endogenous alkaline phosphatase (see the Figure, *a*). Cultivation was performed on mitotically inactivated mouse fibroblasts (or feeder). Immunofluorescence analysis of the ICGi052-A (passage 15) and ICGi052-B (passage 16) cell lines for pluripotency markers showed the expres-




Characteristics of two patient-specific iPSC lines, ICGi052-A and ICGi052-B.

*a* – morphology of cell colonies in phase contrast and histochemical detection of endogenous alkaline phosphatase (ALP) in iPSCs; *b* – immunofluorescence staining for pluripotency markers: NANOG (green signal), TRA-1-60 (red signal), SSEA-4 (green signal), OCT4 (red signal); *c* – real-time PCR for pluripotency markers (*OCT4*, *SOX2*, *NANOG*) of the ICGi052-A and ICGi052-B iPSC lines, patient PBMCs and HUES9 embryonic stem cell lines; *d* – karyotyping (DAPI-bending) of iPSC lines: ICGi052-A at passage 18 and ICGi052-B at passage 20; *e* – sequenograms of sections of the *MAPT* gene (c.2013T>G) PBMCs from a patient with frontotemporal dementia with parkinsonism-17, ICGi052-A and ICGi052-B iPSC lines and a healthy donor (MAPT-WT) (Sanger sequencing). The polymorphism is marked with a black arrow; *f* – immunofluorescence staining of spontaneously differentiated cells in embryoid bodies for markers of three germ layers: ectoderm (TUBB3 (red signal), MAP2 (green signal)), mesoderm (aSMA (green signal), CD29 (red signal)) and endoderm (keratin 18/KRT18 (red signal), HNF3β/FOXA2 (green signal)). The nuclei are stained with DAPI (4,6-diamino-2-phenylindole) (blue signal); *g* – the result of PCR analysis for episomes and mycoplasma contamination in the obtained cell lines. All scale bars are 100 µm.

**Table 3.** Passport of the ICGi052-A and ICGi052-B cell lines

| | |
|---|---|
| Unique identifier | ICGi052-A, ICGi052-B |
| Alternative cell line names | PD57-6, PD57-7 |
| Institution | Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia |
| Ethical approval | The study was approved by the Research Ethics Committee of the Federal Neurosurgical Center (Novosibirsk, Russia), Protocol No. 1 dated 14 March 2017 |
| Cell type | iPSCs |
| Origin | Human |
| Additional information on the origin of the cell line | Age: 46<br>Gender: F<br>Ethnicity: Caucasian race |
| Original cell type | Peripheral blood mononuclear cells (PBMCs) |
| Date of biomaterial collection | 2021 |
| Reprogramming method | Non-integrating episomal vectors |
| Reprogramming factors | OCT4, SOX2, KLF4, LIN28, L-MYC and mp53DD |
| Clonality | Clonal |
| Genetic modification | No |
| Type of genetic modification | N/A |
| Confirmation of elimination/ silencing of reprogramming transgenes | PCR, not detected |
| Disease | Frontotemporal dementia with parkinsonism-17 |
| Gene/locus | *MAPT:c.2013T>G*, rs63750756 |
| Morphology | Human pluripotent cell-like monolayer colonies |
| Pluripotency | Confirmed in tests for the formation of embryoid bodies |
| Karyotype | 46,XX |
| Check for contamination | Bacteria, fungi and mycoplasma not detected |
| Scope of application | *In vitro* model of frontotemporal dementia with parkinsonism-17 |
| Cultivation method | On the feeder layer of mitotically inactivated mouse embryonic fibroblasts |
| Growth medium | 85 % KnockOut DMEM, 15 % KnockOut Serum Replacement, 0.1 mM NA, 0.1 mM 2-mercaptoethanol, 1 % Penicillin-Streptomycin, GlutaMAX-I (all Thermo Fisher Scientific), 10 ng/ml bFGF (SCI Store) |
| Temperature, ℃ | 37 |
| $CO_2$ concentration, % | 5 |
| $O_2$ concentration, % | 20 |
| Passage method | EDTA 0.5 mM |
| Split ratio | 1:8–1:10 |
| Cryopreservation | 90 % FBS, 10 % DMSO |
| Storage conditions | Liquid nitrogen |
| Registry ID | https://hpscreg.eu/cell-line/ICGi052-A, https://hpscreg.eu/cell-line/ICGi052-B |
| Date of certification/registration | 09/07/2024 |

sion of the surface antigens SSEA-4 and TRA-1-60 and the transcription factors NANOG and OCT4 (see the Figure, *b*). Quantitative real-time PCR (RT-qPCR) of the lines at passage 15 also showed a significant increase in the expression level of the *OCT4*, *NANOG* and *SOX2* genes, comparable to the expression level in the control line of HUES9 embryonic stem cells (HVRDe009-A) (Cowan et al., 2004) (see the Figure, *c*). Karyotype analysis, involving 60 metaphase plates from each cell line, showed that cell lines ICGi052-A at passage 18 and ICGi052-B at passage 20 have normal diploid 46,XX karyotypes (see the Figure, *d*). In order to confirm the presence of a pathogenic polymorphism in the cell lines obtained, sequencing by the Sanger method was carried out and showed that both lines, as well as the patient's PBMCs, carry the substitution c.2013T>G (see the Figure, *e*, SNPs are indicated by black arrows).

The main test for the pluripotency of the cell lines obtained is their ability to differentiate into derivatives of three germ layers, for which spontaneous differentiation through embryoid bodies was performed, followed by immunofluorescence staining for specific markers. Both cell lines were shown to be capable of producing: ectoderm (tubulin β3 (TUBB3/TUJ1), microtubule-associated protein 2 (MAP2)); endoderm (hepatocyte nuclear factor 3 beta (HNF3β/FOXA2), keratin 18 (KRT18)); mesoderm (smooth muscle α-actin (aSMA), surface marker CD29) (see the Figure, *f*).

During cultivation, all lines were subjected to a PCR test for mycoplasma contamination, which confirmed its absence (see the Figure, *g*). It was also shown that the episomal vectors were eliminated by passage 19. STR analysis of short tandem repeats at 25 polymorphic loci at passages 15 (for ICGi052-A) and 16 (for ICGi052-B) showed the identity of patient-specific PBMCs (data available on request from the authors). The passport of the cell lines obtained is shown in Table 3.

## Conclusion

Using reprogramming into a pluripotent state, we have obtained and characterised in detail two lines of iPSCs – ICGi052-A and ICGi052-B – from the PBMCs of a patient with frontotemporal dementia with parkinsonism-17, who has a pathological genetic variant c.2013T>G (rs63750756) in the *MAPT* gene. These cell lines meet all the criteria of pluripotent cells (they have a diploid karyotype, express pluripotency markers and are capable of producing derivatives of three germ layers) and are a unique tool for the *in vitro* study of the development of pathology in neural derivatives obtained by directed differentiation of iPSCs, as well as for testing new pharmacological compounds. The cell lines obtained are registered in the Human Pluripotent Stem Cell Registry of hPSCreg and are also included in the cell collection of the Laboratory of Developmental Epigenetics of IC&G SB RAS and can be used to study the pathological mechanisms of the progression of taupathies.

## References

Britti E., Ros J., Esteras N., Abramov A.Y. Tau inhibits mitochondrial calcium efflux and makes neurons vulnerable to calcium-induced cell death. *Cell Calcium*. 2020;86:102150. DOI 10.1016/j.ceca.2019.102150

Choppa P.C., Vojdani A., Tagle C., Andrin R., Magtoto L. Multiplex PCR for the detection of *Mycoplasma fermentans*, *M. hominis* and *M. penetrans* in cell cultures and blood samples of patients with chronic fatigue syndrome. *Mol. Cell Probes.* 1998;12(5):301-308. DOI 10.1006/mcpr.1998.0186

Cowan C.A., Klimanskaya I., McMahon J., Atienza J., Witmyer J., Zucker J.P., Wang S., Morton C.C., McMahon A.P., Powers D., Melton D.A. Derivation of embryonic stem-cell lines from human blastocysts. *N. Engl. J. Med.* 2004;350(13):1353-1356. DOI 10.1056/NEJMsr040330

Dawson H.N., Cantillana V., Chen L., Vitek M.P. The tau N279K exon 10 splicing mutation recapitulates frontotemporal dementia and parkinsonism linked to chromosome 17 tauopathy in a mouse model. *J. Neurosci*. 2007;27(34):9155-9168. DOI 10.1523/JNEUROSCI.5492-06.2007

D'Souza I., Schellenberg G.D. Arginine/serine-rich protein interaction domain-dependent modulation of a tau exon 10 splicing enhancer: altered interactions and mechanisms for functionally antagonistic FTDP-17 mutations Δ280K and N279K. *J. Biol. Chem.* 2006;281:2460-2469. DOI 10.1074/jbc.M505809200

Esmaeli-Azad B., McCarty J.H., Feinstein S.C. Sense and antisense transfection analysis of tau function: tau influences net microtubule assembly, neurite outgrowth and neuritic stability. *J. Cell Sci*. 1994;107(4):869-879. DOI 10.1242/jcs.107.4.869

Esteras N., Kundel F., Amodeo G.F., Pavlov E.V., Klenerman D., Abramov A.Y. Insoluble tau aggregates induce neuronal death through modification of membrane ion conductance, activation of voltage-gated calcium channels and NADPH oxidase. *FEBS J*. 2021;288(1):127-141. DOI 10.1111/febs.15340

Ghetti B., Oblak A.L., Boeve B.F., Johnson K.A., Dickerson B.C., Goedert M. Invited review: Frontotemporal dementia caused by *microtubule-associated protein tau gene* (*MAPT*) mutations: a chameleon for neuropathology and neuroimaging. *Neuropathol. Appl. Neurobiol*. 2015;41(1):24-46. DOI 10.1111/nan.12213

Grigor'eva E.V., Kopytova A.E., Yarkova E.S., Pavlova S.V., Sorogina D.A., Malakhova A.A., Malankhanova T.B., Baydakova G.V., Zakharova E.Y., Medvedev S.P., Pchelina S.N., Zakian S.M. Biochemical characteristics of iPSC-derived dopaminergic neurons from N370S *GBA* variant carriers with and without Parkinson's disease. *Int. J. Mol. Sci.* 2023;24(5):4437. DOI 10.3390/ijms24054437

Hasegawa M., Smith M.J., Iijima M., Tabira T., Goedert M. FTDP-17 mutations N279K and S305N in tau produce increased splicing of exon 10. *FEBS Lett.* 1999;443(2):93-96. DOI 10.1016/S0014-5793(98)01696-2

Hernández F., Merchán-Rubira J., Vallés-Saiz L., Rodríguez-Matellán A., Avila J. Differences between human and murine tau at the N-terminal end. *Front. Aging Neurosci*. 2020;12:11. DOI 10.3389/fnagi.2020.00011

Iovino M., Agathou S., González-Rueda A., Del Castillo Velasco-Herrera M., Borroni B., Alberici A., Lynch T., O'Dowd S., Geti I., Gaffney D., Vallier L., Paulsen O., Káradóttir R.T., Spillantini M.G. Early maturation and distinct tau pathology in induced pluripotent stem cell-derived neurons from patients with MAPT mutations. *Brain*. 2015;138(11):3345-3359. DOI 10.1093/brain/awv222

Korn L., Speicher A.M., Schroeter C.B., Gola L., Kaehne T., Engler A., Disse P., Fernández-Orth J., Csatári J., Naumann M., Seebohm G., Meuth S.G., Schöler H.R., Wiendl H., Kovac S., Pawlowski M. MAPT genotype-dependent mitochondrial aberration and ROS production trigger dysfunction and death in cortical neurons of patients with hereditary FTLD. *Redox Biol*. 2023;59:102597. DOI 10.1016/j.redox.2022.102597

Liu G., David B.T., Trawczynski M., Fessler R.G. Advances in pluripotent stem cells: history, mechanisms, technologies, and applications. *Stem Cell Rev. Rep*. 2020;16(1):3-32. DOI 10.1007/s12015-019-09935-x

Lynch T., Sano M., Marder K.S., Bell K.L., Foster N.L., Defendini R.F., Sima A.A., Keohane C., Nygaard T.G., Fahn S., Mayeux R., Rowland L., Wilhelmsen K. Clinical characteristics of a family-with chromosome 17-linked disinhibition-dementia-parkinsonism-amyotrophy complex. *Neurology*. 1994;44(10):1878-1884. DOI 10.1212/wnl.44.10.1878

Okita K., Yamakawa T., Matsumura Y., Sato Y., Amano N., Watanabe A., Goshima N., Yamanaka S. An efficient nonviral method to generate integration-free human-induced pluripotent stem cells from cord blood and peripheral blood cells. *Stem Cells.* 2013;31(3):458-466. DOI 10.1002/stem.1293

Ritter M.L., Avila J., García-Escudero V., Hernández F., Pérez M. Frontotemporal dementia-associated N279K tau mutation localizes at the nuclear compartment. *Front. Cell. Neurosci*. 2018;12:202. DOI 10.3389/fncel.2018.00202

Valetdinova K.R., Malankhanova T.B., Zakian S.M., Medvedev S.P. The cutting edge of disease modeling: synergy of induced pluripotent stem cell technology and genetically encoded biosensors. *Biomedicines*. 2021;9(8):960. DOI 10.3390/biomedicines9080960

Wren M.C., Zhao J., Liu C.-C., Murray M.E., Atagi Y., Davis M.D., Fu Y., Okano H.J., Ogaki K., Strongosky A.J., Tacik P., Rademakers R., Ross O.A., Dickson D.W., Wszolek Z.K., Kanekiyo T., Bu G. Frontotemporal dementia-associated N279K tau mutant disrupts subcellular vesicle trafficking and induces cellular stress in iPSC-derived neural stem cells. *Mol. Neurodegener*. 2015;10:46. DOI 10.1186/s13024-015-0042-7

**Conflict of interest.** The authors declare no conflict of interest.
Received July 15, 2024. Revised August 28, 2024. Accepted September 2, 2024.

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2024;28(7):688-695

DOI 10.18699/vjgb-24-77

# Comparative analysis of the primary structure and production of recombinant poly(ADP-ribose)polymerase 1 of long-lived *Heterocephalus glaber*

K.N. Naumenko [1], A.R. Nurislamov [2], K.D. Nazarov[1], V.S. Fishman [2], A.A. Popov [1], I.O. Petruseva [1], A.N. Evdokimov[1], O.I. Lavrik [1, 3]

[1] Institute of Chemical Biology and Fundamental Medicine of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
[2] Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
[3] Novosibirsk State University, Novosibirsk, Russia

lavrik@niboch.nsc.ru

**Abstract.** DNA repair is a most important cellular process that helps maintain the integrity of the genome and is currently considered by researchers as one of the factors determining the maximum lifespan. The central regulator of the DNA repair process is the enzyme poly(ADP-ribose)polymerase 1 (PARP1). PARP1 catalyzes the synthesis of poly(ADP-ribose) polymer (PAR) upon DNA damage using nicotinamide adenine dinucleotide ($NAD^+$) as a substrate. This polymer covalently attaches to PARP1, which leads to its dissociation from the complex with damaged DNA and stimulation of the repair process. Despite intensive research on PARP1, its properties as an isolated protein have not been practically studied in mammals that demonstrate a long maximum lifespan, such as, for example, the naked mole rat (*Heterocephalus glaber*). High activity of DNA repair systems is observed in the cells of the naked mole rat, which ensures their high resistance to oxidative stress, as well as to genotoxic effects. The revealed features may be due to the high activity of PARP1 in the cells of the naked mole rat; however, this issue remains poorly understood and, thus, requires more detailed research, including one with the use of isolated protein PARP1 of the naked mole rat, the isolation and characterization of which have not been carried out before. In the present work, the amino acid sequence of PARP1 of the naked mole rat is compared with the amino acid sequences of orthologous proteins of other mammals. In contrast to human PARP1, 13 evolutionarily conservative amino acid substitutions in various functional domains of the protein have been identified in the amino acid sequence of naked mole rat PARP1. Using the cDNA of the naked mole rat's *Parp1* gene, a vector was created for the expression of the target protein in *Escherichia coli* cell culture. For the first time, a detailed description of the procedure for the expression and purification of the recombinant protein PARP1 of the long-lived naked mole rat is presented. In addition, poly(ADP-ribose)polymerase activity of the obtained protein was evaluated. The results presented in this paper are the basis for further detailed characterization of the properties of purified recombinant naked mole rat PARP1.

**Key words:** poly(ADP-ribose)polymerase 1; DNA repair; *Heterocephalus glaber*.

**For citation:** Naumenko K.N., Nurislamov A.R., Nazarov K.D., Fishman V.S., Popov A.A., Petruseva I.O., Evdokimov A.N., Lavrik O.I. Comparative analysis of the primary structure and production of recombinant poly(ADP-ribose)polymerase 1 of long-lived *Heterocephalus glaber*. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2024;28(7):688-695. DOI 10.18699/vjgb-24-77

**Funding.** The work was carried out with the financial support of the RSF (project No. 19-74-10056P).

# Сравнительный анализ первичной структуры и получение рекомбинантной поли(ADP-рибоза)полимеразы 1 долгоживущего *Heterocephalus glaber*

К.Н. Науменко [1], А.Р. Нурисламов [2], К.Д. Назаров[1], В.С. Фишман [2], А.А. Попов [1], И.О. Петрусева [1], А.Н. Евдокимов[1], О.И. Лаврик [1, 3]

[1] Институт химической биологии и фундаментальной медицины Сибирского отделения Российской академии наук, Новосибирск, Россия
[2] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
[3] Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия

lavrik@niboch.nsc.ru

**Аннотация.** Репарация ДНК – важнейший клеточный процесс, который способствует поддержанию целостности генома. В настоящее время эффективная работа систем репарации ДНК рассматривается исследователями как один из ключевых факторов, определяющих максимальную продолжительность жизни. Центральным

регулятором процесса репарации ДНК является фермент поли(ADP-рибоза)полимераза 1 (PARP1), способный синтезировать полимер поли(ADP-рибозы) (PAR) в ответ на повреждение ДНК и присоединять его к белкам-мишеням, в число которых входит и сам PARP1, осуществляя тем самым посттрансляционную модификацию этих белков и регулируя их сродство к ДНК. PARP1 принимает участие и во многих других процессах, ассоциированных с клеточным старением, таких как поддержание целостности теломер и развитие воспалительной реакции. Свойства PARP1 как изолированного белка практически не исследовались у млекопитающих, которые демонстрируют высокую максимальную продолжительность жизни, за исключением человека. Одним из перспективных объектов таких исследований считается голый землекоп (*Heterocephalus glaber*), имеющий экстремально высокую максимальную продолжительность жизни, а также более эффективно функционирующие системы репарации ДНК, которые обеспечивают высокую устойчивость его клеток к воздействию ряда генотоксических агентов, по сравнению с другими мелкими грызунами, например, близкой по размеру и массе тела мышью (*Mus musculus*). В настоящей работе проведено сравнение аминокислотной последовательности PARP1 голого землекопа с аминокислотными последовательностями белков-ортологов других млекопитающих. В отличие от PARP1 человека, в аминокислотной последовательности PARP1 голого землекопа выявлено 13 эволюционно консервативных аминокислотных замен в различных функциональных доменах белка. С использованием поиска в базах данных последовательности кДНК гена *Parp1* голого землекопа и последующего анализа путем выравнивания транскриптомных данных выбрана соответствующая экспрессируемому варианту *Parp1* последовательность кДНК, которая была клонирована с помощью экспрессионного вектора на основе плазмиды pLate31. В результате экспрессии в штамме *Escherichia coli* BL21(DE3)GeneX и очистки, проведенной с использованием трех хроматографических стадий, впервые был получен и охарактеризован функционально активный фермент PARP1 голого землекопа.

**Ключевые слова:** поли(ADP-рибоза)полимераза 1; репарация ДНК; *Heterocephalus glaber*.

## Introduction

Genomic instability is a principal contributor to aging, arising from the accumulation of DNA damage due to exogenous and endogenous factors (López-Otín et al., 2023). DNA repair mechanisms are crucial in mitigating these detrimental effects and preserving genome integrity. The efficiency of DNA repair systems is linked to increased longevity (Schumacher et al., 2021). Consequently, current research focuses on elucidating the mechanisms and functionality of DNA repair systems in mammalian cells with notable life expectancies.

A particularly promising model for such research is the naked mole rat (*Heterocephalus glaber*), which exhibits a significantly longer lifespan compared to the mouse (*Mus musculus*) of similar size and body weight (Buffenstein, 2005; Gorbunova et al., 2014). Comparative studies have demonstrated that naked mole rat cells exhibit heightened DNA repair activity and increased resistance to certain genotoxic agents (methane methyl sulfonate, paraquat, etoposide, etc.) relative to mouse cells (Salmon et al., 2008; MacRae et al., 2015; Evdokimov et al., 2018, 2021). To uncover the underlying reasons for these attributes, studies on isolated mole rat proteins involved in DNA repair are imperative.

Poly(ADP-ribose) polymerase 1 (PARP1) is the central regulator of DNA repair in mammalian cells, catalyzing the synthesis of poly(ADP-ribose) (PAR) using $NAD^+$ as a substrate. This process includes auto-poly(ADP-ribosyl)ation (autoPARylation) and the modification of other acceptor proteins through trans-poly(ADP-ribosyl)ation. PARP1 modulates the activity of repair enzymes, their interaction with damaged DNA, and their recruitment to DNA damage sites (Sinha et al., 2021; Bilkis et al., 2023; Rouleau-Turcotte, Pascal, 2023). Furthermore, PARP1 contributes to the formation of non-membrane compartments, concentrating damaged DNA and repair proteins to enhance the DNA repair process (Singatulina et al., 2019; Leung, 2020; Alemasova, Lavrik, 2022). These functions signify that PARP1 is a crucial factor in regulating DNA repair efficiency and ensuring genomic stability.

An early study examining this relationship found a positive correlation between PARP activity and lifespan across thirteen mammalian species, with humans exhibiting the longest maximum lifespan (Grube, Bürkle, 1992). Subsequent studies of the PARylation reaction kinetics, catalyzed by recombinant PARP1 from humans (*Homo sapiens*) and short-lived gray rats (*Rattus norvegicus*), indicated superior efficiency of the human enzyme (Beneke et al., 2000, 2010).

Additionally, comparative studies of DNA repair efficiency in naked mole rat and mouse cells revealed higher PAR synthesis activity and greater PARP1 protein levels in naked mole rat cells, as evidenced by covalent DNA attachment (Evdokimov et al., 2018). Investigating the properties of naked mole rat PARP1, and comparing them with those from other species, is therefore of significant interest. However, prior studies have not examined naked mole rat PARP1 as an isolated protein due to the unavailability of individual protein preparations.

This study aims to obtain recombinant naked mole rat PARP1 in order to investigate its properties and compare them with PARP1 from other mammals. To achieve this, a comparative analysis of the amino acid sequence of naked mole rat PARP1 with orthologous proteins from other mammals was conducted. Bioinformatics analysis identified the cDNA sequence corresponding to the expressed variant of the naked mole rat *Parp1* gene, which was subsequently cloned into the pLate31 plasmid vector. This facilitated the expression, purification, and characterization of recombinant PARP1 from this long-lived rodent in *E. coli* cells for the first time.

Sequences of oligodeoxynucleotides and primers used in the work

| Name | Sequence (5′-3′) |
|---|---|
| Oligodeoxynucleotide 1 | ggcgataaagttggg |
| Oligodeoxynucleotide 2 | aacgtcagggtcttcc |
| Oligodeoxynucleotide 3 | ggaagaccctgacgttcccaactttatcgcc |
| pLate31-PARP1-For | agaaggagatataactatgatggccgaggcagcggac |
| pLate31-PARP1-Rev | gtggtggtgatggtgatggccccacagggaggacttaaaattgaac |

## Materials and methods

**Oligodeoxynucleotides.** The sequences of the oligodeoxynucleotides used in this study are presented in the Table. Oligodeoxynucleotides 1–3 were synthesized at the Laboratory of Synthetic Biology, Institute of Chemical Biology and Fundamental Medicine of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia (Novosibirsk, Russia). Primers pLate31-PARP1-For and pLate31-PARP1-Rev were synthesized by DNA-Sintez LLC (Moscow, Russia).

**Evolutionary analysis of the primary structure of naked mole rat PARP1.** To identify amino acid substitutions unique to naked mole rat PARP1, multiple sequence alignments of PARP1 amino acid sequences from this species and other mammals were conducted. The analysis included PARP1 orthologs from ten animal species: three rodents (*Fukomys damarensis*, *M. musculus*, *R. norvegicus*) and six other mammals (*H. sapiens*, *Equus caballus*, *Dasypus novemcinctus*, *Loxodonta africana*, *Monodelphis domestica*, *Ornithorhynchus anatinus*). PARP1 amino acid sequences were retrieved from the NCBI and Ensembl databases. Multiple alignment was performed using Clustal Omega (https://www.ebi.ac.uk/jdispatcher/msa/clustalo).

**Cell cultivation.** Naked mole rat skin fibroblasts (NSF8 line) were cultured in αMEM medium supplemented with 15 % FBS, 10 % AmnioMAX, 0.005 μg/ml bFGF, and an antibiotic/antimycotic mixture (Gibco, USA) at 32 °C and 5 % $CO_2$.

**Isolation of total RNA from naked mole rat fibroblasts and preparation of cDNA.** The resulting cell culture was washed with 5 ml of PBS to remove any remaining medium and 1 ml of TRIzol solution (Thermo Fisher Scientific, USA) was added. The cells were resuspended to a homogeneous suspension and transferred to a clean tube. 200 μl of chloroform was added, the tube was incubated for 5 minutes at room temperature and centrifuged for 15 minutes at 16,000 g at 4 °C. After centrifugation, the upper aqueous phase was collected into a clean test tube. The resulting sample was reprecipitated with isopropyl alcohol. The precipitated RNA was dissolved in 200 μl of water and an equal volume of phenol : chloroform : isoamyl alcohol (25 : 24 : 1) was added. After centrifugation, the aqueous phase was collected and the RNA was reprecipitated with ethanol.

To produce *Parp1* cDNA, a reaction mixture with a volume of 10 μl containing 4 μg of total RNA and 100 nmol of oligo dT was incubated for 2 min at 70 °C, after which the reverse transcription buffer and 1 μl of RT-MMLV reverse transcriptase (100 units of activity/μl; Biolabmix, Russia) were added. The reaction was carried out for an hour at 42 °C.

**Construction of a PARP1 expression vector.** A PCR product encoding the translated region of naked mole rat PARP1 cDNA, flanked by specific nucleotide sequences, was generated using the primer pair pLate31-PARP1-For/pLate31-PARP1-Rev (see the Table). The product was reprecipitated with 96 % ethanol, dissolved in 10 μl of LIC buffer, and treated with 1 μl of T4 phage DNA polymerase (1 U/μl; Thermo Fisher Scientific, USA). After thorough mixing and a 5-minute incubation at room temperature, the reaction was halted by adding EDTA to a final concentration of 50 mM. Subsequently, 20 fmol of the linearized pLate31 vector (Thermo Scientific, USA) with complementary "sticky" ends was added, mixed, and the mixture was incubated for 5 minutes at room temperature. The resulting plasmid DNA was used to transform *E. coli* XLBlue cells.

**Determination of the total level of poly(ADP-ribose) synthesized in the autoPARylation reaction.** Reaction mixtures (10 μl) containing 100, 200, or 400 nM recombinant PARP1 protein, 100 nM 32 bp DNA duplex with a single-strand break (from hybridization of oligodeoxynucleotides 1–3, see Table 1), 400 μM $NAD^+$, and [$^{32}$P]-labeled $NAD^+$ (0.4 μCi) were prepared. The reaction commenced with the addition of $NAD^+$ and the mixtures were incubated at 37 °C for 10 minutes. The reaction was terminated by applying the mixture to chromatographic paper (GE Healthcare, USA) pre-treated with 10 % TCA. Unincorporated [$^{32}$P]-$NAD^+$ was removed through successive washes with 5 % TCA and ethanol. The paper was dried, and PAR synthesis levels were assessed via autoradiography using Typhoon FLA 7000 (GE Healthcare, USA).

**Isolation and purification of recombinant PARP1 protein.** *E. coli* BL21(DE3)GeneX cells transformed with the pLate31-PARP1 plasmid were incubated in an autoinduction system (Studier method) using LB medium containing 50 mM $Na_2HPO_4$, 50 mM $KH_2PO_4$, 25 mM $(NH_4)_2SO_4$, 2 mM $MgSO_4$, 0.5 % glycerol, 0.05 % glucose, 0.2 % lactose, and ampicillin (100 μg/ml) for 18 hours at 37 °C. Post-incubation, cells were pelleted by centrifugation at 3,000 g, the supernatant was collected, and the cell pellet was stored at –70 °C.

For cell lysate preparation, the biomass was resuspended in buffer (20 mM Tris-HCl pH 8.0, 10 % glycerol, 2 mM 2-mercaptoethanol, 10 mM imidazole, 0.5 mM PMSF, pro-

tease inhibitors) at 5 ml buffer per 1 g cell biomass. After 20 minutes on ice, an equal volume of buffer (20 mM Tris-HCl pH 8.0, 2 M NaCl, 2 % NP-40, 10 % glycerol, 2 mM 2-mercaptoethanol, 0.5 mM PMSF, protease inhibitors) was added. The suspension was ultrasonically processed at 40 kHz for 20 minutes at 4 °C, followed by centrifugation at 30,000 g for 30 minutes in a Beckman JA 25.50 rotor.

The clarified lysate was passed through a Ni-NTA-agarose column (GE Healthcare, USA), equilibrated with buffer (20 mM Tris-HCl pH 8.0, 1 M NaCl, 2 mM 2-mercaptoethanol, 10 % glycerol, 5 mM imidazole). The column was washed with equilibration buffer and subsequently with buffer (20 mM Tris-HCl pH 8.0, 0.1 M NaCl, 2 mM 2-mercaptoethanol, 10 % glycerol, 5 mM imidazole) until baseline stabilization. Elution was performed using 250 mM imidazole buffer. Fractions containing the target protein were pooled and applied to a heparin-Sepharose column (GE Healthcare, USA). Chromatographic separation was performed isocratically with 0.3 M NaCl buffer for washing off weakly bound proteins and 1 M NaCl buffer for eluting the target protein. Fractions with the target protein were pooled and diluted 10-fold with buffer (20 mM Tris-HCl pH 8.0, 7 mM 2-mercaptoethanol, 10 % glycerol) before application to an ssDNA cellulose column (Sigma, USA). Chromatographic separation on ssDNA cellulose followed similar conditions to heparin-Sepharose separation.

## Results and discussion

To determine whether naked mole rat PARP1 contains evolutionarily conserved amino acid substitutions that could influence its functional properties, we compared its amino acid sequence with those of orthologous proteins in other mammals. The PARP1 sequence is highly conserved among mammals, demonstrating over 90 % homology despite divergence over 150 million years (Fig. 1*a*). Naked mole rat PARP1 retains all functional domains found in other mammals, though several substitutions were identified in highly conserved functional domain sites (Fig. 1*b*). Some of these substitutions also appear in *F. damarensis*, a related species in the Bathyergidae family. These substitutions may impact PARP1's ability to recognize damaged DNA and its catalytic functions.

As a result of searching the cDNA sequence of the naked mole rat *Parp1* gene in databases and subsequent analysis by aligning transcriptomic data for various organs of the naked mole rat (brain: SRS899007; testicles: SRR1959204; liver: ERS1090459) on three alternative *Parp1* templates, the cDNA sequence (NCBI NM_001310226.1 (Bens et al., 2016)) corresponding to the expressed variant of *Parp1* was chosen. We selected this cDNA for amplification and subsequent cloning.

To produce recombinant naked mole rat PARP1 in *E. coli* cells, an expression vector based on the pLate31 plasmid (Thermo Scientific, USA) was used. Specific primers and total cDNA from naked mole rat fibroblasts were used to amplify

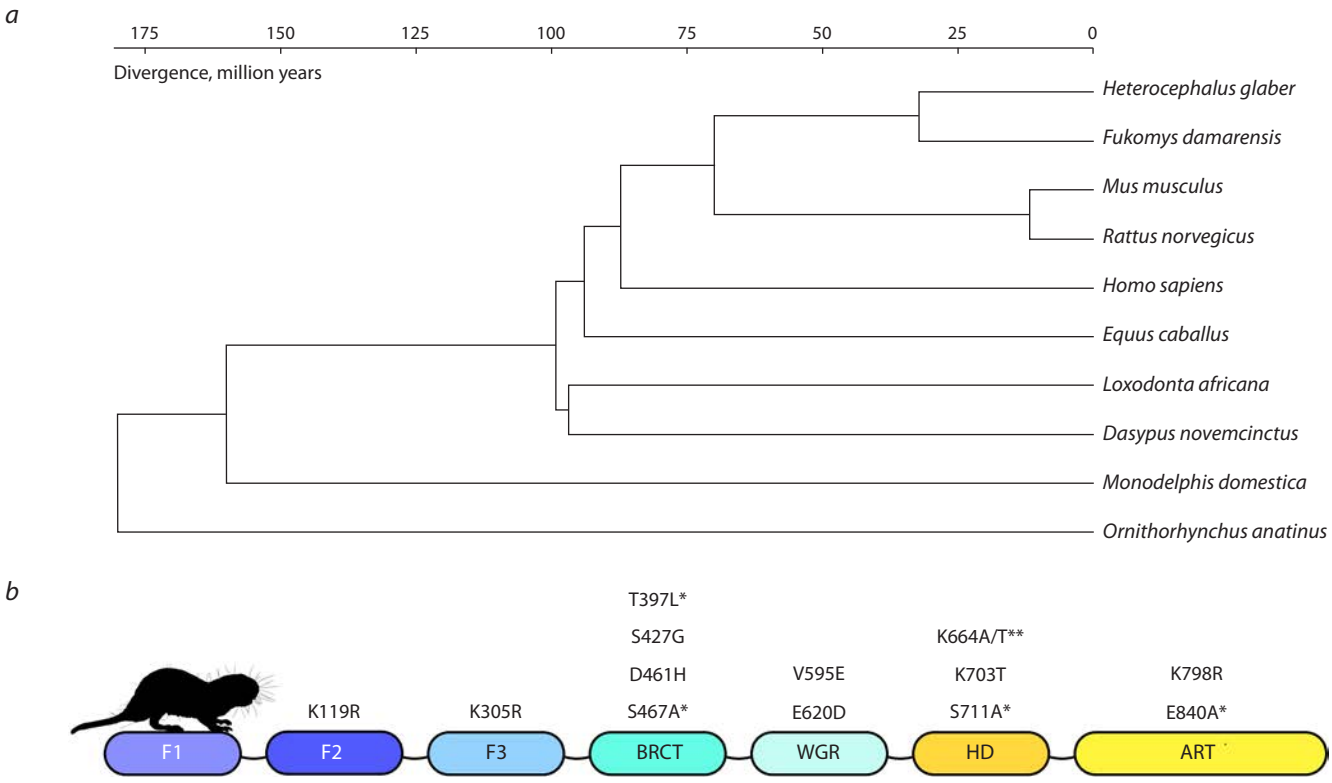



**Fig. 1.** Evolutionary analysis of the primary structure of the PARP1 protein.

*a* – list of mammal species included in the analysis and their divergence; *b* – amino acid substitutions in PARP1 unique to the naked mole rat. * – substitutions also present in the Damaraland mole rat; ** – substitution options from the NCBI and Ensembl databases.

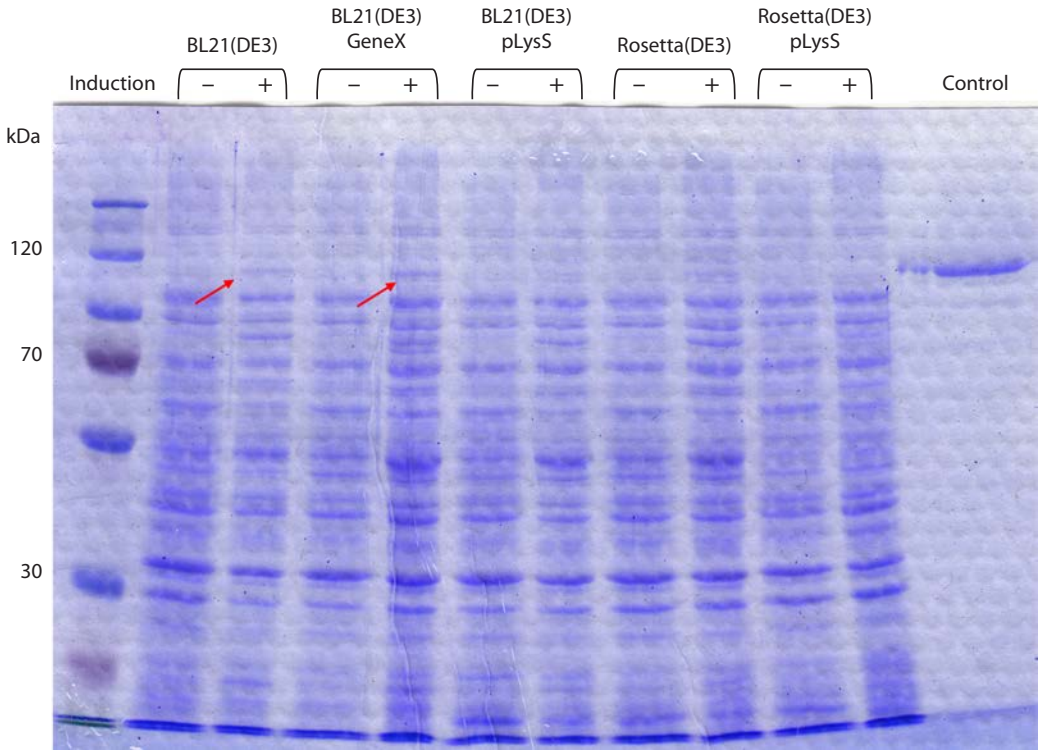



**Fig. 2.** Analysis of PARP1 content in lysates of transformed *E. coli* cells.

Cells were cultured in an autoinduction system at 37 °C with and without lactose induction. The position of the PARP1 protein is indicated by red arrows. Control – recombinant human PARP1 protein.

the coding sequence of PARP1 via PCR. The PCR product was annealed with the linearized pLate31 vector, and *E. coli* XLBlue cells were transformed to amplify plasmid DNA. The absence of errors in the amplified sequence was confirmed by Sanger sequencing.

When searching for optimal conditions, various *E. coli* strains (BL21(DE3), BL21(DE3)plysS, BL21(DE3)GeneX, Rosetta(DE3), and Rosetta(DE3)plysS) were tested for optimal PARP1 expression conditions. Target protein expression was visually detected in induced BL21(DE3) and BL21(DE3) GeneX cells (Fig. 2).

The resulting cultivation conditions were used to produce a preparative amount of biomass from BL21(DE3)GeneX cells transformed with the pLate31-PARP1 vector. The resulting biomass was lysed, followed by treatment in an ultrasonic disintegrator and centrifuged to sediment debris. Next, three chromatographic purification stages were carried out sequentially using columns containing Ni-NTA (Fig. 3*a*), heparin-Sepharose (Fig. 3*b*), and ssDNA cellulose (Fig. 3*c*) as a sorbent (Sukhanova et al., 2004).

The presence of the target protein was monitored by electrophoretic analysis with Laemmli staining (Fig. 4). Fractions containing the purified protein preparation were concentrated using Amicon 10 kDa and the purity of the purified preparation was shown by electrophoresis followed by Coomassie R250 staining.

The protein concentration in the final preparation, determined by the Bradford method, was 0.5 mg/ml. The total yield was 0.3 mg of protein per 10 g of *E. coli* cell biomass.

To test the activity of the resulting recombinant protein, we used an *in vitro* system containing radiolabeled $NAD^+$ and model damaged DNA containing a break and free blunt ends as a cofactor to activate the PAR synthesis reaction catalyzed by PARP1 (Fig. 5). As can be seen from the data presented, the isolated protein has enzymatic activity in the autoPARylation reaction and is suitable for further study of its properties.

Comparative studies of DNA repair systems in naked mole rat and mouse (*M. musculus*) cells showed that naked mole rat cells have more effective base excision repair (BER) and nucleotide excision repair (NER) systems than mouse cells (Evdokimov et al., 2018). PARP1 activity was also significantly higher in long-lived naked mole rat cells compared to short-lived mouse cells (Evdokimov et al., 2018). Future work will involve determining the nature of the interaction between isolated naked mole rat PARP1 and partner proteins in DNA repair, the influence of these proteins on PARP1 activity, and PARP1's affinity for damaged DNA.

Production of recombinant proteins like PARP1 requires selecting optimal conditions for production, isolation, and purification, which can differ from standard methods. The described procedures enabled successful cloning, production in the *E. coli* expression system, and chromatographic purification of naked mole rat PARP1. The proposed chromatographic protein purification procedure, including Ni-NTA chromatography and two "pseudo-affinity" columns under specific buffer and salt conditions, can effectively purify recombinant naked mole rat PARP1.

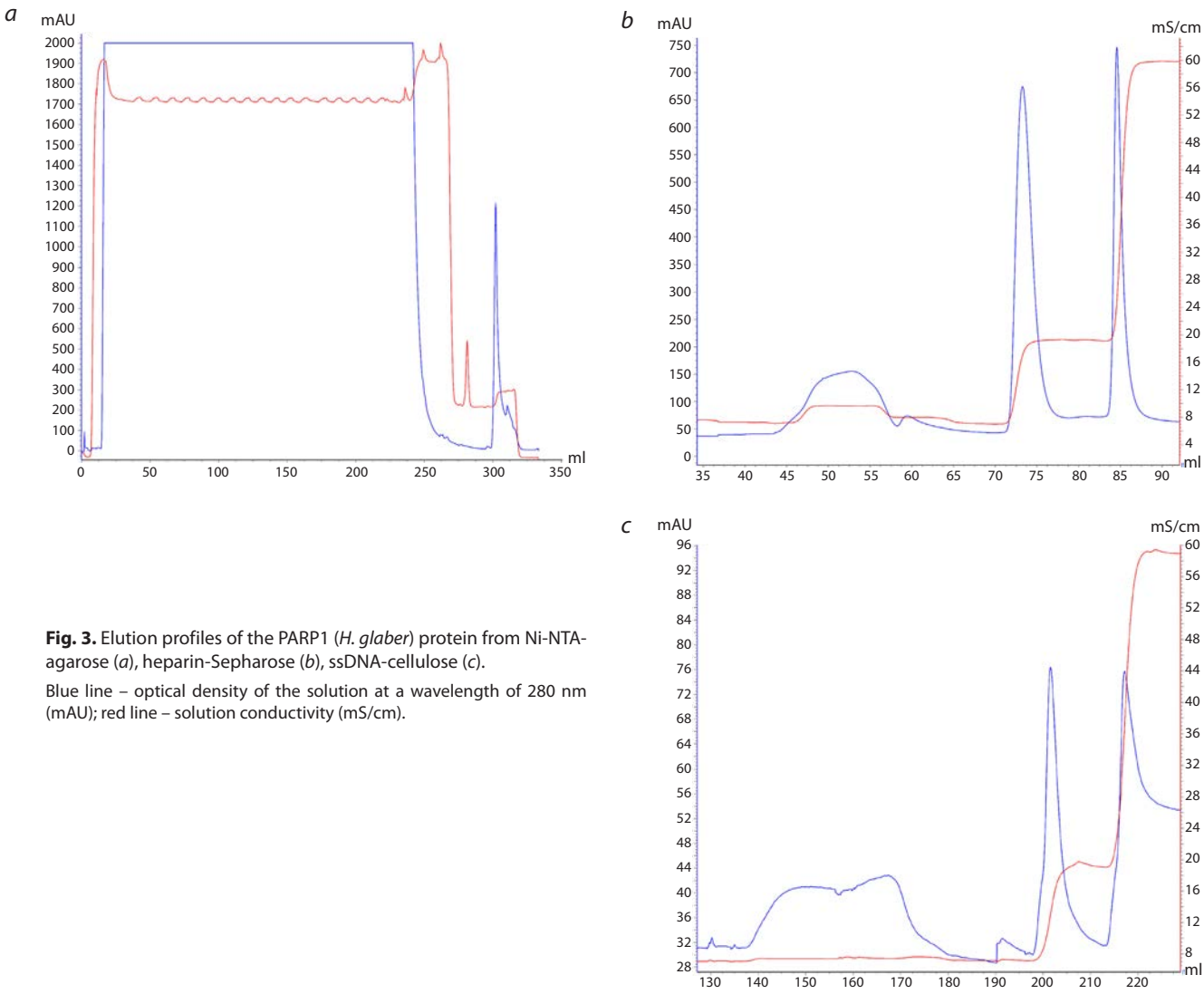



**Fig. 3.** Elution profiles of the PARP1 (*H. glaber*) protein from Ni-NTA-agarose (*a*), heparin-Sepharose (*b*), ssDNA-cellulose (*c*).

Blue line – optical density of the solution at a wavelength of 280 nm (mAU); red line – solution conductivity (mS/cm).

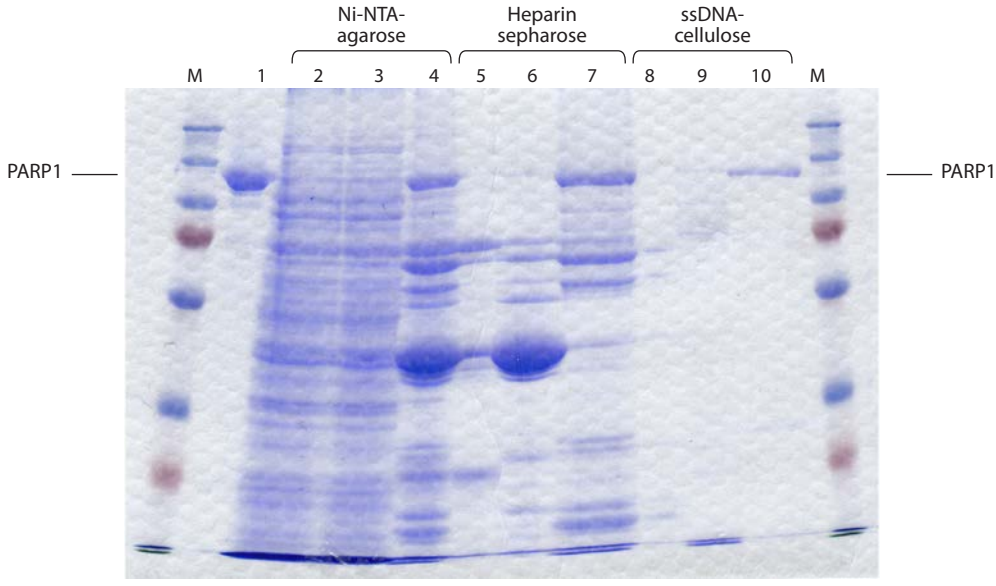



**Fig. 4.** Electropherogram showing the purification stages of recombinant naked mole rat PARP1.

1 – control protein sample; 2–4 – application, breakthrough and elution from Ni-NTA-agarose, respectively; 5 – breakthrough from heparin-Sepharose; 6, 7 – elution from heparin-Sepharose 0.3 M and 1 M NaCl, respectively; 8 – breakthrough from ssDNA cellulose; 9, 10 – elution of the target protein from ssDNA cellulose with 0.3 M and 1 M NaCl, respectively.

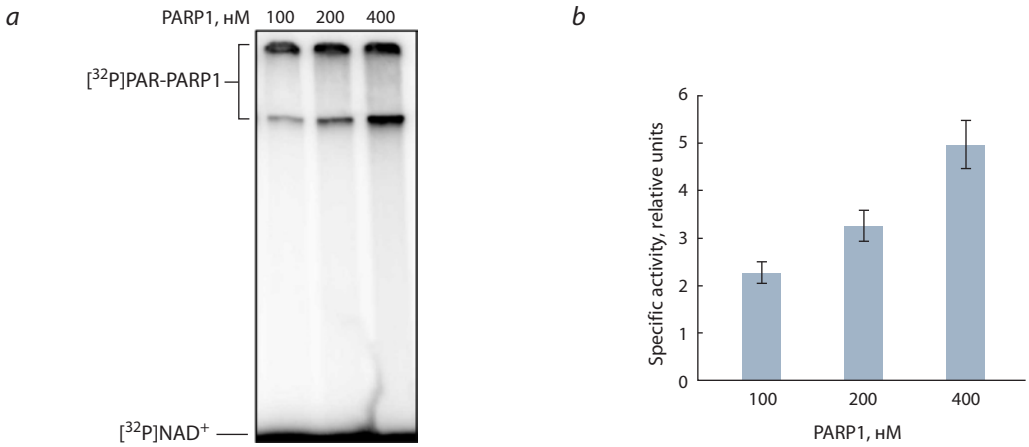



**Fig. 5.** Analysis of the activity of recombinant naked mole rat PARP1 in the autoPARylation reaction.

*a* – autoradiography of 10 % SDS-PAGE, in which the separation of protein modification products was carried out; *b* – diagram summarizing the results of three independent experiments performed using the TCA target method (see Materials and methods).

The comparative analysis of naked mole rat PARP1 and human ortholog protein revealed that substitutions in the ART domain did not affect the catalytic triad, suggesting similar kinetic parameters for PAR synthesis by these enzymes. However, substitutions in other functional domains, particularly those involving changes in residue type near autoPARylation targets (e. g., K305R in the Zn3 domain, D461H in the BRCT domain), may affect the protein's properties. Further studies using mutant forms of naked mole rat PARP1 with these amino acid substitutions are warranted.

## Conclusion

Studying the properties of PARP1 in various long-lived mammals is a promising area of research, as it may contribute to a deeper understanding of the role of DNA repair in aging and how this process is organized in mammalian cells.

Comparison of PARP1 amino acid sequences of two mammals with a long maximum lifespan, the naked mole rat and the human, revealed 13 evolutionarily conserved substitutions in the naked mole rat protein. The impact of these substitutions on the properties and functions of PARP1 remains to be determined. Additionally, cloning of naked mole rat PARP1 was carried out for the first time, along with its production in *E. coli* cells and purification of the recombinant protein using a relatively simple procedure. The recombinant protein's ability to perform the autoPARylation reaction was also assessed. Future work will compare the properties of recombinant PARP1 in long-lived naked mole rats and humans, a task, which has not been previously undertaken.

## References

Alemasova E.E., Lavrik O.I. Poly(ADP-ribose) in condensates: the partnership of phase separation and site-specific interactions. *Int. J. Mol. Sci*. 2022;23(22):14075. DOI 10.3390/ijms232214075

Beneke S., Alvarez-Gonzalez R., Bürkle A. Comparative characterisation of poly(ADP-ribose) polymerase-1 from two mammalian species with different life span. *Exp. Gerontol*. 2000;35(8):989-1002. DOI 10.1016/s0531-5565(00)00134-0

Beneke S., Scherr A.L., Ponath V., Popp O., Bürkle A. Enzyme characteristics of recombinant poly(ADP-ribose) polymerases-1 of rat and human origin mirror the correlation between cellular poly(ADP-ribosyl)ation capacity and species-specific life span. *Mech. Ageing Dev*. 2010;131(5):366-369. DOI 10.1016/j.mad.2010.04.003

Bens M., Sahm A., Groth M., Jahn N., Morhart M., Holtze S., Hildebrandt T.B., Platzer M., Szafranski K. FRAMA: from RNA-seq data to annotated mRNA assemblies. *BMC Genomics*. 2016;17:54. DOI 10.1186/s12864-015-2349-8

Bilkis R., Lake R.J., Cooper K.L., Tomkinson A., Fan H.Y. The CSB chromatin remodeler regulates PARP1- and PARP2-mediated single-strand break repair at actively transcribed DNA regions. *Nucleic Acids Res*. 2023;51(14):7342-7356. DOI 10.1093/nar/gkad515

Buffenstein R. The naked mole-rat: a new long-living model for human aging research. *J. Gerontol. A Biol. Sci. Med. Sci*. 2005;60(11):1369-1377. DOI 10.1093/gerona/60.11.1369

Evdokimov A., Kutuzov M., Petruseva I., Lukjanchikova N., Kashina E., Kolova E., Zemerova T., Romanenko S., Perelman P., Prokopov D., Seluanov A., Gorbunova V., Graphodatsky A., Trifonov V., Khodyreva S., Lavrik O. Naked mole rat cells display more efficient excision repair than mouse cells. *Aging* (*Albany NY*). 2018;10(6):1454-1473. DOI 10.18632/aging.101482

Evdokimov A., Popov A., Ryabchikova E., Koval O., Romanenko S., Trifonov V., Petruseva I., Lavrik I., Lavrik O. Uncovering molecular mechanisms of regulated cell death in the naked mole rat. *Aging* (*Albany NY*). 2021;13(3):3239-3253. DOI 10.18632/aging.202577

Gorbunova V., Seluanov A., Zhang Z., Gladyshev V.N., Vijg J. Comparative genetics of longevity and cancer: insights from long-lived rodents. *Nat. Rev. Genet*. 2014;15(8):531-540. DOI 10.1038/nrg3728

Grube K., Bürkle A. Poly(ADP-ribose) polymerase activity in mononuclear leukocytes of 13 mammalian species correlates with species-specific life span. *Proc. Natl. Acad. Sci. USA*. 1992;89(24):11759-11763. DOI 10.1073/pnas.89.24.11759

Leung A.K.L. Poly(ADP-ribose): a dynamic trigger for biomolecular condensate formation. *Trends Cell Biol*. 2020;30(5):370-383. DOI 10.1016/j.tcb.2020.02.002

López-Otín C., Pietrocola F., Roiz-Valle D., Galluzzi L., Kroemer G. Meta-hallmarks of aging and cancer. *Cell Metab*. 2023;35(1):12-35. DOI 10.1016/j.cmet.2022.11.001.5

MacRae S.L., Croken M.M., Calder R.B., Aliper A., Milholland B., White R.R., Zhavoronkov A., Gladyshev V.N., Seluanov A., Gorbunova V., Zhang Z.D., Vijg J. DNA repair in species with extreme lifespan differences. *Aging* (*Albany NY*). 2015;7(12):1171-1184. DOI 10.18632/aging.100866

Rouleau-Turcotte É., Pascal J.M. ADP-ribose contributions to genome stability and PARP enzyme trapping on sites of DNA damage; paradigm shifts for a coming-of-age modification. *Biol. Chem*. 2023; 299(12):105397. DOI 10.1016/j.jbc.2023.105397

Salmon A.B., Sadighi Akha A.A., Buffenstein R., Miller R.A. Fibroblasts from naked mole-rats are resistant to multiple forms of cell injury, but sensitive to peroxide, ultraviolet light, and endoplasmic reticulum stress. *J. Gerontol. A Biol. Sci. Med. Sci*. 2008;63(3):232-241. DOI 10.1093/gerona/63.3.232

Schumacher B., Pothof J., Vijg J., Hoeijmakers J.H.J. The central role of DNA damage in the ageing process. *Nature*. 2021;592(7856): 695-703. DOI 10.1038/s41586-021-03307-7

Singatulina A.S., Hamon L., Sukhanova M.V., Desforges B., Joshi V., Bouhss A., Lavrik O.I., Pastré D. PARP-1 activation directs FUS to DNA damage sites to form PARG-reversible compartments enriched in damaged DNA. *Cell Rep*. 2019;27(6):1809-1821.e5. DOI 10.1016/j.celrep.2019.04.031

Sinha S., Molla S., Kundu C.N. PARP1-modulated chromatin remodeling is a new target for cancer treatment. *Med. Oncol*. 2021;38(10): 118. DOI 10.1007/s12032-021-01570-2

Sukhanova M.V., Khodyreva S.N., Lavrik O.I. Poly(ADP-ribose) polymerase-1 inhibits strand-displacement synthesis of DNA catalyzed by DNA polymerase beta. *Biochemistry* (*Moscow*). 2004;69(5):558-568. DOI 10.1023/b:biry.0000029855.68502.fa

**Conflict of interest.** The authors declare no conflict of interest.
Received April 23, 2024. Revised July 25, 2024. Accepted July 25, 2024.

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2024;28(7):696-705

DOI 10.18699/vjgb-24-78

# The concept of natural genome reconstruction.

## Part 1. Basic provisions of the "natural genome reconstruction" concept. Changing the genome of hematopoietic stem cells using several natural cellular mechanisms that are inherent in the hematopoietic cell and determine its biological status as "the source of the body's reparative potential"

L.A. Yakubov, O.S. Taranov [1], S.V. Sidorov[2], S.D. Nikonov[3], A.A. Ostanin [4], E.R. Chernykh [4], N.A. Kolchanov [5], S.S. Bogachev [5]

[1] State Scientific Center of Virology and Biotechnology "Vector" of Rospotrebnadzor, Koltsovo, Novosibirsk region, Russia
[2] City Clinical Hospital No. 1, Novosibirsk, Russia
[3] Novosibirsk Tuberculosis Research Institute, Novosibirsk, Russia
[4] Research Institute of Fundamental and Clinical Immunology, Novosibirsk, Russia
[5] Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
labmolbiol@mail.ru

**Abstract.** We present a series of articles proving the existence of a previously unknown mechanism of interaction between hematopoietic stem cells and extracellular double-stranded DNA (and, in particular, double-stranded DNA of the peripheral bloodstream), which explains the possibility of emergence and fixation of genetic information contained in double-stranded DNA of extracellular origin in hematopoietic stem cells. The concept of the possibility of stochastic or targeted changes in the genome of hematopoietic stem cells is formulated based on the discovery of new, previously unknown biological properties of poorly differentiated hematopoietic precursors. The main provisions of the concept are as follows. The hematopoietic stem cell takes up and internalizes fragments of extracellular double-stranded DNA via a natural mechanism. Specific groups of glycocalyx factors, including glycoproteins/proteoglycans, glycosylphosphatidylinositol-anchored proteins and scavenger receptors, take part in the internalization event. The binding sites for DNA fragments are heparin-binding domains and clusters of positively charged amino acid residues that are parts of protein molecules of these factors. Extracellular fragments delivered to the internal compartments of hematopoietic stem cells initiate terminal differentiation, colony formation, and proliferation of hematopoietic precursors. The molecular manifestation of these processes is the emergence and repair of pangenomic single-strand breaks. The occurrence of pangenomic single-strand breaks and restoration of genome (genomic DNA) integrity are associated with activation of a "recombinogenic situation" in the cell; during its active phase, stochastic homologous recombination or other recombination events between extracellular fragments localized in the nucleus and chromosomal DNA are possible. As a result, genetic material of initially extracellular localization either integrates into the recipient genome with the replacement of homologous chromosomal segments, or is transitively present in the nucleus and can manifest itself as a new genetic trait. It is assumed that as a result of stochastic acts of homologous exchange, chromosome loci are corrected in hematopoietic stem cells that have acquired mutations during the existence of the organism, which are the cause of clonal hematopoiesis associated with old age. In this regard, there is a fundamental possibility of changing the hematopoietic status of hematopoietic stem cells in the direction of polyclonality and the original diversity of clones. Such events can form the basis for the rejuvenation of the blood-forming cell system. The results of the laboratory's work indicate that other stem cells in the body capture extracellular DNA fragments too. This fact creates a paradigm for the overall rejuvenation of the body.

**Key words:** extracellular DNA; internalization; single-strand breaks; commitment.

**For citation:** Yakubov L.A., Taranov O.S., Sidorov S.V., Nikonov S.D., Ostanin A.A., Chernykh E.R., Kolchanov N.A., Bogachev S.S. Concept of natural genome reconstruction. Part 1. Basic provisions of the "natural genome reconstruction" concept. Changing the genome of hematopoietic stem cells using several natural cellular mechanisms that are inherent in the hematopoietic cell and determine its biological status as "the source of the body's reparative potential". *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2024;28(7):696-705. DOI 10.18699/vjgb-24-78.

**Funding.** This work was supported by the Ministry of Science and Higher Education of the Russian Federation for the Institute of Cytology and Genetics (state budget-funded project No. FWNR-2022-0016) and by I.N. Zaitseva and A.A. Purtov.

**Acknowledgements.** The authors are grateful to Director-General of Joint Stock Company Clinical Hospital NEUROVITA, Professor, Dr. Med. Sci. A.S. Bryukhovetskiy, for his strong interest in the concept discussed in this paper.

# Концепция природной реконструкции генома.

## Часть 1. Основные положения концепции природной реконструкции генома. Изменение генома гемопоэтических стволовых клеток с использованием нескольких природных клеточных механизмов, имманентно присущих гемопоэтической стволовой клетке и определяющих ее биологический статус как «источник репаративного потенциала организма»

Л.А. Якубов, О.С. Таранов [1], С.В. Сидоров[2], С.Д. Никонов[3], А.А. Останин [4], Е.Р. Черных [4], Н.А. Колчанов [5], С.С. Богачев [5]

[1] Государственный научный центр вирусологии и биотехнологии «Вектор» Роспотребнадзора, р. п. Кольцово, Новосибирская область, Россия
[2] Городская клиническая больница № 1, Новосибирск, Россия
[3] Новосибирский научно-исследовательский институт туберкулеза, Новосибирск, Россия
[4] Научно-исследовательский институт фундаментальной и клинической иммунологии, Новосибирск, Россия
[5] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
labmolbiol@mail.ru

**Аннотация.** Предлагается вниманию цикл статей, доказывающий существование ранее неизвестного механизма взаимодействия гемопоэтической стволовой клетки и экстраклеточной двуцепочечной ДНК (в частности, двуцепочечной ДНК периферического кровяного русла), который объясняет возможность появления и закрепления в гемопоэтических стволовых клетках генетической информации, содержащейся в двуцепочечной ДНК внеклеточного происхождения. Сформулирована концепция возможности стохастического или целенаправленного изменения генома гемопоэтических стволовых клеток, основанная на открытии новых, ранее неизвестных биологических свойств низкодифференцированных гемопоэтических предшественников. Основные положения концепции заключаются в следующих тезисах. Гемопоэтическая стволовая клетка захватывает и интернализует фрагменты экстраклеточной двуцепочечной ДНК естественным природным механизмом. В акте интернализации принимают участие специфические группы факторов гликокаликса, к которым относятся гликопротеины/протеогликаны, гликозилфосфатидилинозитол-заякоренные белки и скавенджер-рецепторы. Сайтами связывания фрагментов ДНК являются гепарин-связывающие домены и кластеры положительно заряженных аминокислотных остатков, входящих в состав белковых молекул указанных факторов. Доставленные во внутренние компартменты гемопоэтических стволовых клеток экстраклеточные фрагменты инициируют терминальную дифференцировку, колониеобразование и пролиферацию предшественников гемопоэза. Молекулярным событием, отражающим эти процессы, является возникновение и репарация пангеномных одноцепочечных разрывов. Процесс возникновения пангеномных одноцепочечных разрывов и восстановление целостности генома (геномной ДНК) сопряжен с активацией в клетке «рекомбиногенной ситуации», во время активной фазы которой возможны стохастическая гомологичная рекомбинация или иные рекомбинационные события между экстраклеточными фрагментами, локализованными в ядре, и ДНК хромосом. Генетический материал исходно экстраклеточной локализации или интегрирует в реципиентный геном с замещением гомологичных хромосомных сегментов, или транзитно присутствует в ядре и может проявляться как новый генетический признак. Предполагается, что в результате стохастических актов гомологичного обмена происходит коррекция локусов хромосом в гемопоэтических стволовых клетках, получивших в ходе существования организма мутации, которые являются причиной клонального гемопоэза, ассоциированного со старостью. В этой связи возникает принципиальная возможность изменения статуса гемопоэза гемопоэтических стволовых клеток в направлении поликлональности и исходного многообразия клонов. Такие события могут составить основу омоложения кровеобразующей системы клеток. Результаты работ свидетельствуют, что другие стволовые клетки организма также захватывают фрагменты экстраклеточной ДНК. Этот факт создает парадигму общего омоложения организма.

**Ключевые слова:** экстраклеточная ДНК; интернализация; одноцепочечные разрывы; коммитирование.

### Double-stranded DNA and its effects on a eukaryotic cell and the organism in general

Double-stranded fragmented extracellular exogenous and endogenous DNA is a participant, inducer and indicator of various processes occurring in the body. First and foremost, exogenous nucleic acids are pathogen-associated molecular patterns that activate different components of the immune system aimed at pathogen removal. Extracellular double-stranded DNA (dsDNA) of endogenous origin can also elicit the body's immune response (Medzhitov, Janeway, 2000; Krieg, 2002; Takeshita, Ishii, 2008; Iwasaki, Medzhitov, 2010; Barbalat et al., 2011; Bode et al., 2011). Fragments of both exogenous and endogenous dsDNA delivered to the cytoplasm of immune cells activate an array of cytosolic DNA sensors and induce the early stages of adaptive immune response (Barber, 2011a, b; Sharma, Fitzgerald, 2011). dsDNA induces autoimmune pro-

cesses Guiducci et al., 2010; Pisetsky, Ullal, 2010; Almqvist et al., 2011; Choubey, 2012; Kaczorowski et al., 2012) and is one of the signals of the bystander effect responsible for transmission of metabolic/catabolic reactions induced in cells exposed to a certain impact (irradiation) to intact bystander cells through the incubation medium (Ermakov et al., 2009). The burst-like increase in plasma DNA concentration has been shown to result in systemic inflammatory response and sepsis (Saukkonen et al., 2007; Castellheim et al., 2009; Arnalich et al., 2010; Kaczorowski et al., 2012). Nucleic acids, including dsDNA, are components of exosomes and are believed to act as a "tuning fork" for the functional state of the organism used by certain cell populations to "fine-tune" physiological and molecular processes in them (Cai et al., 2013; Rashed et al., 2017).

At the cellular level, dsDNA fragments (namely, open double-strand ends of these fragments) internalized by a non-immune cell activate cell cycle arrest and induce repair processes (MacDougall et al., 2007; Zou, 2007). Meanwhile, these fragments become involved in repair under certain conditions, interfering in its correct course; according to the experimental laboratory data, they can be integrated into the recipient genome (Likhacheva et al., 2007, 2008; Dolgova et al., 2012).

Horizontal gene transfer has been studied well in prokaryotes and repeatedly demonstrated for lower metazoans (Andersson, 2005; Soucy et al., 2015; Sibbald et al., 2020); there are proven examples of gene transfer in mammals. Thus, it is known that DNA of destroyed cancer cells is taken up by other susceptible cells in the body, resulting in malignant transformation of these cells. This process is called "genometastasis" (Yakubov et al., 2007; García-Olmo et al., 2012). Evidence for horizontal gene transfer during the uptake of apoptotic bodies and extracellular vesicles has been provided in refs. (Bergsmedh et al., 2001; Holmgren et al., 2002; Sakamoto et al., 2023). There is a study focusing on transfer of extracellular material into an oocyte by spermatozoa that have engulfed high-polymeric DNA from the environment in mussels (Erokhin, Kuznetsov, 2009). Horizontal transfer of a specific unique gene between an insect and a plant has been demonstrated recently (Xia et al., 2021). In 2007, the Nobel Prize was awarded for developing a technology based on homologous recombination of internalized knockout DNA fragments (https://lenta.ru/articles/2007/10/08/nobelmed/).

Several fundamental questions have been consecutively solved as part of the problem of horizonal gene transfer in eukaryotes and humans in particular. The questions related to how the DNA of one organism appeared in another one and the mechanism of DNA spread across the organism have been solved using novel experimental approaches, modern molecular biology techniques. It still remains unclear what is the fixation point of new genetic information in the recipient organism and how a trait associated with it manifests itself as a new biological feature of a new organism.

There are several objective sources and pathways through which foreign DNA appears in the organism: blood manipulations, various transplantations, sexual intercourse, gestation, passage of DNA from food to the organism, viral or bacterial infection, and shared microbiome of the organism.

After it had been proved that peripheral blood of any organism contains a certain quantity of extracellular DNA, the question related to existence of foreign DNA and its spread in the organism was also solved (Anker et al., 1999; Jahr et al., 2001; Laktionov et al., 2004). It is clear that during blood manipulations and transplantation, donor DNA (blood or stroma) necessarily gets into the recipient organism. DNA of damaged spermatozoa, like fetal DNA, gets into the blood and vice versa (Schubbert et al., 1998; Bianchi, Dennis Lo, 2001). Emergence of exogenous DNA through the gastrointestinal tract has been convincingly demonstrated in refs. (Schubbert et al., 1994, 1997; Hohlweg, Doerfler, 2001; Palka-Santini et al., 2003). The gut microbiota is composed of several kilograms of its main constituent species, *Escherichia coli*. Destruction of intestinal cells is accompanied by release of a massive amount of DNA that also gets into the bloodstream. This very phenomenon has underlain methods for diagnosing the state of the gut microflora by analyzing blood serum/plasma that have been developed and launched into clinical practice (https://lenta.ru/articles/2007/10/08/nobelmed/). It means that along with endogenous DNA from cells destroyed by apoptosis or necrosis or DNA of symbiotic microflora, the organism will always contain exogenous DNA, and there are material foundations for the circulation of both foreign and endogenous genetic information in the organism.

Currently, there is almost no discussion on the final question and related ideas about the fixation point of new genetic information and its manifestation as a novel biological trait in scientific literature, since there are no ideas on how these questions can be solved.

It is clear that for a trait to manifest itself, its carrier (namely, dsDNA) needs to be internalized by the cell. There currently exist many publications demonstrating the fact that extracellular DNA fragments are internalized by different cell types; however, it was not until our studies had been published (Petrova et al., 2022; Ritter et al., 2022; Potter et al., 2024) that the mechanism and the ordered structure of factors of this internalization were described. The cited papers revealed that fragments of extracellular dsDNA are internalized by the eukaryotic stem cell via the caveolae-dependent mechanism, with involvement of the heparin-binding domain and clusters of positively charged amino acids of glycoproteins/proteoglycans, glycosylphosphatidylinositol-anchored proteins, and scavenger receptors of the glycocalyx. The total positive charge of the stem cell of different genesis is the key factor for internalization.

A viewpoint is currently being formed assuming that extracellular nucleic acids, including dsDNA, refer to a new type of regulatory system of the organism having complex mechanisms of regulation of cellular process; horizontal gene transfer being one of its manifestations (Ledoux, 1965; Ratajczak et al., 2006; Cocucci et al., 2009; Simons, Raposo, 2009; Camussi et al., 2010; Balaj et al., 2011; Tetta et al., 2011; Ludwig, Giebel, 2012; Ronquist, 2012; Raposo, Stoorvogel, 2013).

## History of the development of the natural genome reconstruction concept

The concept of natural genome reconstruction – the ability to reconstruct pathologically altered chromatin topology (both changes at the nucleotide level and related higher-order changes in chromatin organization) – is the feasibility of introducing molecular changes ("corrections" in the case when it is the cause of a disease) in mutant chromosomal loci (or "healthy-to-healthy" replacement without altering the primary nucleotide sequence) *in vivo*, *ex vivo*, and *in situ*, which employs the principle of a supplementary reconstructive substrate as "genetic material" (fragments of extracellular dsDNA).

The concept is based on five fundamental phenomena of general biology; three of those have recently been discovered and described for the stem cell at the Laboratory of Induced Cellular Processes (Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia).

1. Presence of fragments of extracellular dsDNA in peripheral blood (Anker et al., 1999; Jahr et al., 2001; Laktionov et al., 2004).
2. Ability of stem cells of different genesis to internalize fragments of double-stranded DNA via a natural mechanism (Fig. 1). The following factors are involved in the internalization mechanism: cell surface charge, heparin-binding domains and clusters of positively charged amino acids, glycoproteins/proteoglycans, and scavenger receptors of the glycocalyx. Internalization occurs via the caveolae/clathrin-dependent pathway (Dolgova et al., 2014; Petrova et al., 2022; Ritter et al., 2022).
3. Induction of terminal differentiation of stem cells by dsDNA fragments internalized by intracellular compartments (Fig. 2). Using the mouse model, we have found that extracellular dsDNA fragments represented by PCR fragments, supercoiled plasmid DNA, and fragmented genomic dsDNA are internalized by hematopoietic stem cells (HSCs) via a natural mechanism. Intracellular compartments of c-Kit+/Sca1+/CD34+ HSCs and multipotent descendants can simultaneously contain up to 14,000 copies of ~500 bp long PCR fragments (~0.2 % of the haploid genome). Internalized dsDNA fragments induce terminal differentiation of the progenitor, activate proliferation of descendant cells and colony growth on microcrystalline cellulose. The number of colonies increases by 15–40 % compared to that for untreated samples (Potter et al., 2024). In colonies, most cells (for the mouse model) are terminally differentiated ones in which markers of primitive progenitor cells (c-Kit+/Sca1+ for mice) are lost. A smaller part of cells (up to 15 %) retain markers of immature cells. This fact indicated that upon colony formation, HSCs primarily activated by exposure to Lin–/c-Kit+/Sca1+ divide both asymmetrically (yielding committed hematopoietic progenitor cells) and symmetrically (increasing the pool of poorly differentiated progenitor cells) (Potter et al., 2024). The number of HSCs in the total pool of cells within colonies was several orders of magnitude larger than that in the original sample of bone marrow cells. This fact indicates that the mechanism responsible for amplification (up to 15 % of primitive cells for the mouse model) of a new trait that initially appeared in HSCs of bone marrow cells (0.01 % of primitive cells) has been found.
4. It was found previously that pangenomic single-strand breaks are formed in mammalian embryonic stem cells after induction of their terminal differentiation by treatment with retinoic acid; according to the published data, these breaks are restored within 96 hrs (Vatolin et al., 1997) (Fig. 3). The emergence of pangenomic single-strand breaks causes conversion of nuclear chromatin to its relaxed form. Further repair of single-strand breaks results in restoration of genome integrity and formation of a new 3D chromatin architecture taking into account new torsional strains that are responsible for its new 3D architecture in the selected differentiation direction. It is believed that this genomic "opening and relaxation" is needed to ensure chromatin topology reorganization from the "stem" state to the lineage-committed one (Farzaneh et al., 1982; Johnstone, Williams, 1982). The emergence and repair of pangenomic single-strand breaks is a purely recombination repair process indicating that a "recombinogenic situation" has occurred in the cell (Likhacheva et al., 2008). The molecular machinery of recombination repair involved in chromatin break and repair is unknown.
5. The ability of internalized dsDNA fragments to participate in repair processes (induced either by fragments *per se* or by other factors) occurring in the cell either as a substrate or as the external template for homologous recombination (Dolgova et al., 2014; Potter et al., 2016, 2017, 2024; Ruzanova et al., 2022).

The aforementioned announcement of experimental materials demonstrates that internalization of dsDNA by a HSC induces terminal differentiation (commitment) and activates proliferation progenitors associated with it. We believe that similar to terminal differentiation and emergence of pangenomic single-strand breaks observed in embryonic stem cells, which were demonstrated in ref. (Vatolin et al., 1997), pangenomic single-strand breaks are also induced in HSCs upon commitment, thus leading to the development of a "recombinogenic situation" (Likhacheva et al., 2008). At this very instant, extracellular DNA fragments internalized by stem cells become natural participants of recombinant processes, which are associated with the emergence and repair of these single-strand breaks.

A hypothesis has been put forward that during a "recombinogenic situation" developed in HSCs, different recombination repair events can occur between captured extracellular DNA fragments residing inside the cell, which have induced the emergence of single-strand breaks and the "recombinogenic situation", and chromosomal DNA. These recombination repair events result in alteration of certain chromosomal loci within the repair zone mediated by natural molecular processes because of either homologous exchange or integration of accessory chromatin (e. g., into centromeric or telomeric agglomerates). Transient repair intermediates may appear. These events will lead to integration of extracellular fragments into the recipient genome, or reconstructive replacement of homologous genomic loci, or emergence of transient repair

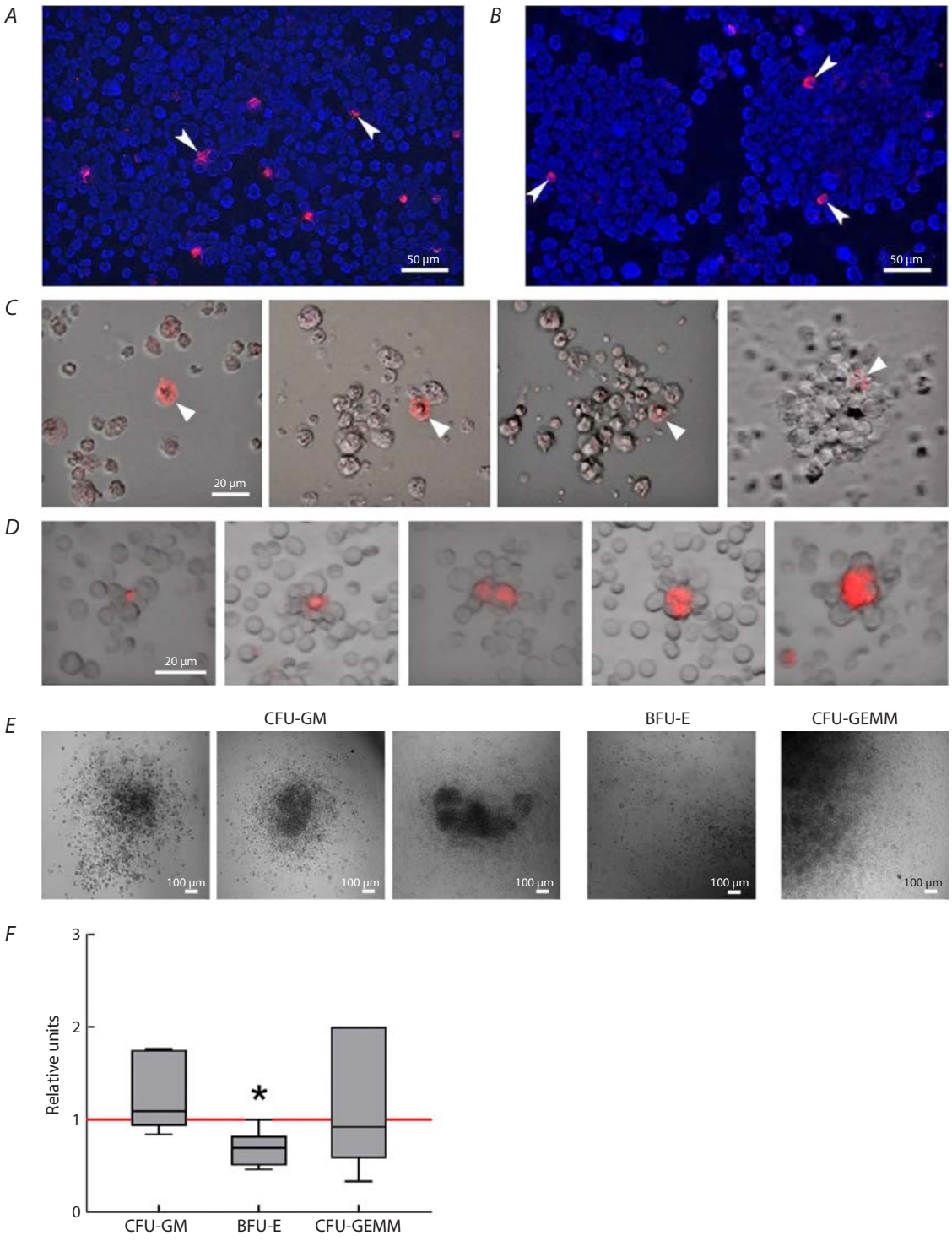



**Fig. 1.** Internalization of TAMRA-labeled DNA probe into stem cells from Krebs-2 carcinoma (*A*) and Epstein–Barr virus-transformed B-cell lymphoma (*B*). Arrows show TAMRA+ cells. One can see two spheres in Figure 1*B*. *C* – formation of spheres in the culture of Epstein–Barr virus-transformed lymphoma cells. A TAMRA+ cell merges with several adjacent cells within 20–30 min. During the observation, all the cells within the visual field are continuously moving chaotically. A strong associate is formed when cells contact the TAMRA+ cell; a free-floating sphere based on it is formed within 8–14 hrs (Dolgova et al., 2016, 2019). *D* – rosettes formed by bone marrow stromal niches, with TAMRA+ cells located in the center (Ritter et al., 2022). *E* – morphologies of the analyzed colonies of the granulocyte-macrophage (CFU-GM), erythroid (BFU-E), and myeloid (CFU-GEMM) hematopoietic lineage upon induction with $hDNA^{gr}$. *F* – stimulation of colony formation of the granulocyte-macrophage (CFU-GM), erythroid (BFU-E), and myeloid (CFU-GEMM) hematopoietic lineages upon induction with $hDNA^{gr}$ compared to control. The values are the ratio between the number of colonies upon induction with the preparation and the number of control colonies assumed to be equal to "1" and denoted with the red line. The data are presented as the median, the interquartile range, and the minimum and maximum values. * – statistically significant differences compared to the control group ($p < 0.05$, Wilcoxon matched pairs test, $n = 6$) (Potter et al., 2024).

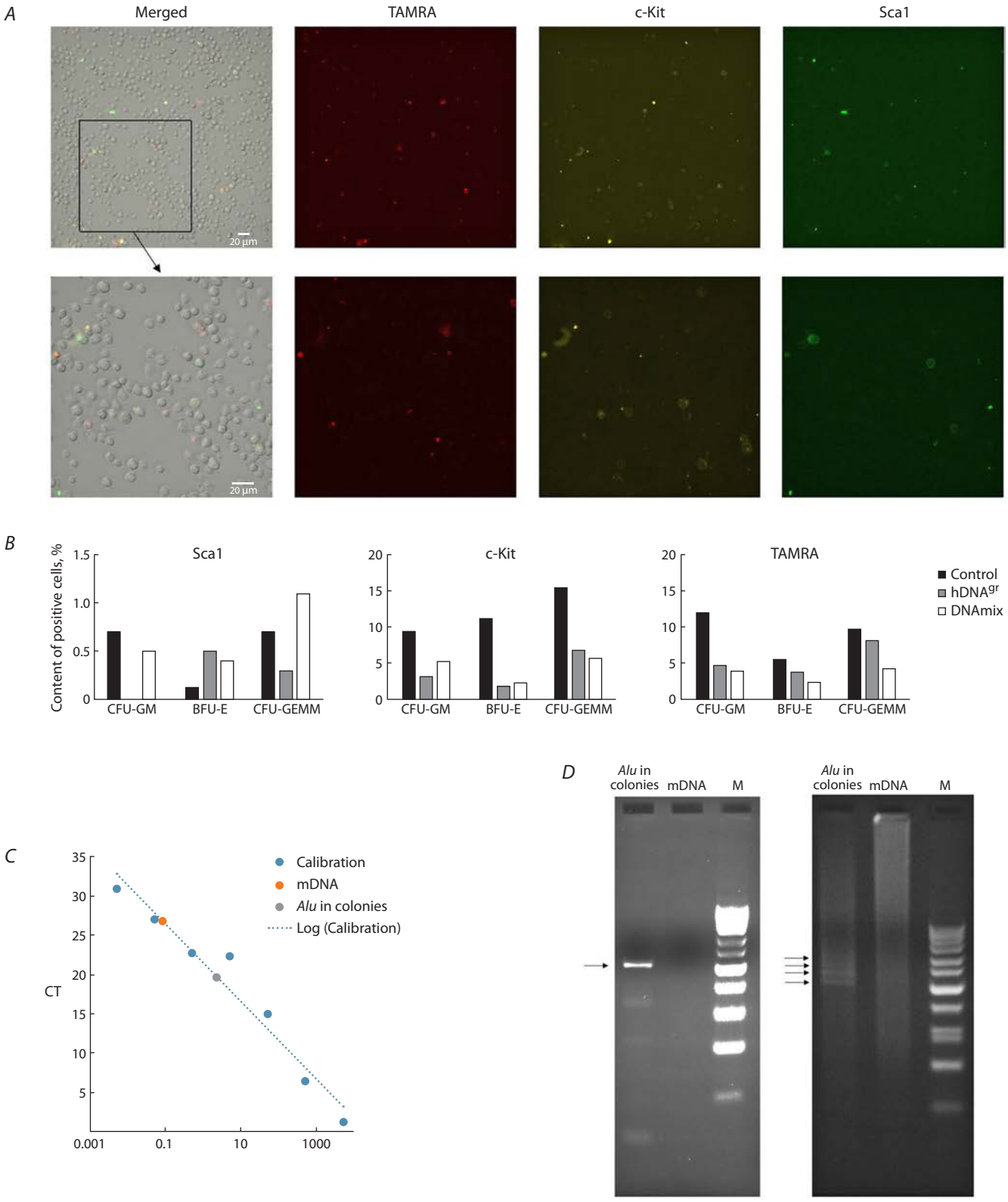



**Fig. 2.** Analysis of internalization of double-stranded DNA probe in HSCs.
*A* – assessment of colocalization of markers of primitive HSCs c-Kit and Sca1 and simultaneously the ability of these cells to internalize TAMRA+ double-stranded DNA probe in intact murine bone marrow. *B* – comparative diagrams of the contents of cells labeled with the respective marker in cell population of an individual colony. The reduced number of cells labelled with stemness markers may indicate that induction with DNA preparation caused stem cell differentiation. *C* – analysis of internalization of double-stranded DNA repeat probe *AluI* in HSCs Sca1/c-Kit by real-time PCR. Calculations indicate that the content of the DNA probe is ~14,000 copies per cell. *D* – electrophoretic mobility of the PCR fragment obtained using primers M13 and DNA template isolated from murine HSCs, induced by hDNA$^{gr}$, and incubated in the presence of human *Alu* fragment. On the left: PCR products from templates "*Alu* in colonies" and murine DNA. On the right: electrophoresis of real-time PCR products from the same templates. Arrows show the PCR products synthesized from concatemerized DNA template *AluI* (Potter et al., 2024).

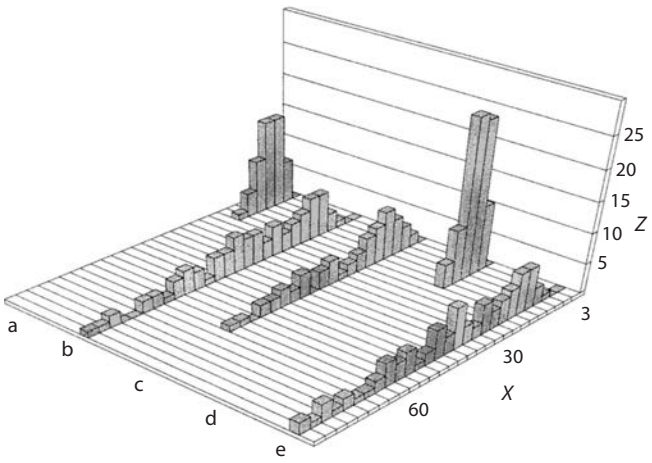



**Fig. 3.** The histogram showing distribution of DNA migration path length of OTF9-63 cells before and after cultivation during different time periods in the presence of retinoic acid.
*X* – DNA migration path length in μm (scale division value, 3 μm); *Z* – number of cells: *a* – undifferentiated cells; *b–d* – differentiation during 24, 36, and 96 hrs, respectively; *e* – undifferentiated embryonic stem cells irradiated with X-rays at a dose of 200 rad. Among cells exposed to radiation, migration path length involves the nucleus size (Vatolin et al., 1997).

intermediates not integrated into the genome, with consequent biological changes at the cellular, tissue, organ, functional system, and organism levels.

It is supposed that nonhomologous integration of extracellular fragments is hardly feasible because there are no molecular foundations for it (namely, nonfunctional double-strand breaks). The feasibility of fixation of nonhomologous extracellular DNA in the genome can be related to the imperfection of the mechanism of single-strand annealing and involvement of microhomologies in this process.

## The concept of natural genome reconstruction

Extracellular dsDNA fragments are taken up and internalized into compartments of HSCs via a natural mechanism. The aforementioned substrate (extracellular dsDNA fragments) is "offered" to a stem cell as extracellular material or under *in vivo*, *ex vivo*, or *in situ* conditions. dsDNA fragments internalized into cellular compartments induce terminal differentiation of hematopoietic stem cells; emergence of pangenomic single-strand breaks that extracellular fragments located inside cells interact with is one of its markers. The pangenomic single-strand breaks ensure the relaxed form of a certain level of nuclear chromatin compaction, thus providing conditions for subsequent chromatin reorganization. Chromatin in this state undergoes mitosis. The first mitotic division results in restoration of the 3D chromatin architecture of the original cell with stem characteristics in one daughter cell. In the second daughter cell, chromatin reorganization ("reprogramming") takes place, and a new 3D architecture in the selected direction of cell differentiation is formed with allowance for a new torsional strain responsible for the new 3D structure of chromatin. A "recombinogenic situation" is induced in the HSC at the instant when pangenomic single-strand breaks appear and are repaired, and conditions for integrating fragments internalized by the cell into chromatin DNA are established. Either replacement of homologous regions or direct integration of extracellular fragments into homologous or non-homologous genomic regions occurs, or transient intermediates are formed. The choice of the pathway and the integration mechanism are unknown; however, homologous exchange between the extracellular template and genomic DNA is primarily suspected. HSCs in which "illegitimate" integration of extracellular nonhomologous DNA has taken place are eliminated from the HSC population. The mechanistic schematics of the aforementioned processes are provided in the thematic experimental sections of this series of studies.

## Conclusion

The proposed concept has been proved in a number of experimental studies united together into a series of studies collectively called "The Concept of Natural Genome Reconstruction". Further articles within this series will focus on questions characterizing early events of interaction between HSCs and extracellular fragments of dsDNA, as well as the implications of this interaction altering both the structural and linear arrangement of extracellular dsDNA fragments, as well as the functional state of target hematopoietic progenitor cells. Genomic changes related to the emergence of extracellular dsDNA fragments in the intracellular space of stem cells will be analyzed. *In vivo* tests will assess the effect of changes in the genome of HSCs caused by interaction with extracellular dsDNA fragments on certain biological parameters of experimental animals (changes in lifespan). Finally, the pleiotropic effect of the impact of extracellular dsDNA on human HSCs will be analyzed in a clinical case study, with special emphasis placed on the shift in the hematopoietic status (oligoclonal/normal).

## References

Almqvist N., Winkler T.H., Mårtensson I.L. Autoantibodies: focus on anti-DNA antibodies. *Self/Nonself*. 2011;2(1):11-18. DOI 10.4161/SELF.2.1.15087

Andersson J.O. Lateral gene transfer in eukaryotes. *Cell. Mol. Life Sci*. 2005;62(11):1182-1197. DOI 10.1007/S00018-005-4539-Z

Anker P., Mulcahy H., Chen X.Q., Stroun M. Detection of circulating tumour DNA in the blood (plasma/serum) of cancer patients. *Cancer Metastasis Rev*. 1999;18(1):65-73. DOI 10.1023/a:1006260 319913

Arnalich F., Maldifassi M.C., Ciria E., Quesada A., Codoceo R., Herruzo R., Garcia-Cerrada C., Montoya F., Vazquez J.J., López-Collazo E., Montiel C. Association of cell-free plasma DNA with perioperative mortality in patients with suspected acute mesenteric ischemia. *Clin. Chim. Acta*. 2010;411(17-18):1269-1274. DOI 10.1016/j.cca.2010.05.017

Balaj L., Lessard R., Dai L., Cho Y.J., Pomeroy S.L., Breakefield X.O., Skog J. Tumour microvesicles contain retrotransposon elements and amplified oncogene sequences. *Nat. Commun*. 2011;2:180. DOI 10.1038/ncomms1180

Barbalat R., Ewald S.E., Mouchess M.L., Barton G.M. Nucleic acid recognition by the innate immune system. *Annu. Rev. Immunol*. 2011;29(1):185-214. DOI 10.1146/annurev-immunol-031210-101340

Barber G.N. Cytoplasmic DNA innate immune pathways. *Immunol. Rev*. 2011a;243(1):99-108. DOI 10.1111/j.1600-065X.2011.01051.x

Barber G.N. Innate immune DNA sensing pathways: STING, AIMII and the regulation of interferon production and inflammatory responses. *Curr. Opin. Immunol*. 2011b;23(1):10-20. DOI 10.1016/ j.coi.2010.12.015

Bergsmedh A., Szeles A., Henriksson M., Bratt A., Folkman M.J., Spetz A.L., Holmgren L. Horizontal transfer of oncogenes by uptake of apoptotic bodies. *Proc. Natl. Acad. Sci. USA*. 2001; 98(11):6407-6411. DOI 10.1073/PNAS.101129998

Bianchi D.W., Dennis Lo Y.M. Fetomaternal cellular and plasma DNA trafficking: the Yin and the Yang. *Ann. NY Acad. Sci*. 2001; 945:119-131. DOI 10.1111/J.1749-6632.2001.TB03872.X

Bode C., Zhao G., Steinhagen F., Kinjo T., Klinman D.M. CpG DNA as a vaccine adjuvant. *Expert Rev. Vaccines*. 2011;10(4): 499-511. DOI 10.1586/erv.10.174

Cai J., Han Y., Ren H., Chen C., He D., Zhou L., Eisner G.M., Asico L.D., Jose P.A., Zeng C. Extracellular vesicle-mediated transfer of donor genomic DNA to recipient cells is a novel mechanism for genetic influence between cells. *J. Mol. Cell Biol*. 2013;5(4):227-238. DOI 10.1093/JMCB/MJT011

Camussi G., Deregibus M.C., Bruno S., Cantaluppi V., Biancone L. Exosomes/microvesicles as a mechanism of cell-to-cell communication. *Kidney Int*. 2010;78(9):838-848. DOI 10.1038/KI.2010.278

Castellheim A., Brekke O.L., Espevik T., Harboe M., Mollnes T.E. Innate immune responses to danger signals in systemic inflammatory response syndrome and sepsis. *Scand. J. Immunol*. 2009; 69(6):479-491. DOI 10.1111/j.1365-3083.2009.02255.x

Choubey D. DNA-responsive inflammasomes and their regulators in autoimmunity. *Clin. Immunol*. 2012;142(3):223-231. DOI 10.1016/ j.clim.2011.12.007

Cocucci E., Racchetti G., Meldolesi J. Shedding microvesicles: artefacts no more. *Trends Cell Biol*. 2009;19(2):43-51. DOI 10.1016/ J.TCB.2008.11.003

Dolgova E.V., Proskurina A.S., Nikolin V.P., Popova N.A., Alyamkina E.A., Orishchenko K.E., Rogachev V.A., Efremov Y.R., Dubatolova T.D., Prokopenko A.V., Chernykh E.R., Ostanin A.A., Taranov O.S., Omigov V.V., Zagrebelniy S.N., Bogachev S.S., Shurdov M.A. "Delayed death" phenomenon: a synergistic action of cyclophosphamide and exogenous DNA. *Gene*. 2012; 495(2):134-145. DOI 10.1016/j.gene.2011.12.032

Dolgova E.V., Alyamkina E.A., Efremov Y.R., Nikolin V.P., Popova N.A., Tyrinova T.V., Kozel A.V., Minkevich A.M., Andrushkevich O.M., Zavyalov E.L., Romaschenko A.V., Bayborodin S.I., Taranov O.S., Omigov V.V., Shevela E.Y., Stupak V.V., Mishinov S.V., Rogachev V.A., Proskurina A.S., Mayorov V.I., Shurdov M.A., Ostanin A.A., Chernykh E.R., Bogachev S.S. Identification of cancer stem cells and a strategy for their elimination. *Cancer Biol. Ther*. 2014;15(10):1378-1394. DOI 10.4161/cbt.29854

Dolgova E.V., Shevela E.Y., Tyrinova T.V., Minkevich A.M., Proskurina A.S., Potter E.A., Orishchenko K.E., Zavjalov E.L., Bayborodin S.I., Nikolin V.P., Popova N.A., Pronkina N.V., Ostanin A.A., Chernykh E.R., Bogachev S.S. Nonadherent spheres with multiple myeloma surface markers contain cells that contribute to sphere formation and are capable of internalizing extracellular double-stranded DNA. *Clin. Lymphoma Myeloma Leuk*. 2016;16(10):563-576. DOI 10.1016/j.clml.2016.06.014

Dolgova E.V., Petrova D.D., Proskurina A.S., Ritter G.S., Kisaretova P.E., Potter E.A., Efremov Y.R., Bayborodin S.I., Karamysheva T.V., Romanenko M.V., Netesov S.V., Taranov O.S., Ostanin A.A., Chernykh E.R., Bogachev S.S. Identification of the xenograft and its ascendant sphere-forming cell line as belonging to EBV-induced lymphoma, and characterization of the status of sphere-forming cells. *Cancer Cell Int*. 2019;19(1):120. DOI 10.1186/s12935-019-0842-x

Erokhin V.E., Kuznetzov A.V. Informative and regulatory roles of dissolved organic matter in seawater. *Morskoi Ekologicheskiy Zhurnal = Marine Ecol. J*. 2009;8(2):64-69 (in Russian)

Ermakov A.V., Konkova M.S., Kostyuk S.V., Egolina N.A., Efremova L.V., Veiko N.N. Oxidative stress as a significant factor for development of an adaptive response in irradiated and nonirradiated human lymphocytes after inducing the bystander effect by low-dose X-radiation. *Mutat. Res*. 2009;669(1-2):155-161. DOI 10.1016/ j.mrfmmm.2009.06.005

Farzaneh F., Zalin R., Brill D., Shall S. DNA strand breaks and ADP-ribosyl transferase activation during cell differentiation. *Nature*. 1982; 300(5890):362-366. DOI 10.1038/300362A0

García-Olmo D.C., Picazo M.G., García-Olmo D. Transformation of non-tumor host cells during tumor progression: theories and evidence. *Expert Opin. Biol. Ther*. 2012;12(Suppl.1):S199-207. DOI 10.1517/14712598.2012.681370

Guiducci C., Gong M., Xu Z., Gill M., Chaussabel D., Meeker T., Chan J.H., Wright T., Punaro M., Bolland S., Soumelis V., Banchereau J., Coffman R.L., Pascual V., Barrat F.J. TLR recognition of self nucleic acids hampers glucocorticoid activity in lupus. *Nature*. 2010;465(7300):937-941. DOI 10.1038/nature 09102

Hohlweg U., Doerfler W. On the fate of plant or other foreign genes upon the uptake in food or after intramuscular injection in mice. *Mol. Genet. Genomics*. 2001;265(2):225-233. DOI 10.1007/S004380100450

Holmgren L., Bergsmedh A., Spetz A.L. Horizontal transfer of DNA by the uptake of apoptotic bodies. *Vox Sang*. 2002;83(Suppl.1): 305-306. DOI 10.1111/J.1423-0410.2002.TB05323.X

Iwasaki A., Medzhitov R. Regulation of adaptive immunity by the innate immune system. *Science*. 2010;327(5963):291-295. DOI 10.1126/science.1183021

Jahr S., Hentze H., Englisch S., Hardt D., Fackelmayer F.O., Hesch R.D., Knippers R. DNA fragments in the blood plasma of cancer patients: quantitations and evidence for their origin from apoptotic and necrotic cells. *Cancer Res*. 2001;61(4):1659-1665. http://www.ncbi.nlm.nih.gov/pubmed/11245480

Johnstone A.P., Williams G.T. Role of DNA breaks and ADP-ribosyl transferase activity in eukaryotic differentiation demonstrated in human lymphocytes. *Nature*. 1982;300(5890):368-370. DOI 10.1038/ 300368A0

Kaczorowski D.J., Scott M.J., Pibris J.P., Afrazi A., Nakao A., Edmonds R.D., Kim S., Kwak J.H., Liu Y., Fan J., Billiar T.R. Mammalian DNA is an endogenous danger signal that stimulates local synthesis and release of complement factor B. *Mol. Med*. 2012;18(1):851-860. DOI 10.2119/molmed.2012.00011

Krieg A.M. CpG motifs in bacterial DNA and their immune effects. *Annual Rev. Immunol*. 2002;20(1):709-760. DOI 10.1146/annurev.immunol.20.100301.064842

Laktionov P.P., Tamkovich S.N., Rykova E.Yu., Bryzgunova O.E., Starikov A.V., Kuznetsova N.P., Sumarokov S.V., Kolomiets S.A., Sevostianova N.V., Vlassov V.V. Extracellular circulating nucleic acids in human plasma in health and disease. *Nucleosides Nucleotides Nucleic Acids*. 2004;23(6-7):879-883. DOI 10.1081/NCN-200026035

Ledoux L. Uptake of DNA by living cells. *Prog. Nucleic Acid Res. Mol. Biol*. 1965;4:231-267. DOI 10.1016/S0079-6603(08)60790-4

Likhacheva A.S., Nikolin V.P., Popova N.A., Dubatolova T.D., Strunkin D.N., Rogachev V.A., Sebeleva T.E., Erofeev I.S., Bogachev S.S., Yakubov L.A., Shurdov M.A. Integration of human DNA fragments into the cell genomes of certain tissues from adult mice treated with cytostatic cyclophosphamide in combination with human DNA. *Gene Ther. Mol. Biol*. 2007;11(2): 185-202

Likhacheva A.S., Rogachev V.A., Nikolin V.P., Popova N.A., Shilov A.G., Sebeleva T.E., Strunkin D.N., Chernykh E.R., Gel'fgat E.L., Bogachev S.S., Shurdov M.A. Involvement of exogenous DNA in the molecular processes in somatic cell. *Informatsionnyy Vestnik VOGiS = The Herald of Vavilov Society for Geneticists and Breeders*. 2008;12(3):426-473 (in Russian)

Ludwig A.K., Giebel B. Exosomes: small vesicles participating in intercellular communication. *Int. J. Biochem. Cell Biol*. 2012; 44(1):11-15. DOI 10.1016/j.biocel.2011.10.005

MacDougall C.A., Byun T.S., Van C., Yee M.C., Cimprich K.A. The structural determinants of checkpoint activation. *Genes Dev*. 2007;21(8):898-903. DOI 10.1101/gad.1522607

Medzhitov R., Janeway C. Innate immune recognition: mechanisms and pathways. *Immunol. Rev.* 2000;173(1):89-97. DOI 10.1034/j.1600-065X.2000.917309.x

Palka-Santini M., Schwarz-Herzke B., Hösel M., Renz D., Auerochs S., Brondke H., Doerfler W. The gastrointestinal tract as the portal of entry for foreign macromolecules: fate of DNA and proteins. *Mol. Genet. Genomics*. 2003;270(3):201-215. DOI 10.1007/S00438-003-0907-2

Petrova D.D., Dolgova E.V., Proskurina A.S., Ritter G.S., Ruzanova V.S., Efremov Y.R., Potter E.A., Kirikovich S.S., Levites E.V., Taranov O.S., Ostanin A.A., Chernykh E.R., Kolchanov N.A., Bogachev S.S. The new general biological property of stem-like tumor cells (Part II: Surface molecules, which belongs to distinctive groups with particular functions, form a unique pattern characteristic of a certain type of tumor stem-like cells). *Int. J. Mol. Sci*. 2022;23(24):15800. DOI 10.3390/ijms232415800

Pisetsky D.S., Ullal A.J. The blood nucleome in the pathogenesis of SLE. *Autoimmun. Rev.* 2010;10(1):35-37. DOI 10.1016/j.autrev.2010.07.007

Potter E.A., Dolgova E.V., Proskurina A.S., Minkevich A.M., Efremov Y.R., Taranov O.S., Omigov V.V., Nikolin V.P., Popova N.A., Bayborodin S.I., Ostanin A.A., Chernykh E.R., Kolchanov N.A., Shurdov M.A., Bogachev S.S. A strategy to eradicate well-developed Krebs-2 ascites in mice. *Oncotarget*. 2016; 7(10):11580-11594. DOI 10.18632/oncotarget.7311

Potter E.A., Dolgova E.V., Proskurina A.S., Efremov Y.R., Minkevich A.M., Rozanov A.S., Peltek S.E., Nikolin V.P., Popova N.A., Seledtsov I.A., Molodtsov V.V., Zavyalov E.L., Taranov O.S., Baiborodin S.I., Ostanin A.A., Chernykh E.R., Kolchanov N.A., Bogachev S.S. Gene expression profiling of tumor-initiating stem cells from mouse Krebs-2 carcinoma using a novel marker of poorly differentiated cells. *Oncotarget*. 2017;8(6):9425-9441. DOI 10.18632/ONCOTARGET.14116

Potter E.A., Dolgova E.V., Proskurina A.S., Ruzanova V.S., Efremov Y.R., Kirikovich S.S., Oshikhmina S.G., Mamaev A.L., Taranov O.S., Bryukhovetskiy A.S., Grivtsova L.U., Kolchanov N.A., Ostanin A.A., Chernykh E.R., Bogachev S.S. Stimulation of mouse hematopoietic stem cells by angiogenin and DNA preparations. *Braz. J. Med. Biol. Res*. 2024;57:e13072. DOI 10.1590/1414-431X 2024E13072

Raposo G., Stoorvogel W. Extracellular vesicles: exosomes, microvesicles, and friends. *J. Cell Biol*. 2013;200(4):373-383. DOI 10.1083/JCB.201211138

Rashed M.H., Bayraktar E., Helal G.K., Abd-Ellah M.F., Amero P., Chavez-Reyes A., Rodriguez-Aguayo C. Exosomes: from garbage bins to promising therapeutic targets. *Int. J. Mol. Sci*. 2017; 18(3): 538. DOI 10.3390/IJMS18030538

Ratajczak J., Wysoczynski M., Hayek F., Janowska-Wieczorek A., Ratajczak M.Z. Membrane-derived microvesicles: important and underappreciated mediators of cell-to-cell communication. *Leukemia*. 2006;20(9):1487-1495. DOI 10.1038/sj.leu.2404296

Ritter G., Dolgova E.V., Petrova D.D., Efremov Y.R., Proskurina A.S., Potter E.A., Ruzanova V.S., Kirikovich S.S., Levites E.V., Taranov O.S., Ostanin A.A., Chernykh E.R., Kolchanov N.A., Bogachev S.S. The new general biological property of stem-like tumor cells (Part I. Peculiarities of the process of the double-stranded DNA fragments internalization into stem-like tumor cells). *Front. Genet*. 2022;13:954395. DOI 10.3389/fgene.2022.954395

Ronquist G. Prostasomes are mediators of intercellular communication: from basic research to clinical implications. *J. Intern. Med*. 2012;271(4):400-413. DOI 10.1111/j.1365-2796.2011.02487.x

Ruzanova V., Proskurina A., Efremov Y., Kirikovich S., Ritter G., Levites E., Dolgova E., Potter E., Babaeva O., Sidorov S., Taranov O., Ostanin A., Chernykh E., Bogachev S. Chronometric administration of cyclophosphamide and a double-stranded DNA-mix at interstrand crosslinks repair timing, called "Karanahan" therapy, is highly efficient in a weakly immunogenic Lewis carcinoma model. *Pathol. Oncol. Res*. 2022;28:1610180. DOI 10.3389/PORE.2022.1610180

Sakamoto Y., Ochiya T., Yoshioka Y. Extracellular vesicles in the breast cancer brain metastasis: physiological functions and clinical applications. *Front. Hum. Neurosci*. 2023;17:1278501. DOI 10.3389/FNHUM.2023.1278501

Saukkonen K., Lakkisto P., Varpula M., Varpula T., Voipio-Pulkki L.M., Pettilä V., Pulkki K. Association of cell-free plasma DNA with hospital mortality and organ dysfunction in intensive care unit patients. *Intensive Care Med*. 2007;33(9):1624-1627. DOI 10.1007/s00134-007-0686-z

Schubbert R., Lettmann C., Doerfler W. Ingested foreign (phage M13) DNA survives transiently in the gastrointestinal tract and enters the bloodstream of mice. *Mol. Gen. Genet*. 1994;242(5): 495-504. DOI 10.1007/BF00285273

Schubbert R., Renz D., Schmitz B., Doerfler W. Foreign (M13) DNA ingested by mice reaches peripheral leukocytes, spleen, and liver via the intestinal wall mucosa and can be covalently linked to mouse DNA. *Proc. Natl. Acad. Sci. USA.* 1997;94(3):961-966. DOI 10.1073/PNAS.94.3.961

Schubbert R., Hohlweg U., Renz D., Doerfler W. On the fate of orally ingested foreign DNA in mice: chromosomal association and placental transmission to the fetus. *Mol. Gen. Genet*. 1998; 259(6):569-576. DOI 10.1007/S004380050850

Sharma S., Fitzgerald K.A. Innate immune sensing of DNA. *PLoS Pathog*. 2011;7(4):e1001310. DOI 10.1371/JOURNAL.PPAT.1001310

Sibbald S.J., Eme L., Archibald J.M., Roger A.J. Lateral gene transfer mechanisms and pan-genomes in eukaryotes. *Trends Parasitol*. 2020;36(11):927-941. DOI 10.1016/J.PT.2020.07.014

Simons M., Raposo G. Exosomes – vesicular carriers for intercellular communication. *Curr. Opin. Cell Biol*. 2009;21(4):575-581. DOI 10.1016/j.ceb.2009.03.007

Soucy S.M., Huang J., Gogarten J.P. Horizontal gene transfer: building the web of life. *Nat. Rev. Genet*. 2015;16(8):472-482. DOI 10.1038/NRG3962

Takeshita F., Ishii K.J. Intracellular DNA sensors in immunity. *Curr. Opin. Immunol*. 2008;20(4):383-388. DOI 10.1016/j.coi.2008.05.009

Tetta C., Bruno S., Fonsato V., Deregibus M.C., Camussi G. The role of microvesicles in tissue repair. *Organogenesis*. 2011;7(2):105-115. DOI 10.4161/org.7.2.15782

Vatolin S.Y., Okhapkina E.V., Matveeva N.M., Shilov A.G., Baiborodin S.I., Philimonenko V.V., Zhdanova N.S., Serov O.L. Scheduled perturbation in DNA during in vitro differentiation of mouse embryo-derived cells. *Mol. Reprod. Dev*. 1997;47(1):1-10. DOI 10.1002/(SICI)1098-2795(199705)47:1<1::AID-MRD1>3.0.CO;2-R

Xia J., Guo Z., Yang Z., Han H., Wang S., Xu H., Yang X., Yang F., Wu Q., Xie W., Zhou X., Dermauw W., Turlings T.C.J., Zhang Y. Whitefly hijacks a plant detoxification gene that neutralizes plant toxins. *Cell*. 2021;184(7):1693-1705.e17. DOI 10.1016/J.CELL.2021.02.014

Yakubov L., Rogachev V., Likhacheva A., Bogachev S., Sebeleva T., Shilov A., Baiborodin S., Petrova N., Mechetina L., Shurdov M., Wickstrom E. Natural human gene correction by small extracellular genomic DNA fragments. *Cell Cycle*. 2007;6(18):2293-2301. DOI 10.4161/cc.6.18.4729

Zou L. Single- and double-stranded DNA: building a trigger of ATR-mediated DNA damage response. *Genes Dev.* 2007;21(8):879-885. DOI 10.1101/gad.1550307

**Conflict of interest.** The authors declare no conflict of interest.
Received June 13, 2024. Revised August 21, 2024. Accepted September 5, 2024.

PLANT GENETICS AND BREEDING
Original article

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2024;28(7):706-715

DOI 10.18699/vjgb-24-79

# Molecular genetic and morphological characteristics of *Micractinium thermotolerans* and *M. inermum* (Trebouxiophyceae, Chlorophyta) from pyroclastic deposits of the Kamchatka Peninsula (Russia)

R.Z. Sushchenko, V.Yu. Nikulin, V.B. Bagmet, A.Yu. Nikulin

Federal Scientific Center of the East Asia Terrestrial Biodiversity of the Far Eastern Branch of the Russian Academy of Sciences, Vladivostok, Russia
allaguvatova@yandex.ru

**Abstract.** During the study of algal diversity in pyroclastic deposits of the Kamchatka Peninsula, *Chlorella*-like green algae strains VCA-72 and VCA-93 were isolated from samples collected from along the Baydarnaya river bed on the Shiveluch volcano in 2018 and at the outlet of thermal vapors along the edge of the caldera on the southern slope of the Gorely volcano in 2020. Identification of the strains was carried out within the framework of an integrative approach using microscopic and molecular genetic methods, including preliminary taxon identification, obtaining nucleotide sequences of the small subunit and the internal transcribed spacer rRNA, reconstruction of phylogenetic trees and secondary structures of the ITS1 and ITS2 rRNA regions. On the phylogenetic tree, strain VCA-93 was clustered in the *Micractinium thermotolerans* species clade. No differences were found when comparing the helical domain models of ITS1 and ITS2 in *M. thermotolerans*. Strain VCA-72 occupied a basal position in the *M. inermum* clade. The secondary structure patterns of the helices of strain VCA-72 were generally similar to those of *M. inermum*, but intraspecific variability was noted, mainly due to substitutions in the apical and lateral loops. Five hCBC substitutions were found in the helical regions of the studied *M. inermum* strains, while no CBC substitutions were found. A detailed analysis of morphology and life cycle allowed us to identify the characteristics of the cells in aging cultures: their size was significantly higher than in vegetative ones and they were pear-shaped, oval, and ellipsoidal with a shallow, wide constriction in the center. In addition, cells with colorless lipid droplets were detected in aging cultures of both species. The ability to synthesize and accumulate lipids indicates the great potential of the strains for the production of biodiesel fuel. A review of the habitats of previous and new findings allowed us to note the ecological plasticity of the studied species. The results obtained complement the information on the biogeography of the species: this is the first record of *M. inermum* for the territory of Russia, and that of *M. thermotolerans*, for the Kamchatka Peninsula.

**Key words:** microalgae; floristic findings; integrative approach; morphology; phylogeny; secondary structure of ITS rRNA.

**For citation:** Sushchenko R.Z., Nikulin V.Yu., Bagmet V.B., Nikulin A.Yu. Molecular genetic and morphological characteristics of *Micractinium thermotolerans* and *M. inermum* (Trebouxiophyceae, Chlorophyta) from pyroclastic deposits of the Kamchatka Peninsula (Russia). *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2024;28(7): 706-715. DOI 10.18699/vjgb-24-79

**Funding.** The research was carried out within the state assignment of Ministry of Science and Higher Education of the Russian Federation (theme No. 124012400285-7).

**Acknowledgements.** The authors are very thankful to colleagues from the Federal Scientific Center of the East Asia Terrestrial Biodiversity FEB RAS, Vladivostok: to S.Yu. Grishin for providing samples from the Shiveluch volcano and to A.A. Goncharov for valuable discussions during the preparation of the manuscript.

# Молекулярно-генетическая и морфологическая характеристика *Micractinium thermotolerans* и *M. inermum* (Trebouxiophyceae, Chlorophyta) из пирокластических отложений полуострова Камчатка (Россия)

Р.З. Сущенко, В.Ю. Никулин, В.Б. Багмет, А.Ю. Никулин

Федеральный научный центр биоразнообразия наземной биоты Восточной Азии Дальневосточного отделения Российской академии наук, Владивосток, Россия
allaguvatova@yandex.ru

**Аннотация.** В ходе исследования разнообразия водорослей пирокластических отложений вулканов Шивелуч и Горелый (полуостров Камчатка) были выделены *Chlorella*-подобные штаммы зеленых водорослей VCA-72 и VCA-93 из проб, отобранных вдоль русла р. Байдарная на вулкане Шивелуч и на выходе термальных паров по краю кальдеры на южном склоне вулкана Горелый в 2018 и 2020 гг. соответственно. Идентификация штаммов выполнялась в рамках комплексного подхода микроскопическими и молекулярно-генетическими методами, включающими предварительную идентификацию, получение нуклеотидных последовательностей малой субъединицы и внутреннего транскрибируемого спейсера рРНК, построение филогенетических деревьев и вторичных структур участков ITS1 и ITS2 рРНК. На филогенетическом древе штамм VCA-93 кластеризовался в видовой кладе *Micractinium thermotolerans*. При сравнении моделей спиральных доменов ITS1 и ITS2 у *M. thermotolerans* различий не выявлено. Штамм VCA-72 занимал базальное положение в кладе *M. inermum*. Модели вторичной структуры спиралей штамма VCA-72 в целом были сходны с моделями для *M. inermum*, однако отмечена внутривидовая вариабельность, обусловленная в основном заменами в верхушечных и боковых петлях. В спиральных участках исследуемых штаммов *M. inermum* обнаружено пять замен hCBC, тогда как замен CBC обнаружено не было. Детальный анализ морфологии и жизненного цикла позволил выявить в стареющих культурах клетки, которые по размеру значительно превышают вегетативные и имеют грушевидную, овальную и эллипсоидную формы с неглубоким широким сужением в центре. Также в стареющих культурах обоих видов были выявлены клетки с бесцветными липидными каплями. Способность синтезировать и накапливать липиды говорит о большом потенциале штаммов для производства биодизельного топлива. Обзор местообитаний предыдущих и новых находок позволяет сделать вывод об экологической пластичности исследуемых видов. Полученные результаты дополняют сведения о биогеографии видов: *M. inermum* обнаружен впервые на территории России, а *M. thermotolerans* – на полуострове Камчатка.

**Ключевые слова:** микроводоросли; флористические находки; комплексный подход; морфология; филогения; вторичная структура ITS рРНК.

## Introduction

*Micractinium* Fresenius (Trebouxiophyceae, Chlorophyta) is a genus comprising *Chlorella*-like green algae, including symbiotic (*M. conductrix* (Brandt) Pröschold et Darienko, *M. tetrahymenae* Pröschold, Pitsch et Darienko) and free-living organisms. The genus currently comprises 24 species (Guiry M.D., Guiry G.M., 2024). *M. pusillum* Fresenius is the type species. It is characterized by a coccoid organization and the formation of colonies of 2–4 cells and bristles (Fresenius, 1858). It was assumed that the ability to form colonies and the presence of bristles are distinctive features of this species (Komárek, Fott, 1983). However, it was shown that these traits only appeared as a protective mechanism in response to co-cultivation with the rotifer *Brachionus calyciflorus* Pallas; in the absence of algophages, the features did not appear (Luo et al., 2005, 2006). In addition, the presence of zooplankton did not always result in the formation of bristles and colonies in *Micractinium* species (Pröschold et al., 2011). For example, *M. conductrix*, *M. inermum* Hoshina et Fujiwara and other species of the genus do not have bristles and are morphologically similar to algae of the genus *Chlorella* (Pröschold et al., 2011; Hoshina, Fujiwara, 2013; Hong et al., 2015).

Homoplastic characters leading to similarities between the genera *Chlorella* and *Micractinium* make it difficult to identify taxa using only morphological data. The use of an integrative approach, combining traditional microscopy methods and molecular phylogenetic analysis, allows to distinguish not only taxa poor in diagnostic characters, but also cryptic species (Komárek et al., 2014; Darienko, Pröschold, 2019).

Members of the genus *Micractinium* are well-known objects of biotechnology research. F. Quintas-Nunes et al. (2023) showed the growth-stimulating effect of exudates of *Micractinium* sp. NFX-FRZ on tomato plants. This could be due to phytohormones synthesized by the algae. *M. inermum* F014 was found to be able to treat radioactive wastewater (Kim et al., 2019). L. Bouarab et al. (2004) found that *M. pusillum* can grow well under mixotrophic conditions. It may provide opportunities for the industrial cultivation of microalgae and achieve high culture densities. Some studies have confirmed the suitability of *Micractinium* sp. for biofuel production (Abou-Shanab et al., 2014; Onay et al., 2014).

During the study on algal diversity of pyroclastic deposits of the Kamchatka Peninsula, *Chlorella*-like strains of green algae, previously identified as species of the genus *Micractinium*, were isolated. The aim of this study was species identification of *Micractinium* representatives using an integrative approach.

## Materials and methods

**Sampling, isolation and cultivation of algal strains.** The materials for this study were *Chlorella*-like clonal cultures of green algae VCA-72 and VCA-93. Strain VCA-72 was isolated from a sample of pyroclastic deposits collected in 2018 along the Baydarnaya river bed on the Shiveluch volcano (56°33.98′ N, 161°8.41′ E). Strain VCA-93 was detected from a sample collected in 2020 at the outlet of thermal vapors along the edge of the caldera on the southern slope of the Gorely volcano, where the temperature of the deposits was ~32 °C, (52°33.306′ N, 158°01.742′ E) (Fig. 1). The biotope on the Gorely volcano was characterized by a lack of vegetation. Melting of nearby snowfields and evaporation of moisture were observed during sampling. Sampling was carried out using classical microbiological methods (Gollerbah, Shtina, 1969).

A soil sample weighing not more than 1 g was inoculated on Petri dishes with sterile liquid modified Waris-H medium (McFadden, Melkonian, 1986; Andersen, 2005) (Supplemen-

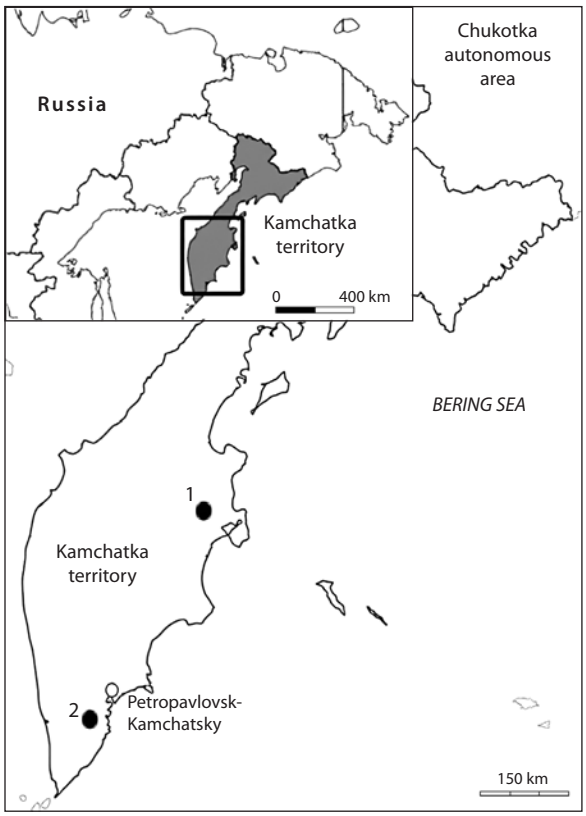



**Fig. 1.** Map of the study area and sampling sites: 1 – Shiveluch volcano, 2 – Gorely volcano.

tary Material 1)[1] and liquid modified Bold Basal Medium with triple nitrogen and vitamins (Starr, Zeikus, 1993; Schlösser, 1997; Andersen, 2005) (Supplementary Material 1) to obtain enrichment cultures. Enrichment cultures were periodically checked for algal growth using an Olympus CK30 inverted microscope (Olympus, Japan) with a maximum magnification of ×400.

Pure cultures were isolated using the micropipette method (Andersen, 2005) and grown in modified Waris-H liquid medium. Algal cultures were maintained at 117–120 lux illumination, 24.9 °C, 16 % humidity, and 16:8 h light:dark cycle.

**Light microscopy, morphological characterization.** The morphology of the strains was examined with Olympus BX 53 (Olympus, Japan), equipped with Nomarski DIC optics. Microphotographs were taken with an Olympus DP 27 camera (Olympus, Japan) at ×1000 magnification. The parameters of 50 vegetative cells were analyzed to identify the boundaries of variation in morphological characteristics for each strain.

**Molecular genetic analysis.** Taxonomic identification of strains was performed by molecular genetic methods, including obtaining nucleotide sequences of the small subunit and internal transcribed spacer rRNA (18S+ITS rRNA; according to the protocol outlined by V.Yu. Nikulin et al. (2023)), construction of phylogenetic trees and secondary structures of the ITS1 and ITS2 rRNA regions.

[1] Supplementary Materials 1 and 2 are available at: https://vavilov.elpub.ru/jour/manager/files/Suppl_Sushchen_Engl_28_7.pdf

For DNA isolation, cell biomass was sampled during the exponential growth phase and concentrated by centrifugation. Total genomic DNA was isolated according to the method of C.S. Echt et al. (1992) with some modifications (Abdullin et al., 2021). Amplification was performed by polymerase chain reaction (PCR) in a T100 Thermal Cycler amplifier (Bio-Rad Laboratories, Inc., USA) with Encyclo Plus kit (Evrogen, Russia), primers 82F (5′-GAAACTGCGAATGGCTC-3′) (López-García et al., 2003) and ITS4R (5′-CCTCCGCT TATTGATATGC-3′) (White et al., 1990). PCR parameters were as follows: initial denaturation at 96 °C for 3 min, followed by 30 cycles including denaturation at 96 °C for 1 min, annealing at 55 °C for 2 min, elongation at 68 °C for 3 min. This was followed by a final elongation at 68 °C for 7 min (Mikhailyuk et al., 2018).

Sequencing was performed using the equipment of the Instrumental Centre of Biotechnology and Gene Engineering of FSCEATB FEB RAS, ABI 3500 genetic analyzer (Applied Biosystems, USA). PCR products were sequenced in both directions with BigDye Terminator sequencing kit v. 3.1 (Applied Biosystems, USA) and the same primers as used for PCR. Additionally, primers SSU528F-800 (5′-CGGT AATTCCAGCTCC-3′) (Hoef-Emden, Melkonian, 2003), 920F (5′-GAAACTTAAAKGAATTG-3′) (Marin et al, 1998), n1400R (5′-GGTAGGAGCGACGGGCGGTGTGTAC-3′) (Marin et al., 2003), and Bd18SF1 (5′-TTTGTACACACCG CCCGTCGC-3′) (Goka et al., 2009) were used. Sequences were assembled with the Staden v.1.4 software package (Bonfield et al., 1995) and compared with other strains available at the National Center for Biotechnology Information (NCBI, USA) using a BLAST search (https://blast.ncbi.nlm.nih.gov/Blast.cgi). The selection of representative sequences for phylogenetic analysis was based on a dataset of green algae of the genus *Micractinium* (Krivina et al., 2023), which included 48 18S+ITS rRNA sequences; 2,418 aligned positions. The sequence of the taxon *Chlorella vulgaris* Beijerinck, representing a phylogenetically distant lineage, was added to the dataset as an outgroup. Sequences were aligned in the SeaView program (Galtier et al., 1996) with manual alignment correction. The best-fit model of nucleotide substitutions for our dataset was determined based on the Akaike Information Criterion (AIC) (Akaike, 1974) in the jModelTest 2.1.1 program (Darriba et al., 2012).

Phylogenetic trees were constructed using the maximum likelihood (ML) method and the Bayesian approach (BI). ML analysis was performed using the RAxML web server v. 7.7.1 (http://embnet.vital-it.ch/raxml-bb/) (Kozlov et al., 2019); BI was performed using the MrBayes 3.1.2 program (Huelsenbeck, Ronquist, 2001). In BI analysis, 5 million generations of Markov chains were created, sampling every 100 generations, i. e., 50,000 samples. The first 25 % of samples (before -lnL values reached a plateau) were excluded from the analysis as "burn-in". Markov chain Monte Carlo convergence (MCMC) to a stationary distribution was assessed visually using the Tracer 1.7.1 program (Rambaut et al., 2018) by plotting posterior probabilities. All ESS values were greater than 200. The stability of ML-derived phylogenetic tree nodes was calculated using the RAxML server using the bootstrap method (Bootstrap Percentage, BP) (Stamatakis et

al., 2008) and by determining the posterior probabilities (PP) in the BI. BP values less than 50 % and PP values less than 0.95 were not considered. Phylogenetic trees were visualized using the FigTree v. 1.4.4 program (Rambaut, 2018).

To confirm strain identification, secondary structures of the ITS1 and ITS2 rRNA regions were compared between phylogenetically related sequences. Secondary structures were constructed based on models developed for representatives of the genus *Micractinium* (Chae et al., 2019; Krivina et al., 2023) using the UNAFold Web Server (http://www.unafold.org/mfold/applications/rna-folding-form.php) (Zuker, 2003) and visualized in the VARNA program (Darty et al., 2009). Next, compensatory and hemicompensatory base substitutions (CBC, hCBC) (Caisová et al., 2013) and other molecular features that distinguish strains were searched for.

## Results and discussion

### Molecular genetic analysis

The sequences of the region comprising 18S–ITS1–5.8S–ITS2 rRNA of strains VCA-93 and VCA-72 were deposited in the GenBank database under accession numbers PP501334 and PP501335, respectively. BLAST searches revealed a high percentage of similarity to the sequences of *Micractinium* sp. (*M. thermotolerans*) (Krivina et al., 2023) ACSSI 332 MT784118 (99.91 %) and *M. inermum* NLP-F014 KF597304 (99.29 %), respectively.

On the phylogenetic tree, strain VCA-93 was clustered in the topologically established species clade of *M. thermotolerans* with strains ACSSI 332 (holotype) and IC-76 (Fig. 2). All three strains were found in the Russian Federation (Kamchatka Peninsula, Chukotski Peninsula (Krivina et al., 2023), and West Siberian Plain (Piligaev et al., 2018), respectively). Related to them is a clade with moderate statistical support (73/0.98; BP/PP) including strains of *Micractinium* sp. from Africa – TvB (isolated from Tiberias hot springs), SH (from a sinkhole near Ein Gedi), CCAP 211/92 (soil sample from Seychelles).

As noted earlier, in contrast to "African" strains, all representatives of *M. thermotolerans* are characterized by the absence of an intron in the second quarter of the 18S gene (Krivina et al., 2023). This is also true for strain VCA-93. The sister species was *M. tetrahymenae* SAG 2587.

Strain VCA-72 occupied a basal position in the moderately supported clade of *M. inermum* (81/0.98) (Fig. 2). In sister position was a clade (89/0.99) composed of closely related strains found in North America, Europe, and Asia: HS26 (Sonora Desert in Arizona, USA) (Ganuza et al., 2016), NLP-F014 (Nakdong River, South Korea) (Park et al., 2015), KM114868 (Weston Park Pond, UK) (Smith et al., 2015), and NIES-2171 (Sendai Botanical Garden, Japan) (Hoshina, Fujiwara, 2013). Thus, there was no geographic structuring in the *M. inermum* clade. The species clades of *M. lacustre*, *M. variabile*, *M. simplicissimum* and one specimen of *M. singularis* were the closest to the *M. inermum* clade.

### ITS1 and ITS2 rRNA secondary structures

The secondary structures of the ITS1 and ITS2 regions of the studied strains corresponded to the generally accepted models developed for eukaryotic organisms, in particular, green algae (Coleman, 2000, 2015). The models of ITS1 and ITS2 helical domains of strain VCA-93 and *M. thermotolerans* strains ACSSI 332 and IC-76 (Supplementary Material 2) were characterized by the absence of substitutions between them.

The presented models of the secondary structure of the helical domains of strain VCA-72 were generally similar to those for *M. inermum* (Fig. 3).

Helix I in ITS1 and helix I, II in ITS2 were monomorphic, but intraspecific variability was observed in all other ITS helices of the compared *M. inermum* strains. Despite this, the conservatism of the structure is due to the predominant localization of substitutions in apical or lateral loops and the presence of hCBCs that preserve base pairing. In terms of the number of nucleotide differences, the ITS1 regions were expectedly less conservative compared to ITS2 (ten versus four differences). The majority of nucleotide substitutions in both spacers (ten substitutions) distinguished our strain from the other four, but there were four substitutions and deletions characterizing specific strains (Fig. 3). Five hCBC substitutions were detected in the helixes (three hCBCs in ITS1: U→C in pos. 12 and 34 of helix III, A→G in pos. 17 of helix IV; two hCBCs in ITS2: C→U in pos. 29 of helix III, U→C in pos. 26 of helix IV), whereas no CBC substitutions were detected. The topology of the basal part of helix IV ITS2 strain KM114868 differed from the others (Fig. 3).

According to the CBC concept (Wolf et al., 2013), the absence of CBC in ITS2 indicates that the compared strains belong to the same species. Thus, based on the results of phylogenetic analysis and modelling of secondary structures, we reliably identified the strains under study: VCA-93 belongs to the species *M. thermotolerans* Krivina, Sinetova, Savchenko, Degtyarev, Tebina et Temraleeva, and VCA-72 belongs to *M. inermum*. The studied genotype of the latter, due to the presence of unique nucleotide substitutions, allowed us to add new data to the pool of molecular diversity of the ITS rRNA region of *M. inermum* species.

### Morphology, reproduction and ecology

***Micractinium thermotolerans* Krivina, Sinetova, Savchenko, Degtyarev, Tebina et Temraleeva** (Fig. 4*a–f*). The cells are spherical, 3.2–6.5 µm in diameter, without bristles (Fig. 4*a*, *b*). Young cells are triangular, ellipsoidal (3.0–5.6 × 3.3–5.9 µm) or irregular. The chloroplast is parietal, cup-shaped with a spherical pyrenoid, covered by starch grains. Reproduction by 2–4 autospores (Fig. 4*c–e*). The sporangium size was 4.4–6.7 µm in diameter. Autospores were uniform in size (up to 2 microns in diameter), triangular or irregular in shape and showed release by rupture of the sporangium cell wall. Cell walls remain visible in culture after release of autospores.

The cell wall is thin, with uniform thickening in older cultures. The cells are pear-shaped, oval and ellipsoidal with a shallow, wide constriction in the center, reaching a length of 7.3–10.5 µm in 6-month cultures (Fig. 4*f*).

The detection of cells significantly larger than mature vegetative cells is consistent with the observations of E. Krivina et al. (2023). They showed that incubation at elevated, but non-lethal temperatures caused the appearance of a population

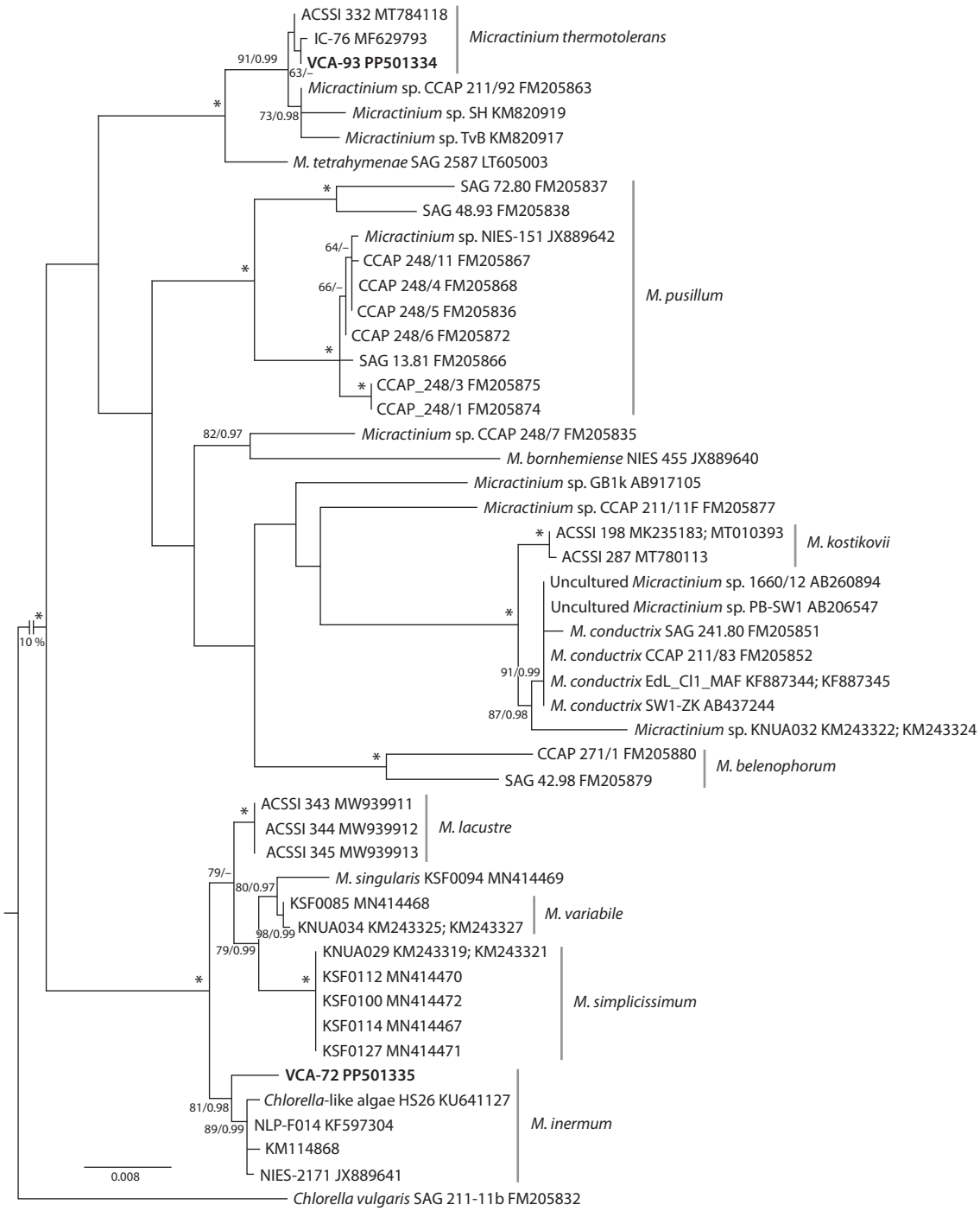



**Fig. 2.** ML tree illustrating the phylogenetic position of strains VCA-93 and VCA-72 (in bold) among members of the genus *Micractinium* based on 18S+ITS rRNA sequence comparison (2,418 aligned positions; GTR+I+G model).

Node supports in ML/BI analyses (BP ≥ 50 % and PP ≥ 0.95) are indicated. Nodes with maximum support (100/1.00) are indicated by asterisks. The branch belonging to the outgroup is shortened (only 10 % of the length is shown). Scale bar is the number of nucleotide substitutions per position.

of single or abnormally dividing giant cells with a diameter of 10.8–19.3 μm in the culture. In our case, the *M. thermotolerans* VCA-93 strain was cultured at 24.9 °C, and the appearance of abnormal cells with lipid droplets was probably a result of culture depletion. It was associated with its aging. It is similar to the results obtained for *M. thermotolerans* ACSSI 332 under nitrogen starvation conditions.

E. Krivina et al. (2023) obtained preliminary data on the fatty acid composition of *M. thermotolerans*. Thus, during the description of the species, the composition of methyl esters of fatty acids of strain ACSSI 332 was revealed (hexadecanoic, 7,10,13-hexadecatrienic, 9,12,15-octadecatrienic, pentadecanoic acids, etc.). It differed significantly from the fatty acid composition of other species of the genus *Micractinium* with

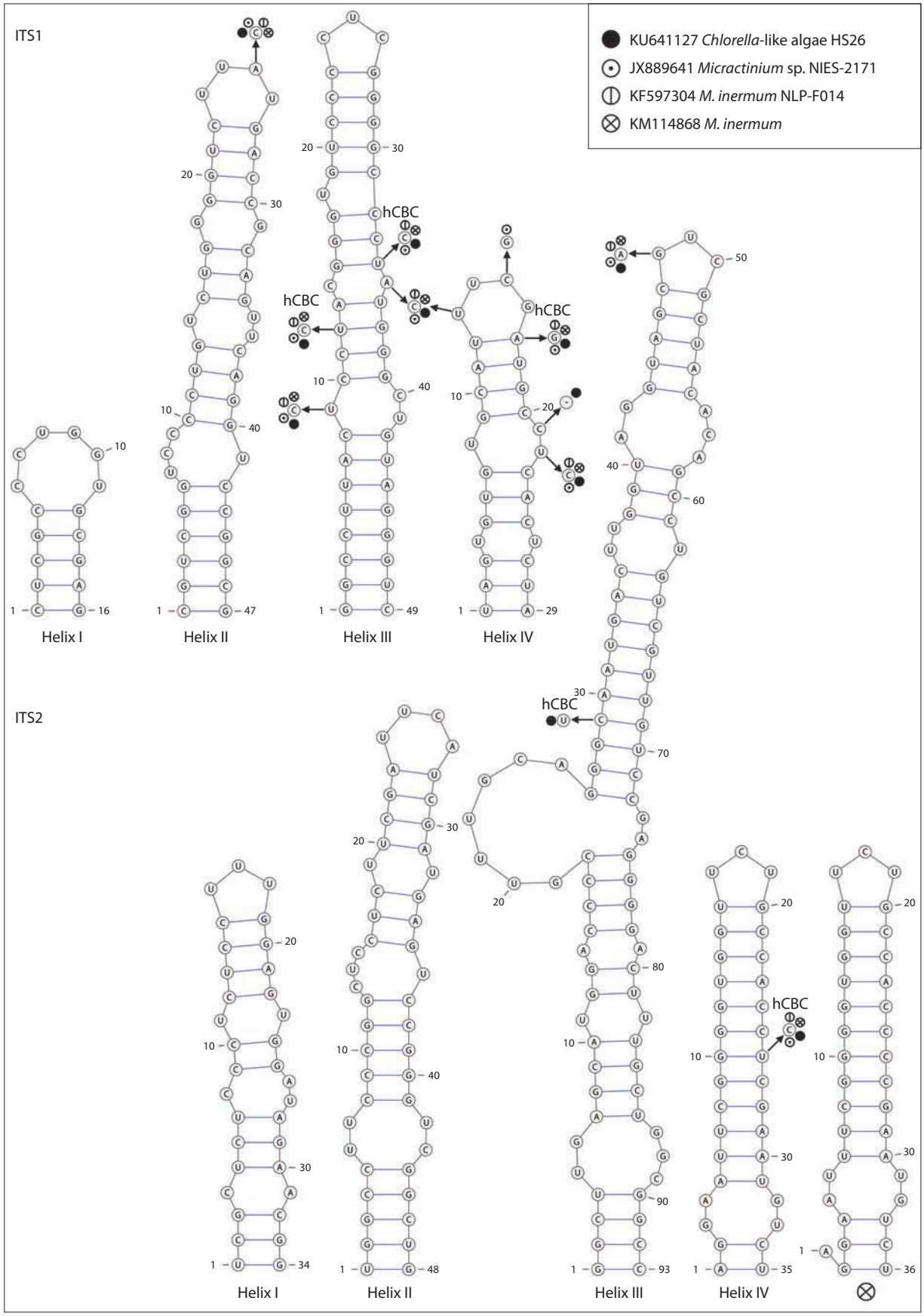



**Fig. 3.** Secondary structure models of the ITS1 and ITS2 helical domains of strain VCA-72.

Nucleotide sequence differences with *M. inermum* strains HS26, NLP-F014, KM114868, and NIES-2171 are indicated by endnotes according to the legend. Helix IV ITS2 of strain KM114868 is shown separately (bottom right) due to topological differences.

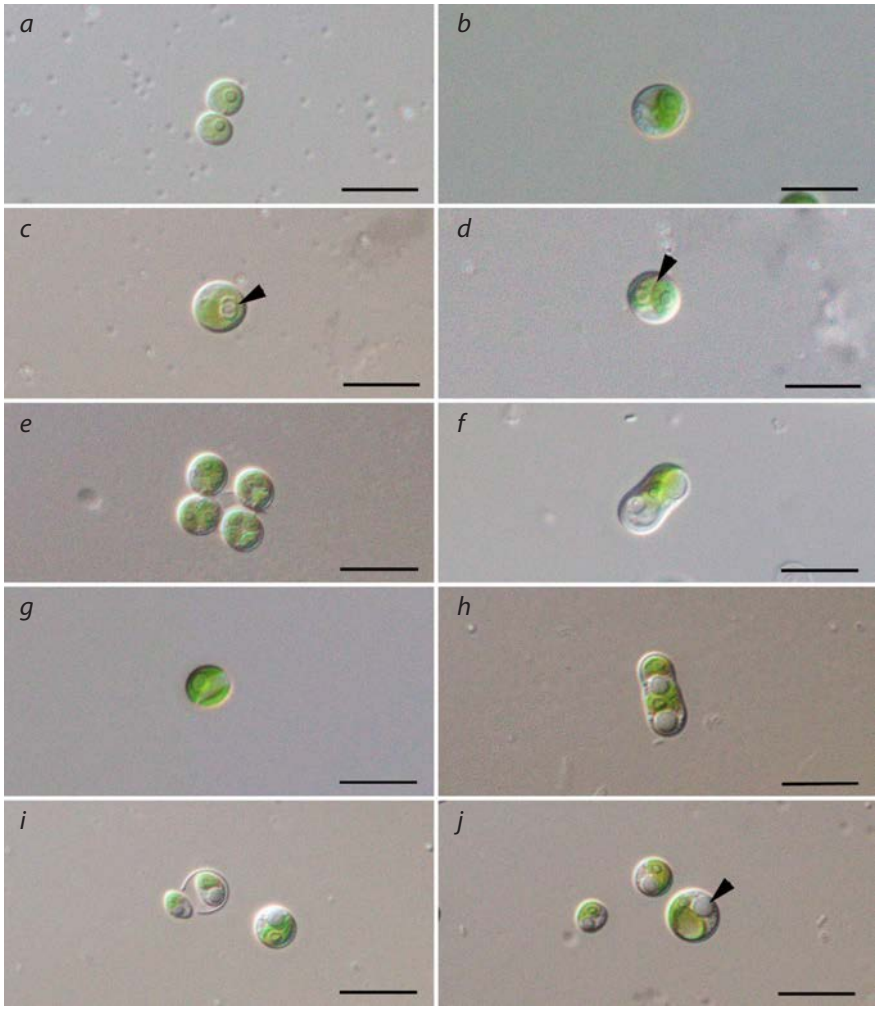



**Fig. 4.** Microphotographs of the strains *M. thermotolerans* VCA-93 (*a–f*) and *M. inermum* VCA-72 (*g–j*).
*a*, *b* – vegetative cells; *c* – presporangial cell (arrow indicates doubling of the pyrenoid); *d* – presporangial cell (arrow indicates doubling of the protoplast); *e* – autosporangia; *f* – ellipsoid cell with a shallow, wide constriction in the center with vacuoles; *g* – vegetative cell; *h* – sporangium in an aging culture; *i* – release of autospores; *j* – cell in an aging culture with a lipid droplet and vacuole (lipid droplet is indicated by an arrow). Scale bar: 10 μm.

greater complexity and diversity (Krivina et al., 2023). The authors noted the biotechnological potential of this species.

Two of the three known strains of *M. thermotolerans* were isolated from extreme habitats: VCA-93 from tephra collected at the outlet of thermal vapors along the edge of the caldera of the Gorely volcano (Kamchatka Peninsula) and ACSSI 332 from a hot spring located on the Chukotka Peninsula (Krivina et al., 2023). At the same time, strain IC-76 was isolated from river sand from the coast of the Ob river (Novosibirsk region) (Piligaev et al., 2018), which may indicate the ecological plasticity of the species.

***Micractinium inermum* Hoshina et Fujiwara** (Fig. 4*g–j*). The solitary cells are spherical (4.3–5.0 μm) (Fig. 4*g*), drop-shaped or ellipsoidal (2.2–4.7 × 3.0–5.0 μm), without bristles. The chloroplast is single, cup-shaped, with a pronounced pyrenoid. Asexual reproduction by two autospores (Fig. 4*i*). Cells in old cultures are spherical (5.7–7.9 μm) or ellipsoidal with a shallow, wide constriction in the center (Fig. 4*h*), 8.4–10.7 μm long and are characterized by the presence of lipid droplets proportional to their size (Fig. 4*i*, *j*).

In the cytoplasm of aging and resting cells, there is an accumulation of single small or large lipid droplets (Andreeva, 1998). It can be colorless or yellow, orange, red. Probably, the color of lipid droplets is associated with carotenoids and their derivatives, which are accumulated in lipid globules. They can be detected by light microscopy in the form of spherical colored bodies in the resting stages of many green algae species (Weiss, 1983). For example, it has been shown that the reddish color of lipid droplets of *Haematococcus pluvialis* Flotow is due to the presence of the fat-soluble carotenoid astaxanthin (Ota et al., 2018). According to our observations, aging cells of *M. inermum* VCA-72 are characterized by the presence of colorless lipid droplets (Fig 4*h–j*).

The detection of lipid droplets in cells is also characteristic of culture depletion, including nitrogen starvation. For example, J. Zhan et al. (2016) found significant changes in the lipid content of *Chlorella* sp. under nitrogen depletion of the culture. Nutrient deficiency, as well as high light intensity and high salt concentration, is an environmental stressor and causes the accumulation of lipids or carbohydrates (Ho et al., 2012; Fernandes et al., 2013; Roleda et al., 2013; Park et al., 2015).

It is known that the *Micractinium* species studied by us are characterized by high growth rates and the ability to synthesize and accumulate lipids (Park et al., 2015; Shi et al., 2019; Krivina et al., 2023). This aspect indicates their great potential for the production of biodiesel (Wijffels, Barbosa, 2010). There have been a number of works aimed at identifying the optimal culture conditions for *M. inermum*, which allow increasing its biological productivity and reducing the cost of maintaining cultures. For example, it has been shown that the maximum lipid productivity of the JL1 strain is achieved by adding glucose to the heterotrophic culture (Shi et al., 2019). The fatty acid profile of algae is of great importance in biodiesel production as it determines the key properties of the fuel (Knothe, 2009). For example, the percentage of oleic acid serves as an indicator of fuel quality (Knothe, 2009). A. Banskota et al. (2024) identified oleic, linoleic and palmitic acids in the biomass of *M. inermum*.

S. Park et al. (2015) suggested the use of wastewater mixture for *M. inermum* NLP-F014 cultivation. The lipid accu-

mulation was up to 40 % under culture depletion conditions. This method can significantly reduce the cost of water and nutrient requirements, and hence cost of cultivation. T. Sydney et al. (2018) treated *M. inermum* culture with ultraviolet B (UVB) to reduce the energy required for cell wall disruption and lipid extraction. This resulted in an increased yield of fatty acid methyl esters. Thus, strains of this species are candidates for biofuel production.

Most publications (Park et al., 2015; Smith et al., 2015; Dickinson et al., 2019; Shi et al., 2019; Banskota et al., 2024) indicate freshwater habitats for *M. inermum*. Probably, small sizes and rapid reproductive rates allow representatives of this species not only to be planktonic in water bodies, but also to survive in terrestrial habitats, including in such extreme biotopes as the volcanic deposits of Kamchatka (pyroclastic deposits along the Baydarnaya river bed on the Shiveluch volcano).

## Conclusion

As a result of the study of algal diversity in the pyroclastic deposits of the Shiveluch and Gorely volcanoes (Kamchatka Peninsula), using an integrative approach, representatives of the genus *Micractinium* were identified. The results obtained complement the information on the secondary structure of the ITS1 and ITS2 rRNA regions, morphology (morphology of cells in aging cultures of *M. inermum* and *M. thermotolerans*), life cycle (the life cycle of *M. inermum* is considered in more detail), ecology (the species are among the primary colonizers of lifeless substrate on the Kamchatka Peninsula; vital activity of *M. thermotolerans* at deposit temperature ~32 °C) and biogeography (*M. inermum* is reported for the first time in Russia, and *M. thermotolerans* is the first finding for the Kamchatka Peninsula) of the discovered species.

## References

Abdullin Sh.R., Nikulin A.Yu., Bagmet V.B., Nikulin V.Yu., Goncharov A.A. New cyanobacterium *Aliterella vladivostokensis* sp. nov. (Aliterellaceae, Chroococcidiopsidales), isolated from temperate monsoon climate zone (Vladivostok, Russia). *Phytotaxa*. 2021;517: 221-233. DOI 10.11646/phytotaxa.527.3.7

Abou-Shanab R.A.I., El-Dalatony M.M., El-Sheekh M.M. Cultivation of a new microalga, *Micractinium reisseri*, in municipal wastewater for nutrient removal, biomass, lipid, and fatty acid production. *Biotechnol. Bioprocess Eng*. 2014;19:510-518. DOI 10.1007/s12257-013-0485-z

Akaike H. A new look at the statistical model identification. *IEEE Trans. Autom. Control*. 1974;19:716-723. DOI 10.1109/TAC.1974.1100705

Andersen R.A. Algal Culturing Techniques. New York: Elsevier Acad. Press, 2005

Andreyeva V.M. Soil and Aerophilic Green Algae (Chlorophyta: Tetrasporales, Chlorococcales, Chlorosarcinales). St. Petersburg: Nauka Publ., 1998 (in Russian)

Banskota A.H., Hui J.P.M., Jones A., McGinn P.J. Characterization of neutral lipids of the oleaginous alga *Micractinum inermum*. *Molecules*. 2024;29:359. DOI 10.3390/molecules29020359

Bonfield J.K., Smith K.F., Staden R. A new DNA sequence assembly program. *Nucleic Acids Res.* 1995;23(24):4992-4999. DOI 10.1093/nar/23.24.4992

Bouarab L., Dauta A., Loudiki M. Heterotrophic and mixotrophic growth of *Micractinium pusillum* Fresenius in the presence of acetate and glucose: effect of light and acetate gradient concentration. *Water Res.* 2004;38(11):2706-2712. DOI 10.1016/j.watres.2004.03.021

Caisová L., Marin B., Melkonian M. A consensus secondary structure of ITS2 in the Chlorophyta identified by phylogenetic reconstruction. *Protist*. 2013;164(4):482-496. DOI 10.1016/j.protis.2013.04.005

Chae H., Lim S., Kim H.S., Choi H.G., Kim J.H. Morphology and phylogenetic relationships of *Micractinium* (Chlorellaceae, Trebouxiophyceae) taxa, including three new species from Antarctica. *Algae*. 2019;34(4):267-275. DOI 10.4490/algae.2019.34.10.15

Coleman A.W. The significance of a coincidence between evolutionary landmarks found in mating affinity and a DNA sequence. *Protist*. 2000;151(1):1-9. DOI 10.1078/1434-4610-00002

Coleman A.W. Nuclear rRNA transcript processing versus internal transcribed spacer secondary structure. *Trends Genet*. 2015;31(3):157-163. DOI 10.1016/j.tig.2015.01.002

Darienko T., Pröschold T. Reevaluation and discovery of new species of the rare genus *Watanabea* and establishment of *Massjukichlorella* gen. nov. (Trebouxiophyceae, Chlorophyta) using an integrative approach. *J. Phycol*. 2019;55:493-499. DOI 10.1111/jpy.12830

Darriba D., Taboada G., Doallo R., Posada D. jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods*. 2012;9: 772. DOI 10.1038/nmeth.2109

Darty K., Denise A., Ponty Y. VARNA: interactive drawing and editing of the RNA secondary structure. *Bioinformatics*. 2009;25:1974-1975. DOI 10.1093/bioinformatics/btp250

Dickinson K.E., Lalonde C.G., McGinn P.J. Effects of spectral light quality and carbon dioxide on the physiology of *Micractinium inermum*: growth, photosynthesis, and biochemical composition. *J. Appl. Phycol*. 2019;31:3385-3396. DOI 10.1007/s10811-019-01880-z

Echt C.S., Erdahl L.A., McCoy T.J. Genetic segregation of random amplified polymorphic DNA in diploid cultivated alfalfa. *Genome*. 1992;35(1):84-87. DOI 10.1139/g92-014

Fernandes B., Teixeira J., Dragone G., Vicente A.A., Kawano S., Bišová K., Vítová M. Relationship between starch and lipid accumulation induced by nutrient depletion and replenishment in the microalga *Parachlorella kessleri*. *Bioresour. Technol*. 2013;144:268-274. DOI 10.1016/j.biortech.2013.06.096

Fresenius G. Beiträge zur Kenntniss mikrokopischer Organismen. *Abh. Senckenberg. Naturforsch. Ges*. 1858;2(2):211-242. DOI 10.5962/bhl.title.2137

Galtier N., Gouy M., Gautier C. Seaview and phylo-win: two graphic tools for sequence alignment and molecular phylogeny. *Comput. Appl. Biosci*. 1996;12:543-548. DOI 10.1093/bioinformatics/12.6.543

Ganuza E., Sellers C.E., Bennett B.W., Carney L.T. A novel treatment protects *Chlorella* at commercial scale from the predatory bacterium *Vampirovibrio chlorellavorus*. *Front. Microbiol*. 2016;7:188348. DOI 10.3389/fmicb.2016.00848

Goka K., Yokoyama J., Une Y., Kuroki T., Suzuki K., Nakahara M., Kobayashi A., Inaba S., Mizutani T., Hyatt A.D. Amphibian chytridiomycosis in Japan: distribution, haplotypes and possible route of entry into Japan. *Mol. Ecol*. 2009;18(23):4757-4774. DOI 10.1111/j.1365-294X.2009.04384.x

Gollerbah M.M., Shtina E.A. Soil Algae. St. Petersburg: Nauka Publ., 1969 (in Russian)

Guiry M.D., Guiry G.M. AlgaeBase. World-wide electronic publication, National University of Ireland, Galway. (Cited on April 15, 2024). Available from: http://www.algaebase.org

Ho S.H., Chen C.Y., Chang J.S. Effect of light intensity and nitrogen starvation on $CO_2$ fixation and lipid/carbohydrate production of an indigenous microalga *Scenedesmus obliquus* CNW-N. *Bioresour. Technol*. 2012;113:244-252. DOI 10.1016/j.biortech.2011.11.133

Hoef-Emden K., Melkonian M. Revision of the genus *Cryptomonas* (Cryptophyceae): a combination of molecular phylogeny and mor-

phology provides insights into a long-hidden dimorphism. *Protist*. 2003;154:371-409. DOI 10.1078/143446103322454130

Hong J.W., Jo S.-W., Cho H.-W., Nam S.W., Shin W., Park K.M., Lee K.I., Yoon H.-S. Phylogeny, morphology, and physiology of *Micractinium* strains isolated from shallow ephemeral freshwater in Antarctica. *Phycol. Res*. 2015;63:212-218. DOI 10.1111/pre.12097

Hoshina R., Fujiwara Y. Molecular characterization of *Chlorella* cultures of the National Institute for Environmental Studies culture collection with description of *Micractinium inermum* sp. nov., *Didymogenes sphaerica* sp. nov., and *Didymogenes soliella* sp. nov. (Chlorellaceae, Trebouxiophyceae). *Phycol. Res*. 2013;31:124-132. DOI 10.1111/pre.12010

Huelsenbeck J.P., Ronquist F. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics.* 2001;17:754-755. DOI 10.1093/bioinformatics/17.8.754

Kim I., Yang H.-M., Park C.W., Yoon I.-H., Seo B.-K., Kim E.-K., Ryu B.-G. Removal of radioactive cesium from an aqueous solution via bioaccumulation by microalgae and magnetic separation. *Sci. Rep*. 2019;9:10149. DOI 10.1038/s41598-019-46586-x

Knothe G. Improving biodiesel fuel properties by modifying fatty ester composition. *Energy Environ. Sci*. 2009;2:2759-2766. DOI 10.1039/B903941D

Komárek J., Fott B. Chlorophyceae (Grünalgen). Ordnung Chlorococcales. In: Huber-Pestalozzi G. (Ed.) Das Phytoplankton des Süsswassers. 7. Teil. 1. Stuttgart: Schweizerbart'sche Verlagsbuchhandlung, 1983;1-1044

Komárek J., Kaštovský J., Mareš J., Johansen J.R. Taxonomic classification of cyanoprokaryotes (cyanobacterial genera), using a polyphasic approach. *Preslia*. 2014;86(4):295-335

Kozlov A.M., Darriba D., Flouri T., Morel B., Stamatakis A. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics*. 2019;35:4453-4455. DOI 10.1093/bioinformatics/btz305

Krivina E., Sinetova M., Savchenko T., Degtyaryov E., Tebina E., Temraleeva A. *Micractinium lacustre* and *M. thermotolerans* spp. nov. (Trebouxiophyceae, Chlorophyta): taxonomy, temperature-dependent growth, photosynthetic characteristics and fatty acid composition. *Algal Res*. 2023;71:103042. DOI 10.1016/j.algal.2023.103042

López-García P., Philippe H., Gail F., Moreira D. Autochthonous eukaryotic diversity in hydrothermal sediment and experimental microcolonizers at the Mid-Atlantic Ridge. *Proc. Natl. Acad. Sci. USA*. 2003;100:697-702. DOI 10.1073/pnas.0235779100

Luo W., Krienitz L., Pflugmacher S., Walz N. Genus and species concept in *Chlorella* and *Micractinium* (Chlorophyta, Chlorellaceae): genotype versus phenotypical variability under ecosystem conditions. *SIL Proceedings*. 2005;29(1):170-173. DOI 10.1080/03680770.2005.11901988

Luo W., Pflugmacher S., Pröschold T., Walz N., Krienitz L. Genotype versus phenotype variability in *Chlorella* and *Micractinium* (Chlorophyta, Trebouxiophyceae). *Protist*. 2006;157:315-333. DOI 10.1016/j.protis.2006.05.006

Marin B., Klingberg M., Melkonian M. Phylogenetic relationships among the Cryptophyta: analyses of nuclear-encoded SSU rRNA sequences support the monophyly of extant plastid-containing lineages. *Protist*. 1998;149:265-276. DOI 10.1016/S1434-4610(98)70033-1

Marin B., Palm A., Klingberg M., Melkonian M. Phylogeny and taxonomic revision of plastid-containing euglenophytes based on SSU rDNA sequence comparisons and synapomorphic signatures in the SSU rRNA secondary structure. *Protist*. 2003;154:99-145. DOI 10.1078/143446103764928521

McFadden G.I., Melkonian M. Use of Hepes buffer for microalgal culture media and fixation for electron microscopy. *Phycologia*. 1986;25:551-557. DOI 10.2216/i0031-8884-25-4-551.1

Mikhailyuk T., Lukešová A., Glaser K., Holzinger A., Obwegeser S., Nyporko S., Friedl T., Karsten U. New taxa of streptophyte algae (Streptophyta) from terrestrial habitats revealed using an integrative approach. *Protist*. 2018;169:406-431. DOI 10.1016/j.protis.2018.03.002

Nikulin V.Yu., Nikulin A.Yu., Gontcharov A.A., Bagmet V.B., Abdullin Sh.R. *Oogamochlamys kurilensis* sp. nov. (Chlorophyta, Volvocales) from the soils of Iturup Island (Sakhalin Region, Russia). *Plants*. 2023;12:3350. DOI 10.3390/plants12193350

Onay M., Sonmez C.A., Oktem H., Yücel M. Thermo-resistant green microalgae for effective biodiesel production: isolation and characterization of unialgal species from geothermal flora of Central Anatolia. *Bioresour. Technol*. 2014;169:62-71. DOI 10.1016/j.biortech.2014.06.078

Ota S., Morita A., Ohnuki S., Hirata A., Sekida S., Okuda K., Ohya Y., Kawano S. Carotenoid dynamics and lipid droplet containing astaxanthin in response to light in the green alga *Haematococcus pluvialis*. *Sci. Rep*. 2018;8:5617. DOI 10.1038/s41598-018-23854-w

Park S., Kim J., Yoon Y., Park Y., Lee T. Blending water-and nutrient-source wastewaters for cost-effective cultivation of high lipid content microalgal species *Micractinium inermum* NLP-F014. *Bioresourc. Technol*. 2015;198:388-394. DOI 10.1016/j.biortech.2015.09.038

Piligaev A.V., Sorokina K.N., Shashkov M.V., Parmon V.N. Screening and comparative metabolic profiling of high lipid content microalgae strains for application in wastewater treatment. *Bioresour. Technol*. 2018;250:538-547. DOI 10.1016/j.biortech.2017.11.063

Pröschold T., Darienko T., Silva P.C., Reisser W., Krienitz L. The systematics of "Zoochlorella" revisited employing an integrative approach. *Environ. Microbiol*. 2011;13:350-364. DOI 10.1111/j.1462-2920.2010.02333.x

Quintas-Nunes F., Brandão P.R., Barreto Crespo M.T., Glick B.R., Nascimento F.X. Plant growth promotion, phytohormone production and genomics of the rhizosphere-associated microalga, *Micractinium rhizosphaerae* sp. nov. *Plants*. 2023;12:651. DOI 10.3390/plants12030651

Rambaut A. *FigTree v.1.4.4*. 2018. http://tree.bio.ed.ac.uk/software/figtree/ (Access date: 01.03.2024)

Rambaut A., Drummond A.J., Xie D., Baele G., Suchard M.A. Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. *Syst. Biol*. 2018;67:901-904. DOI 10.1093/sysbio/syy032

Roleda M.Y., Slocombe S.P., Leakey R.J., Day J.G., Bell E.M., Stanley M.S. Effects of temperature and nutrient regimes on biomass and lipid production by six oleaginous microalgae in batch culture employing a two-phase cultivation strategy. *Bioresour. Technol*. 2013;129:439-449. DOI 10.1016/j.biortech.2012.11.043

Schlösser U.G. Additions to the Culture collection of algae since 1994. *Bot. Acta*. 1997;110:424-429. DOI 10.1111/j.1438-8677.1997.tb00659.x

Shi M., Wei H., Chen Q., Wang X., Zhou W., Liu J. Exploring an isolate of the oleaginous alga *Micractinium inermum* for lipid production: molecular characterization and physiochemical analysis under multiple growth conditions. *J. Appl. Phycol*. 2019;31:1035-1046. DOI 10.1007/s10811-018-1653-5

Smith R.T., Bangert K., Wilkinson S.J., Gilmour D.J. Synergistic carbon metabolism in a fast growing mixotrophic freshwater microalgal species *Micractinium inermum*. *Biomass Bioenergy*. 2015;82:73-86. DOI 10.1016/j.biombioe.2015.04.023

Stamatakis A., Hoover P., Rougemont J. A rapid bootstrap algorithm for the RAxML web servers. *Syst. Biol*. 2008;57:758-771. DOI 10.1080/10635150802429642

Starr R.C., Zeikus J.A. UTEX – the culture collection of algae at the University of Texas at Austin. 1993 list of cultures. *J. Phycol*. 1993;29:1-106. DOI 10.1111/j.0022-3646.1993.00001.x

Sydney T., Marshall-Thompson J.-A., Kapoore R.V., Vaidyanathan S., Pandhal J., Fairclough J.P.A. The effect of high-intensity ultraviolet

light to elicit microalgal cell lysis and enhance lipid extraction. *Metabolites*. 2018;8:65. DOI 10.3390/metabo8040065

Weiss R.L. Fine structure of the snow alga (*Chlamydomonas nivalis*) and associated bacteria. *J. Phycol*. 1983;19:200-204. DOI 10.1111/j.0022-3646.1983.00200.x

White T.J., Bruns T.D., Lee S.B., Taylor J.W. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: PCR Protocols – A Guide to Methods and Application. San Diego, CA, USA: Acad. Press, 1990;315-322

Wijffels R.H., Barbosa M.J. An outlook on microalgal biofuels. *Science*. 2010;329:796-799. DOI 10.1126/science.1189003

Wolf M., Chen S., Song J., Ankenbrand M., Müller T. Compensatory base changes in ITS2 secondary structures correlate with the biological species concept despite intragenomic variability in ITS2 sequences – a proof of concept. *PLoS One*. 2013;8(6):e66726. DOI 10.1371/journal.pone.0066726

Zhan J., Hong Y., Hu H. Effects of nitrogen sources and C/N ratios on the lipid producing potential of *Chlorella* sp. HQ. *J. Microbiol. Biotechnol*. 2016;26:1290-1302. DOI 10.4014/jmb.1512.12074

Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* 2003;31(13):3406-3415. DOI 10.1093/nar/gkg595

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

Received April 18, 2024. Revised June 4, 2024. Accepted June 13, 2024.

PLANT GENETICS AND BREEDING
Original article

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2024;28(7):716-730

DOI 10.18699/vjgb-24-80

# Cytogenetic features of intergeneric amphydiploids and genome-substituted forms of wheat

E.D. Badaeva [1], R.O. Davoyan [2], N.A. Tereshchenko [1], E.V. Lyalina [1], S.A. Zoshchuk [3], N.P. Goncharov [4]

[1] N.I. Vavilov Institute of General Genetics of the Russian Academy of Sciences, Moscow, Russia
[2] National Center of Grain named after P.P. Lukyanenko, Krasnodar, Russia
[3] Engelhardt Institute of Molecular Biology of the Russian Academy of Sciences, Moscow, Russia
[4] Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
katerinabadaeva@gmail.com

**Abstract.** Synthetic intergeneric amphydiploids and genome-substituted wheat forms are an important source for transferring agronomically valuable genes from wild species into the common wheat (*Triticum aestivum* L.) genome. They can be used both in academic research and for breeding purposes as an original material for developing wheat-alien addition and substitution lines followed by translocation induction with the aid of irradiation or nonhomologous chromosome pairing. The chromosome sets and genome constitutions of allopolyploids are usually verified in early hybrid generations, whereas the subsequent fate of these hybrids remains unknown in most cases. Here we analyze karyotypes of five hexa- ($2n = 6x = 42$) and octoploid ($2n = 8x = 56$) amphydiploids of wheat with several species of the *Aegilops*, *Haynaldia*, and *Hordeum* genera, and six genome-substituted wheat–*Aegilops* forms, which were developed over 40 years ago and have been maintained in different gene banks. The analyses involve C-banding and fluorescence *in situ* hybridization (FISH) with pAs1 and pSc119.2 probes. We have found that most accessions are cytologically stable except for Avrodes (genome BBAASS, a hexaploid genome-substituted hybrid of wheat and *Aegilops speltoides*), which segregated with respect to chromosome composition after numerous reproductions. Chromosome analysis has not confirmed the presence of the N genome from *Ae. uniaristata* Vis. in the genome-substituted hybrid Avrotata. Instead, Avrotata carries the D genome. Our study shows that octoploid hybrids, namely AD 7, AD 7147 undergo more complex genome reorganizations as compared to hexaploids: the chromosome number of two presumably octoploid wheat-*Aegilops* hybrids were reduced to the hexaploid level. Genomes of both forms lost seven chromosome pairs, which represented seven homoeologous groups and derived from different parental subgenomes. Thus, each of the resulting hexaploids carries a synthetic/hybrid genome consisting of a unique combination of chromosomes belonging to different parental subgenomes.

**Key words:** genome stabilization; wheat; amphydiploid; *Aegilops*; *Dasypyrum*; *Tritordeum*; genome-substituted forms; karyotype; C-banding; fluorescence *in situ* hybridization.

**For citation:** Badaeva E.D., Davoyan R.O., Tereshchenko N.A., Lyalina E.V., Zoshchuk S.A., Goncharov N.P. Cytogenetic features of intergeneric amphydiploids and genome-substituted forms of wheat. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2024;28(7):716-730. DOI 10.18699/vjgb-24-80

**Funding.** This work was supported by State Budgetary Project No 122022600163-7.

**Acknowledgements.** The authors thank Dr. Andreas Börner (IPK), Gatersleben, Germany, for providing seeds of accession AE 1491 and Dr. I.G. Loskutov (VIR), St.-Petersburg, Russia, for providing seeds of *Tritordeum*.

# Цитогенетические особенности межродовых амфидиплоидов и геномно-замещенных форм пшеницы

Е.Д. Бадаева [1], Р.О. Давоян [2], Н.А. Терещенко [1], Е.В. Лялина [1], С.А. Зощук [3], Н.П. Гончаров [4]

[1] Институт общей генетики им. Н.И. Вавилова Российской академии наук, Москва, Россия
[2] Национальный центр зерна им. П.П. Лукьяненко, Краснодар, Россия
[3] Институт молекулярной биологии им. В.А. Энгельгардта Российской академии наук, Москва, Россия
[4] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
katerinabadaeva@gmail.com

**Аннотация.** Синтетические межродовые гибриды (амфидиплоиды) и геномно-замещенные формы пшеницы – важный источник для переноса хозяйственно ценных генов от диких видов в геном *Triticum aestivum* L. Их используют как для решения теоретических задач, так и в практических целях для получения дополненных или замещенных линий, а также для индукции пшенично-чужеродных транслокаций с помощью облучения или

негомологичной конъюгации хромосом. Хромосомный и геномный состав аллополиплоидных форм обычно верифицируется в ранних гибридных поколениях, часто дальнейшая судьба этих гибридов остается неизученной. В настоящей работе с помощью методов С-дифференциального окрашивания хромосом по Гимза и флуоресцентной гибридизации *in situ* (FISH) с ДНК-зондами pAs1 и pSc119.2 мы провели исследование кариотипов пяти гекса- (2*n* = 6*x* = 42) и октаплоидных (2*n* = 8*x* = 56) геномно-дополненных амфидиплоидов пшеницы с отдельными видами из родов *Aegilops*, *Haynaldia* и *Hordeum*, а также шести гексаплоидных пшенично-эгилопсных геномно-замещенных форм, полученных более 40 лет назад и поддерживаемых в коллекциях разных научно-исследовательских учреждений. Показано, что большинство исследованных форм цитогенетически стабильны, однако Авродес (геном BBAASS) – гексаплоидный геномно-замещенный гибрид пшеницы и *Ae. speltoides*, расщеплялся по хромосомному составу после многих репродукций. Хромосомный анализ не подтвердил ожидаемого геномного состава геномно-замещенной форма Авротата, у которой вместо заявленного N-генома от *Ae. uniaristata* Vis. обнаружен D-геном. В данной работе показано, что октаплоидные формы проходят через более сложные преобразования геномов, чем гексаплоидные: в двух исследованных предположительно октаплоидных амфидиплоидах AD 7, AD 7147 произошла редукция числа хромосом до гексаплоидного уровня. У обеих форм были утрачены семь пар хромосом из разных родительских субгеномов, представляющих все семь гомеологических групп. В результате у них сформировался смешанный (гибридный) геном, состоящий из уникальной комбинации хромосом нескольких родительских субгеномов.

**Ключевые слова:** становление геномов; пшеница; амфидиплоиды; *Aegilops*; *Dasypyrum*; *Tritordeum*; геномно-дополненные формы; геномно-замещенные формы; кариотип; С-бэндинг; флуоресцентная *in situ* гибридизация.

## Introduction

Common wheat *Triticum aestivum* L. is one of the most important crops. It ranks third to rice and maize in grain global production (Biodiversity, 2024). It is thought that common wheat arose about 8–10 MY BP in northwestern Iran, near Caspian Sea, as a result of hybridization between a tetraploid wheat and wild goat grass *Aegilops tauschii* Coss. followed by spontaneous chromosome duplication (Kihara, 1975; Dvořák et al., 1998; Feldman, 2001; Feldman, Levy, 2023). Such crosses might have occurred repeatedly and involve different parental wheat and *Aegilops* forms growing in the same region (Hirosawa et al., 2004; Luo et al., 2007). In turn, the resulting hexaploid wheats might cross to each other and to other species, thereby extending and enriching the gene pool of the novel crop (Feldman, 2001; Wang et al., 2013).

Common wheat is more flexible than cultivated tetraploid species (Dubcovsky, Dvořák, 2007); therefore, it is better suited to new environment when spreading to new areas. It is also characterized by better adaptability, higher yield, larger grains, and easier threshing as compared to hulled tetraploid wheat (Tadesse et al., 2016). The addition of the D genome from *Ae. tauschii* conferred grain qualities appropriate for the production of bread, one of the staples in human nutrition. Owing to these advantages, common wheat rapidly penetrated from its center of origin to the neighboring areas; then, to Europe Asia and Africa; and, finally, to North and South America and Australia. It gradually replaced hulled tetra- and hexaploid wheat species. Having been cultivated for over eight millennia, it occupied vast regions with diverse soil and climate conditions.

Meanwhile, intense breeding for high yield, which involved a limited number of founder varieties, narrowed considerably the gene pool of common wheat (Martynov et al., 2006; Girma, 2017; Feldman, Levy, 2023) in the past century. The task of gene pool expansion and search for new donors of commercially valuable traits is increasingly important (Bespalova, 2015). Wild Crop Relatives (WCR) are considered to be among the most promising donors of new genes for wheat improvement (Prohens et al., 2017; Sharma M.P. et al., 2020; Sharma S. et al., 2021). Species of the *Aegilops* L. genus, wheat relatives, possess many agronomically valuable traits that can be used in wheat breeding: pest resistance, drought tolerance, high micronutrient content, and others (Gill et al., 1986; Monneveux et al., 2000; Schneider et al., 2008; Molnár-Láng et al., 2015; Olivera et al., 2018; Kishii, 2019; Kumar et al., 2019). The close phylogenetic relationship between the *Triticum* L. and *Aegilops* genera facilitas successful transfer of genetic material transfer between them, as plasmon and two of the three common wheat nuclear sub-genomes, B and D, have been inherited from *Aegilops* species (Kihara, 1975; Tsunewaki, 1996).

Nevertheless, the direct gene transfer from *Aegilops* to wheat is a difficult task. Several approaches have been suggested to improve the efficiency of alien genetic material transfer. One of them involves crosses between wheat and a target species, chromosome doubling in the $F_1$, and developing wheat-alien addition and substitution lines. These lines are then used for inducing wheat-alien translocations (Peng et al., 2011; Zhang P. et al., 2015; Kishii, 2019; Kroupin et al., 2019). For instance, this approach was used to obtain wheat addition and substitution lines with rye (Gill, Kimber, 1974), barley (Islam, Shepherd, 1990; Cabrera et al., 1995; Molnár-Láng et al., 2000; Trubacheeva et al., 2009), *Aegilops* (Friebe et al., 1992, 2000; Logojan, Molnár-Láng, 2000; Molnár-Láng et al., 2014), *Haynaldia villosa* (L.) Schur (syn. *Dasypyrum villosum* (L.) P. Candargy) (Minelli et al., 2005), *Thinopyrum* Á. Löve (syn. *Elytrigia* Desv.) (Schulz-Schaeffer, Friebe, 1992; Linc et al., 2012; Kroupin et al., 2019), and other cereals. A number of allopolyploid hybrids between various tetraploid wheat species and *Ae. tauschii* have been obtained at CIMMYT, Mexico (Kishii, 2019; Aberkane et al., 2020). Pedigree analyses indicate that the

genetic material of *Aegilops*, mainly *Ae. tauschii*, as well as *Ae. umbellulata* Zhuk. and *Ae. ventricosa* Tausch is present in over 1,350 varieties and 9,000 elite lines of common wheat (Martynov et al., 2015), and their ratio is still increasing.

In addition to commercial breeding, synthetic allopolyploids are extensively used in studies of processes accompanying hybrid genome formation (Özkan et al., 2001; Kashkush et al., 2002; Levy, Feldman, 2004). Addition and substitution lines obtained from such allopolyploids were successfully used for the establishing of genetic relationships (homoeology) of chromosomes of different cereal species (Dhaliwal et al., 1990; Cabrera et al., 1995; Friebe et al., 1995a, b, 2000; Badaeva et al., 2018). However, these studies were primarily focused on processes occurring at early stages of allopolyploid formation, whereas their fate remained unknown in most cases.

Another approach was proposed by Dr. E.G. Zhirov. It is based on the development of genome-substituted common wheat forms, in which their D genome is substituted by the genome of a diploid *Aegilops* or of other cereal species (Zhirov, Ternovskaya, 1984; Davoyan R.O. et al., 2012). The first step of the production of these forms involved the extraction of the tetraploid BBAA component from common wheat cv. Avrora. The resulting tetra-component, tetraAvrora, was crossed to a diploid *Aegilops* species, whose genome was expected to replace common wheat D genome. The hybrids were treated with colchicine to double the chromosome number and obtain fertile amphydiploids. In spite of the fact that some genome-substituted forms obtained by E.G. Zhirov were cytologically characterized and are still used as donors of resistance genes in the breeding of common wheat and triticale (×*Triticosecale* Wittm.) (Davoyan R.O., Zhirov, 1995; Davoyan E.R. et al., 2012, 2023; Davoyan R.O. et al., 2019), most of these hybrids have not been analyzed by C-banding.

This article is aimed in cytogenetic verification of intergeneric synthetic amphydiploids and genome-substituted accessions of common wheat obtained over 30 years ago and maintained in gene banks of different institutions using C-banding and (for some hybrids) fluorescence *in situ* hybridization (FISH).

## Materials and methods

Experiments were conducted with the following artificial genome-substituted hybrids and intergeneric amphydiploids shown in the Table.

Six genome-substituted forms were raised by Dr. E.G. Zhirov at the Lukyanenko Research Institute of Agriculture, Krasnodar, more than 40 years ago. Their detailed description is provided in Zhirov's Dr. Sci. thesis "Wheat genomes: study and reconstruction" (Kyiv, Institute of Plant Physiology and Genetics, National Academy of Sciences, Ukraine, 1989). Two wheat–*Aegilops* amphydiploids were obtained by G.B. Piralov (1976) at the Institute of Genetics and Breeding, Academy of Sciences of the Azerbaijan SSR, Baku. One was accidently found in the collection of the Institute of Cultivated Plants, IPK, Gatersleben, Germany. Its origin is unknown. The hybrid between emmer and *Haynaldia villosa* was synthesized by P.M. Zhukovsky (1944). The amphydiploid of durum wheat and wild barley *Hordeum chilense* Roem. & Schult. was developed in Spain in the early 1980s (Martin, Sanchez-Mongelaguna, 1982; Fernández, Jouve, 1984).

Karyotypes were analyzed by the conventional Giemsa C-banding protocol (Badaeva et al., 1994). *Tritordeum* was additionally analyzed by FISH (Badaeva et al., 2017) with

Material examined

| Name | Cross | 2*n* | Expected genome constitution | Gene bank |
|---|---|---|---|---|
| Avrodes | *T. aestivum* × *Ae. speltoides* | 42 | BBAASS | NCG |
| Avrosis | *T. aestivum* × *Ae. sharonensis* | 42 | $BBAAS^{sh}S^{sh}$ | |
| Avrolata | *T. aestivum* × *Ae. umbellulata* | 42 | BBAAUU | |
| Avrotica | *T. aestivum* × *Ae. mutica* | 42 | BBAATT | ICG |
| Avrodata* | *T. aestivum* × *Ae. caudata* | 42 | BBAACC | |
| Avrotata* | *T. aestivum* × *Ae. uniaristata* | 42 | BBAANN | NCG |
| AD 7* | *T. ispahanicum* × *Ae. cylindrica* | 56 | $BBAAD^cD^cC^cC^c$ | ICG |
| AD 7147* | Amphydiploid 4*x* wheat and *Ae. ventricosa* | 56 | $BBAAD^vD^vN^vN^v$ | |
| AE 1491* | Not known | ? | ? | IPK |
| *Haynatricum*, K-38259 | *Haynaldia villosa* × *T. dicoccum* | 42 | BBAAVV | ICG |
| *Tritordeum martinii*, K-7997 | *T. durum* × *Hordeum chilense* | 42 | $BBAAH^cH^c$ | VIR |

Note. * Amphydiploids and genome-substituted accessions with unproved chromosome numbers, genome constitutions, or chromosome sets. NCG – National Center of Grain named after P.P. Lukyanenko, Krasnodar, Russia; ICG – Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia; IPK – Leibniz-Institut für Pflanzengenetik und Kulturpflanzenforschung, Gatersleben, Germany; VIR – N.I. Vavilov All-Russian Institute of Plant Genetic Resources (VIR), St. Petersburg, Russia.

DNA probes pAs1 (Rayburn, Gill, 1986) and pSc119.2 (Bedbrook et al., 1980). Wheat chromosomes were identified according to B.S. Gill et al. (1991), and chromosomes of other species, according to the nomenclatures proposed in (Dhaliwal et al., 1990; Cabrera et al., 1995; Friebe et al., 1995a, 2000; Linc et al., 1999; Badaeva et al., 2008, 2011, 2015a; Liu et al., 2010; Adonina et al., 2015; Molnár et al., 2016; Danilova et al., 2017; Said et al., 2021).

## Results and discussion

### Genome-substituted forms

#### *Avrodes*

Avrodes was cytogenetically proven to be hexaploid form in which the D genome is replaced by genome S from *Ae. speltoides* Tausch (Figs. 1, 2). Avrodes is cytologically unstable, and its chromosome numbers and combinations of the A, B, and S genome chromosomes vary among individual plants.

The plants examined had seven or eight A genome chromosome pairs. All plants had 1A, 2A, 4A, 5A, 6A, and 7A. Chromosome 2A of Avrodes differed from 2A of Avrora in having clear telomeric and terminal C bands. Unlike other chromosomes of the A genome, 6A was present as tetrasome, one pair substituting 6S. Most Avrodes plants had only one 7A pair, but two had an additional copy, substituting 7B (monosomic 7A/7B substitution; Fig. 2*c*). The karyotype of one plant lacked chromosome 3A, which had been replaced by an additional 3S pair.

Only three of seven chromosomes of the B genome, namely, 2B, 3B and 6B were present in all Avrodes plants examined. The 1BL:1RS wheat–rye translocation inherited from Avrora was seen in all plants, but the translocated chromosome was present in one or two copies (monosomic 1BL:1RS/1S substitution), or it was modified by a translocation of an unidentified fragment onto the distal portion of the long 1B arm (Fig. 2*a*, red arrow).

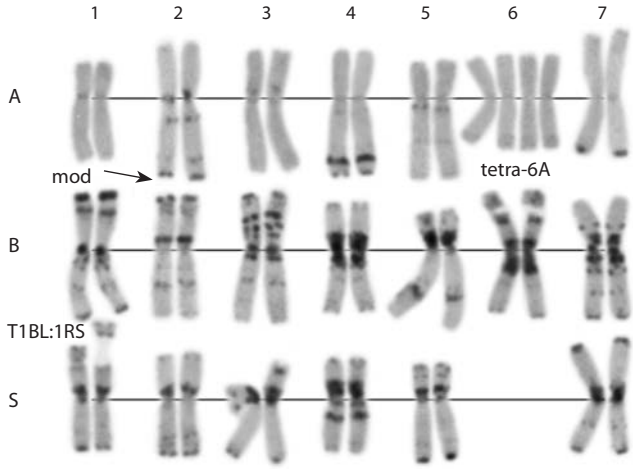



**Fig. 1.** C-banding karyotype of the Avrodes genome-substituted form.
A, B, S – genomes; 1–7 – homoeologous groups.

Some plants were nulli4B-tetra4S (Figs. 2*a*, *b*) and others, nulli5B-tetra4S (Figs. 2*a*, *b*), where the two 5S chromosome pairs showed different C-banding patterns (Fig. 2*c*, green arrows). One pair matched exactly chromosome 5S of *Ae. speltoides*, and the other, designated as 5S*, was shorter, and it lacked the large telomeric band in the long arm (Fig. 3). Note that just this modified chromosome pair passed to Avrodes-derived elite accessions resistant to stripe or yellow rust (*Puccinia striiformis* Westend. f. sp. *tritici Eriks*.) (Davoyan E.R. et al., 2023).

The S genomes of different Avrodes plants was represented by 12 to 16 chromosomes, but chromosome 6S was always missing (Fig. 2). Only 2S and 7S were present only in the disomic state. Chromosomes 3S, 4S, and 5S were present as di- or tetrasomics, where 3S replaced homoeologs of the A genome, and 4S and 5S, of the B genome. Most plants had

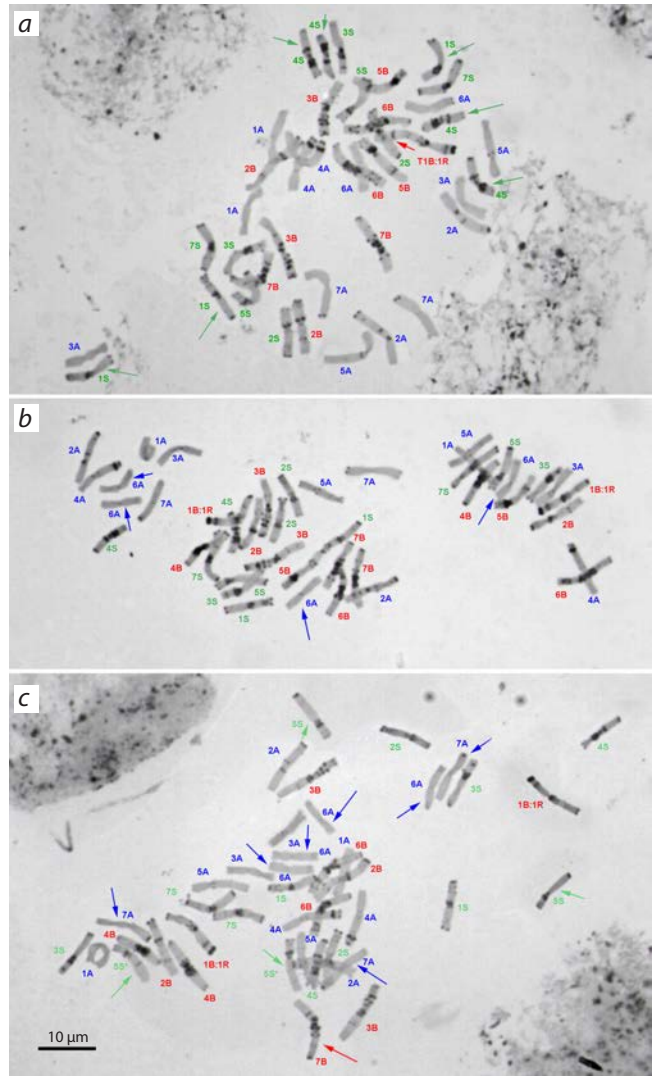



**Fig. 2.** C-banded metaphase plates in plants of the Avrodes genome-substituted form with different chromosome combinations.
Arrows indicate mono-, tri-, and tetrasomic chromosomes belonging to different genomes: red arrows – B genome; blue – A; and green – S.

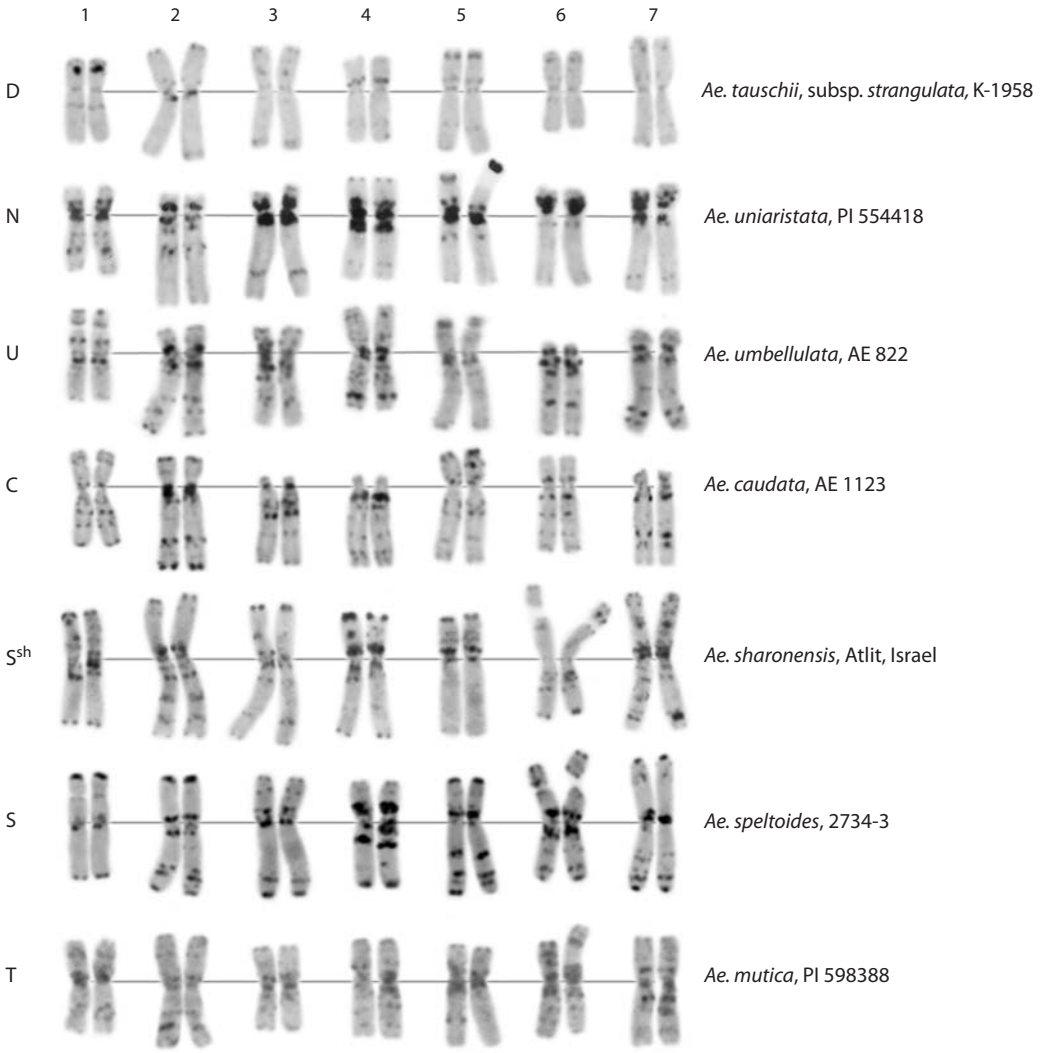



**Fig. 3.** C-banded karyotypes of diploid *Aegilops* species supposedly or actually involved in the development of genome-substituted wheat forms.

"Type specimens" of species not involved in the development of forms examined are shown for reference. The D–T genome symbols are indicated on the left; species names and origin/accession vouchers, on the right.

one 1S pair, and only two had an additional chromosome 1S, which replaced 1BL:1RS (monosomic 1S/1BL:1RS substitution, Fig. 2*a*).

The significant cytological instability of Avrodes also manifested itself in an abnormal meiotic chromosome pairing, in particular, high frequency of multivalents, formerly reported by R.O. Davoyan et al. (2012, 2019). This high frequency may be due to both the presence of genes suppressing *Ph1* (the gene regulating homoeologous chromosome pairing) in the S genome (Dvořák et al., 2006), and occurrence of intergenomic B/S or A/S substitutions in most plants, which have three or four copies of some *Ae. speltoides* chromosomes. The genomic instability of Avrodes may also be contributed by gametocidal genes, located on chromosomes 2S and 6S in *Ae. speltoides* (Tsujimoto, Tsunewaki, 1988; King J. et al., 2018; Said et al., 2024). It is worth noting that we found only one of the gametocidal chromosomes, 2S, whereas 6S had been lost.

### *Avrosis*

Avrosis is a hexaploid in which the D genome is replaced by the $S^{sh}$ genome from *Ae. sharonensis*. Eig. The presence of the A, B, and $S^{sh}$ genomes was proven by cytogenetic analysis, including C-banding (Fig. 4). Like Avrodes, Avrosis bears the 1BL:1RS wheat–rye translocation. However, the C-banding patterns of 2A, 2B, 3B, and 5B chromosomes of these forms differed from each other. Unlike Avrodes, Avrosis is cytologically stable: all plants examined had identical chromosome composition and C-banding patterns.

Chromosome T1B:1R was the only exception. In some plants, the distal portion of the short arm was deleted. The chromosomes of the $S^{sh}$ genome showed the morphology and heterochromatin distribution typical of *Ae. sharonensis* (Fig. 4). However, the direct parental accessions of Avrosis had not been indicated by originator, and we could not reveal chromosome changes associated with polyploidization.

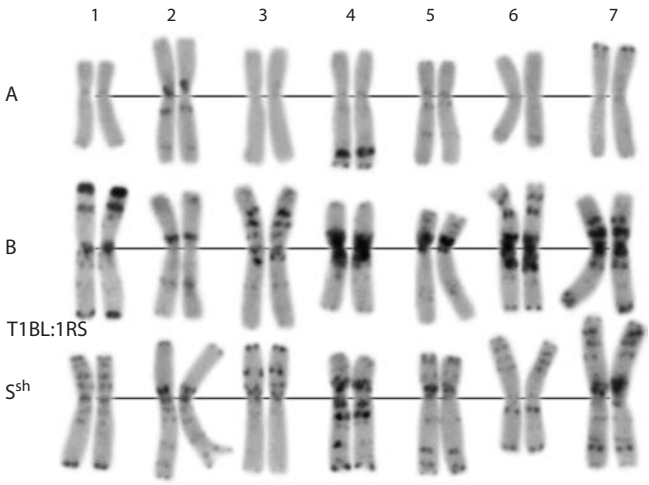



**Fig. 4**. Karyotype of the genome-substituted form Avrosis.
A, B, $S^{sh}$ – genomes; 1–7 – homologous groups.

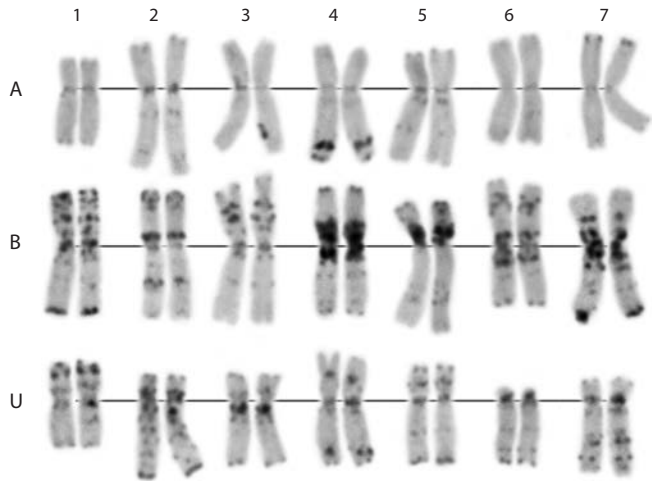



**Fig. 5.** Karyotype of the genome-substituted form Avrolata.
A, B, U – genomes; 1–7 – homologous groups.

In contrast to Avrodes, Avrosis was used in breeding programs solely as a donor of powdery mildew (*Blumeria graminis* (DC.) Speer f. sp. *tritici* Marshal) resistance (Zhirov, Ternovskaya, 1993), although *Ae. sharonensis* possesses many agronomically valuable traits (Olivera, Steffenson, 2009; Millet et al., 2014). The main difficulty in Avrosis use is that the $S^{sh}$ genome hosts highly efficient gametocidal genes *Gc* (Tsujimoto, Tsunewaki, 1984, 1988; Said et al., 2024). They induce the lethality of gametes that have lost the $4S^{sh}$ chromosome, bearing this gene. As a result, the $4S^{sh}$ chromosome is preferentially transmitted to gametes (Miller et al., 1982; King I. et al., 1991).

Nevertheless, some scientists succeeded in obtaining wheat × *Ae. sharonensis* introgression lines for chromosomes of other homoeologous groups, in particular, $1S^{sh}$ and $5S^{sh}$ (Millet et al., 2014). considering these results, we can hope that other $S^{sh}$ chromosomes can be transmitted to the progeny and the genetic potential of Avrodes in common wheat breeding is far from being exhausted.

### *Avrolata*

Avrolata is a hexaploid wheat in which the D genome is replaced by the U genome of *Ae. umbellulata*. It is cytologically stable, like Avrosis. All its plants had identical genome constitutions and banding patterns. We found no chromosome rearrangements in the accession studied. C-banding analysis confirmed the presence of the A, B, and U genomes in its karyotype (Fig. 5). In contrast to Avrodes and Avrosis, Avrolata did not bear the 1BL:1RS wheat–rye translocation; rather, it had the normal 1B chromosome.

The C-banding patterns of chromosomes belonging to the A and B genomes were generally similar to those of Avrosis, and the U chromosomes showed morphologies and banding patterns typical of *Ae. umbellulata* (Figs. 3, 5). As the parental *Ae. umbellulata* form was unknown, we could not assess putative changes of the U genome chromosomes of this hybrid.

The lack of 1BL:1RS in the karyotype of Avrolata may be due to the fact that cv. Avrora was *ab initio* heterogeneous for the presence of this translocation, and the direct parent of Avrolata belonged to the biotype lacking it. It is conceivable that durum wheat chromosome 1B survived recurrent crosses in the extraction of the Avrora tetra-component.

As reported in (Davoyan E.R. et al., 2012; Davoyan R.O. et al., 2012), Avrolata, along with Avrodes, is a source of novel genes for leaf rust (*Puccinia triticina* Rob. ex Desm. f. sp. *tritici* Eriks.) resistance. It is known that *Ae. umbellulata*, which was the source of the U genome in Avrolata, is extensively used in common wheat breeding, especially in the United States, as donor of the *Lr9* leaf rust resistance gene (Friebe et al., 1996b; McIntosh et al., 2013). Pedigree analysis shows that the ratio of varieties obtained with the use of *Ae. umbellulata* constantly increases and constitutes 25–29 % in 2000s (Martynov et al., 2015). Although *Lr9* had been detected in Avrolata, it was not found in its progeny (Davoyan E.R. et al., 2012). Apparently, the resistance in the derived accessions was determined by a novel *Lr* gene or a couple of unidentified genes. Avrolata was also employed in the breeding of other crops. A molecular study demonstrated the transmission of chromosomes 1U and 2U to the progeny of Avrolata crosses with winter hexaploid triticale (Orlovskaya et al., 2015).

### *Avrotica*

Avrotica is a genome-substituted form, whose parents were common wheat cv. Avrora and *Ae. mutica* Boiss. (syn. *Amblyopyrum muticum* (Boiss.) Eig, T genome). Cytogenetic analysis proved that Avrotica bears chromosomes of wheat A and B genomes and the T genome of *Ae. mutica*. However, in contrast to previously considered genome-substituted forms, Avrotica has a more complex combination of parental chromosomes.

Specifically, its karyotype maintained two chromosomes of the D genome, 1D and 3D, but lacked wheat 1A and

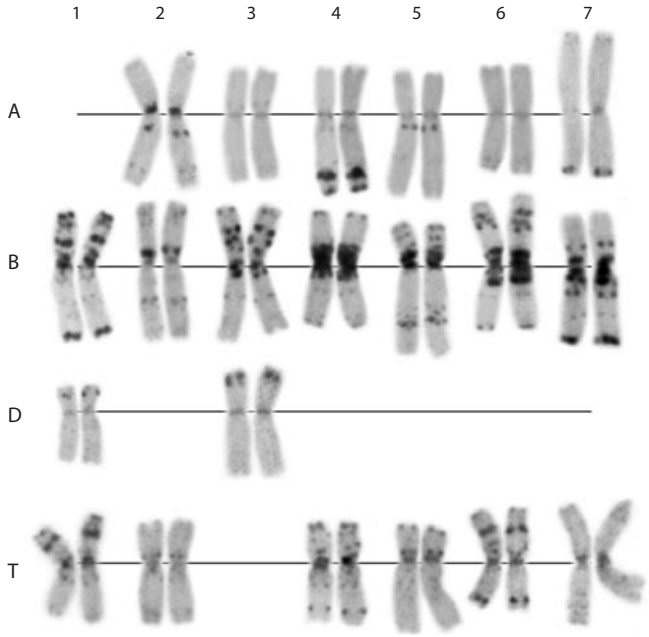



**Fig. 6.** Karyotype of the genome-substituted form Avrotica.
A, B, D, T – genomes; 1–7 – homoeologous groups.

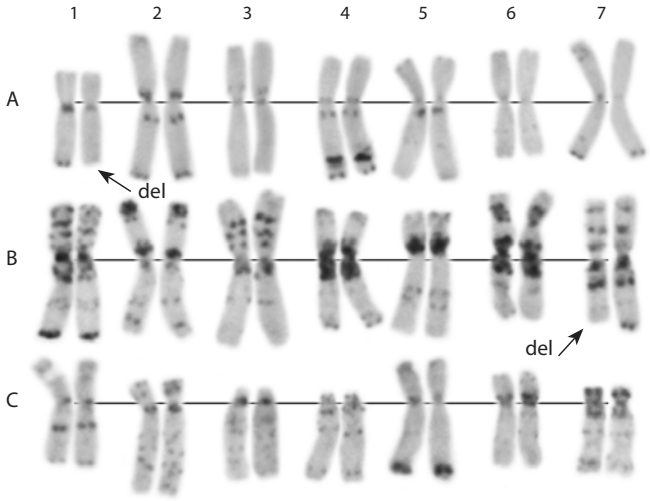



**Fig. 7.** Karyotype of the genome-substituted form Avrodata.
A, B, C – genomes; 1–7 – homoeologous groups. The arrow indicates a terminal deletion/translocation involving long arms of 1A and 7B chromosomes.

*Ae. mutica* 3T (Fig. 6). Thus, the alien genome is represented by only six chromosome pairs. Like Avrolata, Avrotica did not possess the wheat–rye 1BL:1RS translocation, although the C-banding patterns of other chromosomes were similar to those of Avrodes. We could not compare T chromosomes with those of the parental *Ae. mutica* accession, because the originators had not indicated the source of the latter. It should be noted that the homologous T chromosomes of the amphydiploid showed identical banding patterns, whereas the diploid species is highly polymorphic; in particular, is characterized by heteromorphism of homologs (Friebe et al., 1996a).

Although Avrotica is found to be rust resistant (Davoyan R.O. et al., 2012, 2019), this trait has not been transferred to common wheat. A Chinese team (Liu et al., 2015) produced a powdery mildew resistant incomplete amphydiploid of cv. Chinese Spring with *Ae. mutica* and an addition line for chromosome 7T. The allopolyploid had the complete set of the T genome chromosomes but lacked the pair of wheat chromosome 7B.

Another team crossed common wheat cvs. Chinese Spring and Pavon 76 to *Ae. mutica* accession bearing genes – suppressors of the *Ph1* locus (King J. et al., 2017). The $F_1$ hybrids were twice or thrice backcrossed to the parental cultivar. The plants were scored for alien introgressions by SNP genotyping. Genotypes with single introgressions were used to produce di-haploid plants. This procedure yielded 67 homozygous and stably inheritable introgression lines involving six of the seven *Ae. mutica* chromosomes (King J. et al., 2019). The team failed to obtain introgression lines for chromosome 3T, which was absent from Avrotica as well. It is reasonable to conjecture that this chromosome bears genes adversely affecting the viability and/or fertility of the *T. aestivum* × *Ae. mutica* allopolyploid; for this reason, plants carrying 3T were abandoned by selection in early hybrid generations.

### *Avrodata*

The pedigree of Avrodata indicates that it was obtained by crossing common wheat Avrora and *Ae. caudata* L. (syn. *Ae. markgrafii* (Greuter) Hammer). Cytological analysis confirmed the presence of the A and B wheat genomes and the C genome of *Ae. caudata* (Figs. 3, 7). All plants examined had identical chromosome sets, but chromosome rearrangements were detected in some of them (Fig. 7). They may have been induced in wheat–*Ae. caudata* crosses by gametocidal genes located on chromosome 3C (Endo, Tsunewaki, 1975).

The Avrodata lacks the 1BL:1RS wheat–rye translocation, and C-banding patterns of most chromosomes of the A and B wheat genomes (e. g., 2A, 4A, 5A, 6A, 1B, 2B, 5B, 6B, 7B) differed from the corresponding chromosomes of other genome-substituted forms obtained with cv. Avrora. In particular, the banding pattern of chromosome 7B was more similar to 7B of durum rather than common wheat. These observations suggest that Avrodata had been obtained from another parental wheat form or that the extraction of the tetra-component from Avrora resulted in the transmission of only part of A and B genome chromosomes of common wheat. The presence of unbalanced chromosome rearrangements in Avrodata plants shows that this genome-substituted form is cytologically unstable. No information on the use of this accession in breeding has been reported.

A genome-substituted amphydiploid of common wheat cv. Alcedo and *Ae. caudata* was synthesized in Germany

(Blüthner et al., 1988). The octoploid amphydiploid and addition lines developed on its basis were analyzed by C-banding, and meiotic chromosome pairing was also studied (Blüthner et al., 1988; Friebe et al., 1992). No deviations in C-banding patterns caused by chromosome rearrangements were detected, although numerous aberrations were noted in meiosis in all studied lines (Friebe et al., 1992). The banding pattern deviations observed in some wheat chromosomes were attributed to putative involvement of other wheat varieties in its origin.

The poor use of Avrodata in breeding may be due to the difficulty of the transmission of C genome material to common wheat associated with (1) a large number of species-specific chromosome rearrangements found in *Ae. caudata* (Danilova et al., 2017; Gong et al., 2017; Grewal et al., 2020) and (2) the presence of gametocidal genes on *Ae. caudata* chromosomes.

### *Avrotata*

We found that Avrotata is a cytologically stable hexaploid form. Its karyotype contains the A and B wheat genomes but no chromosomes corresponding to the N genome of *Ae. uniaristata* Vis have been detected (Figs. 3, 8). The third Avrotata genome showed the greatest similarity to the D genome of diploid *Ae. tauschii* subsp. *strangulata* Eig. (Fig. 3), which differs from the wheat D genome in C-banding patterns of chromosomes 3D and 6D. Presumably, the third Avrotata genome, $D^t$, is a mix of chromosomes derived from diploid *Ae. tauschii* and the D genome of common wheat, but this assumption cannot be proven by C-banding, because orthologous chromosomes of these genomes are closely similar.

*Ae. uniaristata* is tolerant to aluminum, and British scientists synthesized a hybrid between Chinese Spring and *Ae. uniaristata* to transmit this trait to the common wheat genome. This hybrid was employed in the development of several addition lines (Miller et al., 1995). The scientists showed that aluminum tolerance is controlled by chromosome 3N (Iqbal et al., 2000b). Analyses of the lines by *in situ* hybridization (Iqbal et al., 2000a) and later by C-banding (Badaeva et al., 2011) confirmed that they bear *Ae. uniaristata* chromosomes. These data allowed the cytological and genetic classifications of chromosomes of the N genome to be brought into compliance.

The mapping of RFLP markers on *Ae. uniaristata* chromosomes showed that they had been considerably rearranged with regard to homoeologous wheat chromosomes owing to the N genome-specific translocations and inversions (Iqbal et al., 2000b). The deep structural rearrangements of *Ae. uniaristata* chromosomes over the course of speciation were confirmed by the results of chromosome painting with oligo probe cocktail specific to each of the seven homoeologous groups of Triticeae (Li et al., 2020). It is reasonable to assume that the divergence of homoeologous wheat and *Ae. uniaristata* chromosomes impedes the transfer of genetic material between species, including the development of stable viable amphydiploids and genome-substituted forms. Unfortunately, no available data on the cytological verification of the genome constitution of Avrotata during early stages of its development have been reported. For this reason, we cannot decide whether the absence of the N genome from Avrotata was determined by the difficulty in the development of the form itself or the D genome replaced the N over the course of material propagation.

## Wheat–*Aegilops* amphydiploids

### *Amphydiploid AD 7*

AD 7 is a spontaneous amphydiploid of tetraploid wheat *T. ispahanicum* Heslot (genome BBAA) and tetraploid *Ae. cylindrica* Host ($D^cD^cC^cC^c$). The ancestral form of

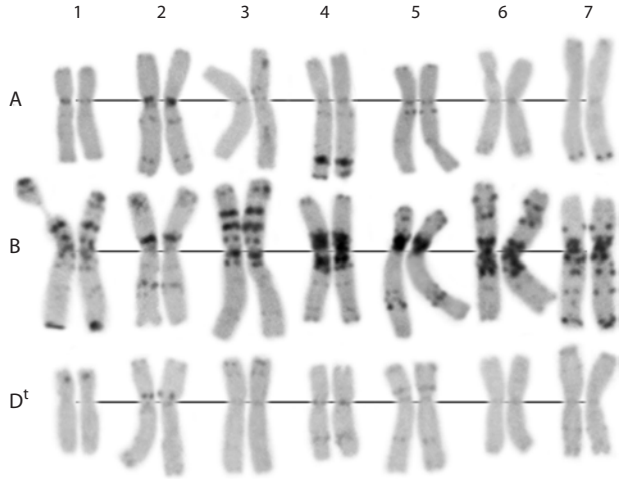



**Fig. 8.** Karyotype of the genome-substituted form Avrotata.
A, B, $D^t$ – genomes; 1–7 – homoeologous groups.

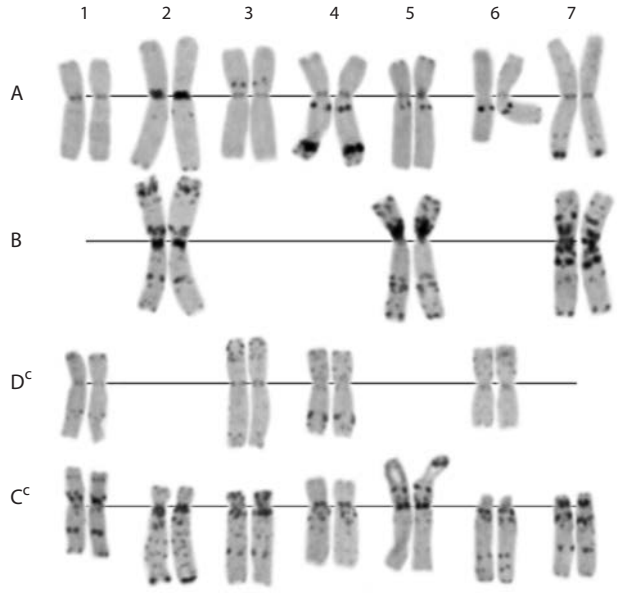



**Fig. 9.** Karyotype of the AD 7 amphydiploid.
A, B, $D^c$, $C^c$ – genomes; 1–7 – homoeologous groups.

AD 7 was an octoploid 2*n* = 8*x* = 56 with the genome constitution BBAAD$^{c}$D$^{c}$C$^{c}$C$^{c}$ (Mustafaev, Piralov, 1981). C-banding analysis confirmed the origin of the accession from *Ae. cylindrica* but showed that chromosome number of amphydiploid was reduced to hexaploid level.

Complete set of the wheat A-genome and *Ae. cylindrica* C$^{c}$-genome chromosomes were preserved in AD 7. The third genome proved to be mixed. It combined chromosomes of the wheat B genome and *Ae. cylindrica* D$^{c}$ genome, so that all the seven homologous groups were represented: 1D$^{c}$1D$^{c}$ 2B2B 3D$^{c}$3D$^{c}$ 4D$^{c}$4D$^{c}$ 5B5B 6D$^{c}$6D$^{c}$ 7B7B (Fig. 9). The chromosomes of the D$^{c}$ genome showed banding patterns typical of *Ae. cylindrica* (Linc et al., 1999; Badaeva et al. 2002). Some plants were monosomic for chromosome 6A (2*n* = 41). No chromosome rearrangements were found in the plants studied.

### *Amphydiploid AD 7147*

Amphydiploid AD 7147 was obtained by G.R. Piralov (1976) by crossing tetraploid wheat and *Ae. ventricosa* (Mustafaev, Piralov, 1981). The chromosome number doubled spontaneously; as supposed by G.R. Piralov, owing to the fusion of unreduced gametes. Regular chromosome pairing yielding 28 bivalents was observed in the meiosis of the original 56-chromosome amphydiploid. C-banding analysis of the AD 7147 confirmed that its origin from tetraploid wheat (genome BBAA) and *Ae. ventricosa* (genome D$^{v}$D$^{v}$N$^{v}$N$^{v}$) (Fig. 10). However, the C-banding patterns of the A and B genome chromosomes differed from those typical of durum wheat, being closer to *T. carthlicum* Nevski or the European variety of emmer *T. dicoccum* Schrank ex Schübl. (Badaeva et al., 2015b).

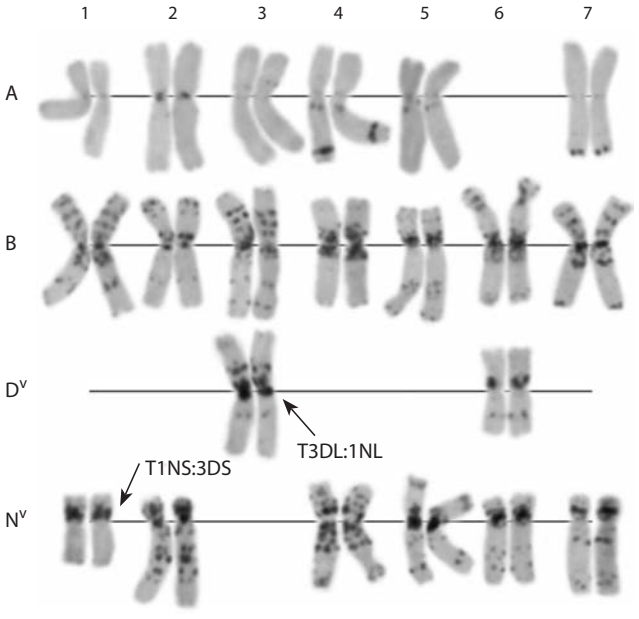



**Fig. 10.** Karyotype of the AD 7147 amphydiploid.
A, B, D$^{v}$, N$^{v}$ – genomes; 1–7 – homologous groups.

We found that AD 7147 bears a 1N$^{v}$:3D$^{v}$ translocation, most likely, inherited from the parental *Aegilops* accession. This translocation is common in natural *Ae. ventricosa* populations (Badaeva et al., 2002, 2011). As in the previous amphydiploid, the chromosome number in AD 7147 was reduced to hexaploid level as a result of a loss of one "hybrid" genome. In this case, though, the wheat B genome remained intact, 3N$^{v}$ was lost from the N$^{v}$ genome, and 6A, from the wheat A genome. Thus, the chromosome number reduction in the hybrid involved mainly the D$^{v}$ genome of *Ae. ventricosa*, of which only two chromosome pairs were preserved: 3D$^{v}$ (in the form of two translocated chromosomes T1N$^{v}$:3D$^{v}$) and 6D$^{v}$.

*Ae. ventricosa* is tetraploid species. Presently, it is extensively employed in wheat breeding as donor of pest resistance genes (Dosba, Doussinault, 1978; Garcia-Olmedo et al., 1984; Delibes et al., 1987, 1988). The gene cluster *Sr38/Lr37/Yr17*, inherited from *Ae. ventricosa* (Tanguy et al., 2005), had been mapped on chromosome 2A (Bariana, McIntosh, 1994). Pedigree analysis (Martynov et al. 2015) shows that this introgression is present in more than 34–37 % of modern common wheat varieties, mostly of European origin. The introgression originates from the French VPM-1 breeding line, which was produced by Maia in 1967 by crossing common wheat cv. Marne and a synthetic amphydiploid *Ae. ventricosa* × *T. persicum* Vav. (syn. *T. carthlicum* Nevski) (Dosba et al., 1978). Apparently, the genome constitution of this amphydiploid is similar to that of the original AD 7147 accession, but we cannot test the modern constitution of the French hybrid. We have no information on the use of AD 7146 in wheat breeding either, but, by way of analogy with the French *Ae. ventricosa* × *T. persicum*, it could be a promising donor of agronomically important genes.

### *Amphydiploid AE 1491*

Synthetic hexaploid amphydiploid AE 1491 was accidently identified among *Aegilops* accessions from the gene bank of the Leibniz Institute of Plant Genetics and Crop Plant Research (Germany). Analysis of chromosome morphology and the C-banding patterns (Fig. 11) brought us to suggestion that it is a hybrid of tetraploid *Ae. ventricosa* (genome D$^{v}$D$^{v}$N$^{v}$N$^{v}$) and einkorn wheat, presumably, *T. boeoticum* Boiss. (genome A$^{b}$A$^{b}$) or *T. monococcum* L. (A$^{m}$A$^{m}$). AE 1491 carried a 1N$^{v}$:3D$^{v}$ translocation, and it is conceivable that it was also present in the parental *Ae. ventricosa* accession.

No cases of aneuploidy, significant changes of C-banding patterns in comparison to the parental species (Badaeva et al., 2002, 2015a), or new variants of chromosome structural rearrangements were detected. An amphydiploid *T. aegilopoides* Link (syn. *T. boeoticum*) × *Ae. ventricosa* was produced and studied by (Siddiqui 2009; Siddiqui et al., 2009), but we do not know whether it corresponds to our accession.

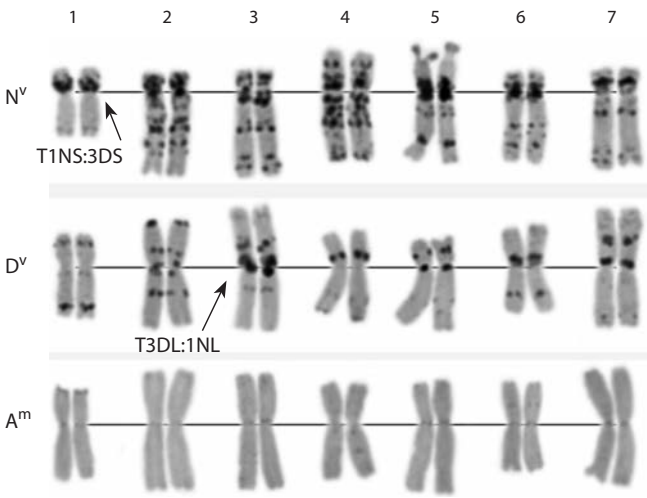



**Fig. 11.** Karyotype of the AE 1491 amphydiploid.
$N^v$, $D^v$, $A^m$ – genomes; 1–7 – homologous groups.

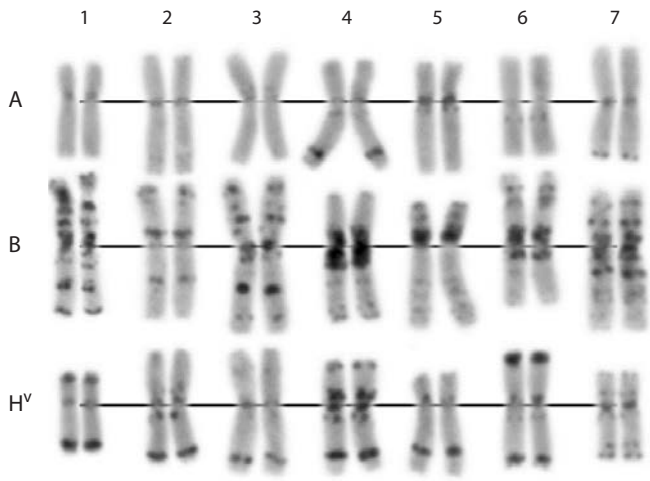



**Fig. 12.** Karyotype of *Haynatricum*.
A, B, $H^v$ – genomes; 1–7 – homoeologous groups.

## Wheat amphydiploids

### *Haynatricum Zhuk.*

Amphydiploids of wheat and *Dasypyrum villosum* (syn. *Haynaldia villosa*) were successfully produced by scientists from different countries starting from the 19th–early 20th century. Crosses to various wheat species, mostly tetraploids (*T. dicoccoides* (Körn. ex Asch. & Graebn.) Schweinf., *T. dicoccum*, *T. turgidum* L., *T. aethiopicum* Jakubz., *T. durum* Desf., *T. araraticum* Jakubz., *T. timopheevii* (Zhuk.) Zhuk.) or, less often, hexaploids (spelt and common wheat) (Pace et al. 2011) were undertaken. Our *T. dicoccum* × *D. villosum* amphydiploid has been developed by P.M. Zhukovsky and named *Haynatricum* Zhuk. (syn. *Triticum* ×*turgidovillosum* Tschermak) (Zhukovsky, 1944). It is maintained in the VIR gene bank under accession number K-38259.

The accession was shown to bear the entire sets of wheat A and B genome chromosomes and the $H^v$ genome chromosomes of *D. villosum* (Fig. 12). The C-banding patterns of wheat chromosomes were similar to those of the Transcaucasian group of cultivated emmer (Badaeva et al. 2015). It is likely that the parental form of this allopolyploid was *T. dicoccum* accession from Armenia, Azerbaijan, or neighboring regions of Turkey or Iran. All *Haynatricum* plants examined were euploid ($2n = 6x = 42$). No chromosomal rearrangements were detected. This fact, along with the absence of notable C-banding changes, points to a high cytological stability of the accession, which was obtained nearly 85 years ago.

*D. villosum* is a good donor of genes for disease resistance. Its amphydiploids and substitution and addition lines derived therefrom are broadly used in wheat breeding in China (Huang et al., 2007; Zhang W. et al., 2013) and other countries. Our accession differs from them in the distribution of heterochromatin blocks on chromosomes of wheat and *D. villosum* and therefore it is genetically different and may contain a different set of resistance genes.

### *Tritordeum martinii* A. Pujadas

The amphydiploid of durum wheat and wild barley *H. chilense* was synthesized in the early 1980s as a bridge for transferring agronomically useful genes from barley to wheat (Martin, Sanchez-Mongelaguna, 1982). Its karyotype was examined in detail by C-banding (Cabrera et al., 1995) and FISH with various DNA probes (Prieto et al., 2004; Martín, Cabrera 2005).

Analyses of *Tritordeum* chromosomes by C-banding (Fig. 13*a*) and FISH with pAs1 (green) and pSc119.2 (red) (Fig. 13*b*) probes confirmed the presence of the A, B, and $H^c$ genomes. Their C-banding and FISH patterns did not differ from those described in the literature. No aneuploidy or chromosome rearrangements were detected, which pointed to a good cytological stability of the accession.

## Conclusion

This feature is of great importance for the preservation and propagation of the allopolyploid, which is presently considered to be a new promising man-made crop (De Caro et al. 2024).

The results of the study of genome-substituted and synthetic genome-added amphydiploids of wheat and species of the *Aegilops*, *Dasypyrum*, and *Hordeum* genera bring us to the conclusions that:

- The chromosome sets of allopolyploids having 42 chromosomes are more stable than those of octoploids; however,
- Hexaploid forms containing related genomes (B–S and Avrodes) can remain cytologically unstable over many generations. The cytological instability manifests itself in the heterogeneity of the chromosome sets of plants,

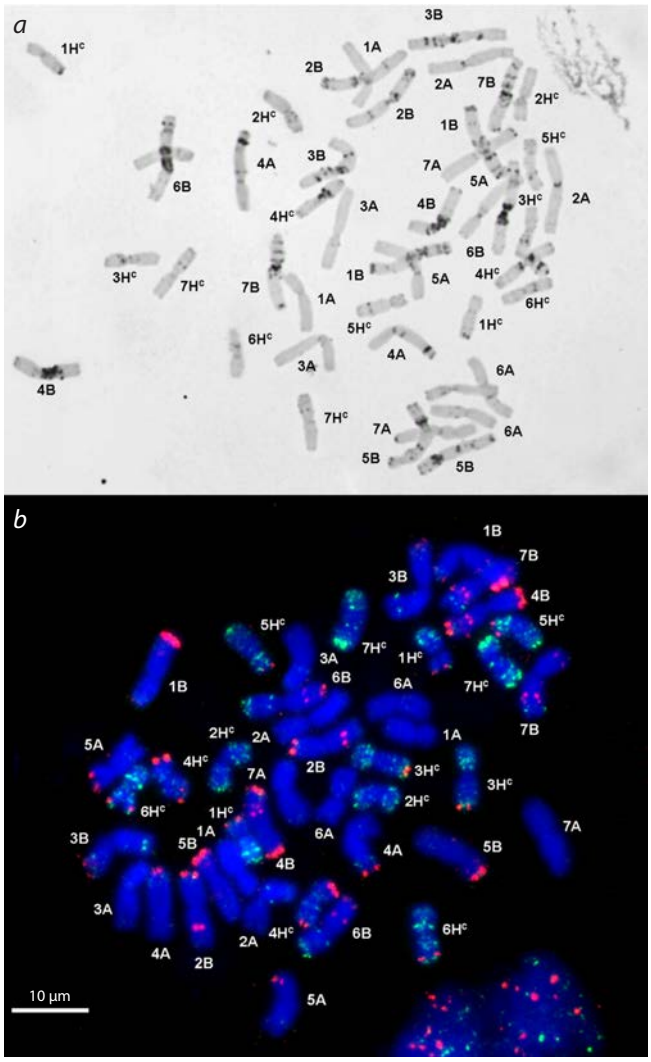



**Fig. 13.** C-banded metaphase plate (*a*) and the distribution of probes pAs1 (green) and pSc119.2 (red) on *Tritordeum* chromosomes (*b*).

Chromosome designations: 1A–7A – wheat A genome; 1B–7B – wheat B genome; $H^c$ – *H. chilense* genome.

monosomic or disomic substitutions of chromosomes belonging to related genomes, and more frequent chromosome rearrangements.

- The inconsistency of the chromosome sets in some genome-substituted forms indicates that the use of such materials in breeding and phylogenetic studies should be preceded by their thorough cytological verification.
- The most significant rearrangements of parental genomes were found in octoploids. They included chromosome number reduction to hexaploid level. The elimination involves chromosomes of different genomes (depending on the polyploid origin), covering all seven homoeologous groups.
- Although synthetic amphydiploids lag behind modern wheat cultivars in manifestation, their genes encoding mono- and polygenic characters may be more efficient than common wheat genes. The gene pool of synthetic wheats may provide new genes for resistance to biotic (Goncharov et al., 2020) and abiotic (Mahmood et al., 2023) stress factors for improving cultivar of common wheat.

## References

Aberkane H., Payne T., Kishi M., Smale M., Amri A., Jamora N. Transferring diversity of goat grass to farmers' fields through the development of synthetic hexaploid wheat. *Food Secur*. 2020;12(5):1017-1033. DOI 10.1007/s12571-020-01051-w

Adonina I.G., Goncharov N.P., Badaeva E.D., Sergeeva E.M., Petrash N.V., Salina E.A. (GAA)*n* microsatellite as an indicator of the A genome reorganization during wheat evolution and domestication. *Comp. Cytogenet*. 2015;9(4):533-547. DOI 10.3897/CompCytogen.v9i4.5120

Badaeva E.D., Badaev N.S., Gill B.S., Filatenko A.A. Intraspecific karyotype divergence in *Triticum araraticum* (Poaceae). *Plant Syst. Evol*. 1994;192(1-2):117-145. DOI 10.1007/BF00985912

Badaeva E.D., Amosova A.V., Muravenko O.V., Samatadze T.E., Chikida N.N., Zelenin A.V., Raupp W.J., Friebe B., Gill B.S. Genome differentiation in *Aegilops*. 3. Evolution of the D-genome cluster. *Plant Syst. Evol*. 2002;231(1-4):163-190. DOI 10.1007/s006060200018

Badaeva E.D., Dedkova O.S., Koenig J., Bernard S., Bernard M. Analysis of introgression of *Aegilops ventricosa* Tausch. genetic material in a common wheat background using C-banding. *Theor. Appl. Genet*. 2008;117(5):803-811. DOI 10.1007/s00122-008-0821-4

Badaeva E.D., Dedkova O.S., Zoshchuk S.A., Amosova A.V., Reader S., Bernard M., Zelenin A.V. Comparative analysis of the N-genome in diploid and polyploid *Aegilops* species. *Chromosome Res*. 2011;19(4):541-548. DOI 10.1007/s10577-011-9211-x

Badaeva E.D., Amosova A.V., Goncharov N.P., Macas J., Ruban A.S., Grechishnikova I.V., Zoshchuk S.A., Houben A. A set of cytogenetic markers allows the precise identification of all A-genome chromosomes in diploid and polyploid wheat. *Cytogenet. Genome Res*. 2015a;146(1):71-79. DOI 10.1159/000433458

Badaeva E.D., Dedkova O.S., Pukhalskyi V.A., Zelenin A.V. Chromosomal changes over the course of polyploid wheat evolution and domestication. In: Ogihara Y., Takumi S., Handa H. (Eds) Advances in Wheat Genetics: From Genome to Field. Tokyo: Springer, 2015b: 83-89. DOI 10.1007/978-4-431-55675-6_9

Badaeva E.D., Ruban A.S., Aliyeva-Schnorr L., Municio C., Hesse S., Houben A. In situ hybridization to plant chromosomes. In: Liehr T. (Ed.) Fluorescence In Situ Hybridization (FISH): Application guide. Ser.: Springer Protocols Handbooks. Berlin; Heidelberg: Springer, 2017;477-494. DOI 10.1007/978-3-662-52959-1_49

Badaeva E.D., Ruban A.S., Shishkina A.A., Sibikeev S.N., Druzhin A.E., Surzhikov S.A., Dragovich A.Yu. Genetic classification of *Aegilops columnaris* Zhuk. ($2n = 4x = 28$, $U^cU^cX^cX^c$) chromosomes based on FISH analysis and substitution patterns in common wheat × *Ae. columnaris* introgressive lines. *Genome*. 2018;61(2):131-143. DOI 10.1139/gen-2017-0186

Bariana H.S., McIntosh R.A. Characterisation and origin of rust and powdery mildew resistance genes in VPM1 wheat. *Euphytica*. 1994;76(1):53-61. DOI 10.1007/BF00024020

Bedbrook R.J., Jones J., O'Dell M., Thompson R.J., Flavell R.B. A molecular description of telomeric heterochromatin in *Secale* species. *Cell*. 1980;19(2):545-560. DOI 10.1016/0092-8674(80)90529-2

Bespalova L.A. Broadening the genepool as the major factor of the third Green Revolution in wheat breeding. *Vestnik Rossiskoi Akademii Nauk = Herald of the Russian Academy of Sciences*. 2015;85(1):9-11. DOI 10.7868/S086958731501003X (in Russian)

Biodiversity. Facts and figures on food and biodiversity. 2024 [cited 2024, 11 July]. Available from: https://idrc-crdi.ca/en/research-in-action/facts-figures-food-andbiodiversity

Blüthner W.-D., Schubert V., Mettin D. Instability in amphiploids and backcross derivatives of a *Triticum aestivum* × *Ae. caudata* cross. In: Miller T.E., Koebner R.M.D. (Eds) Proceedings of the 7th International Wheat Genetics Symposium. Cambridge,1988;209-213

Cabrera A., Friebe B., Jiang J., Gill B.S. Characterization of *Hordeum chilense* chromosomes by C-banding and in situ hybridization using highly repeated DNA probes. *Genome*. 1995;38(3):435-442. DOI 10.1139/g95-057

Danilova T.V., Akhunova A.R., Akhunov E.D., Friebe B., Gill B.S. Major structural genomic alterations can be associated with hybrid speciation in *Aegilops markgrafii* (Triticeae). *Plant J*. 2017;92(2):317-330. DOI 10.1111/tpj.13657

Davoyan E.R., Davoyan R.O., Bebyakina I.V., Davoyan O.R., Zubanova Y.S., Kravchenko A.M., Zinchenko A.N. Identification of a leaf-rust resistance gene in species of *Aegilops* L., synthetic forms, and introgression lines of common wheat. *Russ. J. Genet. Appl. Res*. 2012;2(4):325-329. DOI 10.1134/S2079059712040041

Davoyan E.R., Bebyakina I.V., Davoyan R.O., Boldakov D.M., Badaeva E.D., Adonina I.G., Salina E.A., Zinchenko A.N., Zubanova Yu.S. A study of bread wheat lines from crosses with the synthetic form Avrodes in regard to their yellow rust resistance. *Biotehnologiya i Selektsiya Rastenii = Plant Biotechnology and Breeding.* 2023;6(3): 25-34. DOI 10.30901/2658-6266-2023-3-o4 (in Russian)

Davoyan R.O., Bebyakina I.V., Davoyan O.R., Zinchenko A.N., Davoyan E.R., Kravchenko A.M., Zubanova Y.S. The use of synthetic forms in preservation and exploitation of the gene pool of wild common wheat relatives. *Russ. J. Genet. Appl. Res.* 2012;2(6):480-485. DOI 10.1134/S2079059712060044

Davoyan R.O., Bebyakina I.V., Davoyan E.R., Mikov D.S., Zubanova Yu.S., Boldakov D.M., Badaeva E.D., Adonina I.G., Salina E.A., Zinchenko A.N. The development and study of common wheat introgression lines derived from the synthetic form RS7. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2019;23(7):827-835. DOI 10.18699/VJ19.556 (in Russian)

Davoyan R.O., Zhirov E.G. Genome-substituted form Avrodes as the source of soft wheat plant resistance to leaf rust and powdery mildew. *Selkohozyaistvennaya Biologiya = Agricultural Biology*. 1995; 30(1):98-101 (in Russian)

De Caro S., Venezia A., Di Stasio L., Danzi D., Pignone D., Mamone G., Iakomino G. Tritordeum: promising cultivars to improve health. *Foods*. 2024;13(5):661. DOI 10.3390/foods13050661

Delibes A., Lopez-Braña I., Mena M., García-Olmedo F. Genetic transfer of resistance to powdery mildew and of an associated biochemical marker from *Aegilops ventricosa* to hexaploid wheat. *Theor. Appl. Genet*. 1987;73(4):605-608. DOI 10.1007/BF00289201

Delibes A., Doussinault G., Mena M., López-Braña I., García-Olmedo F. Eyespot resistance gene *Pch-1* from *Aegilops ventricosa* is associated with a different chromosome in wheat line H-93-70 than the resistance factor in "Roazon" wheat. *Theor. Appl. Genet.* 1988; 76(4):573-576. DOI 10.1007/BF00260911

Dhaliwal H.S., Friebe B., Gill K.S., Gill B.S. Cytogenetic identification of *Aegilops squarrosa* chromosome additions in durum wheat. *Theor. Appl. Genet*. 1990;79(6):769-774. DOI 10.1007/BF00224243

Dosba F., Doussinault G. Obtention of wheat lines with favorable agronomical characteristics of *Aegilops ventricosa*. *Ann. Amelior. Plant.* 1978;28(1):27-44

Dosba F., Tanguy A.M., Rivoal R. Extraction, identification and utilization of the addition lines *T. aestivum - Ae. ventricosa*. In: Ramanujan S. (Ed.) Proceedings of the 5th International Wheat Genetics Symposium, 23–28 Febr. New Delhi, 1978;332-337

Dubcovsky J., Dvořák J. Genome plasticity a key factor in the success of polyploid wheat under domestication. *Science*. 2007;316(5833): 1862-1866. DOI 10.1126/science.1143986

Dvořák J., Luo M.C., Yang Z.L., Zhang H.B. The structure of the *Aegilops tauschii* genepool and the evolution of hexaploid wheat. *Theor. Appl. Genet.* 1998;97(4):657-670. DOI 10.1007/s001220050942

Dvořák J., Deal K.R., Luo M.C. Discovery and mapping of wheat *Ph1* suppressors. *Genetics*. 2006;174(1):17-27. DOI 10.1534/genetics.106.058115

Endo T.R., Tsunewaki K. Sterility of common wheat with *Aegilops triuncialis* cytoplasm. *Heredity*. 1975;66(1):13-18. DOI 10.1093/oxfordjournals.jhered.a108562

Feldman M. Origin of cultivated wheat. In: Bonjean A.P., Angus W.J. (Eds) The World Wheat Book: A history of wheat breeding. London: Intersept Ltd, 2001;3-56

Feldman M., Levy A.A. Wheat Evolution and Domestication. Springer: Cham, 2023. DOI 10.1007/978-3-031-30175-9

Fernández J.A., Jouve N. Giemsa C-banding of the chromosomes of *Hordeum chilense* and its amphiploid × *Triticum turgidum* conv. *durum*. *Zeitschrift fur Pflanzenzuchtung = J. Plant Breed*. 1984;93(3): 212-221. DOI 10.1007/BF00032990

Friebe B., Schubert V., Blüthner W.D., Hammer K. C-banding pattern and polymorphism of *Aegilops caudata* and chromosomal constitutions of the amphiploid *T. aestivum - Ae. caudata* and six derived chromosome addition lines. *Theor. Appl. Genet.* 1992;83(5):589-596. DOI 10.1007/BF00226902

Friebe B., Jiang J., Tuleen N., Gill B.S. Standard karyotype of *Triticum umbellulatum* and the characterization of derived chromosome addition and translocation lines in common wheat. *Theor. Appl. Genet.* 1995a;90(1):150-156. DOI 10.1007/BF00221010

Friebe B., Tuleen N.A., Gill B.S. Standard karyotype of *Triticum searsii* and its relationship with other S-genome species and common wheat. *Theor. Appl. Genet.* 1995b;91(2):248-254. DOI 10.1007/BF00220885

Friebe B., Badaeva E.D., Hammer K., Gill B.S. Standard karyotypes of *Aegilops uniaristata, Ae. mutica, Ae. comosa* subspecies *comosa* and *heldreichii* (Poaceae). *Plant Syst. Evol*. 1996a;202(3):199-210. DOI 10.1007/BF00983382

Friebe B., Jiang J., Raupp W.J., McIntosh R.A., Gill B.S. Characterization of wheat-alien translocations conferring resistance to diseases and pests: current status. *Euphytica*. 1996b;91(1):59-87. DOI 10.1007/BF00035277

Friebe B., Qi L.L., Nasuda S., Zhang P., Tuleen N.A., Gill B.S. Development of a complete set of *Triticum aestivum-Aegilops speltoides* chromosome addition lines. *Theor. Appl. Genet.* 2000;101(1):51-58. DOI 10.1007/s001220051448

Garcia-Olmedo F., Delibes A., Sanchez-Monge R. Transfer of resistance to eyespot disease from *Aegilops ventricosa* to wheat. In: Breeding for Disease Resistance and Oat Breeding: Proceedings of the EUCARPIA Cereal Section Meeting, 28 Feb.–1 Mar. Weihenstephan, 1984;6:156-168

Gill B.S., Kimber G. The Giemsa C-banded karyotype of rye. *Proc. Natl. Acad. Sci. USA.* 1974;71(4):1247-1249. DOI 10.1073/pnas.71.4.1247

Gill B.S., Raupp W.J., Sharma H.C., Browder L.E., Hatchett J.H., Harvey T.L., Moseman J.G., Waines J.G. Resistance in *Aegilops squarrosa* to wheat leaf rust, wheat powdery mildew, greenbug, and Hessian fly. *Plant Dis.* 1986;70:553-556. DOI 10.1094/PD-70-553

Gill B.S., Friebe B., Endo T.R. Standard karyotype and nomenclature system for description of chromosome bands and structural aberrations in wheat (*Triticum aestivum*). *Genome*. 1991;34(5):830-839. DOI 10.1139/g95-030

Girma E. Genetic erosion of wheat (*Triticum* spp.): concept, research results and challenges. *J. Nat. Sci. Res.* 2017;7(23):72-81

Goncharov N.P., Boguslavsky R.L., Orlova E.A., Belousova M.Kh., Aminov N.Kh., Konovalov A.A., Kondratenko E.Ya., Gultyaeva E.I. Leaf rust resistance in wheat amphidiploids. *Pisma v Vavilovskii Zhurnal Genetiki i Selektsii = Letters to Vavilov Journal of Genetics and Breeding*. 2020;6(3):95-106. DOI 10.18699/Letters2020-6-14 (in Russian)

Gong W., Han R., Li H., Song J., Yan H., Li G., Liu A., Cao X., Guo J., Zhai S., Cheng D., Zhao Z., Liu C., Liu J. Agronomic traits and molecular marker identification of wheat–*Aegilops caudata* addition lines. *Front. Plant Sci.* 2017;8:1743. DOI 10.3389/fpls.2017.01743

Grewal S., Othmeni M., Walker J., Hubbart-Edwards S., Yang C.-y., Scholefield D., Ashling S., Isaac P., King I.P., King J. Development of wheat-*Aegilops caudata* introgression lines and their characterization using genome-specific KASP markers. *Front. Plant Sci*. 2020; 11:606. DOI 10.3389/fpls.2020.00606

Hirosawa S., Takumi S., Ishii T., Kawahara T., Nakamura C., Mori N. Chloroplast and nuclear DNA variation in common wheat: insight into the origin and evolution of common wheat. *Genes Genet. Syst*. 2004;79(5):271-282. DOI 10.1266/ggs.79.271

Huang D.-h., Lin Z.-s., Chen X., Zhang Z.-y., Chen C.-c., Cheng S.-h., Xin Z.-y. Molecular characterization of a *Triticum durum-Haynaldia villosa* amphiploid and its derivatives for resistance to *Gaeumannomyces graminis* var. *tritici*. *Agricult. Sci. China*. 2007;6(5):513-521. DOI 10.1016/S1671-2927(07)60077-7

Iqbal N., Reader S.M., Caligari P.D.S., Miller T.E. Characterization of *Aegilops uniaristata* chromosomes by comparative DNA marker analysis and repetitive DNA sequence in situ hybridization. *Theor. Appl. Genet.* 2000a;101(8):1173-1179. DOI 10.1007/s001220051594

Iqbal N., Reader S.M., Caligari P.D.S., Miller T.E. The production and characterization of recombination between chromosome 3N of *Aegilops uniaristata* and chromosome 3A of wheat. *Heredity*. 2000b; 84(4):487-492. DOI 10.1046/j.1365-2540.2000.00706.x

Islam A.K.M.R., Shepherd K.W. Incorporation of barley chromosomes into wheat. In: Bajaj Y.P.S. (Ed.) Wheat. Biotechnology in Agriculture and Forestry. Berlin: Springer, 1990;128-151. DOI 10.1007/978-3-662-10933-5_8

Kashkush K., Feldman M., Levy A.A. Gene loss, silencing and activation in a newly synthesized wheat allotetraploid. *Genetics*. 2002; 160(4):1651-1659. DOI 10.1093/genetics/160.4.1651

Kihara H. Origin of cultivated plants with special reference to wheat. *Seiken Ziho*. 1975;25/26:1-24

King I.P., Miller T.E., Koebner R.M.D. Determination of the transmission frequency of chromosome $4^{Sl}$ of *Aegilops sharonensis* in a range of wheat genetic backgrounds. *Theor. Appl. Genet.* 1991; 81(4):519-523. DOI 10.1007/BF00219443

King J., Grewal S., Yang C.-y., Hubbart S.., Scholefield D., Ashling S., Edwards K.J., Allen A.M., Burridge A., Bloor C., Davassi A., da Silva G.J., Chalmers K., King I.P. A step change in the transfer of interspecific variation into wheat from *Amblyopyrum muticum*. *Plant Biotechnol. J*. 2017;15(2):217-226. DOI 10.1111/pbi.12606

King J., Grewal S., Yang C.-y., Hubbart E.S., Scholefield D., Ashling S., Harper J.A., Allen A.M., Edwards K.J., Burridge A.J., King I.P. Introgression of *Aegilops speltoides* segments in *Triticum aestivum* and the effect of the gametocidal genes. *Ann. Bot.* 2018;121(2):229-240. DOI 10.1093/aob/mcx149

King J., Newell C., Grewal S., Hubbart-Edwards S., Yang C.-y., Scholefield D., Ashling S., Stride A., King I.P. Development of stable homozygous wheat/*Amblyopyrum muticum* (*Aegilops mutica*) introgression lines and their cytogenetic and molecular characterization. *Front. Plant Sci.* 2019;10:34. DOI 10.3389/fpls.2019.00034

Kishii M. An update of recent use of *Aegilops* species in wheat breeding. *Front. Plant Sci*. 2019;10:585. DOI 10.3389/fpls.2019.00585

Kroupin P.Yu., Divashuk M.G., Karlov G.I Gene resources of perennial wild cereals involved in breeding to improve wheat crop (review). *Sel'skokhozyaystvennaya Biologiya = Agricultural Biology.* 2019;54(3):409-425. DOI 10.15389/agrobiology.2019.3.409eng

Kumar A., Kapoor P., Chunduri V., Sharma S., Garg M. Potential of *Aegilops* sp. for improvement of grain processing and nutritional quality in wheat (*Triticum aestivum*). *Front. Plant Sci.* 2019;10:308. DOI 10.3389/fpls.2019.00308

Levy A.A., Feldman M. Genetic and epigenetic reprogramming of the wheat genome upon allopolyploidization. *Biol. J. Linn. Soc.* 2004; 82(4):607-613. DOI 10.1111/j.1095-8312.2004.00346.x

Li G., Zhang T., Yu Z., Wang H., Yang E., Yang Z. An efficient Oligo-FISH painting system for revealing chromosome rearrangements and polyploidization in Triticeae. *Plant J*. 2020;105(4):978-993. DOI 10.1111/TPJ.15081

Linc G., Friebe B.R., Kynast R.G., Molnar-Lang M., Köszegi B., Sutka J., Gill B.S. Molecular cytogenetic analysis of *Aegilops cylindrica* Host. *Genome*. 1999;42(3):497-503. DOI 10.1139/gen-42-3-497

Linc G., Sepsi A., Molnar-Lang M. A FISH karyotype to study chromosome polymorphisms for the *Elytrigia elongata* E genome. *Cytogenet. Genome Res*. 2012;136(2):138-144. DOI 10.1159/000334835

Liu C., Li G.-R., Sehgal K.S., Jia J.-Q., Yang Z.-J., Friebe B., Gill B.S. Genome relationships in the genus *Dasypyrum*: evidence from molecular phylogenetic analysis and in situ hybridization. *Plant Syst. Evol*. 2010;288(3-4):149-156. DOI 10.1007/s00606-010-0319-9

Liu C., Li G.-R., Gong W.-P., Li G.-Y., Han R., Li H.-S., Song J.-M., Liu A.-F., Cao X.-Y., Chu X.-S., Yang Z.-J., Huang C.-Y., Zhao Z.-D., Liu J.-J. Molecular and cytogenetic characterization of a powdery mildew-resistant wheat-*Aegilops mutica* partial amphiploid and addition line. *Cytogenet. Genome Res.* 2015;147(2-3):186-194. DOI 10.1159/000443625

Logojan A.A., Molnár-Láng M. Production of *Triticum aestivum* - *Aegilops biuncialis* chromosome additions. *Cereal Res. Commun.* 2000;28(3):221-222. DOI 10.1007/BF03543597

Luo M.-C., Yang Z.-L., You F.M., Kawahara T., Waines J.G., Dvořák J. The structure of wild and domesticated emmer wheat populations, gene flow between them, and the site of emmer domestication. *Theor. Appl. Genet.* 2007;114(6):947-959. DOI 10.1007/s00122-006-0474-0

Mahmood Y.A., DeSilva J., King I.P., King J., Foulkes M.J. Leaf photosynthesis traits and associations with biomass and drought tolerance in amphidiploid and ancestral wheat genotypes. *Eur. J. Agronomy*. 2023;147:126846. DOI 10.1016/j.eja.2023.126846

Martín A., Cabrera A. Cytogenetics of *Hordeum chilense*: current status and considerations with reference to breeding. *Cytogenet. Genome Res*. 2005;109(1-3):378-384. DOI 10.1159/000082423

Martin A., Sanchez-Mongelaguna E. Cytology and morphology of the amphiploid *Hordeum chilense* × *Triticum turgidum* conv. *durum*. *Euphytica*. 1982;31(1):261-268. DOI 10.1007/BF00028329

Martynov S.P., Dobrotvorskaya T.V., Pukhalskiy V.A. Dynamics of genetic diversity in winter common wheat *Tritium aestivum* L. cultivars released in Russia from 1929 to 2005. *Russ. J. Genet*. 2006; 42(10):1137-1147. DOI 10.1134/S1022795406100061

Martynov S.P., Dobrotvorskaya T.V., Mitrofanova O.P. Genealogical analysis of the use of aegilops (*Aegilops* L.) genetic material in wheat (*Triticum aestivum* L.). *Russ. J. Genet*. 2015;51(9):855-862. DOI 10.1134/S1022795415090070

McIntosh R.A., Yamazaki Y., Dubkovsky G., Rogers J., Morris C.F., Appels R., Xia X.C. Catalogue of Gene Symbols for Wheat. The 12th International Wheat Genetics Symposium, 8–13 Sept. 2013. Yokohama, Japan, 2013

Miller T.E., Hutchinson J., Chapman V. Investigation of a preferentially transmitted *Aegilops sharonensis* chromosome in wheat. *Theor. Appl. Genet.* 1982;61(1):27-33. DOI 10.1007/BF00261506

Miller T.E., Reader S.M., Mahmood A., Purdie K.A., King I.P. Chromosome 3N of *Aegilops uniaristata* – a source of tolerance to high levels of aluminium for wheat. In: Li S., Xin Z.Y. (Eds) Proceeding of the 8th International Wheat Genetics Symposium, 20–25 July 1993. Beijing: China Agricult. Sci. Press, 1995;1037-1042

Millet E., Manisterski J., Ben-Yehuda P., Distelfeld A., Deek J., Wan A., Chen X., Steffenson B.J. Introgression of leaf rust and stripe rust resistance from Sharon goatgrass (*Aegilops sharonensis* Eig) into bread wheat (*Triticum aestivum* L.). *Genome*. 2014;57(6):309-316. DOI 10.1139/gen-2014-0004

Minelli S., Ceccarelli M., Mariani M., De Pace C., Cioninia P.G. Cytogenetics of *Triticum* × *Dasypyrum* hybrids and derived lines. *Cytogenet. Genome Res*. 2005;109(1-3):385-392. DOI 10.1159/000082424

Molnár I., Vrána J., Burešová V., Cápal P., Farkas A., Darkó É., Cseh A., Kubaláková M., Molnár-Láng M., Doležel J. Dissecting the U, M, S and C genomes of wild relatives of bread wheat (*Aegilops* spp.) into chromosomes and exploring their synteny with wheat. *Plant J*. 2016;88(3):452-467. DOI 10.1111/tpj.13266

Molnár-Láng M., Linc G., Logojan A., Sutka J. Production and meiotic pairing behaviour of new hybrids of winter wheat (*Triticum aestivum*) × winter barley (*Hordeum vulgare*). *Genome*. 2000;43(6):1045-1054. DOI 10.1139/gen-43-6-1045

Molnár-Láng M., Molnár I., Szakács É., Linc G., Bedö Z. Production and molecular cytogenetic identification of wheat-alien hybrids and introgression lines. In: Tuberosa R., Graner A., Frison E. (Eds) Genomics of Plant Genetic Resources. Vol. 1. Managing, Sequencing and Mining Genetic Resources. New York: Springer, 2014; 255-284

Molnár-Láng M., Ceoloni C., Doležel J. (Eds) Alien Introgression in Wheat. Cytogenetics, Molecular Biology, and Genomics. Switzerland: Springer, 2015. DOI 10.1007/978-3-319-23494-6

Monneveux P., Zaharieva M., Rekika D. The utilisation of *Triticum* and *Aegilops* species for the improvement of durum wheat. In: Royo C., Nachit M., Di Fonzo N., Araus J.L. (Eds) Durum Wheat Improvement in the Mediterranean Region: New Challenges. Zaragoza: Ciheam, 2000;71-81

Mustafaev I.D., Piralov G.R. Some aspects of interrelations between tetraploid wheat species and *Aegilops ventricosa* Tausch. *Sel'skokhozyaystvennaya Biologiya = Agricultural Biology.* 1981;16(2): 223-228 (in Russian)

Olivera P.D., Rouse M.N., Jin Y. Identification of new sources of resistance to wheat stem rust in *Aegilops* spp. in the tertiary genepool of wheat. *Front. Plant Sci.* 2018;9:1719. DOI 10.3389/fpls.2018.01719

Olivera P.D., Steffenson B.J. *Aegilops sharonensis*: origin, genetics, diversity, and potential for wheat improvement. *Botany*. 2009;87(8): 740-756. DOI 10.1139/B09-040

Orlovskaya O.A., Leonova I.N., Adonina I.G., Salina E.A., Khotyleva L.V., Shumny V.K. Molecular cytogenetic analysis of triticale and wheat lines with introgressions of the genetic material of Triticeae tribe species. *Russ. J. Genet. Appl. Res*. 2016;6(5):527-536. DOI DOI 10.1134/S2079059716050087

Özkan H., Levy A.A., Feldman M. Allopolyploidy-induced rapid genome evolution in the wheat (*Aegilops-Triticum*) group. *Plant Cell*. 2001;13(8):1735-1747. DOI 10.1105/tpc.13.8.1735

Pace C.D., Vaccino P., Cionini P.G., Pasquini M., Bizzarri M., Qualset C.O. *Dasypyrum*. In: Cole C. (Ed.) Wild Crop Relatives: Genomic and Breeding Resources Cereals. Berlin: Springer, 2011; 185-292

Peng J.H., Sun D., Nevo E. Domestication evolution, genetics and genomics in wheat. *Mol. Breed.* 2011;28(3):281. DOI 10.1007/s11032-011-9608-4

Piralov G.R. The results of hybridization of wheat with aegilops, rye, *Haynaldia* and wheatgrass. In: Genetics and Breeding in Azerbaijan. Vol. 1. Baku, 1976;136-137 (in Russian)

Prieto P., Martin A., Cabrera A. Chromosomal distribution of telomeric and telomeric-associated sequences in *Hordeum chilense* by in situ hybridization. *Hereditas*. 2004;141(2):122-127. DOI 10.1111/j.1601-5223.2004.01825.x

Prohens J., Gramazio P., Plazas M., Dempewolf H., Kilian B., Díez M.J., Fita A., Herraiz F.J., Rodríguez-Burruezo A., Soler S., Knapp S., Vilanova S. Introgressiomics: a new approach for using crop wild relatives in breeding for adaptation to climate change. *Euphytica*. 2017;213(7):158. DOI 10.1007/s10681-017-1938-9

Rayburn A.L., Gill B.S. Isolation of a D-genome specific repeated DNA sequence from *Aegilops squarrosa*. *Plant Mol. Biol. Rep.* 1986;4(2): 102-109. DOI 10.1007/BF02732107

Said M., Holušová K., Farkas A., Ivanizs L., Gaál E., Cápal P., Abrouk M., Martis-Thiele M.M., Kalapos B., Bartoš J., Friebe B., Doležel J., Molnár I. Development of DNA markers from physically mapped loci in *Aegilops comosa* and *Aegilops umbellulata* using single-gene FISH and chromosome sequences. *Front. Plant Sci*. 2021;12:1136. DOI 10.3389/fpls.2021.689031

Said M., Gaál E., Farkas A., Molnár I., Bartoš J., Doležel J., Cabrera A., Endo T.R. Gametocidal genes: from a discovery to the application in wheat breeding. *Front. Plant Sci.* 2024;15:1396553. DOI 10.3389/fpls.2024.1396553

Schneider A., Molnár I., Molnár-Láng M. Utilisation of *Aegilops* (goatgrass) species to widen the genetic diversity of cultivated wheat. *Euphytica*. 2008;163(1):1-19. DOI 10.1007/s10681-007-9624-y

Schulz-Schaeffer J., Friebe B. Karyological characterization of a partial amphiploid, *Triticum turgidum* L. var. *durum* × *Agropyron intermedium* (Host) P.B. *Euphytica*. 1992;62(2):83-88. DOI 10.1007/BF00037932

Sharma M., Punya, Gupta B.B. Role of wild relatives for development of climate-resilient varieties. In: Salgotra R.K., Zargar S.M. (Eds) Rediscovery of Genetic and Genomic Resources for Future Food Security. Singapore: Springer, 2020;303-314. DOI 10.1007/978-981-15-0156-2_11

Sharma S., Schulthess A.W., Bassi F.M., Badaeva E.D., Neumann K., Graner A., Özkan H., Werner P., Knüpffer P., Kilian B. Introducing beneficial alleles from plant genetic resources into the wheat germplasm. *Biology*. 2021;10(10):982. DOI 10.3390/biology10100982

Siddiqui K. Induced mutations in *Triticum aegilopoides*, *Aegilops ventricosa* and their synthetic allopolyploid. *Hereditas*. 2009;73:45-50. DOI 10.1111/j.1601-5223.1973.tb01066.x

Siddiqui K., Ingversen J., Køie B. Inheritance of protein patterns in a synthetic allopolyploid of *Triticum monococcum* (AA) and *Aegilops ventricosa* ($DDM^{v}M^{v}$). *Hereditas*. 2009;72:205-214. DOI 10.1111/j.1601-5223.1972.tb01044.x

Tadesse W., Amri A., Ogbonnaya F.C., Sanchez-Garcia M., Sohail Q., Baum M. Wheat. In: Singh M., Upadhyaya H.D. (Eds) Genetic and Genomic Resources for Grain Cereals Improvement. San Diego: Acad. Press, 2016;81-124

Tanguy A.-M., Coriton O., Abélard P., Dedryver F., Jahier J. Structure of *Aegilops ventricosa* chromosome $6N^{v}$, the donor of wheat genes *Yr17*, *Lr37*, *Sr38*, and *Cre5*. *Genome*. 2005;48(3):541-546. DOI 10.1139/g05-001

Trubacheeva N.V., Efremova T.T., Badaeva E.D., Kravtsova L.A., Belova L.I., Devyatkina E.P., Pershina L.A. Production of alloplasmic and euplasmic wheat-barley ditelosomic substitution lines $7H^{1}L^{mar}$(7D) and analysis of the 18S/5S mitochondrial repeat in these lines. *Russ. J. Genet*. 2009;45(12):1438-1443. DOI 10.1134/S1022795409120059

Tsujimoto H., Tsunewaki K. Gametocidal genes in wheat and its relatives. I. Genetic analyses in common wheat of a gametocidal gene derived from *Aegilops speltoides*. *Can. J. Genet. Cytol*. 1984;26(1): 78-84. DOI 10.1139/g84-013

Tsujimoto H., Tsunewaki K. Gametocidal genes in wheat and its relatives. III. Chromosome location and effects of two *Aegilops speltoi-*

*des*-derived gametocidal genes in common wheat. *Genome*. 1988; 30(2):239-244. DOI 10.1139/g88-041

Tsunewaki K. Plasmon analysis as the counterpart of genome analysis. In: Jauhar P.P. (Ed.) Methods of Genome Analysis in Plant: Their Merrits and Piffals. Boca Ration: CRC Press, 1996;271-299.

Wang J., Luo M.-C., Chen Z., You F.M., Wei Y., Zheng Y., Dvorak J. *Aegilops tauschii* single nucleotide polymorphisms shed light on the origins of wheat D-genome genetic diversity and pinpoint the geographic origin of hexaploid wheat. *New Phytologist*. 2013;198(3): 925-937. DOI 10.1111/nph.12164

Zhang P., Dundas I.S., McIntosh R.A., Xu S.S., Park R.F., Gill B.S., Friebe B. Wheat–*Aegilops* introgressions. In: Molnár-Láng M., Ceoloni C., Doležel J. (Eds) Alien Introgression in Wheat. Cytogenetics, Molecular Biology, and Genomics. Switzerland: Springer, 2015;221-244. DOI 10.1007/978-3-319-23494-6

Zhang W., Zhang R., Feng Y., Bie T., Chen P. Distribution of highly repeated DNA sequences in *Haynaldia villosa* and its application in the identification of alien chromatin. *Chin. Sci. Bull*. 2013;58(8): 890-897. DOI 10.1007/s11434-012-5598-9

Zhirov E.G., Ternovskaya T.K. Genome engineering in wheat. *Vestnik Sel'skokhozyaystvennykh Nauk = Herald of Agricultural Sciences*. 1984;10:58-66 (in Russian)

Zhirov E.G., Ternovskaya T.K. Transfer of the chromosome conferring mildew resistance from *Aegilops sharonensis* Eig into *Triticum aestivum* L. *Genetika = Genetics* (*Moscow*). 1993;29(4):639-645 (in Russian)

Zhukovsky P.M. Studies on hybridization and immunity of plants. *Trudy Moskovskoi Selskohozyaistvennoi Akademii imeni K.A. Timirjazeva = Proceedings of the Timiryazev Moscow Agricultural Academy*. 1944;6:3-48 (in Russian)

**Conflict of interest.** The authors declare no conflict of interest.
Received July 17, 2024. Revised August 1, 2024. Accepted August 7, 2024.

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2024;28(7):731-743

DOI 10.18699/vjgb-24-81

# The potential of the amaranth collection maintained at VIR in the context of global plant breeding and utilization trends

D.V. Sokolova, A.E. Solovieva, A.M. Zaretsky, T.V. Shelenga

Federal Research Center the N.I. Vavilov All-Russian Institute of Plant Genetic Resources (VIR), St. Petersburg, Russia
dianasokol@bk.ru

**Abstract.** Amaranth is an ancient crop of the family Amaranthaceae, but it is fairly new to Russia. Its seeds and leaf biomass contain a high-quality gluten-free protein, fatty acids, squalene (a polyunsaturated hydrocarbon), flavonoids, vitamins, and minerals. A comprehensive study of amaranth, enhancement of its breeding, and development of new cultivars will contribute to food quality improvement through the use of plant raw materials enriched for wholesome and highly nutritious components. At present, selection and hybridization still remain the main amaranth breeding techniques. Meanwhile, mutation breeding and polyploidy have been successfully employed to increase its seed yield and protein content. The genes encoding amaranth proteins have been used to produce transgenic plants of potato, bread wheat, and maize. Despite the great potential of amaranth, little research has been dedicated to the study of its genomics, concentrating mainly on the identification of its species diversity. Targets of breeding practice for amaranth include such characteristics as large size and nonshattering of seeds, short stem, earliness, high yield, cold hardiness, synchronized maturation, resistance to pests and diseases, and high nutritional value, including the content and quality of protein, lipids, squalene, and bioactive compounds. A unique collection of amaranth maintained at the N.I. Vavilov All-Russian Institute of Plant Genetic Resources (VIR) currently incorporates 570 accessions from various countries. For 70 years it has been replenished with local varieties, commercial cultivars, and wild species supplied by collecting missions, research centers, botanical gardens, genebanks, and experimental breeding stations from all over the world. Long-standing studies have resulted in the formation of trait-specific groups of accessions, with high yields of seeds and leaf biomass, earliness, cold hardiness, high protein content in seeds and biomass, short stems, and resistance to seed shattering, earmarked for vegetable or ornamental purposes. The gene pool of amaranth preserved at VIR can provide unlimited opportunities for breeding and meet the needs of the country's population, enriching the human diet with ingredients produced from such a health-friendly and useful crop.

**Key words:** *Amaranthus* L.; valuable traits; breeding trends; species diversity; VIR's amaranth collection.

**For citation:** Sokolova D.V., Solovieva A.E., Zaretsky A.M., Shelenga T.V. The potential of the amaranth collection maintained at VIR in the context of global plant breeding and utilization trends. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2024;28(7):731-743. DOI 10.18699/vjgb-24-81

**Funding.** This research was supported by the Russian Science Foundation under Project No. 24-26-00218 (Interaction of Phenotypic and Biochemical Characteristics of Grain Amaranth with Squalene Content in Seeds under the Conditions of the Northwest of the Russian Federation).

# Потенциал коллекции амаранта ВИР в свете мировых тенденций использования и селекции

Д.В. Соколова, А.Е. Соловьева, А.М. Зарецкий, Т.В. Шеленга

Федеральный исследовательский центр Всероссийский институт генетических ресурсов растений им. Н.И. Вавилова (ВИР), Санкт-Петербург, Россия
dianasokol@bk.ru

**Аннотация.** Амарант – древняя культура семейства Амарантовые (Amaranthaceae). Для России это достаточно новая сельскохозяйственная культура. В семенах и листовой биомассе содержатся высококачественный безглютеновый белок, жирные кислоты, полиненасыщенный углеводород сквален, флавоноиды, витамины и минералы. Комплексное изучение амаранта, развитие его селекции и создание новых сортов являются крайне важным направлением для решения проблемы повышения качества пищевой продукции путем использования растительного сырья, обогащенного полезными и высокопитательными компонентами. На сегодняшний день основными методами селекционной работы с амарантом остаются отбор и гибридизация. Методы мутационной селекции и полиплоидии были успешно использованы для увеличения урожайности семян и содержания белка. С помощью генов, кодирующих белки амаранта, созданы трансгенные растения картофеля, мягкой пшеницы и кукурузы. Несмотря на большой потенциал амаранта, изучению его геномики посвящено не много исследова-

ний, направленных главным образом на идентификацию видового разнообразия. В направления селекционной работы с амарантом входят такие признаки, как «крупность и неосыпаемость семян», «низкорослость», «скороспелость», «высокая урожайность», «холодостойкость», «синхронность созревания», «устойчивость к вредителям и болезням», «высокая питательная ценность»: содержание и качество белка, липидов, сквалена, биологически активных соединений. Уникальная коллекция амаранта Всероссийского института генетических ресурсов растений им. Н.И. Вавилова (ВИР) включает 570 образцов из различных стран мира. На протяжении 70 лет она пополнялась местными, селекционными сортами и дикими видами за счет экспедиций, поступлений из научно-исследовательских институтов, ботанических садов, генбанков и опытных селекционных станций всего мира. В результате многолетнего изучения были сформированы признаковые группы образцов с высокой урожайностью семян и листовой биомассы, скороспелые и холодостойкие, с повышенным содержанием белка в семенах и биомассе, низкорослые, устойчивые к осыпанию семян, овощного и декоративного направления использования. Сохраняемый в ВИР генофонд амаранта способен предоставлять неограниченные возможности для селекции и восполнять нужды населения страны, обогащая питательный рацион продуктами из этой здоровой и полезной культуры.

**Ключевые слова:** *Amaranthus* L.; ценные признаки; направления селекции; видовое разнообразие; коллекция амаранта ВИР.

## Introduction

The industrialization of agricultural production and the consolidation of the integrated global market ensured sustainable growth of worldwide food supplies induced by higher crop yields. Meanwhile, the innovation processes in agriculture, based on the release of high-yielding crop cultivars and the development of agricultural technologies, dealt with only some of the staple crops (soybean, wheat, rice, maize, sunflower, etc.). The resulting depletion of agricultural biodiversity and gradual replacement of minor crops posed a potential threat to global food security (Khoury et al., 2014; Dawson et al., 2019). The abovementioned crops provide enough calories, but they are deficient in essential amino acids, minerals, and vitamins for maintaining a wholesome and well-balanced human diet, leading to the "hidden" malnutrition of more than two billion people worldwide, whose daily diet consists almost entirely of these crops (Cheng et al., 2015).

Throughout its existence, humankind has used around 3,000 plant species, but only about 150 of them are cultivated commercially (Mangelsdorf, 1966). Other sources report that about 30,000 plant species in the world are edible, but only 7,000 of them are used for food (Ramdwar et al., 2017). Integration of a wide range of "forgotten crops" to improve dietary diversity can improve the quality of human nutrition in many countries and ensure their food security (Mayes et al., 2011; Ebert, 2014; Joshi D.C. et al., 2018).

Amaranth is one of the plants with the potential to become an alternative grain crop on the global scale (Das, 2016). The objective of this study was to make a historical review of amaranth breeding, characterize the genetic diversity of amaranth preserved in the VIR collection, and contemplate its prospects for domestic breeding practice.

Amaranth is an ancient crop, described in botany as gen. *Amaranthus* L., subfam. Amaranthoideae, fam. Amaranthaceae, ord. Caryophyllales. The genus *Amaranthus* L. includes, according to various sources, from 60 to 87 species, being one of the top ten taxonomically most complex crops. Along with buckwheat and quinoa, it represents a small group of "pseudocereals" (Saunders, Becker, 1984; Teutonico, Knorr, 1985).

Most of its species are wild or weedy. The grain species of amaranth are *Amaranthus cruentus* and *A. hypochondriacus* from Central and North Americas, and *A. caudatus* of South American origin (Covas, 1994). The history of amaranth cultivation in these regions dates back 5,000 to 7,000 years. According to archaeological records describing the materials collected in the northwest of Argentina, the age of the discovered seeds was dated to the beginning of the Middle Holocene (5,500 to 6,000 BC). The oldest finds of amaranth were registered in a cave in the highland area of Peñas de la Cruz, Department of Antofagasta de la Sierra (3,665 MASL). These data indicated a much earlier use of amaranth – 10,000 and 7,000 BC (Arreguez et al., 2013).

Amaranth was also a very important food source for the population of pre-Hispanic South America (Chagaray, 2005). From ancient times, the Aztec and Mayan tribes used it as a grain crop, second in importance only to maize and legumes (Sauer, 1967; Smith M.E., 1996). Amaranth leaves were also consumed. Mixing whole or ground amaranth grain to prepare bread, porridges, and cakes was quite popular, as well as its ceremonial use in temples. The Aztecs made small figurines of gods from amaranth dough and ate them as part of their rituals (De Montellano, 1990). Most likely, it was the reason why Spanish conquerors strictly prohibited amaranth consumption and cultivation in the early 16th century, which led to its falling into obsolescence for many years (Saunders, Becker, 1984).

The revival of interest in amaranth in the late 20th century was associated with the efforts to study its unique biochemical characteristics, multipurpose utilization prospects, and the C4 photosynthesis mechanism typical of amaranth (Venskutonis, Kraujalis, 2013; Magomedov, Chirkova, 2015). For Russian agriculture it is a fairly new crop, with its enormous potential for growth intensity, productivity, and high complete protein content in seeds and leaf biomass (Kononkov et al., 1999). Therefore, a comprehensive study of amaranth, its improvement through breeding, and development of new cultivars will contribute a great deal to the task of raising the quality of human nutrition by means of employing plant materials enriched with useful and highly nutritious components.

Classification of the genus *Amaranthus* is hampered by the absence of species-specific and qualitative identifying characters, a wide range of phenotypic variability between species, and introgression and hybridization between weedy and cultivated species (Sauer, 1967; Hauptli, Jain 1978). Many researchers have applied morphological, biochemical, molecular and cytogenetic methods to assess the level of interspecies phylogenetic relationships (Murray, 1940; Costea et al., 2001; Das, 2012; Akin-Idowu et al., 2016). Most of the authors agree that all grain amaranths descended from their weedy progenitor *A. hybridus*.

The cultivated grain species have close relationships among themselves, but *A. hypochondriacus* ($2n = 32$) and *A. caudatus* ($2n = 32$) are more closely related to each other than to *A. cruentus* ($2n = 34$). A specific feature of *A. cruentus* is the presence of one copy of chromosome 2, resulting in a haploid number of $n = 17$ (Singh et al., 2023).

A majority of the *Amaranthus* representatives are annual herbaceous plants with claret-colored or yellow-green leaves and inflorescences. The anatomical and morphological diversity of amaranth plants pertains to species specificity and growing conditions. Plant height varies from 40 cm to 5 m. The stem is usually erect, striated, and prolifically leafy. Plants of some species exhibit a sprawling semi-accumbent shape. The degree of branching in amaranth plants can be weak, medium, or strong. The plant habit is formed by a combination of features: the position and branching of the main stem, its size, and the shape of the inflorescence. The leaves are stipule-free, alternate or opposite, differing in leaf and margin shapes. The average number of leaves on a plant can reach 250, with the leaf surface area being ca. 7,500–8,000 $cm^2$. The inflorescence is a compound panicle of varying shape, density, and color. The flowers are small, actinomorphic, dioecious, less often bisexual, clustered in the leaf axils. The androecium consists of 5 stamens; the gynoecium, of 3, less often 4 carpels. The superior ovary is unilocular (Das, 2016).

Amaranth is recognized as a "superfood" because of its nutraceutical value, i. e., the content of high-quality gluten-free protein, unsaturated fatty acids, dietary fibers, flavonoids, vitamins (thiamine, riboflavin, ascorbic acid, and niacin), and minerals (calcium, magnesium, and copper, plus sodium, iron, phosphorus, and zinc) (Kononkov et al., 1999; Grobelnik-Mlakar et al., 2009; Palombini et al., 2013; Joshi D.C. et al., 2018; Soriano Garcıa et al., 2018; Sokolova et al., 2021). Its seeds contain methionine (15.8 mg/g total protein) and lysine (55.8 mg/g total protein), ensuring the crop's higher nutritional value compared to most cereals (Tang, Tsao, 2017).

The amount of lipids in amaranth seeds varies greatly across the species and genotypes, ranging within 1.9–9.7 %. Palmitic, oleic, linoleic and linolenic fatty acids are present in high amounts, accounting aggregately for over 90 % of the total fatty acid content. Amaranth seed oil has proved its therapeutic effect.

The fatty acid composition of amaranth oil is almost similar to that of cereals, but there is a difference: it contains relatively high levels of squalene ($C_{30}H_{50}$), a polyunsaturated hydrocarbon (Bressani, 1994). Squalene has a wide range of applications in medicine – as an adjuvant in vaccines, or an immunomodulator and antioxidant in complex therapy against a number of diseases, such as diabetes and coronary heart disease – and is also used in cosmetics (Gonor et al., 2006; Huang et al., 2009). There is convincing evidence that squalene reduces the risk of cancer development and controls cholesterol levels in the human organism (Miettinen, Vanhanen, 1994; Rao et al., 1998; Smith T.J., 2000). The ever increasing interest in this compound is explained by its combined therapeutic effect: antioxidant, hypolipidemic, antitoxic, and antidiabetic (Magomedov et al., 2017).

## Uses of amaranth

Cultivated amaranth species are divided into two main groups according to the ways of their utilization: food (vegetable and grain products), and feed. Besides, there are uses less known to the public: ornamental, pharmaceutical, and construction material production (Fig. 1). Such division is arbitrary enough: one and the same cultivar can be used as both feed and food (grain), while the leaves of younger plants belonging to all cultivated species can be consumed fresh as salad ingredients (Ruth et al., 2021; Sokolova et al., 2021).

Initially, amaranth was cultivated for its edible seeds (Central and South America, and mountainous areas in Asia) and as a green vegetable crop (Africa, South Asia, and Southeast Asia). Vegetable amaranth species are widely used for food in India, in the countries of Asia and Southeast Asia, and in Africa, but they are little known in North and South Americas.

Leaves, shoots, and tender juicy stems of vegetable amaranths are used to prepare sauces, soups, or vegetable stews. Young leaves of grain amaranth are also consumed as leafy vegetables. The claret-colored leaves of *A. cruentus* serve as raw material for the production of teas enriched with amaranth. Amaranth seeds are digestible both in their whole (porridges, cereals, and candies) and milled form (bread, pasta, or baked products) (Das, 2016).

Oil is extracted mainly from the seeds of two amaranth species: *A. cruentus* and *A. hypocondriacus*; its content varies within 4.8–8.1 % (He, Corke, 2003; Gamel et al., 2007). The yield of amaranth oil exceeds that of most cereals, but is inferior to oil crops (Ayorinde, 1989; Leon-Camacho et al., 2001). H.P. He and H. Corke (2003) studied the oil content in 104 samples of 30 amaranth species: this indicator showed significant variability, depending on the genotype, growing environments, and the effect of abiotic factors. Moreover, wild forms of amaranth matched its cultivars in their oil content levels, thus confirming their value for breeders (Table 1).

The use of amaranth for animal feed implies utilization of the aboveground plant biomass. However, protein-rich amaranth seeds are also included in feed mixtures. The yield of green biomass averages 85–103 t/ha, with the dry matter yield of 15.7–16.7 t/ha (Abbasi et al., 2012; Shadi et al., 2020). The biomass yield of *A. hypochondriacus* cultivars in the northern regions of China can reach 130 t/ha, with the dry matter yield of 20 t/ha (Sun G.Q. et al., 2017). H. Shadi et al. (2020) reported a higher level of nondegradable digestible


**Fig. 1.** The uses of amaranth (created by the authors).

protein in amaranth silage than in the one made from maize, which attests to its higher efficiency. High crude protein and low lignin content, low nitrate and oxalic acid levels account for the high potential of amaranth silage as a feed for ruminants (Sleugh et al., 2001; Rezaei et al., 2009). However, amaranth leaf biomass contains antinutrients, such as trypsin inhibitors, saponins, alkaloids, and oxalates – the feature that reduces the crop's nutritional value and requires appropriate improvement through breeding (Cheeke et al., 1981).

Many amaranths demonstrate noticeable ornamental properties (Sauer, 1967). A new trend in amaranth utilization is the use of its lignified leading shoots in the construction business to manufacture various wood-based panels (Evon et al., 2021).

Amaranth cultivation technologies today are quite different from the agricultural practices of early civilizations, when neither machinery nor advanced techniques were available. Modern agricultural mechanization methods make it possible to achieve higher yields and profits.

## Amaranth improvement through breeding

Cultivated amaranth is characterized by its rich genetic diversity and environmental plasticity. It is predominantly an

**Table 1.** The content of oil and squalene in the seeds of different amaranth species, according to (He, Corke, 2003)

| Species | Oil, % | Squalene, mg/g of seed |
|---|---|---|
| *A. rudis* | 8.25 | 4.75 |
| *A. blitum* | 6.96 | 2.93 |
| *A. spinosus* | 6.45 | 1.89 |
| *A. powelli* | 6.15 | 2.66 |
| *A. retroflexus* | 5.79 | 2.46 |
| *A. albus* | 5.68 | 2.44 |
| *A. dubius* | 5.30 | 2.03 |
| *Amaranthus* sp. | 5.29 | 2.56 |
| *A. viridis* | 5.00 | 2.12 |
| *A. hybridus* | 4.66 | 2.59 |
| *A. hypochondriacus* | 4.58 | 2.55 |
| *A. tricolor* | 4.50 | 2.39 |
| *A. cruentus* | 3.21 | 1.31 |

autogamous crop, but its percentage of outbreeding is 5–39 %, a level sufficient to ensure gene flow among populations (Hauptli, Jain, 1985). The crop's diversified mechanism of reproduction is feasible due to the ratio and distribution of male and pistillate flowers in its inflorescences. Obligate allogamous species include *A. tuberculatus*, *A. palmeri*, *A. arenicola*, and *A. rudis*, all of them dioecious.

For amaranth, breeding trends depend on the ways of its utilization. Negative features of this crop, limiting its commercial cultivation and requiring improvement by breeders, include: too small seeds, great plant height, asynchronous seed ripening, seed shattering, size and rigidity of the stem holding a large and heavy inflorescence, and dense structure of the latter (Kauffman, 1984). The great height of grain amaranth plants (1.8–2.5 m) is a negative trait, making its harvesting difficult. Such plant forms are prone to lodging and need to be supported, which leads to an increase in cultivation costs. Plant height variability within amaranth cultivars allows breeders to select short-stemmed plants and use them in crosses.

Breeding work with grain amaranths should be targeted mainly at such useful traits as large size and nonshattering of seeds, short stem, earliness, high yield, cold hardiness, synchronized maturation, resistance to pests and diseases, and high nutritional value, including the content and quality of protein, lipids, and bioactive compounds. A desirable target for breeders working with vegetable amaranths is greater bushiness to ensure the possibility of repeated cutting (Sreelathakumary, Peter, 1993).

## Conventional breeding techniques

### Selection

This approach was most popular in the United States and India, where selection from local populations resulted in the release of amaranth cultivars still in use today. Germplasm lines from the working collection of the Rodale Research Center, Pennsylvania, USA, became the progenitors of the majority of amaranth cultivars developed in the United States and China (Stallknecht, Schulz-Schaefer, 1993). The first *A. cruentus* lines registered by the Crop Science Society of America were Montana-3, with white seeds and high yield, and Montana-5, combining the traits of Montana-3 with synchronized ripening (Schulz-Schaeffer et al., 1989a, b). Later, the lodging-resistant cv. Amont was obtained through selection from Montana-3 at Montana State University (Schulz-Schaeffer et al., 1991). Charles S. Kauffman from the Rodale Research Center presented the results of his research and successful selection of grain amaranths for at a number of important traits: seed size, synchronized maturation, increased protein content in seeds, and resistance to shattering and pests. It was also there that K-432, a semi-dwarf line with plant height not exceeding 92 cm, was produced by means of selection among short-stemmed forms of amaranth (Kauffman, 1992).

High-yielding, mid-season, and dwarf forms suitable for mechanized harvesting were selected by screening local grain amaranth accessions from the Mexican plant genetic resources collection (Espitia, 1992). Three well-known high-yielding cultivars of *A. caudatus*, Oscar Blanco, Noel Vietmeyer, and Alan Garcia, were developed at the University of Cuzco, Peru. Having achieved wide distribution, they are currently cultivated commercially over hundreds of hectares. In Kenya, an improved line of *A. hypochondriacus* served as a source for the local cultivar Jumla (Kauffman, Weber, 1990; Joshi B.D., Rana, 1991).

An example of effective germplasm utilization in India was the development of the grain amaranth cultivar Annapurna in 1984 at the NBPGR Regional Station in Shimla, India; it was obtained as a pure line of *A. hypochondriacus* from source material of local origin (Joshi B.D. et al., 1983). The average seed yield of cv. Annapurna is 2.25 t/ha, and the protein content is 15 %. The same station released cv. Durga, resistant to lodging, major diseases and pests, and early-ripening cvs. Gujarat Amaranth-1, Gujarat Amaranth-2, Kapilasa, and Suvarna (Raiger, Bhandari, 2012).

The experiments conducted in 1977–1988 in Minnesota, USA, reported the highest grain yield of 1.72 t/ha for promising cultivars (Myers, Putnam, 1988). Modern cultivars developed by domestic breeders demonstrate the yields of 2.35 t/ha (cv. Karakula) and 2.09 t/ha (cv. Voronezhsky) (State Register for Selection Achievements Admitted for Usage, 2023). It should be taken into account that the present-day agricultural practice for amaranth cultivation is quite different from earlier practices, when neither machinery nor technologies were available. Joint efforts of plant breeders and agricultural technologists make it possible to achieve higher yields and profits.

The early-ripening and cold-hardy amaranth cultivar Frant (*A. cruentus*) was released by the N.I. Vavilov All-Russian Institute of Plant Genetic Resources (VIR) after selecting single-stemmed red-leaved plant forms up to 1.2 m tall from a local population of Indian origin, inbreeding, and free pollination of their linear progeny (Patent, 2022).

### Hybridization

Hybridization is known to be the most widespread and effective breeding technique to obtain new gene combinations. One of the first to classify the interspecies hybridization within the genus *Amaranthus* was M.J. Murray (1940). He structured amaranth species according to the arrangement of male flowers in the inflorescences, and made a lot of crosses between monoecious and dioecious species. T.N. Khoshoo and M. Pal (1972) during their studies succeeded in hybridizing *A. hypochondriacus* (served as a pollinator) with *A. hybridus* and *A. caudatus*. The $F_1$ hybrids of *A. hypochondriacus* × *A. hybridus* had the highest pollen fertility. A weighty contribution to understanding the amaranth gene pool availability was made by E.J. Greizerstein and L. Poggio (1992, 1995) who analyzed the meiotic configuration in 13 different spontaneous amaranth hybrids.

Generally, hybridization is most effective with *A. hypochondriacus* and *A. hybridus*, since these species are close in their evolutionary development and contain the same chromosome number ($2n = 32$). For example, crossing *A. hypochondriacus* with a Pakistani accession of *A. hybridus*

at the Nebraska Agricultural Experiment Station resulted in the release of a high-yielding cultivar of grain amaranth, Plainsman (PI 358322), widely distributed across the United States. Its characteristics include earliness, high productivity, and the plant height of 1.5–1.8 m (Baltensperger et al., 1992).

Hybridization helped to transfer useful traits from wild amaranth species. For example, the traits of wild *A. powellii* were implanted into the breeding lines of *A. cruentus* and *A. hypochondriacus* to reduce seed shattering. Hybrids with the wild dioecious species *A. cannabinus* exhibited increased seed size. Resistance of *A. hybridus* to herbicides was transferred into the breeding lines of *A. hypochondriacus* and *A. cruentus* (Brenner et al., 2000).

Interspecies hybridization between grain amaranth species and vegetable ones often results in hybrids with teratological manifestations, high pollen sterility, and chromosomal aberrations, evidencing the existence of a significant incompatibility barrier between them (Mohindeen, Irulappan, 1993). Intraspecies hybridization within *A. hypochondriacus* failed to reveal any heterotic effect in the progeny. A similar result was observed for crosses among representatives of *A. cruentus*. Heterosis, however, was registered for interspecies crosses between *A. cruentus* and *A. hypochondriacus*, resulting in a statistically significant increase in the progeny's leaf biomass (Lehmann et al., 1991). M.G. Stetter et al. (2016) developed an effective technique for producing intra- and interspecies amaranth hybrids, which included immersing inflorescences in a water bath at 45 °C for 10 min to emasculate male flowers, as well as the SNP markers for their identification.

Male sterility is reproductive failure in some plants, where the male organs in hermaphrodite flowers are nonfunctional and produce nonviable pollen grains. It is widely used by plant breeders and commercial producers of hybrids. For amaranths, cytoplasmic male sterility (CMS) is a rare phenomenon, identified only in one species, *A. hypochondriacus* (Peters, Jain 1987; Brenner, 1993). Kenyan breeders (Gudu, Gupta 1988) pinpointed twenty plants with male sterility in a population of cv. Jumla. As a result of a long-term research, David Brenner from Iowa State University, USA, registered the first CMS line of amaranth, DB 199313, and selected a sterility maintainer for it (Brenner, 2019). The first CMS-based amaranth hybrid is likely to be expected in the nearest future.

Diallelic crosses among six genotypes of *A. hypochondriacus* ($F_1$ and $F_2$) were made to analyze protein content in amaranth seeds (Pandey, Pal, 1985). The resulting hybrids exceeded the average parental value in the studied character, and the hybrids from three of those crosses surpassed the best parent. These results confirmed the positive effect of hybridization on the breeding process aimed at higher protein content in seeds, an important trait for amaranth.

### Mutation breeding

The diversity of genetic combinations can be increased through both classical hybridization and mutagenesis. The frequency of spontaneous mutations is fairly low, and plant breeders induce them artificially with physical or chemical mutagens. This approach proved its efficiency for crop improvement. Mutation breeding of amaranths for higher seed quality and quantity has mainly been based on the use of radiation mutagenesis. For example, the radiation method (175 Gray) was used to develop mutant lines of *A. cruentus*, which reliably demonstrated a stable 1,000 seed weight increase in the $M_4$ and $M_5$ generations (Gajdošová et al., 2007). In 2009, mutants were obtained from the local Peruvian cultivar Selection Ancash, with higher concentrations of micronutrients and improved bioavailability due to a decrease in the content of phytic acid (Gómez-Pando et al., 2009). The 2022 studies resulted in the production of six mutant amaranth lines with resistance to soil salinity (Kpochemè et al., 2022). A team of Russian researchers used sodium azide treatment to develop salt-tolerant amaranth forms promising for further breeding; their seeds showed an increase in protein content by 52 % and that of linolenic acid by 25 % (Taipova et al., 2022).

### Polyploidy

Polyploidy is regarded as an important evolutionary process for many crop species. Artificially induced polyploidy is the most rapid method to produce new genotypes. It was as early as 50 years ago that a number of plant breeders from various countries started attempting to raise the productivity of grain and vegetable amaranths by inducing polyploidy with colchicine (Behera et al., 1974; Madhusoodanan, Pal, 1984; Sun Y., Yue, 1993). They described some morphological and phenological features in the produced plants: shortening and thickening of the stem, an increase in seed size by 42–159 %, and the onset of flowering occurring one week later than usual. Notably, tetraploids of *A. caudatus* manifested an increase in the content of protein (by 60 %), amino acids, lysine, and threonine. The results showed that polyploidy in amaranth led to an increase in grain size without a decrease in productivity or nutritional value, confirming the value of this method for breeding programs.

## Genetic engineering and molecular genetics methods

The progress in molecular biology over the past decades has considerably added to the knowledge required for plant genetic diversity management, contributing to the significant advancement of molecular genetics methods in breeding practice. Marker-assisted selection (MAS), being one of such methods, increased the efficiency of selection for a specific trait, and genetic engineering made it possible to transfer a gene from one plant organism to another.

Despite the great potential of amaranth, little research has been dedicated to the study of its genomics, concentrating mainly on the identification of its species diversity (Table 2). A genome-wide association study (GWAS) identified associations among specific phenotypes and genomic variants in 10 qualitative traits of amaranth (Jamalluddin et al., 2022). A total of 22 associated markers for inflorescence, leaf, petiole and stem pigmentation were identified on 16 chromosomes in 16 amaranth species. These SNP markers are sources of

valuable genetic information for those engaged in phenotyping different amaranth species and improving their cultivars. However, no reports have yet appeared concerning the development of markers for valuable biochemical parameters in amaranth.

Russian researchers (Shcherban, Stasyuk, 2020) undertook to analyze the polymorphism of the gene encoding the squalene synthase (*SQS*) enzyme in a number of grain and vegetable amaranths. They reported a low level of polymorphism and conservatism of the main functional domains in the gene's coding part. The data obtained by the authors can help to select grain amaranth forms with higher squalene concentration in seeds.

In 1997, a case study of cv. Azteca demonstrated the results of an Agrobacterium-mediated transformation in *A. hypochondriacus* (Jofre-Garfias et al., 1997). The authors were the first to develop a regeneration technique and an Agrobacterium-mediated system for the crop's transformation, using them to analyze the expression of the light-harvesting chlorophyll *a/b*-binding (*Lhcb*) protein gene promoter in transgenic plants. A team of Indian scientists studied the potential of Agrobacterium-mediated genetic transformation in *A. tricolor* introducing Ti-plasmid-based constructs with transgenes that enhanced resistance to biotic stresses caused by fungal pathogens, viruses, and pests (Pal et al., 2013). Their efforts resulted in the creation of a reproducible genetic transformation protocol that could be used to produce amaranth plants resistant to biotic factors. U. Munusamy et al. (2013) pioneered in reporting a successful Agrobacterium-mediated transformation of flowers in the inflorescence of *A. hypochondriacus*. This accomplishment expanded the possibilities of amaranth improvement, as it is not always possible to achieve differentiation of the shoots from a transformed hypocotyl callus (Murugan, Sathishkumar, 2016). For *A. cruentus*, a successful Agrobacterium-mediated transformation from epicotyl explants was reported by Russian authors (Taipova et al., 2020), with the efficiency percentage of 4 %.

Genome-editing tools can also serve to enhance gene efficiency in other crops. For example, protein-coding genes of amaranth have been used to produce transgenic plants of potato, bread wheat, and maize. A 1992 publication (Raina, Datta, 1992) reported successful molecular cloning of *AmA1*, a protein-coding gene from amaranth seeds with a balanced amino acid composition. Later, a team of Indian scientists succeeded in incorporating this gene into potato, which increased the total protein content in tubers by 60 % (Chakraborty et al., 2000). It is noteworthy that the transgenic potato manifested enhanced photosynthetic activity and higher leaf biomass, with an additional positive effect on the overall yield (Chakraborty et al., 2010).

The same gene of amaranth was used to transform bread wheat, increasing its content of essential amino acids, since bread wheat is known to have a severe deficiency in lysine, threonine, and tyrosine (Tamás et al., 2009). Scientists from the Mexican Center for Research and Advanced Studies (Centro de Investigación y de Estudios Avanzados) used the 11S globulin DNA from *A. hypochondriacus* to transform the genotype of tropical maize, producing transgenic maize plants with a superexpressed 11S globulin gene which encoded one of the storage proteins in amaranth seeds. As a result, the total protein content in maize seeds showed a 32 % increase (Rascon-Cruz et al., 2004).

Advances in genetic transformation will make it possible to enhance various traits of grain amaranth through genome editing in the nearest future.

**Table 2.** The use of genetic technologies in amaranth studies

| Genetic material | Research objective | Method | Authors |
|---|---|---|---|
| 33 accessions of grain amaranths | Species identification | RAPD | Transue et al., 1994 |
| 41 accessions of 4 amaranth species | Species identification | SNP | Maughan et al., 2011 |
| 348 accessions of 37 species | Genetic diversity evaluation | SSR | Suresh et al., 2014 |
| *A. hypochondriacus* | QTL mapping | SNP | Lightfoot et al., 2017 |
| 18 accessions of grain amaranths | Species identification, and phylogeny specification | RAPD, ISSR | Лиманская и др., 2017 |
| 30 accessions of *Amaranthus* spp. | Analysis of 15 phenotypic characteristics | RAPD | Oduwaye et al., 2019 |
| 188 accessions of vegetable, grain and weedy species | GWAS-analysis of morphological character | SNP | Nguyen et al., 2019 |
| *A. cruentus* cv Arusha | Studying the role of specific genes in phytic acid synthesis | Genome assembly on the chromosome level | Ma et al., 2021 |
| 188 accessions of 18 amaranth species | GWAS analysis of morphological characters, such as leaf, stem and inflorescence shape, size and color | SNP | Jamalluddin et al., 2022 |

## Potential of VIR's amaranth collection for breeding

As of 2023, 35 amaranth cultivars were listed in the State Register of the Russian Federation (State Register for Selection Achievements Admitted for Usage, 2023). The "oldest" among those, cv. Cherginsky, dates back to 1995; it represents the most numerous group of cultivars earmarked for animal feed purposes (17 in total). The grain group consists of three cultivars, the ornamental group of 10, and the vegetable one of 5. It is worth mentioning that domestic breeding achievements for the amaranth crop are insufficient both in quantity and in the diversity of uses.

The unique collection of amaranth maintained at VIR, currently holding 570 accessions from various countries, is unmatched in the world (Fig. 2).

The first accession, Sirukeerai (*Amaranthus* sp., PC-1), arrived from Bangalore Nursery and Gardens, India, in 1955. The collection was subsequently supplied with local cultivars and wild species by numerous collecting missions and shipments from various research centers, botanical gardens, genebanks, and breeding stations all over the world. The largest numbers of accessions came from Mexico, the USA, Germany, and India (Fig. 3).

*A. cruentus* accounts for 80 % of the amaranth collection (106 accessions), followed by *A. hypochondriacus* (89 accessions), *A. caudatus* (88), *Amaranthus* sp. (86), *A. hybridus* (51), and *A. tricolor* (41) (Fig. 4). Most of the species in the collection are monoecious. Accessions of *A. tuberculatus* and *A. palmeri* are dioecious.

The amaranth germplasm collection preserved at VIR undergoes comprehensive studies: useful agronomic traits of the accessions are evaluated, including biochemical indicators, morphological descriptions are produced, and utilization areas are specified. The identified valuable biotypes are grouped into trait-specific collections. On the basis of the data procured during long-term research, the amaranth accessions have been distributed into groups according to their best features: high seed yield, high leaf biomass, increased protein content in seeds, short stem, earliness, cold hardiness, resistance to seed shattering, and fitness for vegetable or ornamental use.

VIR has been conducting a study of the amino acid composition in the leaf biomass of cultivated vegetable and grain amaranths as well as wild species. A case study of accessions belonging to 12 different amaranth species identified 18 free amino acids, with eight of them essential (Sokolova et al., 2021). A number of grain amaranth accessions representing *A. caudatus*, *A. cruentus* and *A. hypochondriacus* were recognized as promising sources of highly balanced amino acid composition in green biomass. It was found within the same study that weedy amaranth species had significant potential in terms of their therapeutic effects on the human organism due to high phenolic and lysine contents in their leaves. An accession of *A. blitum* (syn. *A. lividus*) (PC-31, India) was identified as the best for its content of ascorbic acid (90.2 mg/100 g), saccharides, organic acids, phenolic compounds, and fatty acids, as well as for its capability to accumulate up to 90.53 mg/100 g of lysine, and significant amounts of tyrosine, tryptophan, and cystine.

One of the problems with amaranth cultivation in Russia is the crop's thermophilic nature. The optimum temperature range for seed germination is 20–25 °C (Kononkov, Sergeeva, 2011). That is why amaranth can be grown mainly in the southern regions of Russia. Hence, there is a need to develop cold-hardy cultivars. For many years the amaranth collection of VIR has been assessed for cold hardiness under the conditions of Northwest Russia. Genotypes have been selected for their tolerance to low temperatures, and ability to produce mature seeds within a shorter period. The result of these efforts is cv. Frant released by VIR, with its ability to form seeds in 90 days and provide three cuttings of green biomass for tea production in the environments of Leningrad Province.

## Conclusion

Amaranth is rapidly becoming more and more popular in Russia. Of late, it has received a lot of attention from researchers, medical experts, and crop producers due to its di-

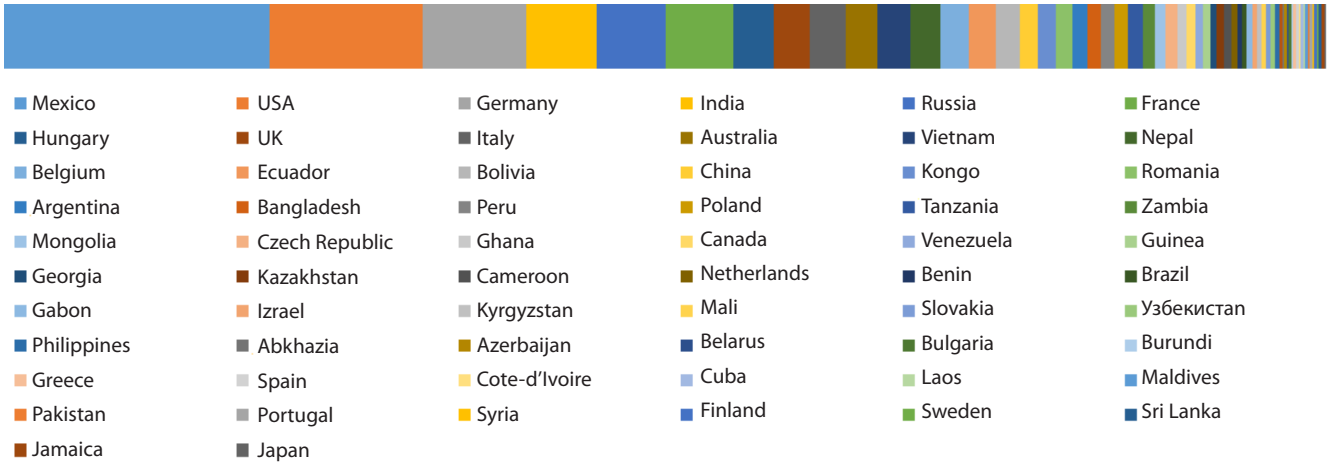



**Fig. 2.** Origin of amaranth accessions maintained at VIR.

Sirukeerai, *Amaranthus* sp.
(PC-1, India)

*A. caudatus* L.
(PC-146, Germany)

*A. cruentus* L.
(PC-94, USA)

*A. caudatus* L.
(PC-150, Greece)

*A. cruentus* L.
(PC-218, Mexico)

*A. cruentus* L.9/5
(PC-289, Mexico)


*A. tricolor* L.
(PC-321, India)

*A. blitum* L.
(PC-12, India)

Франт, *A. cruentus* L.
(PC-318, Russia)

**Fig. 3.** Amaranth accessions maintained at VIR (the photos were taken in the fields of Pushkin and Pavlovsk Laboratories of VIR, St. Petersburg).

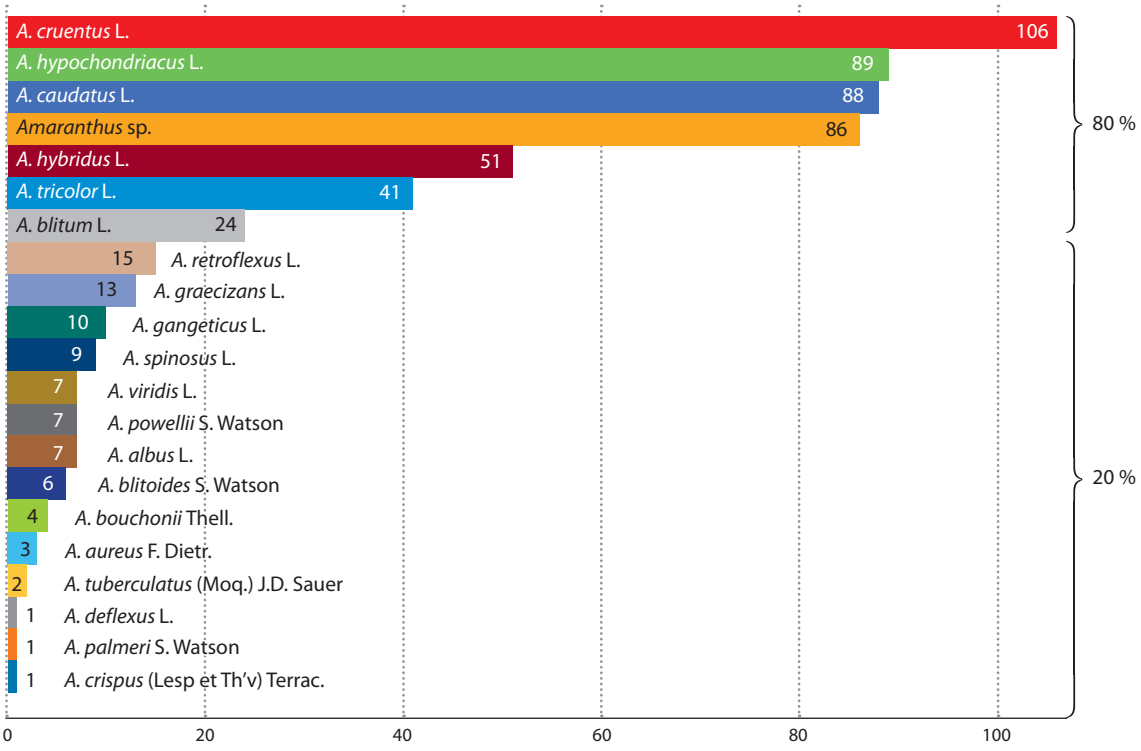



**Fig. 4.** Amaranth species diversity maintained at VIR.

verse uses, unique biochemical composition, and therapeutic potential. A wide range of genetic variability demonstrated by local cultivars of some amaranth species opens up great prospects for its improvement by both conventional and advanced breeding methods.

The amaranth collection maintained at VIR, with its nearly seventy-year history, is unique in its origin and diversity. It harbors trait-specific groups of accessions useful for all prioritized breeding trends. The crop's genetic diversity is highly promising for breeding practice and intensive research in the light of modern knowledge and technologies. Long-term comprehensive studies made it possible to identify amaranth accessions that may be recommended for inclusion in breeding programs. It should be highlighted that the amaranth gene pool preserved at VIR is capable of providing unlimited opportunities for breeding and meeting the needs of the country's population, enriching the diet with health-friendly and wholesome products.

## References

Abbasi D., Rouzbehan Y., Rezaei J. Effect of harvest date and nitrogen fertilization rate on the nutritive value of amaranth forage (*Amaranthus hypochondriacus*). *Anim. Feed Sci. Technol. Animal*. 2012;171:6-13. DOI 10.1016/j.anifeedsci.2011.09.014

Akin-Idowu P., Gbadegesin M., Orkpeh U., Ibitoye D., Odunola O. Characterization of grain amaranth (*Amaranthus* spp.) germplasm in south west Nigeria using morphological, nutritional, and random amplified polymorphic DNA (RAPD) analysis. *Resources*. 2016;5(1):6. DOI 10.3390/resources5010006

Arreguez G.A., Martínez J.G., Ponessa G. *Amaranthus hybridus* L. ssp. *hybridus* in an archaeological site from the initial mid-holocene in the southern argentinian Puna. *Quat. Int.* 2013;307:81-85. DOI 10.1016/j.quaint.2013.02.035

Ayorinde F.O. Determination of fatty acid composition of *Amaranthus* species. *J. Am. Oil Chem. Soc.* 1989;66:1812-1814

Baltensperger D.D., Weber L.E., Nelson L.A. Registration of 'Plainsman' grain amaranths. *Crop Sci*. 1992;32:1510-1511. DOI 10.2135/cropsci1992.0011183X003200060047x

Behera B., Tripathy A., Patnaik S.N. Histological analysis of colchicines-induced deformities and cytochimeras in *Amaranthus caudatus* and *A. dubius*. *J. Heridity*. 1974;65:179-184

Brenner D.M. Hybrid seeds for increased amaranth grain yield. *Legacy*. 1993;6:9-11

Brenner D.M. Registration of DB 199313, cytoplasmic male sterile grain amaranth genetic stock. *J. Plant Regist*. 2019;13:251-253. DOI 10.3198/jpr2018.06.0042crgs

Brenner D.M., Baltensperger D.D., Kulakow P.A., Lehmann J.W., Myers R.L., Slabbert M.M., Sleugh B.B. Genetic resources and breeding in *Amaranthus*. In: Janick J. (Ed.) Plant Breeding Reviews. Wiley, New York, 2000;19;227-285. DOI 10.1002/9780470650172.ch7

Bressani R. Composition and nutritional properties of amaranth. In: Paredes-Lopez O. (Ed.). Amaranth, Biology, Chemistry and Technology. Chap. 10. Boca Raton: CRC Press, 1994. DOI 10.1201/9781351069601-10

Chagaray A. Estudio de factibilidad del cultivo del amaranto. Dirección Provincial de programación del desarrollo Ministerio de producción y desarrollo Gobierno de la provincia de Catamarca. Catamarca, Argentina, 2005

Chakraborty S., Chakraborty N., Datta A. Increased nutritive value of transgenic potato by expressing a nonallergenic seed albumin gene from *Amaranthus hypochondriacus*. *Proc. Natl. Acad. Sci. USA*. 2000;97:3724-3729. DOI 10.1073/pnas.050012697

Chakraborty S., Chakraborty N., Agrawal L., Ghosh S., Narula K., Shekhar S., Naikb P.S., Pandec P.C., Chakrborti S.K., Datta A. Next-generation protein-rich potato expressing the seed protein gene *AmA1* is a result of proteome rebalancing in transgenic tuber. *Proc. Natl. Acad. Sci. USA*. 2010;107:17533-17538. DOI 10.1073/pnas.1006265107

Cheeke P.R., Carlsson R., Kohler G.O. Nutritive value of leaf protein concentrates prepared from *Amaranthus* species. *Can. J. Anim. Sci.* 1981;61:199-204. DOI 10.4141/cjas81-026

Cheng A., Mayes S., Dalle G., Demissew S., Massawe F. Diversifying crops for food and nutrition security – a case of teff. *Biol. Rev.* 2015;92:188-198. DOI 10.1111/brv.12225

Costea M., Sanders A., Waines G. Preliminary results towards revision of the *Amaranthus* hybridus species complex (Amaranthaceae). *Sida*. 2001;19:931-974

Covas G. Perspectivas del cultivo de los amarantos en la republica Argentina. 1993. https://www.semanticscholar.org/paper/Perspectivas-del-cultivo-de-los-amarantos-en-la-Covas/c2d32274738dbf8eb28950d2f077f6809c3b132f

Das S. Taxonomical observation on the grain amaranths and new varieties of *Amaranthus cruentus* (Amaranthaceae). *Nord. J. Bot*. 2012; 30:412-420. DOI 10.1111/j.1756-1051.2011.01383.x

Das S. *Amaranthus*: A Promising Crop of FUTURE. Springer, Singapore, 2016. DOI 10.1007/978-981-10-1469-7

Dawson I.K., Park S.E., Attwood S.J., Jamnadass R., Powell W., Sunderland T., Carsan S. Contributions of biodiversity to the sustainable intensification of food production. *Global Food Secur*. 2019;21:23-37. DOI 10.1016/j.gfs.2019.07.002

De Montellano B.R.O. Aztec medicine, health, and nutrition. Rutgers University Press, Great Britain, 1990

Ebert A. Potential of underutilized traditional vegetables and legume crops to contribute to food and nutritional security, income and more sustainable production systems. *Sustainability*. 2014;6:319-335. DOI 10.3390/su6010319

Espitia E. Amaranth germplasm development and agronomic studies in Mexico. *Food Rev. Int.* 1992;8:71-86. DOI 10.1080/87559129209540930

Evon P., de Langalerie G., Labonne L., Merah O., Talou T., Ballas S., Véronèse T. Low-density insulation blocks and hardboards from amaranth (*Amaranthus cruentus*) Stems, a new perspective for building applications. *Coatings.* 2021;11:349. DOI 10.3390/coatings11030349

Gajdošová A., Libiaková G., Fejér J. Improvement of selected *Amaranthus* cultivars by means of mutation induction and biotechnological approaches. 2007. In: Ochatt S., Mohan Jain S. (Eds.). Breeding of Neglected and Under-utilized Crops, Spices and Herbs. Edenbridge Ltd. USA, 2007;151-169

Gamel T.H., Mesallam A.S., Damir A.A., Shekib L.A., Linssen J.P. Characterization of amaranth seed oils. *J. Food Lipids*. 2007;14: 323-334

Gómez-Pando L.R., Eguiluz A., Jiménez J., Falconi J., Aguilar E., Shu Q. Barley (*Hordeun vulgare*) and kiwicha (*Amaranthus caudatus*) improvement by mutation induction in Peru. In: Shu Q.Y. (Ed.). Induced Plant Mutation in the Genomics Era. Food and Agriculture Organization of the United Nations, Rome, 2009;330-332

Gonor K.V., Pogozheva A.V., Derbeneva S.A., Maltsev G.Yu., Trushina E.N., Mustaphina O.K. The influence of a diet with including amaranth oil antioxidant and immune status in patients with ischemic heart disease and hyperlipoproteidemia. *Voprosy Pitaniya = Problems of Nutrition*. 2006;75(6):30-33. (in Russian)

Greizerstein E.J., Poggio L. Estudios citogeneticos de seis hibridos interespecificos de *Amaranthus* (Amaranthaseae). *Darwiniana*. 1992; 31:159-165

Greizerstein E.J., Poggio L. Meiotic studies of spontaneous hybrids of *Amaranthus*: genome analysis. *Plant Breed*. 1995;114:448-450

Grobelnik-Mlakar S., Turinek M., Jakop M., Bavec M., Bavec F. Nutrition value and use of grain amaranth: potential future application in bread making. *Agricultura*. 2009;6:43-53

Gudu S., Gupta V.K. Male-sterility in the grain amaranth (*Amaranthus hypochondriacus* ex-Nepal) variety Jumla. *Euphytica*. 1988;37:23-26. DOI 10.1007/BF00037218

Hauptli H., Jain S.K. Biosystematics and agronomic potential of some weedy and cultivated amaranths. *Theor. Appl. Genet.* 1978;52:177-185. DOI 10.1007/bf00282575

Hauptli H., Jain S. Genetic variation in outcrossing rate and correlated floral traits in a population of grain amaranth (*Amaranthus cruentus* L.). *Genetica*. 1985;66:21-27. DOI 10.1007/bf00123602

He H.P., Corke H. Oil and squalene in *Amaranthus* grain and leaf. *J. Agric. Food Chem*. 2003;51:7913-7920. DOI 10.1021/jf030489q

Huang Z.R., Lin Y.K., Fang J.Y. Biological and pharmacological activities of squalene and related compounds: potential uses in cosmetic dermatology. *Molecules*. 2009;14:540-554. DOI 10.3390/molecules14010540

Jamalluddin N., Massawe F.J., Mayes S., Ho W.K., Symonds R.C. Genetic diversity analysis and marker-trait associations in *Amaranthus* species. *PLoS One*. 2022;17:0267752. DOI 10.1371/journal.pone.0267752

Jofre-Garfias A.E., Villegas-Sepúlveda N., Cabrera-Ponce J.L., Adame-Alvarez R.M., Herrera-Estrella L., Simpson J. Agrobacterium-mediated transformation of *Amaranthus hypochondriacus*: light- and tissue-specific expression of a pea chlorophyll a/b-binding protein promoter. *Plant Cell Rep*. 1997;16:847-852. DOI 10.1007/s002990050332

Joshi B.D., Rana R.S. Grain Amaranths: the Future Food Crops. New Delhi: National Bureau of Plant Genetic Resources, 1991

Joshi B.D., Mehra K.L., Sharma S.D. Cultivation of grain amaranth in the north-western hills. *Indian Farming*. 1983;32:34-37

Joshi D.C., Sood S., Hosahatti R., Kant L., Pattanayak A., Kumar A., Yadav D., Stetter M.G. From zero to hero: the past, present and future of grain amaranth breeding. *Theor. Appl. Genet.* 2018;131: 1807-1823. DOI 10.1007/s00122-018-3138-y

Kauffman C.S. Thoughts on the development of improved varieties of grain amaranth. In: Proceedings Third Amaranth Conference, Grain Amaranth: Expanding Consumption through Improved Cropping, Marketing and Crop Development. Rodale Press, USA, 1984

Kauffman C.S. Realizing the potential of grain amaranth. *Food Rev. Int.* 1992;8:5-21. DOI 10.1080/87559129209540927

Kauffman C.S., Weber L.E. Grain amaranth. In: Advances in New Crops. Portland: Timber Press, 1990;127-139

Khoshoo T.N., Pal M. Cytogenetic pattern in *Amaranthus*. *Chromosomes Today.* 1972;3:259-267

Khoury C.K., Bjorkman A.D., Dempewolf H., Ramirez-Villegas J., Guarino L., Jarvis A., Rieseberg L.H., Struik P.C. Increasing homogeneity in global food supplies and the implications for food security. *Proc. Natl. Acad. Sci. USA*. 2014;111:4001-4006. DOI 10.1073/pnas.1313490111

Kononkov P.F., Gins V.K., Gins M.S. Amaranth is a Promising Crop of the 21st Century. Moscow, RUDN University Publ., 1999 (in Russian)

Kononkov P.F., Sergeeva V.A. Amaranth – the valuable vegetable and forage crops multifaceted use. *Agrarnyy Vestnik Urala = Agrarian Bulletin of the Urals*. 2011;4:63-64 (in Russian)

Kpochemè A.O.E.K., Hotegni N.F., Missihoun A.A., Gnanvi B.N., Atou R., Wouyou A., Montcho D., Gandonou C.B., Agbangla C., Ahoton L. Morphological characterization of *Amaranthus cruentus* L. mutant lines derived from local and preferred Amaranthus cultivar. *Int. J. Biol. Chem. Sci.* 2022;16:1554-1569. DOI 10.4314/ijbcs.v16i4.16

Lehmann J.W., Clark R.L., Frey K.J. Biomass heterosis and combining ability in interspecific and intraspecific matings of grain amaranths. *Crop Sci.* 1991;31:1111-1116. DOI 10.2135/cropsci1991.0011183x003100050004x

Leon-Camacho M., Garcia-Gonzalez D.L., Aparicio R. A detailed and comprehensive study of amaranth (*Amaranthus cruentus* L.) oil fatty profile. *Eur. Food Res. Technol*. 2001;213:349-355. DOI 10.1007/s002170100340

Lightfoot D.J., Jarvis D.E., Ramaraj T., Lee R., Jellen E.N., Maughan P.J. Single-molecule sequencing and Hi-C-based proximity-guided assembly of amaranth (*Amaranthus hypochondriacus*) chromosomes provide insights into genome evolution. *BMC Biology*. 2017;15:74. DOI 10.1186/s12915-017-0412-4

Lymanska S.V., Miroshnichenko L.A., Goptsiy T.I., Korneeva O.S. Polymorphism of RAPD and ISSR markers in grain amaranth spe-

cies. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2017;21(2):189-197. DOI 10.18699/VJ17.236

Ma X., Vaistij F.E., Li Y., Van Rensburg W.S.J., Harvey S., Bairu M.W., Venter S.L., Mavengahama S., Ning Z., Graham I.A., Deynze A.V., Peer Y.V., Denby K.J. A chromosome-level *Amaranthus cruentus* genome assembly highlights gene family evolution and biosynthetic gene clusters that may underpin the nutritional value of this traditional crop. *Plant J*. 2021;107:613-628. DOI 10.1111/tpj.15298

Madhusoodanan K.J., Pal M. Autotetraploids in *Amaranthus* tricolor Linn. *Indian J. Genet*. 1984;44:181-185

Magomedov I.M., Chirkova T.V. Amaranth – past, present and future. *Uspekhi Sovremennogo Yestestvoznaniya = Advances in Current Natural Sciences*. 2015;1:1108-1113 (in Russian)

Magomedov I.M., Chirkova A.I., Chirkova T.V. The role of biopeptides and antioxidants from amaranth grain in the prevention of chronic human diseases. *Prakticheskaya Fitoterapiya = Practical Phytotherapy.* 2017;2:49-54 (in Russian)

Mangelsdorf P.C. Genetic potentials for increasing yields of food crops and animals. *Proc. Natl. Acad. Sci. USA*. 1966;56:370-375. DOI 10.1073/pnas.56.2.370

Maughan P., Smith S., Fairbanks D., Jellen E. Development, characterization, and linkage mapping of single nucleotide polymorphisms in the grain amaranths (*Amaranthus* sp.). *Plant Gen.* 2011;4:92. DOI 10.3835/plantgenome2010.12.0027

Mayes S., Massawe F.J., Alderson P.G., Roberts J.A., Azam-Ali S.N., Hermann M. The potential for underutilized crops to improve security of food production. *J. Exp. Bot*. 2011;63:1075-1079. DOI 10.1093/jxb/err396

Miettinen T.A., Vanhanen H. Serum concentration and metabolism of cholesterol during rapeseed oil and squalene feeding. *Am. J. Clin. Nutr*. 1994;59(2):356-363. DOI 10.1093/ajcn/59.2.356

Mohindeen H.K., Irulappan I. Improvement in amaranths. In: Chadha K.L., Kalloo G. (Eds.) Advances in Horticulture: Vegetable Crops. New Delhi: Malhotra Publishing House, 1993

Munusamy U., Abdullah S.N.A., Aziz M.A., Khazaai H. Female reproductive system of *Amaranthus* as the target for *Agrobacterium*-mediated transformation. *Adv. Biosci. Biotechnol*. 2013;4:188-192. DOI 10.4236/abb.2013.42027

Murray M.J. The genetics of sex determination in the family Amaranthaceae. *Genetics*. 1940;25:409-431. DOI 10.1093/genetics/25.4.409

Murugan S.B., Sathishkumar R. Establishment of high frequency callus induction and genetic transformation in neglected leafy vegetable *Amaranthus trisis*. *Austin J. Biotechnol. Bioeng.* 2016;3:1058

Myers R.L., Putnam D.H. Growing grain amaranth as a specialty crop. Center for alternative crops and products, university of Minnesota. 1988. https://conservancy.umn.edu/items/d6b7c0a1-6c3e-45a8-bf8d-08ab79a5cdb6 (date of access: 28.06.2024)

Nguyen D.C., Tran D.S., Tran T.T.H., Ohsawa R., Yoshioka Y. Genetic diversity of leafy amaranth (*Amaranthus tricolor* L.) resources in Vietnam. *Breed. Sci.* 2019;69:640-650. DOI 10.1270/jsbbs.19050

Oduwaye O.A., Ayo-Vaughan M.A., Porbeni J.B.O., Oyelakin O.O. Genetic diversity in Amaranth (*Amaranthus* spp.) based on phenotypic and RAPD markers. *Nigerian J. Biotechnol*. 2019;36:62-68. DOI 10.4314/njb.v36i1.9

Pal A., Swain S.S., Das A.B., Mukherjee A.K., Chand P.K. Stable germ line transformation of a leafy vegetable crop amaranth (*Amaranthus tricolor* L.) mediated by *Agrobacterium tumefaciens*. *In Vitro Cell. Dev. Biol.-Plant*. 2013;49:114-128. DOI 10.1007/s11627-013-9489-9

Palombini S.V., Claus T., Maruyama S.A., Gohara A.K., Souza A.H.P., Souza N.E., Visentainer J.V., Gomes S.T.M., Matsushita M. Evaluation of nutritional compounds in new amaranth and quinoa cultivars. *Food Sci. Technol*. 2013;33:339-344. DOI 10.1590/s0101-20612013005000051

Pandey R.M., Pal M. Genetics of grain protein in *Amaranthus*. *Crop Improv*. 1985;12:55-58

Patent, 2022. https://gossortrf.ru/registry/gosudarstvennyy-reestr-selektsionnykh-dostizheniy-dopushchennykh-k-ispolzovaniyu-tom-1-sorta-rasteni/frant-amarant-metelchatyy/ (date of access: 04.11.2024) (in Russian)

Peters I., Jain S. Genetics of grain amaranths. Gene-cytoplasmic male-sterility. *J. Heredity*. 1987;78:251-256. DOI 10.1093/oxfordjournals.jhered.a110377

Raiger H.L., Bhandari D.C. Underutilized crops: varieties released in India. All India Coordinated Research Network on Underutilized Crops. New Daihi: NBPGR, 2012

Raina A., Datta A. Molecular cloning of a gene encoding a seed-specific protein with nutritionally balanced amino acid composition from *Amaranthus*. *Proc. Natl. Acad. Sci. USA*. 1992;89:11774-11778. DOI 10.1073/pnas.89.24.11774

Ramdwar M.N., Chadee S.T., Stoute V.A. Estimating the potential consumption level of amaranth for food security initiatives in Trinidad, West Indies. *Cogent Food Agric*. 2017;3:1321475. DOI 10.1080/23311932.2017.1321475

Rao C.V., Newmark H.L., Reddy B.S. Chemopreventive effect of squalene on colon cancer. *Carcinogenesis*. 1998;19:287-290. DOI 10.1093/carcin/19.2.287

Rascón-Cruz Q., Sinagawa-García S., Osuna-Castro J.A., Bohorova N., Paredes-López O. Accumulation, assembly, and digestibility of amarantin expressed in transgenic tropical maize. *Theor. Appl. Genet.* 2004;108:335-342. DOI 10.1007/s00122-003-1430-x

Rezaei J., Rouzbehan Y., Fazaeli H. An assessment of digestibility and protein quality of the fresh and ensiled amaranth forage according to CNCPS. *Iranian J. Anim. Sci.* 2009;40:31-38

Ruth O.N., Unathi K., Nomali N., Chinsamy M. Underutilization versus nutritional-nutraceutical potential of the *Amaranthus* food plant: a mini-review. *Appl. Sci.* 2021;11:6879. DOI 10.3390/app11156879

Sauer J.D. The grain amaranths and their relatives: a revised taxonomic and geographic survey. *Ann. Mo. Bot. Gard.* 1967;54:103-137. DOI 10.2307/2394998

Saunders R.M., Becker R. *Amaranthus*: a potential food and feed resource. In: Pomeranz Y. (Ed.). Advances in Cereal Science and Technology. St. Paul: American Association of Cereal Chemists, 1984; 357-396

Schulz-Schaeffer J., Stallknecht G.F., Baldridge D.E., Larson R.A. Registration of Montana-3 grain amaranth germplasm. *Crop Sci*. 1989a;29:244-245

Schulz-Schaeffer J., Webb D.M., Baldridge D.E., Stallknecht G.F., Larson R.A. Registration of Montana-5 grain amaranth germplasm. *Crop Sci.* 1989b;29:1581

Schulz-Schaeffer J., Baldridge D.E., Bowman H.F., Stallknecht G.F., Larson R.A. Registration of 'Amont' grain amaranth. *Crop Sci*. 1991;31:482-483

Shadi H., Rouzbehan Y., Rezaei J., Fazaeli H. Yield, chemical composition, fermentation characteristics, in vitro ruminal variables, and degradability of ensiled amaranth (*Amaranthus hypochondriacus*) cultivars compared with corn (*Zea mays*) silage. *Transl. Anim. Sci.* 2020;4:1-12. DOI 10.1093/tas/txaa180

Shcherban A.B., Stasyuk A.I. Polymorphism of the squalene synthase gene (*SQS*) in different species of amaranth (*Amaranthus* L.). *Russ. J. Genet.* 2020;56:298-306. DOI 10.31857/S0016675820030145

Singh A., Mahato A.K., Maurya A., Rajkumar S., Singh A.K., Bhardwaj R., Kaushik S.K., Kumar S., Gupta V., Singh K., Singh R. Amaranth Genomic Resource Database: an integrated database resource of Amaranth genes and genomics. *Front. Plant Sci.* 2023;14:1203855. DOI 10.3389/fpls.2023.1203855

Sleugh B.B., Moore K.J., Brummer E.C., Knapp A.D., Russell J., Gibson L. Forage value of various amaranth species at different harvest dates. *Crop Sci.* 2021;41:466-472. DOI 10.2135/cropsci2001.412466x

Smith M.E. The Aztecs. Blackwell, Oxford, 1996

Smith T.J. Squalene: potential chemopreventive agent. *Expert Opin. Invest. Drugs*. 2000;9:1841-1848. DOI 10.1517/13543784.9.8.1841

Sokolova D., Zvereva O., Shelenga T., Solovieva A. Comparative characteristics of the amino acid composition in amaranth accessions from the VIR collection. *Turk. J. Agric. For.* 2021;45:68-78. DOI 10.3906/tar-2007-7

Soriano-Garcıa M., Arias-Olguín I.I., Montes J.P.C., Ramırez D.G.R., Silvestre Mendoza Figueroa J., Flores-Valverde E., Rodrguez M.R.V. Nutritional functional value and therapeutic utilization of amaranth. *J. Analytical Pharm. Res.* 2018;7(5):596-600. DOI 10.15406/japlr.2018.07.00288

Sreelathakumary I., Peter K.V. Amaranth-*Amaranthus* spp. In: Kallo G., Bergh B.O. (Eds.). Genetic Improvement of Vegetable Crops. Oxford: Pergamon Press, 1993; 315-323

Stallknecht G.F., Schulz-Schaefer J.R. Amaranth rediscovered. In: Janick J., Simon J.E. (Eds.). New Crops. New York: Wiley, 1993

State Register of Selection Achievements Admitted for Usage, 2023. https://gossortrf.ru/publication/reestry.php (date of access: 24.04.2024). (in Russian)

Stetter M.G., Zeitler L., Steinhaus A., Kroener K., Biljecki M., Schmid K.J. Crossing methods and cultivation conditions for rapid production of segregating populations in three grain amaranth species. *Front. Plant Sci*. 2016;7:816. DOI 10.3389/fpls.2016.00816

Sun Y., Yue S. Research on polyploid grain amaranth – a preliminary study on selection of grain amaranth with character of bigger seed (in Chinese). In: Yue S. (Ed.). The Research and Development of Grain Amaranth in China. Beijing, China, 1993

Sun G.Q., Ma J., Du W., Wang Y., Li S.L., Xiong Y., Yu X., Lei X., Yabin M.L. Effects of dietary supplementation of *Amaranthus hypochondriacus* L. on Ruminal fermentation, blood indicators and performance of dairy cows. *Chin. J. Animal Nutr*. 2017;29:1652-1660

Suresh S., Chung J.-W., Cho G.-T., Sung J.-S., Park J.-H., Gwag J.-G., Baek H.-J. Analysis of molecular genetic diversity and population structure in *Amaranthus* germplasm using SSR markers. *Plant Biosyst. – Int. J. Dealing Aspects Plant Biosyst*. 2014;148:635-644. DOI 10.1080/11263504.2013.788095

Taipova R.M., Musin K.G., Kuluev B.R. Agrobacterium-mediated transformation of *Amaranthus cruentus* L. epicotils. *Zhurnal Sibirskogo Federal'nogo Universiteta. Seriya: Biologiya = Journal of Siberian Federal University Biology*. 2020;13:179-187. DOI 10.17516/1997-1389-0292 (in Russian)

Taipova R.M., Nesterov V.N., Rozentsvet O.A., Kuluev B.R. Changes in the content of proteins and lipids and in the state of the antioxidant system in mutant forms of *Amaranthus cruentus* L. *Trudy po Prikladnoy Botanike, Genetike i Selektsii = Proceedings on Applied Botany, Genetics and Breeding*. 2022;183:76-85. DOI 10.30901/2227-8834-2022-1-76-85 (in Russian)

Tamás C., Kisgyörgy B.N., Rakszegi M., Wilkinson M.D., Yang M.-S., Láng L., Tamas L., Bedő Z. Transgenic approach to improve wheat (*Triticum aestivum* L.) nutritional quality. *Plant Cell Rep*. 2009;28:1085-1094. DOI 10.1007/s00299-009-0716-0

Tang Y., Tsao R. Phytochemicals in quinoa and amaranth grains and their antioxidant, anti-inflammatory, and potential health beneficial effects: A review. *Mol. Nutr. Food Res.* 2017;61:1600767. DOI 10.1002/mnfr. 201600767

Teutonico R.A., Knorr D. Amaranth: composition, properties and applications of a rediscovered food crop. *Food Technol*. 1985;1:49-60

Transue D.K., Fairbanks D.J., Robison L.R., Andersen W.R. Species identification by RAPD analysis of grain amaranth genetic resources. *Crop Sci*. 1994;34:1385. DOI 10.2135/cropsci1994.0011183x003400050044x

Venskutonis P.R., Kraujalis P. Nutritional components of amaranth seeds and vegetables: A review on composition, properties, and uses. *Compr. Rev. Food Sci. Food Saf.* 2013;12:381-412. DOI 10.1111/1541-4337.12021

**Conflict of interest.** The authors declare no conflict of interest.
Received May 8, 2024. Revised July 30, 2024. Accepted July 31, 2024.

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2024;28(7):744-751

DOI 10.18699/vjgb-24-82

# Substantia nigra alterations in mice modeling Parkinson's disease

I.N. Rozhkova[1], S.V. Okotrub[1], E.Yu. Brusentsev [1], T.A. Rakhmanova[1, 2], D.A. Lebedeva[1], V.S. Kozeneva[1, 2], N.A. Shavshaeva[1, 2], N.V. Khotskin[1], S.Ya. Amstislavsky[1]

[1] Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
[2] Novosibirsk State University, Novosibirsk, Russia

amstis@yandex.ru

**Abstract.** Parkinson's disease (PD) is an age-related neurodegenerative pathology of the central nervous system. The well-known abnormalities characteristic of PD are dysfunctions in the nigrostriatal system including the substantia nigra of the midbrain and the striatum. Moreover, in PD persons, alpha-synucleinopathy is associated with abnormalities in the dopaminergic brain system. To study the mechanisms of this pathology, genetic models in mice have been designed. Transgenic mice of the B6.Cg-Tg(Prnp-SNCA*A53T)23Mkle/J strain (referred to as B6.Cg-Tg further in the text) possess the A53T mutation in the human alpha-synuclein *SNCA* gene. The density of neurons in the prefrontal cortex, hippocampus, substantia nigra and striatum in B6.Cg-Tg mice was assessed in our previous work, but the dopaminergic system was not studied there, although it plays a key role in the development of PD. The aim of the current study was to investigate motor coordination and body balance, as well as dopaminergic neuronal density and alpha-synuclein accumulation in the substantia nigra in male B6.Cg-Tg mice at the age of six months. Wild-type mice of the same sex and age, siblings of the B6.Cg-Tg mice from the same litters, lacking the *SNCA* gene with the A53T mutation, but expressing murine alpha-synuclein, were used as controls (referred to as the wild type further in the text). Motor coordination and body balance were assessed with the rota-rod test; the density of dopaminergic neurons and accumulation of alpha-synuclein in the substantia nigra were evaluated by the immunohistochemical method. There was no difference between B6.Cg-Tg mice and WT siblings in motor coordination and body balance. However, accumulation of alpha-synuclein and a decrease in the number of dopaminergic neurons in the substantia nigra were found in the B6.Cg-Tg mouse strain. Thus, the mice of the B6.Cg-Tg strain at the age of six months have some symptoms of the onset of PD, such as the accumulation of mutant alpha-synuclein and a decrease in the number of dopaminergic neurons in the substantia nigra. Taken together, the results obtained in our work qualify the B6.Cg-Tg strain as a pertinent model for studying the early stage of human PD already at the age of six months.

**Key words:** mice; Parkinson's disease; motor coordination; dopaminergic brain system; alpha-synuclein.

**For citation:** Rozhkova I.N., Okotrub S.V., Brusentsev E.Yu., Rakhmanova T.A., Lebedeva D.A., Kozeneva V.S., Shavshaeva N.A., Khotskin N.V., Amstislavsky S.Ya. Substantia nigra alterations in mice modeling Parkinson's disease. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2024;28(7):744-751. DOI 10.18699/vjgb-24-82

**Funding.** The work was carried out with the support of RNF No. 23-25-00123.

**Acknowledgements.** The work was carried out using the equipment of the Shared Use Center "Center for Genetic Resources of Laboratory Animals" of the Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, supported by the Ministry of Education and Science of Russia (unique project identifier RFMEFI62119X0023). Microscopic work was carried out in the Shared Use Center for Microscopic Analysis of Biological Objects of the Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences.

**Contribution of the authors.** The idea of the work and the planning of the experiment – S.Ya.A., I.N.R., E.Yu.B., conducting behavioral experiments – I.N.R., V.S.K., S.N.A., N.V.H., processing the results – S.V.O., E.Yu.B., D.A.L., T.A.R., V.S.K., S.Ya.A., writing and editing of the manuscript – S.Ya.A., E.Yu.B., V.S.K., S.N.A.

# Изменения в черной субстанции головного мозга у мышей, моделирующих болезнь Паркинсона

И.Н. Рожкова[1], С.В. Окотруб[1], Е.Ю. Брусенцев [1], Т.А. Рахманова[1, 2], Д.А. Лебедева[1], В.С. Козенева[1, 2], Н.А. Шавшаева[1, 2], Н.В. Хоцкин[1], С.Я. Амстиславский[1]

[1] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
[2] Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия

amstis@yandex.ru

**Аннотация.** Болезнь Паркинсона (БП) – возрастная нейродегенеративная патология центральной нервной системы. Наиболее характерными нарушениями при БП считаются аномалии в нигростриарной системе, включающей в себя черную субстанцию среднего мозга и полосатое тело. При БП аномалии дофаминергической системы мозга сопровождаются альфа-синуклеинопатией. Для изучения механизмов возникновения данной патологии созданы генетические модели на мышах. Трансгенные мыши линии B6.Cg-Tg(PrNp-SNCA*A53T)23Mkle/J (далее по тексту B6.Cg-Tg) имеют мутацию A53T в гене *SNCA* альфа-синуклеина человека. В нашей предыдущей работе оценена плотность нейронов в префронтальной коре, гиппокампе, черной субстанции и полосатом теле у мышей этой линии, однако дофаминергическая система мозга, которая играет ключевую роль в развитии БП, изучена не была. Целью настоящего исследования стало изучение координации движений и баланса тела, а также плотности дофаминовых нейронов и накопления альфа-синуклеина в черной субстанции самцов мышей линии B6.Cg-Tg в возрасте шести месяцев. В качестве контроля использованы сибсы, у которых не было экспрессии гена *SNCA* с мутацией A53T и экспрессировался мышиный альфа-синуклеин (далее по тексту – дикий тип; wild type, WT), того же пола и возраста, из тех же самых выводков, что и исследуемые мыши B6.Cg-Tg. Координация движений и баланс тела были изучены с помощью теста «рота-род»; плотность дофаминовых нейронов и накопление альфа-синуклеина в черной субстанции оценены иммуногистохимическим методом. Полученные результаты показывают, что мыши B6.Cg-Tg не имеют отличий по координации движений и баланса тела от контроля – сибсов дикого типа. Однако у мышей B6.Cg-Tg в черной субстанции были обнаружены накопление альфа-синуклеина и уменьшение числа дофаминовых нейронов. Таким образом, мыши линии B6.Cg-Tg в возрасте шести месяцев имеют симптомы начала развития БП, такие как накопление мутантного альфа-синуклеина и уменьшение числа дофаминовых нейронов в черной субстанции. Полученные в этом исследовании результаты позволяют характеризовать линию B6.Cg-Tg в качестве адекватной модели для изучения ранней стадии БП человека уже в возрасте шести месяцев.
**Ключевые слова:** мыши; болезнь Паркинсона; координация движений; дофаминергическая система мозга; альфа-синуклеин.

## Introduction

Parkinson's disease (PD) is the second most common neurodegenerative disorder in humans. It manifests primarily in idiopathic and sporadic forms (Beitz, 2014; Tran et al., 2020; Bidesi et al., 2021). PD's main symptoms include impaired motor function, muscle rigidity, resting tremors, and bradykinesia (Halliday et al., 2006; Beitz, 2014). Non-motor symptoms such as sleep disturbances, neuropsychiatric conditions, and cognitive deficits are also prevalent (Beitz, 2014; Hayes, 2019). Age-related changes in various brain areas are common in PD (Halliday et al., 2006; Dickson et al., 2009).

This pathology is characterized by a variety of abnormalities related to the synthesis of neurotransmitters in the brain and thefunctioning of their receptors (Jellinger, 1991; Deutch et al., 2006). The most common brain pathology associated with PD involves the nigrostriatal pathway, which encompasses the substantia nigra (SNC) in the midbrain and the striatum (STR) (Dickson et al., 2009; Beitz, 2014; Hayes, 2019). This pathway is crucial for motor control, as it activates dopaminergic neurons that regulate flexor and extensor muscle movements (Korchounov et al., 2010).

Another key feature of PD is the accumulation of alpha-synuclein in brain neurons, particularly in the SNC and STR (Dickson et al., 2009; Burre et al., 2018; Lai et al., 2021). Alpha-synuclein regulates synaptic activity, including dopamine synthesis, transport, and storage (Burre et al., 2018; Bidesi et al., 2021). Mutations in the alpha-synuclein gene, such as A53T and A30P, result in abnormal protein folding and aggregation (Polymeropoulos et al., 1997; Spillantini et al., 1997). These changes lead to neuronal loss and alpha-synuclein accumulation, both hallmark features of PD (Dickson et al., 2009; Venda et al., 2010; Poewe et al., 2017).

Animal models are used to study PD mechanisms and identify potential treatment strategies. These models are typically classified as toxic or genetically engineered (Grigoryan, Bazyan, 2007; Korolenko et al., 2020). Studies have shown that damage to neurons in the STR leads to simultaneous flexor and extensor muscle impairment (Stern, 1966; Kato, Kimura, 1992). In PD models, motor deficits are often accompanied by alpha-synucleinopathy (Dickson, 2018).

Commonly used PD models include transgenic mice expressing mutant forms of the human alpha-synuclein (*SNCA*) gene, such as the A30P and A53T mutations (Unger et al., 2006; Grigoryan, Bazyan, 2007; Korolenko et al., 2020). These models display motor impairments that correlate with nigrostriatal degeneration (Chia et al., 2020). Transgenic mice with the A30P mutation in the *SNCA* gene typically exhibit a milder form of BP phenotype compared to mice with the A53T mutation, which makes it challenging to study behavioral characteristics and disorders in different brain structures (Van der Putten et al., 2000; Lee et al., 2001; Crabtree, Zhang, 2012). However, results can vary based on the genetic background, affecting the extent to which PD traits are expressed, how and under which promoter the gene was transferred, and on other factors (Crabtree, Zhang, 2012).

Various techniques are employed to investigate PD in mouse models (Graham, Sidhu, 2010; Tikhonova et al., 2020; Langley et al., 2021). In particular, the rotarod test (RR) is widely used to assess coordination of movement and body balance (Graham, Sidhu, 2010; Seo et al., 2020). However, the results of testing for the key features characterizing PD may vary if the same transgene is expressed on a different genetic background. The review article by D.M. Crabtree and J. Zhang (2012) describes in detail how different genetic

factors affect the manifestation of the pathology emphasizing the role of genetic background on which the transgenic model was designed and the promoter under which the transgene was embedded. The study of Paumier et al. (2013) indicates that transgenic mice that express human alpha-synuclein with the A53T mutation at the age of two months exhibit a longer latency time before the drop from a rotating rod during the rotarod test (which evaluates coordination of movement and body balance) compared to the wild-type siblings. Results of other studies exploiting the transgenic mouse strains with the same A53T mutation were either opposite (Graham, Sidhu, 2010) or demonstrated no differences in the RR test results between the wild-type control animals and the transgenic ones even at the age of nine months (Liu et al., 2018) depending on the genetic background on which the transgenic mouse strain was obtained.

When studying mice modeling PD, immunohistochemical methods are also used to study various brain structures (Tang et al., 2017; Korolenko et al., 2020; Langley et al., 2021). The focus of these studies is on the SNC and STR, which are significantly altered in PD (Burre et al., 2018; Lai et al., 2021).

Transgenic hemizygous mice of the B6.Cg-Tg(PrNp-SNCA*A53T)23Mkle/J strain (hereinafter referred to as B6.CgTg), which express the A53T mutation in the human alpha-synuclein *SNCA* gene, were first produced at the Jackson Laboratory (USA) (https://www.jax.org/strain/006823). This gene is not expressed in all individuals (Unger et al., 2006), and therefore, among siblings, there may be transgenic as well as wild-type mice. The model replicates the behavioral features and reproduces the symptoms of synucleinopathy, specifically, the age-related neurodegenerative changes characteristic of PD (Pupyshev et al., 2018; Korolenko et al., 2020; Seo et al., 2020; Zhang et al., 2022).

In studies of B6.Cg-Tg mice as PD models, C57BL/6J mice are almost always selected and used as a control group (Pupyshev et al., 2018; Seo et al., 2020). The maternal environment has a significant impact on offspring, mediated through epigenetic processes, both during the gestation and the early postnatal period (Case et al., 2010; Nicholas, Ozanne, 2019; Wu, Dean, 2020). Selection of an appropriate control group is an essential part of the experimental design as it allows to avoid variables that may influence the estimated parameters characterizing PD. In particular, maternal factors may affect prenatal (Wu, Dean, 2020) and early postnatal development, especially during the weaning period (Case et al., 2010). To minimize the maternal influence, it was recommended to compare siblings, especially in the studies involving transgenic and knockout animals (Holmdahl, Malissen, 2012; Chen et al., 2020). Thus, the study conducted on mice of the B6.Cg-Tg strain needs wild-type siblings as controls.

The nigrostriatal pathway of the brain, which has been implicated in the development of Parkinson's disease, has not been sufficiently explored in B6.Cg-Tg mice that model human PD with the A53T mutation in the *SNCA* gene. In our previous study (Rozhkova et al., 2023), we investigated the density of neurons in the prefrontal cortex, hippocampus, SNC, and STR of B6.Cg-Tg mice at an early stage of this pathology, specifically at the age of six months. However, the dopaminergic system, which is crucial for the development of PD, was not examined.

The aim of this work was to compare male mice of the B6.Cg-Tg strain with wild-type (WT) siblings by the following parameters: 1) coordination of movements and body balance; 2) accumulation of alpha-synuclein; and 3) the density of dopamine neurons in the substantia nigra.

## Materials and methods

**Experimental animals.** Male siblings resulting from mating C57BL/6J females (4) with hemizygous B6.Cg-Tg males (4) were used in this experiment. In total, four litters were obtained, and genotyping of the offspring was performed. The experimental group consisted of animals carrying the A53T mutation (B6.Cg-Tg), while the remaining animals served as wild-type (WT) controls. Five hemizygous B6.Cg-Tg males and 10 WT males, which lacked the A53T mutation in the *SNCA* gene, were used.

All animals were housed in the SPF-vivarium at the Institute of Cytology and Genetics SB RAS (Novosibirsk, Russia) in individually ventilated OptiMice cages (Animal Care, USA), measuring 34.3 × 29.2 × 15.5 cm. They were kept at 22–24 °C with a humidity level of 40–50 % and an inverted 12:12-hour day-night cycle (sunrise at 3 a. m.). Fractionated birch chips (TU 16.10.23-001-0084157135-2019) were used as bedding. The animals had free access to standardized chew ("Delta Feeds" LbK 120 R-22, GOST 34566-2019, BioPro, Russia) and purified water ("Severjanka", Ecoproject, Russia).

All experiments were approved by the Bioethics Committee of the Institute of Cytology and Genetics SB RAS (protocol No. 145, March 29, 2023) and complied with the European Convention for the Protection of Vertebrate Animals used for Experimental and Other Scientific Purposes.

**Study of motor coordination and body balance.** At six months of age, five male B6.Cg-Tg mice with the A53T mutation and 10 wild-type controls from the same litters were tested. Motor coordination and body balance were assessed using the rotarod (RR) test. Two days before testing, the animals were isolated in clean individual OptiMice cages (34.3 × 29.2 × 15.5 cm). All equipment was sanitized using a 6 % hydrogen peroxide solution before testing.

The accelerated RR test is widely used to assess movement coordination and balance, particularly in studies of neurodegenerative conditions such as PD (Seo et al., 2020). The Ugo Basile 47600 device (Ugo Basile, Italy) consists of five 5.7 cm-wide tracks with 3 cm drums at a height of 16 cm, separated by flat round parts. The device has dimensions of 27.94 × 43.18 × 38.10 cm, weighs 6.4 kg, and accelerates between 2–80 rpm. The rotarod device was programmed to increase speed linearly from 5 to 40 rpm over 300 seconds. Three sessions were conducted for each mouse, with one-minute breaks between sessions. The time taken for the mouse to fall off the rod was recorded for each run. After

the RR test, the brains of five B6.Cg-Tg mice (confirmed for the A53T mutation) and five wild-type controls from the same litters were studied. Animals were randomly selected from four litters, and brain analysis was conducted following intracardiac perfusion, as described below.

**Intracardiac perfusion.** Brain tissue fixation was performed following the method described by Rozhkova et al. (2023). Mice were anesthetized with intraperitoneal injections of medetomidine hydrochloride (Meditin, 0.01 mg/kg; Api-San, Russia) and zoletil (Zoletil, 50 mg/kg; Virbac, France) 10 minutes later. Perfusion was carried out using 15 ml of phosphate-buffered saline (PBS), followed by 15 mL of 10 % formalin. The brain was removed and placed in a 30 % sucrose solution based on PBS with 5 mL of 10 % formalin for dehydration and fixation. After two weeks being kept at +4 °C, the tissue samples sank to the bottom of the tube. The samples were then embedded in O.C.T Tissue-Tek (Sakura Finetek, USA) and frozen at –70 °C in an MDF-594 horizontal low-temperature freezer (Sanyo, Japan).

**Preparation of frozen brain sections.** Brain sections from the substantia nigra (SNC) were prepared at a distance of –3.08 to –3.52 mm from bregma, following the G. Paxinos and K. Franklin atlas (Paxinos, Franklin, 2012). Frozen sections, 10 μm thick, were prepared using an HM550 OP Cryotome (Thermo Scientific, USA) at –25 °C and placed on PCI-coated adhesive glass slides with polished edges (CITOTEST, China).

**Immunohistochemical analysis.** The samples were stained according to the antibody kit manufacturers' protocols. Before staining, sections were dehydrated and then rehydrated for five minutes in PBS. Heat-induced antigen retrieval was performed using 10 mM alkaline citrate buffer (pH 9) at 95 °C for 15 minutes in a TW-2.02 water bath (Elmi, Latvia). Thereafter sections were removed from the buffer and cooled to room temperature. Samples were washed three times in PBS-Tween buffer: PBS supplemented with 0.1 % Tween-20 (P9416-100 mL; Sigma-Aldrich, USA) for 15 min. Each section was then covered with Protein Block buffer (ab64226; Abcam, UK) for 30 min, followed by the excess liquid removal according to the manufacturer's recommendations.

After washing procedure and exposure to Protein Block buffer, 50 μL of primary antibody were added; the samples were left overnight at 4 °C in a humidified dark chamber. The concentrations of antibodies used were 1:450 for anti-Tyrosine Hydroxylase (TH) – anti-TH (ab6211; Abcam, UK). For alpha-synuclein detection, 50 μL of alpha-Synuclein Antibody primary antibody (NB110-61645, dilution, Novus Biologicals, Littleton, CO, USA) were added at a concentration of 1:600 and left for 36 h at 4 °C in a humidified dark chamber. The sections were then washed in PBS-Tween buffer for 15 min, excess liquid was removed, 50 μL of the secondary antibody Goat Anti-Rabbit IgG H&L AF488 (ab150077; Abcam, UK) were added at a concentration of 1:700. The samples were left in a humidified dark chamber for two hours at 4 °C. Thereafter, the samples were washed in PBS-Tween buffer for 15 min, excess liquid was removed, and the samples were placed in ProLong, Glass Antifade Mountant medium (Thermo P36982; Thermo Fisher Scientific, USA). When assessing alpha-synuclein in neurons, in order to identify neurons, 80 μL of DAPI (Maric et al., 2021) were additionally added to the sections for 15 min and then washed twice with PBS for 3 min. After adding antibodies, the sections were placed in a humidified dark chamber.

**Neuronal density analysis.** The density of antibody-labeled neurons was assessed using an Axio Imager.M2 microscope (Carl Zeiss, Germany) equipped with a Zeiss Axiocam 506 mono camera. Neuronal counts were performed using ImageJ software (National Institutes of Health, USA), and neuronal density was calculated as the number of neurons per cubic millimeter ($mm^3$), as described in Rozhkova et al. (2023).

**Statistical analysis.** Data were analyzed using STATISTICA v. 12.0 software (StatSoft, Inc., USA). Results are presented as medians (Me) with the first (q1) and third (q3) quartiles – Me [Q1;Q3]. Behavioral data (the median of three sessions) and neuron density were compared using the Mann–Whitney U-test. Statistical significance was set at $p < 0.05$.

## Results

Data from the RR test are presented in Figure 1. Statistical analysis using the Mann–Whitney test showed no significant differences in latency between B6.Cg-Tg mice and wild-type controls. Neuronal density data for TH-labeled neurons in the substantia nigra are shown in Figure 2. The analysis revealed a significantly lower density of dopamine neurons in B6.Cg-Tg mice compared to wild-type controls ($p < 0.05$), with $0.92 \times 10^5$ [$0.86 \times 10^5$; $0.93 \times 10^5$] versus $1.25 \times 10^5$ [$1.00 \times 10^5$; $1.26 \times 10^5$] neurons, respectively. Data on the density of neurons labeled for alpha-synuclein in the substantia nigra are shown in Figure 3. Statistical analysis

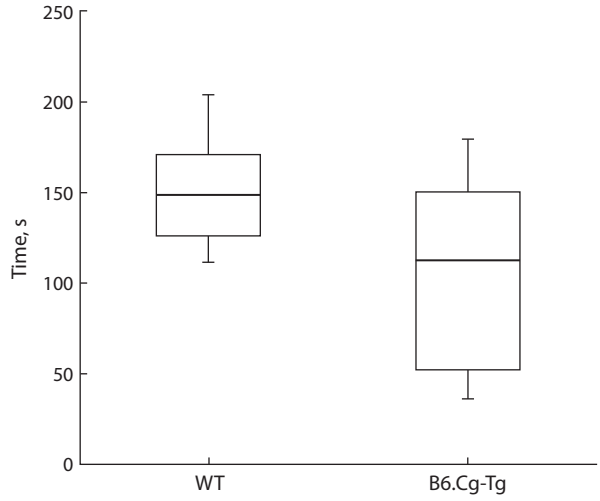



**Fig. 1.** Latency before the drop in the rotarod test for male offspring of B6.Cg-Tg mice and their wild-type siblings at the age of six months.

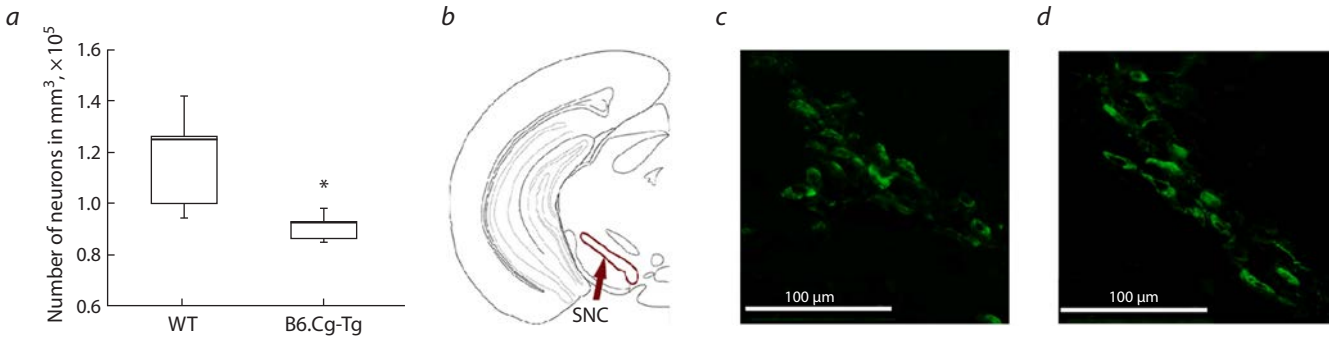



**Fig. 2.** Density of dopaminergic neurons in the substantia nigra (SNC), the neurons were labeled with antibodies against tyrosine hydroxylase – TH.

*a* – number of neurons per mm$^3$; *b* – schematic representation of the region of interest in the brain. Microphotographs of the sections in this region: *c* – wild type (WT); *d* – B6.Cg-Tg. * $p < 0.05$.

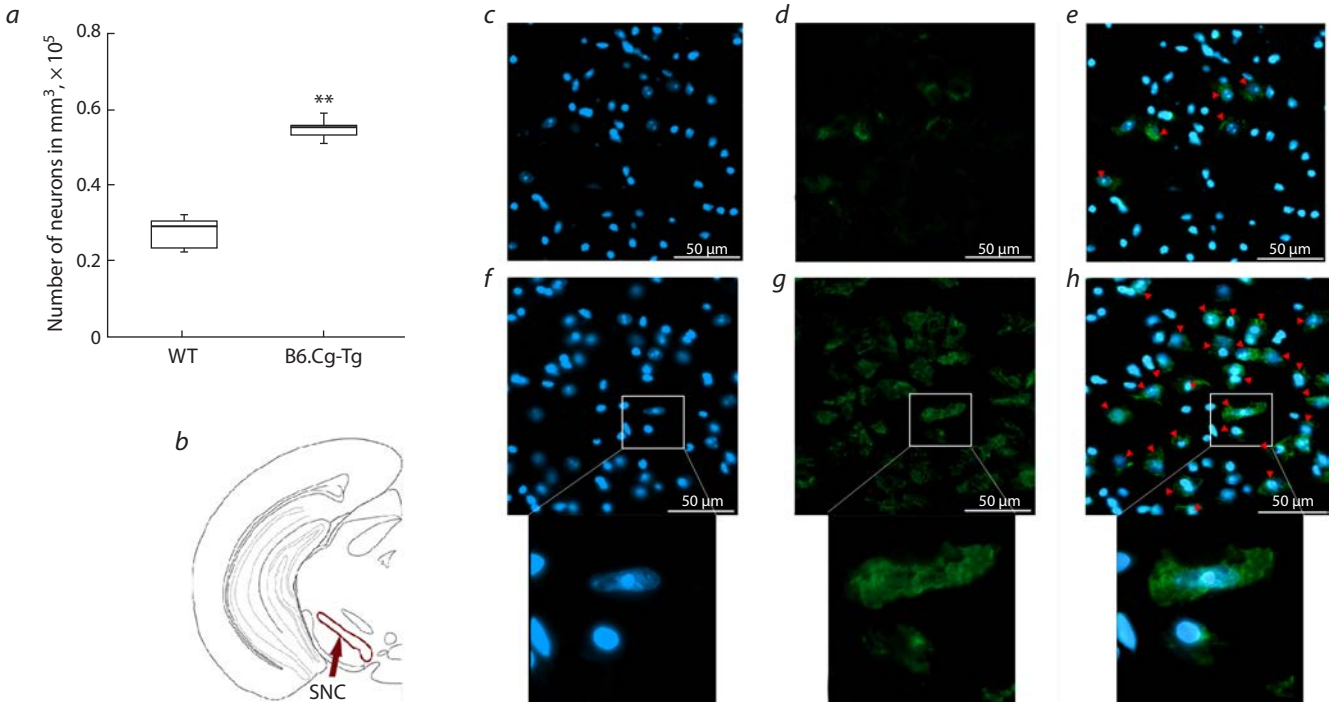



**Fig. 3.** Density of neurons with alpha-synuclein in the substantia nigra (SNC), neurons are labeled with antibodies against alpha-synuclein, neuronal nuclei are stained with DAPI.

*a* – number of neurons per mm$^3$; *b* – schematic representation of the region of interest in the brain. Microphotographs of the sections in this region: *c*–*e* – wild type, *f*–*h* – B6.Cg-Tg; *c*, *f* – DAPI; *d*, *g* – alpha-synuclein; *e*, *h* – merged images. Arrowheads indicate neurons with alpha-synuclein inclusions. Figures in the white boxes represent high-magnification images.
** $p < 0.01$.

indicated a significant increase in alpha-synuclein-positive neurons in B6.Cg-Tg mice compared to wild-type controls ($p < 0.01$), with $0.55 \times 10^5$ [$0.53 \times 10^5$; $0.56 \times 10^5$] neurons versus $0.29 \times 10^5$ [$0.23 \times 10^5$; $0.31 \times 10^5$].

## Discussion

The rotarod test is widely used to evaluate motor activity and function in mice with neurodegenerative disorders (Graham, Sidhu, 2010; Oaks et al., 2013; Seo et al., 2020). In mice with the A53T mutation, motor activity changes typically begin between 2 and 12 months of age. This is linked to the synthesis of alpha-synuclein in brain neurons (Unger et al., 2006; Graham, Sidhu, 2010; Wang et al., 2022). Therefore, behavioral testing during this time is suitable (Zhang et al., 2019). In one study, A53T-a-Syn mice of various ages underwent nine rotarod sessions over three days, and the average results were compared (Oaks et al., 2013). The findings revealed worse motor coordination in A53T-a-Syn mice than in control animals as early as two and four months old (Oaks et al., 2013).

Similar results were seen in two-month-old mice from other transgenic lines carrying the *SNCA* gene and the A53T

mutation, also using the rotarod test (Zhang et al., 2019). However, in older mice, these motor differences leveled out or were even reversed (Graham, Sidhu, 2010). In this study, as well as in the J.H. Seo et al. (2020) report, no differences in coordination and balance were found between six-month-old B6.Cg-Tg mice and wild-type littermates in the accelerated rotarod test. Depending on the Parkinson's disease (PD) model, motor impairments can emerge either earlier (Oaks et al., 2013; Zhang et al., 2019) or later (Graham, Sidhu, 2010; Seo et al., 2020). In our results, early-stage B6.Cg-Tg mice did not show motor deficits, similar to human PD where motor symptoms often arise later in life as the disease progresses (Halliday et al., 2006).

It is well established that Parkinson's disease causes disruptions in the nigrostriatal pathway (Dickson et al., 2009; Beitz, 2014; Hayes, 2019). Damage to these brain structures serves as a key marker of PD in both humans and animal models (Kato, Kimura, 1992; Unger et al., 2006; Beitz, 2014; Taguchi et al., 2020). Our previous research showed a reduction in the number of neurons in the substantia nigra of male B6.Cg-Tg mice (Rozhkova et al., 2023). In this study, we also observed fewer dopaminergic neurons in this region compared to wild-type mice. Similar results were reported in BAC-SNCA$^{A53T/-}$ mice, where a reduction in TH-positive neurons in the substantia nigra was observed (Taguchi et al., 2020). Additionally, A53T-Tg mice showed a reduced response to therapeutic dopamine treatments (Unger et al., 2006). Dopaminergic neurons in the substantia nigra are critical for regulating motor activity (Schultz et al., 1983). Despite this, our six-month-old B6.Cg-Tg mice did not exhibit motor deficits, although these may develop later (Chia et al., 2020).

The current study also found more alpha-synuclein-positive neurons in the substantia nigra of B6.Cg-Tg mice than in WT littermates, which mirrors findings in human PD cases (Dickson, 2018; Bae et al., 2021). This is a common trait in various mouse models of PD (Vander Putten et al., 2000; Taguchi et al., 2020; Wang et al., 2022). A similar result was found in certain transgenic *SNCA* mouse strains with the A53T mutation, where alpha-synuclein oligomers increased in the substantia nigra starting at three months old (Taguchi et al., 2020; Wang et al., 2022). In Pitx3-A53T-a-Syn mice aged six to 18 months, alpha-synuclein accumulation in the substantia nigra increased compared to C57BL controls. This was linked to significant degeneration of parvalbumin-positive neurons (Zheng et al., 2022). The connection between alpha-synuclein inclusions and neuron death is well established (Kalia et al., 2013).

Toxic alpha-synuclein oligomers can disrupt protein expression and endoplasmic reticulum function (Kalia et al., 2013). Mutant alpha-synuclein buildup in brain neurons is likely key to PD symptoms. Both the A53T and A30P mutations cause alpha-synuclein to become less soluble (Grigoryan, Bazyan, 2007). Our findings show that an increase in alpha-synuclein-positive neurons in the substantia nigra of B6.Cg-Tg mice is already evident by six months of age.

## Conclusion

In this study, we characterized six-month-old B6.Cg-Tg (PrNp-SNCA*A53T)23Mkle/J mice for markers important for Parkinson's disease manifestation, such as motor coordination and body balance, as well as the density of dopamine neurons and alpha-synuclein neurons in the substantia nigra. Six-month-old B6.Cg-Tg mice show early signs of Parkinson's disease. These include the accumulation of mutant alpha-synuclein and a reduced number of dopaminergic neurons in the substantia nigra. Our results suggest that these mice can serve as a suitable model for studying the early stages of Parkinson's disease in humans. This model offers valuable insights for future research on disease onset and potential treatments.

## References

Bae Y.J., Kim J.M., Sohn C.H., Choi J.H., Choi B.S., Song Y.S., Nam Y., Cho S.J., Jeon B., Kim J.H. Imaging the substantia nigra in Parkinson disease and other Parkinsonian syndromes. *Radiology*. 2021;300:260-278. DOI 10.1148/radiol.2021203341

Beitz J.M. Parkinson's disease: a review. *Front. Biosci.* (*Schol. Ed.*). 2014;6:65-74. DOI 10.2741/S415

Bidesi N.S., Vang Andersen I., Windhorst A.D., Shalgunov V., Herth M.M. The role of neuroimaging in Parkinson's disease. *J. Neurochem.* 2021;159:660-689. DOI 10.1111/jnc.15516

Burre J., Sharma M., Sudhof T.C. Cell biology and pathophysiology of α-synuclein. *Cold. Spring. Harb. Perspect. Med.* 2018;8:a024091. DOI 10.1101/cshperspect.a024091

Case L.K., Del Rio R., Bonney E.A., Zachary J.F., Blankenhorn E.P., Tung K.S., Teuscher C. The postnatal maternal environment affects autoimmune disease susceptibility in A/J mice. *Cell Immunol.* 2010; 260:119-127. DOI 10.1016/j.cellimm.2009.10.002

Chen D., Liu Y., Shu G., Chen C., Sullivan D.A., Kam W.R., Hann S., Fowler M., Warman M.L. Ocular manifestations of chordin-like 1 knockout mice. *Cornea*. 2020;39:1145-1150. DOI 10.1097/ICO.0000000000002371

Chia S.J., Tan E.K., Chao Y.X. Historical perspective: models of Parkinson's disease. *Int. J. Mol. Sci.* 2020;21:2464. DOI 10.3390/ijms21072464

Crabtree D.M., Zhang J. Genetically engineered mouse models of Parkinson's disease. *Brain Res. Bull.* 2012;88:13-32. DOI 10.1016/j.brainresbull.2011.07.019

Deutch A.Y. Striatal plasticity in parkinsonism: dystrophic changes in medium spiny neurons and progression in Parkinson's disease. *J. Neural Transm. Suppl.* 2006;70:67. DOI 10.1007/978-3-211-45295-0_12

Dickson D.W. Neuropathology of Parkinson disease. *Parkinsonism Relat. Disord.* 2018;46:S30-S33. DOI 10.1016/j.parkreldis.2017.07.033

Dickson D.W., Braak H., Duda J.E., Duyckaerts C., Gasser T., Halliday G.M., Hardy J., Leverenz J.B., Del Tredici K., Wszolek Z.K., Litvan I. Neuropathological assessment of Parkinson's disease: refining the diagnostic criteria. *Lancet Neurol.* 2009;8:1150-1157. DOI 10.1016/S1474-4422(09)70238-8

Graham D.R., Sidhu A. Mice expressing the A53T mutant form of human alpha-synuclein exhibit hyperactivity and reduced anxiety-like behavior. *J. Neurosci. Res.* 2010;88:1777-1183. DOI 10.1002/jnr.22331

Grigoryan G.A., Bazyan A.S. The experimental models of Parkinson's disease in animals. *Uspekhi Fiziologicheskikh Nauk = Progress in Physiological Science.* 2007;38:80-88 (in Russian)

Halliday G.M., Del Tredici K., Braak H. Critical appraisal of brain pathology staging related to presymptomatic and symptomatic cases of sporadic Parkinson's disease. *J. Neural Transm. Suppl.* 2006;70:99. DOI 10.1007/978-3-211-45295-0_16

Hayes M.T. Parkinson's disease and parkinsonism. *Am. J. Med.* 2019; 132:802-807. DOI 10.1016/j.amjmed.2019.03.001

Holmdahl R., Malissen B. The need for littermate controls. *Eur. J. Immunol.* 2012;42:45-47. DOI 10.1002/eji.201142048

Jellinger K.A. Pathology of Parkinson's disease. *Mol. Chem. Neuropath.* 1991;14:153-197. DOI 10.1007/bf03159935

Kalia L.V., Kalia S.K., McLean P.J., Lozano A.M., Lang A.E. α-Synuclein oligomers and clinical implications for Parkinson disease. *Ann. Neurol.* 2013;73:155-169. DOI 10.1002/ana.23746

Kato M., Kimura M. Effects of reversible blockade of basal ganglia on a voluntary arm movement. *J. Neurophysiol.* 1992;68:1516-1534. DOI 10.1152/jn.1992.68.5.1516

Korchounov A., Meyer M.F., Krasnianski M. Postsynaptic nigrostriatal dopamine receptors and their role in movement regulation. *J. Neural Transm.* (*Vienna*). 2010;117:1359-1369. DOI 10.1007/s00702-010-0454-z

Korolenko T.A., Shintyapina A.B., Belichenko V.M., Pupyshev A.B., Akopyan A.A., Fedoseeva L.A., Russkikh G.S., Vavilin V.A., Tenditnik M.V., Lin C-L., Amstislavskaya T.G., Tikhonova M.A. Early Parkinson's disease-like pathology in a transgenic mouse model involves a decreased Cst3 mRNA expression but not neuroinflammatory response in the brain. *Med. Univer.* 2020;3:66-78. DOI 10.2478/medu-2020-0008

Lai T.T., Kim Y.J., Nguyen P.T., Koh Y.H., Nguyen T.T., Ma H.I., Kim Y.E. Temporal evolution of inflammation and neurodegeneration with alpha-synuclein propagation in Parkinson's disease mouse model. *Fron. Int. Neurosci.* 2021;15:715190. DOI 10.3389/fnint.2021.715190

Langley M.R., Ghaisas S., Palanisamy B.N., Ay M., Jin H., Anantharam V., Kanthasamy A., Kanthasamy A.G. Characterization of non-motor behavioral impairments and their neurochemical mechanisms in the MitoPark mouse model of progressive neurodegeneration in Parkinson's disease. *Exp. Neurol.* 2021;341:113716. DOI 10.1016/j.expneurol.2021.113716

Lee F.J., Liu F., Pristupa Z.B., Niznik H.B. Direct binding andfunctional coupling of alpha-synuclein to the dopamine transporters accelerate dopamine-induced apoptosis. *FASEB J.* 2001;15:916-926. DOI 10.1096/fj.00-0334com

Liu Q., Xu Y., Wan W., Ma Z. An unexpected improvement in spatial learning and memory ability in alpha-synuclein A53T transgenic mice. *J. Neural. Transm.* (*Vienna*). 2018;125(2):203-210. DOI 10.1007/s00702-017-1819-3

Maric D., Jahanipour J., Li X.R., Singh A., Mobiny A., Van Nguyen H., Sedlock A., Grama K., Roysam B. Whole-brain tissue mapping toolkit using large-scale highly multiplexed immunofluorescence imaging and deep neural networks. *Nat. Commun.* 2021;12:1550. DOI 10.1038/s41467-021-21735-x

Nicholas L.M., Ozanne S.E. Early life programming in mice by maternal overnutrition: mechanistic insights and interventional approaches. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 2019;374:20180116. DOI 10.1098/rstb.2018.0116

Oaks A.W., Frankfurt M., Finkelstein D.I., Sidhu A. Age-dependent effects of A53T alpha-synuclein on behavior and dopaminergic function. *PLoS One*. 2013;8:e60378. DOI 10.1371/journal.pone.0060378

Paumier K.L., Sukoff Rizzo S.J., Berger Z., Chen Y., Gonzales C., Kaftan E., Li L., Lotarski S., Monaghan M., Shen W., Stolyar P., Vasilyev D., Zaleska M., D Hirst W., Dunlop J. Behavioral characterization of A53T mice reveals early and late stage deficits related to Parkinson's disease. *PLoS One*. 2013;8(8):e70274. DOI 10.1371/journal.pone.0070274.

Paxinos G., Franklin K. Mouse Brain in Stereotaxic Coordinates. 4th ed. San Diego: Acad. Press, 2012

Poewe W., Seppi K., Tanner C.M., Halliday G.M., Brundin P., Volkmann J., Schrag A.E., Lang A.E. Parkinson disease. *Nat. Rev. Dis. Primers*. 2017;3:17013. DOI 10.1038/nrdp.2017.13

Polymeropoulos M.H., Lavedan C., Leroy E., Ide S.E., Dehejia A., Dutra A., Pike B., Root H., Rubenstein J., Boyer R., Stenroos E.S., Chandrasekharappa S., Athanassiadou A., Papapetropoulos T., Johnson W.G., Lazzarini A.M., Duvoisin R.C., Di Iorio G., Golbe L.I., Nussbaum R.L. Mutation in the alpha-synuclein gene identified in families with Parkinson's disease. *Science*. 1997;276:2045-2047. DOI 10.1126/science.276.5321.2045

Pupyshev A.B., Korolenko T.A., Akopyan A.A., Amstislavskaya T.G., Tikhonova M.A. Suppression of autophagy in the brain of transgenic mice with overexpression of A53T-mutant α-synuclein as an early event at synucleinopathy progression. *Neurosci. Lett.* 2018;672: 140-144. DOI 10.1016/j.neulet.2017.12.001

Rozhkova I.N., Okotrub S.V., Brusentsev E.Yu., Rakhmanova T.A., Lebedeva D.A., Kozeneva V.S., Khotskin N.V., Amstislavsky S.Ya. Analysis of behavior and brain neuronal density in B6.Cg-Tg(Prnp-SNCA*A53T)23Mkle/J mice modeling, a Parkinson's disease. *Rossiyskiy Fiziologicheskiy Zhurnal imeni I.M. Sechenova = Russian Journal of Physiology.* 2023;59:1633-1647. DOI 10.31857/S0869813923090091 (in Russian)

Schultz W., Ruffieux A., Aebischer P. The activity of pars compacta neurons of the monkey substantia nigra in relation to motor activation. *Exp. Brain Res.* 1983;51:377-387. DOI 10.1016/0304-3940(84)90456-7

Seo J.H., Kang S.W., Kim K., Wi S., Lee J.W., Cho S.R. Environmental enrichment attenuates oxidative stress and alters detoxifying enzymes in an A53T α-synuclein transgenic mouse model of Parkinson's disease. *Antioxidants* (*Basel*). 2020;9:928. DOI 10.3390/antiox9100928

Spillantini M.G., Schmidt M.L., Lee V.M., Trojanowski J.Q., Jakes R., Goedert M. Alpha-synuclein in Lewy bodies. *Nature*. 1997;388: 839-840. DOI 10.1038/42166

Stern G. The effects of lesions in the substantia nigra. *Brain*. 1966;89: 449-478. DOI 10.1093/brain/89.3.449

Taguchi T., Ikuno M., Hondo M., Parajuli L.K., Taguchi K., Ueda J., Sawamura M., Okuda S., Nakanishi E., Hara J., Uemura N., Hatanaka Y., Ayaki T., Matsuzawa S., Tanaka M., El-Agnaf O.M.A., Koike M., Yanagisawa M., Uemura M.T., Yamakado H., Takahashi R. α-Synuclein BAC transgenic mice exhibit RBD-like behaviour and hyposmia: a prodromal Parkinson's disease model. *Brain*. 2020;143:249-265. DOI 10.1093/brain/awz380

Tang H., Gao Y., Zhang Q., Nie K., Zhu R., Gao L., Feng S., Wang L., Zhao J., Huang Z., Zhang Y., Wang L. Chronic cerebral hypoperfusion independently exacerbates cognitive impairment within the pathopoiesis of Parkinson's disease via microvascular pathologys. *Behav. Brain Res.* 2017;333:286-294. DOI 10.1016/j.bbr.2017.05.061

Tikhonova M.A., Tikhonova N.G., Tenditnik M.V., Ovsyukova M.V., Akopyan A.A., Dubrovina N.I., Amstislavskaya T.G., Khlestkina E.K. Effects of grape polyphenols on the life span and neuroinflammatory alterations related to neurodegenerative Parkinson disease-like disturbances in mice. *Molecules*. 2020;25:5339. DOI 10.3390/molecules25225339

Tran J., Anastacio H., Bardy C. Genetic predispositions of Parkinson's disease revealed in patient-derived brain cells. *NPJ Parkinsons Dis.* 2020;6:8. DOI 10.1038/s41531-020-0110-8

Unger E.L., Eve D.J., Perez X.A., Reichenbach D.K., Xu Y., Lee M.K., Andrews A.M. Locomotor hyperactivity and alterations in dopamine neurotransmission are associated with overexpression of A53T mutant human alpha-synuclein in mice. *Neurobiol. Dis.* 2006;21:431-443. DOI 10.1016/j.nbd.2005.08.005

Van der Putten H., Wiederhold K.H., Probst A., Barbieri S., Mistl C., Danner S., Kauffmann S., Hofele K., Spooren W.P., Ruegg M.A., Lin S., Caroni P., Sommer B., Tolnay M., Bilbe G. Neuropathology in mice expressing human alpha-synuclein. *J. Neurosci.* 2000;20: 6021-6029. DOI 10.1523/JNEUROSCI.20-16-06021.2000

Venda L., Cragg S., Buchman V.L., Wade-Martins R. α-Synuclein and dopamine at the crossroads of Parkinson's disease. *Trends Neurosci*. 2010;12:559-568. DOI 10.1016/j.tins.2010.09.004

Wang Y., Sun Z., Du S., Wei H., Li X., Li X., Shen J., Chen X., Cai Z. The increase of α-synuclein and alterations of dynein in A53T transgenic and aging mouse. *J. Clin. Neurosci.* 2022;96:154-162. DOI 10.1016/j.jocn.2021.11.002

Wu D., Dean J. Maternal factors regulating preimplantation development in mice. *Curr. Top. Dev. Biol.* 2020;140:317-340. DOI 10.1016/bs.ctdb.2019.10.006

Zhang Yu., Wu Q., Zhang L., Wang Q., Yang Z., Liu J., Feng L. Caffeic acid reduces A53T α-synuclein by activating JNK/Bcl-2-mediated autophagy *in vitro* and improves behaviour and protects dopaminergic neurons in a mouse model of Parkinson's disease. *Pharmacol. Res.* 2019;150:104538. DOI 10.1016/j.phrs.2019.104538

Zhang Yu., Wu Q., Ren Y., Zhang Y., Feng L. A53T α-synuclein induces neurogenesis impairment and cognitive dysfunction in line M83 transgenic mice and reduces the proliferation of embryonic neural stem cells. *Brain Res. Bull.* 2022;182:118-129. DOI 10.1016/j.brainresbull.2022.02.010

Zheng M., Liu Y., Xiao Z., Jiao L., Lin X. Tau knockout and α-synuclein A53T synergy modulated parvalbumin-positive neurons degeneration staging in substantia nigra pars reticulata of Parkinson's disease-liked model. *Front. Aging Neurosci.* 2022;13:784665. DOI 10.3389/fnagi.2021.784665

**Conflict of interest.** The authors declare no conflict of interest.
Received February 12, 2024. Revised July 29, 2024. Accepted August 2, 2024.

ANIMAL GENETICS
Original article

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2024;28(7):752-758

DOI 10.18699/vjgb-24-83

# Spatial genetic characterization of the red fox (*Vulpes vulpes*) in the area between the Alps and the Central Dinaric Mountains

A. Devedžić[1], F. Urzi[2], B. Pokorny[3], G. Vengušt[4], D.Ž. Vengušt[4], F. Janžekovič[1], L. Velić[5], T. Eterović[5], B.K. Stroil[6], E. Bužan[2, 3] ✉

[1] Faculty of Natural Sciences and Mathematics, University of Maribor, Maribor, Slovenia
[2] Faculty of Mathematics, Natural Sciences and Information Technologies, University of Primorska, Koper, Slovenia
[3] Faculty of Environmental Protection, Velenje, Slovenia
[4] Veterinary Faculty, University of Ljubljana, Ljubljana, Slovenia
[5] Faculty of Veterinary Medicine, University of Sarajevo, Sarajevo, Bosnia and Herzegovina
[6] Institute of Genetic Engineering and Biotechnology, University of Sarajevo, Sarajevo, Bosnia and Herzegovina
✉ elena.buzan@upr.si

**Abstract.** Red fox, *Vulpes vulpes*, is a globally distributed species characterized by its high adaptability to diverse habitats and a broad range of food resources. This remarkable adaptability has allowed the red fox to thrive in various environments, from urban areas to remote wilderness. In this study, we used a set of microsatellite markers for the comparative genetic analysis of red fox populations from two countries. We included populations from the Eastern Alps and the northern Dinaric Mountains in Slovenia, as well as the Central Dinaric Mountains in Bosnia and Herzegovina. We successfully isolated DNA and genotyped 118 red fox samples. Our analyses, which included Bayesian clustering techniques, revealed a weak genetic differentiation among the studied populations. However, it is noteworthy that statistically significant differences in estimates of genetic differentiation were only apparent when comparing the populations between the two countries. Further spatial genetic clustering analyses provided additional insights, unveiling a differentiation into four genetic clusters. These clusters comprised two distinct groups in Bosnia and Herzegovina and two in Slovenia. This pattern of differentiation suggests that isolation by distance is a key factor influencing the genetic structure of the red fox in this studied region. Additionally, our findings highlighted that populations from the Alps and northern Dinaric Mountains exhibit higher genetic diversity and observed heterozygosity compared to their counterparts in the Central Dinaric Mountains. The genetic diversity is also notable when compared to other European red fox populations. Studying genetic diversity is crucial for the resilience and adaptability of populations, ensuring their survival amid environmental changes and human-induced pressures.
**Key words:** red fox; *Vulpes vulpes*; microsatellites; genetic diversity; Dinaric Mountains.

**For citation:** Devedžić A., Urzi F., Pokorny B., Vengušt G., Vengušt D.Ž., Janžekovič F., Velić L., Eterović T., Stroil B.K., Bužan E. Spatial genetic characterization of the red fox (*Vulpes vulpes*) in the area between the Alps and the Central Dinaric Mountains. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2024;28(7):752-758. DOI 10.18699/vjgb-24-83

**Funding.** This study was funded by: (i) the Slovenian Research and Innovation Agency (ARIS; programme groups P1–03861, P1–0403); (ii) the STEPCHANGE European Union's Horizon 2020 Research and Innovation Program under grant agreement No. 101006386.

**Acknowledgements.** We thank all the collaborators and founders who help accelerate collaboration with hunters as citizen scientists in wildlife monitoring and biodiversity research. We thank many hunters, foresters and other colleagues for collecting and providing the samples.

**Authors' contribution.** Conceptualization: E.B.; sample providing: B.P., D.Ž.V., G.V., L.V., T.E., B.K.S.; laboratory and statistical analyses: A.D., F.U., E.B.; writing – draft preparation: A.D., E.B., F.U.; writing – review and editing: A.D., E.B., B.P., F.U., F.J., B.K.S; funding: E.B. All authors have read and agreed to the published version of the manuscript.

# Пространственная генетическая характеристика красной лисицы (*Vulpes vulpes*) в районе между Альпами и центральными Динарскими горами

А. Деведжич[1], Ф. Урзи[2], Б. Покорны[3], Г. Венгушт[4], Д.Ж. Венгушт[4], Ф. Янжекович[1], Л. Велич[5], Т. Этерович[5], Б.К. Стройл[6], Э. Бужан[2, 3] ✉

[1] Факультет естественных наук и математики, Университет Марибора, Марибор, Словения
[2] Факультет математики, естественных наук и информационных технологий, Приморский университет, Копер, Словения
[3] Факультет охраны окружающей среды, Веленье, Словения
[4] Ветеринарный факультет, Университет Любляны, Любляна, Словения
[5] Факультет ветеринарной медицины, Университет Сараево, Сараево, Босния и Герцеговина
[6] Институт генной инженерии и биотехнологии, Университет Сараево, Сараево, Босния и Герцеговина
✉ elena.buzan@upr.si

**Аннотация.** Красная лисица, *Vulpes vulpes*, – широко распространенный вид, характеризующийся высокой адаптивностью к различным средам обитания и широким спектром пищевых ресурсов. Удивительная адаптивность позволила красной лисице выживать в самых разнообразных условиях, от городских территорий до отдаленных диких мест. С использованием набора микросателлитных маркеров нами проведен сравнительный генетический анализ популяций красной лисицы из двух стран. В исследование включены популяции из Восточных Альп и северных Динарских гор в Словении, а также из центральных Динарских гор в Боснии и Герцеговине. Мы успешно выделили ДНК и генотипировали 118 образцов красной лисицы. Анализы, включавшие байесовские методы кластеризации, показали слабую генетическую дифференциацию между исследуемыми популяциями. Однако надо отметить, что статистически значимые различия в оценках генетической дифференциации были очевидны лишь при сравнении популяций между двумя странами. Дополнительные пространственные генетические анализы выявили разделение на четыре генетических кластера, которые включали две отдельные группы в Боснии и Герцеговине и две – в Словении. Такая картина дифференциации предполагает, что ключевым фактором, влияющим на генетическую структуру красной лисицы в данном регионе, является изоляция по расстоянию. Кроме того, наши результаты показали, что популяции из Альп и северных Динарских гор обладают более высокими генетическим разнообразием и наблюдаемой гетерозиготностью по сравнению с популяциями из центральных Динарских гор. Генетическое разнообразие заметно также при сравнении с другими европейскими популяциями красной лисицы. Изучение генетического разнообразия имеет важное значение для устойчивости и адаптируемости популяций, обеспечивая их выживание в условиях экологических изменений и антропогенного давления.

**Ключевые слова:** красная лисица; *Vulpes vulpes*; микросателлиты; генетическое разнообразие; Динарские горы.

## Introduction

Red fox (*Vulpes vulpes*) is recognized as one of the most widespread terrestrial predators in Europe, potentially ranking among the most prevalent mammals (Hoffmann et al., 2004). Its extensive distribution across diverse European environments exposes the species to a spectrum of environmental and climatic conditions, contributing to variations in life strategies and fitness among populations (Nowak, 1999). Red fox exhibits adaptability across various habitats, including forests, tundra, prairies, deserts, mountains, agricultural lands, and urban areas. It is often regarded as a pest due to its perceived detrimental effects on prey populations and its capacity to transmit various diseases. In many European countries, it is considered an important game species (Atterby et al., 2015).

The species is native both in Slovenia and Bosnia and Herzegovina (B&H). In Slovenia, red fox is the most common and predominant wild carnivore, and populations are widespread from the Adriatic coast to the Prekmurje region. In the period 2014–2023, between 10,055 (in 2014) and 14,707 (in 2021) individuals were hunted annually, and reported roadkill in the same period was in the range from 686 (in 2023) to 1,177 (in 2019) individuals, respectively. Moreover, every year between approx. 200 and 300 red foxes were found dead either due to diseases (mainly sarcoptic mange) or unknown reasons (OSLIS, 2024). In B&H, red fox occupies a variety of habitats with an estimated density of around 0.5 individuals per km$^2$. Considering hunting data, it is estimated that about 24,000 foxes live in B&H, distributed over the entire country (Nemet, 2018).

Red fox populations exhibit pronounced lack of mitochondrial DNA (mtDNA) genetic structuring on a wide spatiotemporal scale (Frati et al., 1998; Teacher et al., 2011; Edwards et al., 2012; Kutschera et al., 2013; McDevitt et al., 2022). The lack of phylogeographic structuring in red fox, based on mtDNA markers, has been attributed to its persistence outside the traditional refugia areas during the last glacial maximum (LGM), but several lines of evidence suggest survival of red fox phylogenetic lineages in southern refugia (Kutschera et al., 2013). Southern European regions generally have less species connectivity than northern ones, but populations still show low differentiation (Frati et al., 1998; Kirschning et al., 2007; Gachot-Neveu et al., 2009; Edwards et al., 2012). Other studies based on nuclear DNA markers confirmed low genetic diversity of red foxes in Europe, including Poland (Mullins et al., 2014), United Kingdom (Atterby et al., 2015), and Scandinavia (Norén et al., 2015). These studies also revealed significant genetic structure at smaller geographic scales, suggesting the formation of distinct subpopulations within regions.

Red fox genetic diversity, structure, and gene flow between populations are influenced by internal factors such as vagility and dispersion, and external factors like landscape and environment. Landscape features, such as rivers or mountain ranges, may restrict gene flow and impact genetic diversity of subpopulations (Manel et al., 2003; Kirschning et al., 2007; Valvo, 2011; Sommer S. et al., 2013; Galov et al., 2014; Balkenhol et al., 2015). Historical and current factors, including past population bottlenecks, range expansions, landscape features, and habitat fragmentation, collectively influence red fox population genetic diversity and structure (Frati et al., 1998; Teacher et al., 2011; Edwards et al., 2012; Kutschera et al., 2013; Statham et al., 2014; McDevitt et al., 2022).

Two geographically wider and comparative studies (Zecchin et al., 2019; McDevitt et al., 2022) have already studied the genetic structure of red foxes in Slovenia. Zecchin et al. (2019) revealed that red foxes in Slovenia belong to a unified group together with the Croatian population, and that there is no spatially and temporally extensive phylogeographic structure of the species within Europe. Similarly, McDevitt et al. (2022) showed that the Slovenian red fox population belongs to the "Central Europe" cluster together with populations from Croatia and Serbia. Both studies used nuclear markers (microsatellites and single nucleotide polymorphisms, respectively) and underlined the importance of genomic data in identification of refugia regions and post-glacial expansions across Europe (e. g. the Balkans), while at the same time providing necessary insights on red fox genetic diversity and structure.

Red fox from Croatia was individually studied by Galov et al. (2014). They found a high degree of genetic diversity among individuals, indicating a wide range of genetic varia-

tion within the population. Despite mitochondrial haplotype diversity, which is among the highest of all European red fox populations, the study showed a remarkable lack of population structure, indicating extensive gene flow and interbreeding of red foxes in the country. The study by Kirschning et al. (2007) included red foxes from Serbia, and different mitochondrial haplotypes were found, demonstrating genetic structuring within the population. However, as with the neighbouring population from Croatia (Galov et al., 2014), a relatively low degree of genetic differentiation was found on nuclear genetic markers between different geographical regions within Serbia.

To date, there has been no research on the genetic structure of red fox in Bosnia and Herzegovina. Therefore, the primary aim of our study was to conduct an analysis of the genetic variability and structure of red fox populations in B&H in comparison with the Slovenian population. More broadly, we also attempted to establish conclusions regarding the connections with neighbouring populations from Croatia and Serbia.

## Materials and methods

**Study area and sampling.** The study area included territories of B&H and Slovenia (Supplementary Material 1[1]), covering an area of around 66,189 km$^2$ (Nemet, 2018; Hunting Association of Slovenia, 2021).

Animals used in the study were either legally harvested during the hunting season or collected as roadkill or natural death. No animal was shot or otherwise killed for the purposes of this study solely. Samples of harvested animals were collected by hunters or wildlife researchers immediately after harvest between 2019 and 2021. Tissue samples were preserved in 70 % ethanol and blood samples were stored at –20 °C until analysis. In line with our objectives, samples included in the analysis were collected in Slovenia ($n$ = 59) and B&H ($n$ = 59). The Faculty of Veterinary Medicine, University of Sarajevo and the Veterinary Faculty, University of Ljubljana, which oversee the rabies surveillance initiative in both countries, sent only negative samples for genetic analysis at the Molecular Ecology Laboratory at the University of Primorska.

**DNA extraction and quality control.** The extraction of DNA from collected samples was performed using the peqGOLD Blood & Tissue DNA Mini Kit (VWR International, LLC, Austria), following the manufacturer's instructions. The concentration and purity of DNA obtained were measured with a 3.0 Qubit Fluorometer using Invitrogen™ – Qubit™ dsDNA BR Assay Kit (Life Technologies, Carlsbad, CA, USA). Nineteen microsatellites (Supplementary Material 2), originally identified and screened in canine genome studies in red foxes and domestic dogs (Richman et al., 2001; Kukekova et al., 2004, 2007), were amplified in three multiplex PCR reactions with DNA Thermo Cycler (Applied Biosystems). Amplification was carried out with ready-to-use KAPA2G Fast Multiplex Mix (Kapa Biosystems) in 15 µl of the reaction mixture containing 5 µl of template DNA (~25 ng DNA), and 0.2 mM final concentration for each primer used in the set. The amplification was performed under the following conditions: initial denaturation step at 95 °C for 3 minutes, followed by 30 cycles of denaturation for 35 seconds, annealing at 58 °C for 35 seconds, extension at 72 °C for 35 seconds; a final extension step at 72 °C for 10 minutes.

The fragment analysis was performed on a SeqStudio sequencer (Thermo Fisher Scientific) using the GeneScan LIZ500 (−250) standard (Applied Biosystems). The results were validated with the software GENEMAPPER v.5.0 (Applied Biosystems). Null alleles and heterozygosity deficiency were assessed using FreeNA (Chapuis, Estoup, 2007), a program based on Dempster's algorithm, to avoid Hardy–Weinberg equilibrium (HWE) deviation. GENEPOP 4.7 software (Rousset, 2008) was utilised to perform the exact test for heterozygosity deficiency, calculating the deviation from HWE and inbreeding coefficient ($F_{IS}$) estimates. Statistical significance was set at $p < 0.05$.

Genetic diversity parameters, including the number and richness of alleles, were computed using GENETIX 4.05.2 (Belkhir et al., 2004), FSTAT 2.9.4 (Goudet, 1995), and GENEPOP 4.7 (Rousset, 2008). Pairwise fixation index ($F_{ST}$) between populations was assessed using the Weir and Cockerham algorithm (1984) in GENEPOP 4.7, employing 1,000 permutations for statistical rigour.

STRUCTURE 2.3.4 (Falush et al., 2003) was employed to determine population structure, assessing the probability of genetic interbreeding between individuals. The model considered the unique origin of each ancestor for a specific allele, providing the probability (Q) that each subject belongs to a particular cluster or group. For STRUCTURE, 10 independent cycles for each K (number of clusters) were conducted between 1 and 10, using a Markov chain model with 1,000,000 Markov Chain Model Monte Carlo (MCMC) iterations and 100,000 burn-in iterations for each cycle. We applied a mixed model featuring independent allele frequencies. Depending on the specific populations under investigation, we opted for the independent allele frequencies model, if the allele frequencies among distinct populations differ accordingly. The program STRUCTURE Harvester v0.6.94 (Earl, Vonholdt, 2014) was used to combine results and determine the most optimal K based on ΔK developed by Evanno et al. (2005), with results recorded using CLUMPP (Jakobsson, Rosenberg, 2007) and DISTRUCT (Rosenberg, 2004).

Discriminant analysis of principal components (DAPC) was employed to identify genetic clusters within the dataset using the Adegenet package (Jombart, 2008) in R 3.5.1 software (R Development Core Team, 2011). DAPC, a multivariate statistical method, reduces genetic data dimensionality through principal component analysis (PCA) and uses discriminant analysis to identify genetic clusters or subpopulations.

For spatial population structure analysis, a visual user interface (GUI) was developed in the R environment using Geneland 4.9.2 (Guillot et al., 2005). Geneland, a Bayesian method, employs MCMC simulations to estimate the number of genetic populations (K) and assigns individuals to populations based on genetic similarity, considering spatial autocorrelation and isolation by distance.

Arlequin ver. 3.5.2.2 (Excoffier, Lischer, 2010) was used for the analysis of molecular variance (AMOVA), providing techniques to test the genetic differences between individuals and populations and between the optimal number of clusters identified by STRUCTURE (K = 2). Finally, we tested isolation by distance (IBD) patterns within all genetic populations

[1] Supplementary Materials 1–6 are available at:
https://vavilovj-icg.ru/download/pict-2024-28/appx25.pdf

through the Mantel test, comparing genetic and geographical distances. Euclidean distances were calculated using R 3.5.1 software (R Development Core Team, 2011) and the Adegenet library (Jombart, 2008). Significance was determined through 999 Monte Carlo simulations, evaluating the correlation between Edwards distances and Euclidean geographic distances.

## Results

We visualised fox genotypes with GeneMapper v.5.0 software. Six monomorphic markers and six markers with high null allele frequencies (>20 %) were excluded. Monomorphic markers lack variation in DNA sequence length, while null alleles result in non-amplification of expected alleles in PCR tests. We than proceeded with statistical analyses on the remaining seven markers (V402, Vv-C01.424, VVM189, Vv-REN169O18, FH2541, Vv-INU055, and Vv-C08.618). The purpose of these exclusions is to ensure the accuracy of the results and prevent any unfair impact on the genetic variability parameters of the populations under study.

### Genetic diversity within B&H and Slovenian red fox populations

Our findings revealed lower genetic diversity in the B&H population, i. e. evident in lower observed heterozygosity and allelic richness ($H_O$ = 0.648, $AR$ = 7.538) compared to the Slovenian population ($H_O$ = 0.770, $AR$ = 9.724). The genetic diversity in B&H was also lower than genetic diversity of red fox across Austria, Croatia, Italy, and Slovenia ($H_O$ = 0.75, $AR$ = 12.6) as reported by Zecchin et al. (2019). However, this difference could be a result of a higher number of microsatellites used in Zecchin et al. study, which is three times higher than in the present one. Conversely, compared to the analysed Slovenian population, the observed heterozygosity was lower but allelic richness was higher in this regional study of red foxes. The observed heterozygosity in B&H foxes was lower than observed heterozygosity of foxes from Poland ($H_O$ = 0.729, $AR$ = 4.871), even though they used only microsatellite markers (Mullins et al., 2014). In contrast, the Slovenian population displayed higher $H_O$ and $AR$ values compared to Polish population. Finally, both B&H and Slovenian populations had greater observed heterozygosity and allelic richness than red foxes from the United Kingdom (Atterby et al., 2015; 11 microsatellite loci: $H_O$ = 0.543, $AR$ = 5.033).

### Genetic differentiation between B&H and Slovenian red foxes

Spatial genetic structure and genetic clustering depend on the software and analysis employed. For instance, STRUCTURE (Fig. 1) and Geneland identified two clusters (K = 2): one includes red fox from B&H and the other one samples from Slovenia (Supplementary Materials 3 & 4). On the contrary, DAPC analysis suggested four clusters (K = 4) (Fig. 2), dividing each countries' populations into two additional clusters. These discrepancies stem from different methodologies and assumptions inherent to each approach in population genetics analysis.

STRUCTURE estimates ancestry proportions and admixture based on genotypes, relying on assumptions like linkage equilibrium and HWE. DAPC employs dimensionality reduction and discriminant analysis for membership assignment, emphasising differences between clusters. Geneland combines spatial and genetic clustering. All three methods confirmed spatial differentiation at the regional level, with a weaker structure or admixture identified by STRUCTURE and DAPC, possibly due to long-distance dispersal of red foxes. The assumption of HWE was violated in STRUCTURE and Geneland analyses, affecting quantitative estimates. However, due to geographic or ecological constraints, there may be subpopulations with limited gene flow, which geographic information system (GIS) image (Fig. 3) clearly revealed by combining our microsatellite data and sampling coordinates in both populations.


**Fig. 1.** Genetic structure of red fox populations in Slovenia (1) and B&H (2), determined by STRUCTURE.

Each individual is represented by a line proportionally partitioned into colour segments corresponding to its membership in particular cluster. K is the number of clusters.

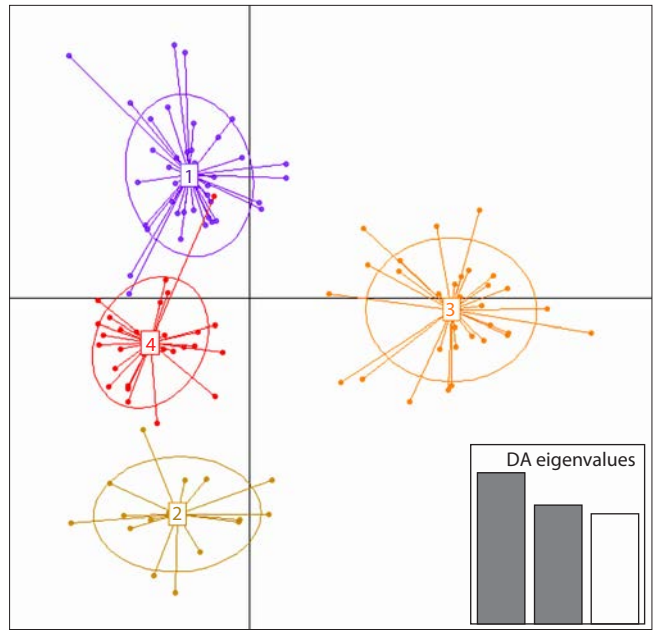



**Fig. 2.** The DAPC scatter plot shows the presence of four clusters (K = 4).

Cluster 1 includes red fox samples from Slovenia, while the Cluster 3 includes red fox samples from B&H. In the Clusters 2 and 4 we can find samples from both countries (Supplementary Material 5).

DAPC revealed that red foxes from Slovenia form a distinct cluster (Cluster 1), while Clusters 2, 3, and 4 are shared between both countries. Cluster 3, predominant in B&H individuals, displayed the most genetic variation. Cluster 2 was genetically closer to B&H, and Cluster 4 showed high similarity to Cluster 1, suggesting potential admixture between B&H and Slovenian red fox populations. Historical gene flow from Slovenia to the Balkans was supported by previous studies, aligning with post glacial colonisation patterns (Sommer R.S., Nadachowski, 2006).

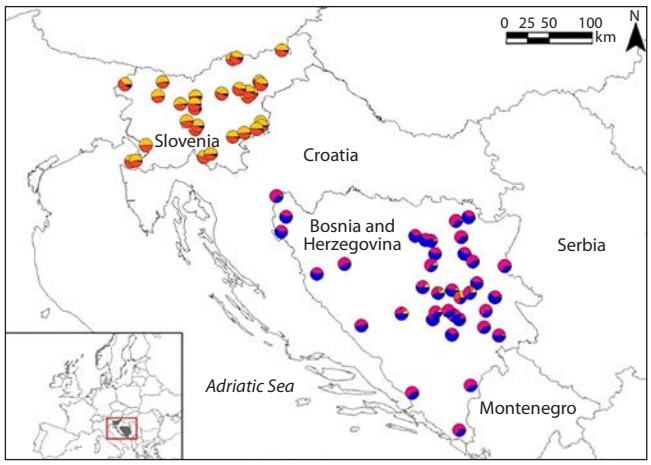



**Fig. 3.** Genetic structure of red fox in two studied countries based on the CLUMPP Q values and harvesting locations (K = 4 clusters).

Q values represent the proportion of individuals that are assigned to each cluster shown by GIS mapping (Supplementary Material 6).

AMOVA indicated genetic differentiation among populations and substantial within-individual variation. The $F_{ST}$ value of 0.068 suggested significant but not highly differentiated populations, with 6.9 % of genetic variation attributed to differences between populations ($p < 0.001$).

IBD analysis demonstrated a positive correlation ($p < 0.001$, $R^2 = 0.13$; this correlation was higher than the randomly simulated $p$-value of 0.001) between genetic and geographic distances for B&H and Slovenian red fox populations, indicating differentiation due to limited dispersal, landscape barriers, habitat fragmentation, resource distribution, interactions with other species or human activities. However, we acknowledge the influence of factors like distance between B&H and Slovenia (almost 300 km). Understanding how habitat changes affect mobile species like red fox, especially in terms of genetic structure, remains an on-going challenge. Landscape history, barriers to movement, and migration rates can leave lasting genetic patterns and impact current population structures. Overall, our findings emphasise the importance of considering diverse analytical approaches and acknowledging the influence of geographical factors on genetic structure in red fox populations.

## Discussion

We successfully revealed genetic structuring of the red fox populations between Alps and the Central Dinaric Mts. and showed weak genetic differentiation of the species in the studied area. Our findings are consistent with previous ones, which revealed the presence of a genetically unified red fox population in Slovenia (Zecchin et al., 2019; McDevitt et al., 2022). The DAPC analysis divided the Slovenian red foxes into two clusters, indicating the existence of possible subpopulations and admixture between the North Dinaric Mts. and the Prekmurje region, but this could also be a consequence of isolation by distance. Red foxes from B&H also belong to one STRUCTURE cluster, but the DAPC analysis divided them into two genetic clusters: first including the area of the Central Dinaric Mts. and second in the northern parts of the country. The observed genetic structure by DAPC analysis can be attributed to various factors. The territory of B&H contains significant natural and anthropogenic barriers, such as mountain ranges and rivers, potentially leading to population isolation and genetic drift effects. The Dinaric Mts. along the western border with Croatia as well as internal mountain ranges may limit gene flow between Croatia and B&H. Galov et al. (2014) demonstrated the limiting effects of Istrian narrow land bridge and mountain altitudes above 1,000 m on fox migration in Croatia, which can also be the case for areas in B&H and Slovenia with many topographical similarities in studied habitats.

Although our results suggest limited gene flow between Slovenian and B&H red fox populations, the absence of Croatian samples hinders relevant conclusions. However, Zecchin et al. (2019) found a unified genetic structure of the fox in Croatia and Slovenia. Despite this, we can assume that there are some possible barriers to gene flow among foxes in the studied regions such as rivers like Sava, a tributary of the Danube, which forms the northern border between B&H and Croatia, and the Drina, which flows north and forms part of the eastern border between B&H and Serbia. The influence of rivers on red fox movement was also discussed in the population genetic analysis of Serbian red foxes by Kirschning et al. (2007). The study highlighted the role of the two major rivers, the Danube and the Tisza, which flow through Serbia, as potential barriers to fox migration. There are also anthropogenic barriers that lead to genetic isolation of foxes in B&H, such as the highway that crosses Croatia and runs through the north-western part of the border between B&H and Croatia.

The existence of two genetic groups of red fox between Alps and northern Dinaric Mts. and other in the Central Dinaric Mts. was previously showed by Zecchin et al. (2019), who identified separated clusters in Friuli Venezia Giulia region as a border area in which circulating individuals are genetically more like those from Slovenia and Croatia than to those of the remaining areas of north-eastern Italy. As a support to our results, a phylogenetic analysis of the mitochondrial cytochrome b and D-loop by Statham et al. (2014) identified a clear differentiation between Italian red foxes and the red fox populations from the Balkans and Eastern Europe. Our results indicated that the B&H population belongs to the Balkan cluster according to mtDNA but additional analysis is needed to confirm this assumption.

The limited number of genetic markers used in our study likely influenced the Bayesian genetic structure analysis, resulting in an inability to distinctly separate subgroups by STRUCTURE and Geneland analysis (Falush et al., 2003). This limitation becomes more evident when compared to the clearer subgroup differentiation observed in the DAPC analysis. On the other hand, the landscape barriers such as rivers and lower mountains might not restrict gene flow among foxes in Slovenia, Croatia, and B&H. Indeed, in the studied area there is adequate habitat connectivity for red fox, therefore it seems that landscape characteristics do not cause important barriers to gene flow, resulting in lack of population differentiation in this species (Kirschning et al., 2007). This can also be confirmed by weak phylogeographic structuring of red fox on much larger geographic scales, such as within the Holarctic lineage (Kutschera et al., 2013).

Despite Galov et al. (2014) identified significant mtDNA genetic structure of red fox on the Istrian Peninsula, which borders coastal Slovenia, our data did not reveal similar genetic differentiation in the Slovenian population. This could likely be due to stable connectivity with neighbouring Italian population and the rest of the Slovenian population. However, it might also be due to variety of markers used in both studies.

Red foxes in Bosnia and Herzegovina have lower genetic diversity comparing to Slovenian and other European populations, which could be a consequence of a recent population decline due to rabies epidemics, compared to Slovenia and Croatia, where oral vaccination programmes (ORV) of foxes led to the control of the rabies epidemic (Rabies Bulletin Europe, 2017). Indeed, since the implementation of the ORV campaign in Slovenia (in 1988) and Croatia (in 2011), the number of rabies cases have been decreasing consistently with no cases detected since 2015 (Bedeković et al., 2016; Picot et al., 2017). The last rabies case in Slovenia was recorded in 2013, whereas in B&H was present until 2020. There, the lack of economic resources has led to a lack of continuous ORV, resulting in continued rabies hotspots in B&H (Tasioudi et al., 2014; Lojkić et al., 2021). Non-suitable disease management may lead to decrease in population size and influence the genetic diversity in the B&H red fox population, both in terms of allelic richness and observed heterozygosity compared to Slovenian population.

The statistically significant *p*-values for the $F_{IS}$ estimates in Slovenia and B&H red fox population ($p = 0.37$ and $p = 0.13$, respectively) can be associated with historical bottlenecks due to diseases but also with possible regional isolation of populations (which was not revealed in our study, possible due to limited number of markers), leading to lower observed heterozygosity.

The results of AMOVA suggested existing genetic differences among populations. High genetic variation among individuals within sampled populations suggests substantial gene flow between them. However, the high level of genetic diversity within populations may suggest that maintaining habitat connectivity is an important factor in promoting gene flow and maintaining genetic variation within populations. These values demonstrate statistically significant differentiation between observed locations; however, it is crucial to acknowledge that the interpretation can vary based on factors like the particular genetic markers employed, the sampling strategy adopted, and the genetic structure inherent in the populations.

The significant IBD values among populations indicate a pattern of isolation by distance, attributed to missing data from neighbouring populations in Croatia. Assuming that the distance between the closest locations in the two countries exceeds several hundred kilometres, the observed genetic divergence is probably consequence of an isolation by distance. Considering also the DAPC analysis, which also confirmed shared genetic makeup between populations despite the distance, as has also been demonstrated in a study from Poland (Mullins et al., 2014).

Our findings are consistent with those of Teacher et al. (2011), who reported a small degree of isolation by distance, in mitochondrial control region, and a wide-scale absence of phylogeographic structure based on cytochrome b data, in red foxes in Western Europe. The species' versatility and capacity to adapt to a broad range of habitats, along with a comparatively high degree of dispersal in both males and females, are suggested to be the main reasons for the low level of genetic structure observed (Teacher et al., 2011).

## Conclusion

This study enhanced our understanding of the red fox genetic structure in the region encompassing the Eastern Alps and the Dinaric Mountains. In alignment with previous research, we found the absence of a pronounced genetic structure within the red fox population in this area. Populations from Slovenia and B&H showed low level of genetic differentiation, with significant differences found only when comparing populations in countries.

## References

Atterby H., Allnutt T.R., MacNicoll A.D., Jones E.P., Smith G.C. Population genetic structure of the red fox (*Vulpes vulpes*) in the UK. *Mamm. Res*. 2015;60:9-19. DOI 10.1007/s13364-014-0209-6

Balkenhol N., Cushman S.A., Storfer A., Waits L.P. Introduction to landscape genetics – concepts, methods, applications. In: Landscape Genetics. Ch. 1. John Wiley & Sons Ltd., 2015;1-8. DOI 10.1002/9781118525258.ch01

Bedeković T., Šimić I., Krešić N., Lojkić I., Mihaljević Ž., Sučec I., Janković I.L., Hostnik P. Evaluation of ELISA for the detection of rabies virus antibodies from the thoracic liquid and muscle extract samples in the monitoring of fox oral vaccination campaigns. *BMC Vet. Res*. 2016;12(1):76. DOI 10.1186/s12917-016-0701-0

Belkhir K., Borsa P., Chikhi L., Raufaste N., Bonhomme F. GENETIX 4.05. Population Genetics Software for Windows TM. Univ. of Montpellier, 2004

Chapuis M.P., Estoup A. Microsatellite null alleles and estimation of population differentiation. *Mol. Biol. Evol*. 2007;24(3):621-631. DOI 10.1093/molbev/msl191

Earl D.A., Vonholdt B.M. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour*. 2014;4:359-361. DOI 10.1007/s12686-011-9548-7

Edwards C.J., Soulsbury C.D., Statham M.J., Ho S.Y., Wall D., Dolf G., Iossa G., Baker P.J., Harris S., Sacks B.N., Bradley D.G. Temporal genetic variation of the red fox, *Vulpes vulpes*, across western Europe and the British Isles. *Quat. Sci. Rev.* 2012;57:95-104. DOI 10.1016/j.quascirev.2012.10.010

Evanno G., Regnaut S., Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 2005;14:2611-2620

Excoffier L., Lischer H.E.L. 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

Falush D., Stephens M., Pritchard J.K. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*. 2003;164(4):1567-1587. DOI 10.1093/genetics/164.4.1567

Frati F., Hartl G.B., Lovari S., Delibes M., Markov G. Quaternary radiation and genetic structure of the red fox *Vulpes vulpes* in the Mediterranean Basin, as revealed by allozymes and mitochondrial DNA. *J. Zool*. 1998;245(1):43-51. DOI 10.1111/j.1469-7998.1998.tb00070.x

Gachot-Neveu H., Lefevre P., Roeder J.J., Henry C., Poulle M.L. Genetic detection of sex-biased and age-biased dispersal in a population of wild carnivore, the red fox, *Vulpes vulpes*. *Zoolog. Sci*. 2009;26(2):145-152. DOI 10.2108/zsj.26.145

Galov A., Sindičić M., Andreanszky T., Čurković S., Dežđek D., Slavica A., Hartl G.B., Krueger B. High genetic diversity and low po-

pulation structure in red foxes (*Vulpes vulpes*) from Croatia. *Mamm. Biol*. 2014;79(1):77-80. DOI 10.1016/j.mambio.2013.10.003

Goudet J. FSTAT (version 1.2): a computer program to calculate F-statistics. *J. Hered*. 1995;86(6):485-486. DOI 10.1093/oxfordjournals.jhered.a111627

Guillot G., Mortier F., Estoup A. GENELAND: a computer package for landscape genetics. *Mol. Ecol. Notes*. 2005;5(3):712-715. DOI 10.1111/j.1471-8286.2005.01031.x

Hoffmann M., Macdonald D.W., Sillero-Zubiri C. Canids: Foxes, Wolves, Jackals and Dogs. Status Survey and Conservation Action Plan. Gland: IUCN, 2004

Hunting Association of Slovenia. Information centre data. Available online: https://www.lovska-zveza.si/wp-content/uploads/2022/05/Letno-porocilo-LZS-2021-za-splet.pdf. Accessed on 5 January 2024

Jakobsson M., Rosenberg N.A. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*. 2007;23(14):1801-1806. DOI 10.1093/bioinformatics/btm233

Jombart T. *adegenet*: a R package for the multivariate analysis of genetic markers. *Bioinformatics*. 2008;24(11):1403-1405. DOI 10.1093/bioinformatics/btn129

Kirschning J., Zachos F.E., Cirovic D., Radovic I.T., Hmwe S.S., Hartl G.B. Population genetic analysis of Serbian red foxes (*Vulpes vulpes*) by means of mitochondrial control region sequences. *Biochem. Genet*. 2007;45(5-6):409-420. DOI 10.1007/s10528-007-9082-1

Kukekova A.V., Trut L.N., Oskina I.N., Kharlamova A.V., Shikhevich S.G., Kirkness E.F., Aguirre G.D., Acland G.M. A marker set for construction of a genetic map of the silver fox (*Vulpes vulpes*). *J. Hered*. 2004;95(3):185-194. DOI 10.1093/jhered/esh033

Kukekova A.V., Trut L.N., Oskina I.N., Johnson J.L., Temnykh S.V., Kharlamova A.V., Shepeleva D.V., Gulievich R.G., Shikhevich S.G., Graphodatsky A.S., Aguirre G.D., Acland G.M. A meiotic linkage map of the silver fox, aligned and compared to the canine genome. *Genome Res*. 2007;17(3):387-399. DOI 10.1101/gr.5893307

Kutschera V.E., Lecomte N., Janke A., Selva N., Sokolov A.A., Haun T., Steyer K., Nowak C., Hailer F. A range-wide synthesis and timeline for phylogeographic events in the red fox (*Vulpes vulpes*). *BMC Evol. Biol*. 2013;13(1):114. DOI 10.1186/1471-2148-13-114

Lojkić I., Šimić I., Bedeković T., Krešić N. Current status of rabies and its eradication in Eastern and Southeastern Europe. *Pathogens*. 2021;10(6):742. DOI 10.3390/pathogens10060742

Manel S., Schwartz M.K., Luikart G., Taberlet P. Landscape genetics: combining landscape ecology and population genetics. *Trends Ecol. Evol*. 2003;18(4):189-197. DOI 10.1016/S0169-5347(03)00008-9

McDevitt A.D., Coscia I., Browett S.S., Ruiz-González A., Statham M.J., Ruczyńska I., Panek M., Korolev A., Okhlopkov I.M., Saveljev A.P., Pokorny B., Flajšman K., Harrison S.W.R., Lobkov V., Ćirović D., Mullins J., Pertoldi C., Randi E., Sacks B.N., Kowalczyk R., Wójcik J.M. Next-generation phylogeography resolves post-glacial colonization patterns in a widespread carnivore, the red fox (*Vulpes vulpes*), in Europe. *Mol. Ecol*. 2022;31(3):993-1006. DOI 10.1111/mec.16276

Mullins J., McDevitt A.D., Kowalczyk R., Ruczyńska I., Górny M., Wójcik J.M. The influence of habitat structure on genetic differentiation in red fox populations in north-eastern Poland. *Acta Theriol*. 2014;59(3):367-376. DOI 10.1007/s13364-014-0180-2

Nemet A. Evaluation of the Sensitivity of the System of Passive Control of Rabies in Bosnia and Herzegovina. Master thesis. Univ. of Sarajevo, 2018

Norén K., Statham M.J., Ågren E.O., Isomursu M., Flagstad Ø., Eide N.E., Berg T.B.G., Bech-Sanderhoff L., Sacks B.N. Genetic footprints reveal geographic patterns of expansion in Fennoscandian red foxes. *Global Change Biol*. 2015;21(9):3299-3312. DOI 10.1111/gcb.12922

Nowak R.M. Walker's Mammals of the World. Johns Hopkins Univ. Press, 1999

OSLIS (Osrednji Slovenski Lovsko Informacijski Sistem). Available online: http://oslis.gozdis.si/. Accessed on 6 May 2024

Picot V., Rasuli A., Abella-Rider A., Saadatian-Elahi M., Aikimbayev A., Barkia A., Benmaiz S., Bouslama Z., De Balogh K., Dehove A., Davlyatov F., Farahtaj F., Gongal G., Gholami A., Imnadze P., Issad M., Khoufi S., Nedosekov V., Rafila A., Rich H., Soufi A., Tuychiev J., Vranjes N., Vodopija R., Zaouia I., Nel L. The Middle East and Eastern Europe rabies Expert Bureau (MEEREB) third meeting: Lyon, France (7–8 April, 2015). *J. Infect. Public Health*. 2017;10(6):695-701. DOI 10.1016/j.jiph.2017.03.005

R Development Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria, 2011. http://www.R-project.org

Rabies Bulletin Europe. Available online: https://www.who-rabies-bulletin.org/news/re-emergence-rabies-bosnia-herzegovina. Accessed on December 2017

Richman M., Mellersh C.S., André C., Galibert F., Ostrander E.A. Characterization of a minimal screening set of 172 microsatellite markers for genome-wide screens of the canine genome. *J. Biochem. Biophys*. 2001;47(1-2):137-149. DOI 10.1016/s0165-022x(00)00160-3

Rosenberg N.A. DISTRUCT: a program for the graphical display of population structure. *Mol. Ecol. Notes*. 2004;4(1):137-138. DOI 10.1046/j.1471-8286.2003.00566.x

Rousset F. GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Mol. Ecol. Resour*. 2008;8(1):103-106. DOI 10.1111/j.1471-8286.2007.01931.x

Sommer R.S., Nadachowski A. Glacial refugia of mammals in Europe: evidence from fossil records. *Mammal Rev*. 2006;36(4):251-265. DOI 10.1111/j.1365-2907.2006.00093.x

Sommer S., Mcdevitt A.D., Balkenhol N. Landscape genetic approaches in conservation biology and management. *Conserv. Genet*. 2013;14(2):249-251. DOI 10.1007/s10592-013-0473-z

Statham M.J., Murdoch J., Janecka J., Aubry K.B., Edwards C.J., Soulsbury C.D., Berry O., Wang Z., Harrison D., Pearch M., Tomsett L., Chupasko J., Sacks B.N. Range-wide multilocus phylogeography of the red fox reveals ancient continental divergence, minimal genomic exchange and distinct demographic histories. *Mol. Ecol*. 2014;23(19):4813-4830. DOI 10.1111/mec.12898

Tasioudi K.E., Iliadou P., Agianniotaki E.I., Robardet E., Liandris E., Doudounakis S., Tzani M., Tsaroucha P., Picard-Meyer E., Cliquet F., Mangana-Vougiouka O. Recurrence of animal rabies, Greece, 2012. *Emerg. Infect. Dis*. 2014;20(2):327-329. DOI 10.3201/eid2002.130473

Teacher A.G., Thomas J.A., Barnes I. Modern and ancient red fox (*Vulpes vulpes*) in Europe show an unusual lack of geographical and temporal structuring, and differing responses within the carnivores to historical climatic change. *BMC Evol. Biol*. 2011;11:214. DOI 10.1186/1471-2148-11-214

Valvo G. Applications of Landscape Genetics for Wildlife Conservation and Management. Doctoral thesis. Univ. of Padua, Department of Animal Sciences, 2011

Weir B.S., Cockerham C.C. Estimating F-statistics for the analysis of population structure. *Evolution*. 1984;38:1358-1370

Zecchin B., De Nardi M., Nouvellet P., Vernesi C., Babbucci M., Crestanello B., Bago Z., Bedeković T., Hostnik P., Milani A., Donnelly C.A., Bergelloni L., Lorenzetto M., Cattoli G. Genetic and spatial characterization of the red fox (*Vulpes vulpes*) population in the area stretching between the Eastern and Dinaric Alps and its relationship with rabies and canine distemper dynamics. *PLoS One*. 2019;14(3):e0213515. DOI 10.1371/journal.pone.0213515

**Conflict of interest.** The authors declare no conflict of interest.
Received July 8, 2024. Revised August 19, 2024. Accepted August 27, 2024.

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2024;28(7):759-769

DOI 10.18699/vjgb-24-84

# Detailed cytogenetic analysis of three duck species (the northern pintail, mallard, and common goldeneye) and karyotype evolution in the family Anatidae (Anseriformes, Aves)

V.R. Beklemisheva [1] ✉, K.V. Tishakova [1], S.A. Romanenko [1], D.A. Andreushkova[1], V.A. Yudkin [2, 3], E.A. Interesova [2, 4], F. Yang [5], M.A. Ferguson-Smith [6], A.S. Graphodatsky [1], A.A. Proskuryakova [1]

[1] Institute of Molecular and Cellular Biology of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
[2] Institute of Systematics and Ecology of Animals of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
[3] Novosibirsk State University, Novosibirsk, Russia
[4] Tomsk State University, Tomsk, Russia
[5] School of Life Sciences and Medicine, Shandong University of Technology, Zibo, China
[6] Cambridge Resource Center for Comparative Genomics, University of Cambridge, Cambridge, United Kingdom

✉ bekl@mcb.nsc.ru

**Abstract.** Galliformes and Anseriformes are two branches of the Galloanserae group, basal to other Neognathae. In contrast to Galliformes, Anseriformes have not been thoroughly researched by cytogenetic methods. This report is focused on representatives of Anseriformes and the evolution of their chromosome sets. Detailed cytogenetic analysis (G-banding, C-banding, and fluorescence *in situ* hybridization) was performed on three duck species: the northern pintail (*Anas acuta*, 2*n* = 80), the mallard (*A. platyrhynchos*, 2*n* = 80), and the common goldeneye (*Bucephala clangula*, 2*n* = 80). Using stone curlew (*Burhinus oedicnemus*, 2*n* = 42, Charadriiformes) chromosome painting probes, we created homology maps covering macrochromosomes and some microchromosomes. The results indicated a high level of syntenic group conservation among the duck genomes. The two Anas species share their macrochromosome number, whereas in *B. clangula*, this number is increased due to fissions of two ancestral elements. Additionally, in this species, the presence of massive heterochromatic blocks in most macroautosomes and sex chromosomes was discovered. Localization of clusters of ribosomal DNA and telomere repeats revealed that the duck karyotypes contain some microchromosomes that bear ribosomal RNA genes and/or are enriched for telomere repeats and constitutive heterochromatin. Dot plot (D-GENIES) analysis confirmed the established view about the high level of syntenic group conservation among Anatidae genomes. The new data about the three Anatidae species add knowledge about the transformation of macro- and sex chromosomes of Anseriformes during evolution.

**Key words:** *Anas acuta*; *Anas platyrhynchos*; *Bucephala clangula*; *Burhinus oedicnemus*; comparative chromosome painting; constitutive heterochromatin; ribosomal genes; telomere.

**For citation:** Beklemisheva V.R., Tishakova K.V., Romanenko S.A., Andreushkova D.A., Yudkin V.A., Interesova E.A., Yang F., Ferguson-Smith M.A., Graphodatsky A.S., Proskuryakova A.A. Detailed cytogenetic analysis of three duck species (the northern pintail, mallard, and common goldeneye) and karyotype evolution in the family Anatidae (Anseriformes, Aves). *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2024;28(7):759-769. DOI 10.18699/vjgb-24-84

**Funding.** This research was supported by Russian Science Foundation grant No. 23-74-01016 (to P.A.A.).

**Acknowledgements.** The authors express their gratitude to Ju. Butakova (Laboratory of Animal Cytogenetics, IMCB SB RAS, Novosibirsk, Russia) for her comprehensive assistance, D.V. Yudkin for helping to obtain the tissue samples, and V.T. Volobuev for establishing the stone curlew cell line. The authors gratefully acknowledge the resources provided by the Molecular and Cellular Biology core facility at the IMCB SB RAS. The English language was corrected and certified by shevchuk-editing.com. We thank two anonymous reviewers for their insightful comments and suggestions.

# Комплексный цитогенетический анализ кариотипов трех видов уток (шилохвость, кряква и обыкновенный гоголь) и эволюция кариотипов у представителей семейства Anatidae (Anseriformes, Aves)

В.Р. Беклемишева [1] ✉, К.В. Тишакова [1], С.А. Романенко [1], Д.А. Андреюшкова[1], В.А. Юдкин [2, 3], Е.А. Интересова [2, 4], Ф. Янг [5], М.А. Фергюсон-Смит [6], А.С. Графодатский [1], А.А. Проскурякова [1]

[1] Институт молекулярной и клеточной биологии Сибирского отделения Российской академии наук, Новосибирск, Россия
[2] Институт систематики и экологии животных Сибирского отделения Российской академии наук, Новосибирск, Россия
[3] Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия

[4] Томский государственный университет, Томск, Россия
[5] Школа медицины и наук о жизни, Шандонский технологический университет, Цзыбо, Китай
[6] Кембриджский ресурсный центр сравнительной геномики, Кембриджский университет, Кембридж, Великобритания
bekl@mcb.nsc.ru

**Аннотация.** Курообразные (Galliformes) и гусеобразные (Anseriformes) – две ветви группы Galloanserae, базальные по отношению к остальным Neognathae. В сравнении с Galliformes, эволюция кариотипов Anseriformes недостаточно изучена. Настоящее исследование посвящено представителям гусеобразных и изменению их хромосомных наборов в ходе эволюции. Была получена подробная информация о кариотипах (G-, C-бэндинг, флуоресцентная гибридизация *in situ*) трех видов уток: шилохвости (*Anas acuta*, 2*n* = 80), кряквы (*Anas platyrhynchos*, 2*n* = 80) и обыкновенного гоголя (*Bucephala clangula*, 2*n* = 80). С использованием зондов, разработанных на основе сортированных хромосом авдотки (*Burhinus oedicnemus*, 2*n* = 42, Charadriiformes), были выявлены районы гомологии на макрохромосомах и части микрохромосом уток. Изученные виды рода *Anas* имеют одинаковое число макрохромосом, при этом у *B. clangula* число крупных хромосом увеличено за счет двух разрывов предковых элементов. В отличие от представителей *Anas*, у этого вида обнаружены массивные гетерохроматиновые блоки в большинстве крупных макроаутосом и в половых хромосомах. Данные хромосомного пэйнтинга дополнены информацией о локализации рибосомной ДНК и амплификации теломерных повторов. Сравнительный анализ геномов с помощью приложения D-GENIES подтвердил высокий уровень консерватизма синтенных групп у Anatidae. Полученные результаты расширили представление о преобразованиях макро- и половых хромосом Anseriformes в ходе эволюции.
**Ключевые слова:** *Anas acuta*; *Anas platyrhynchos*; *Bucephala clangula*; сравнительный хромосомный пэйнтинг; конститутивный гетерохроматин; рибосомные гены; теломера.

## Introduction

Aves is one of the most specialized and species-rich groups of amniotes. Avian genomes are characterized by a relatively small size (1.4 Gb on average) (Gregory, 2023), a high degree of conserved synteny, and a relatively small proportion of repetitive DNA (Delany et al., 2000; Ellegren, 2013; Campagna, Toews, 2022). Birds' diploid chromosome number (2*n*) varies from 40 to 142 owing to the presence of numerous microchromosomes (Christidis, 1990; Griffin et al., 2007; Degrandi et al., 2020). Despite the growing popularity of birds as a genomic research object, many species remain understudied both at the genomic and cytogenetic level. According to the Bird Chromosome Database, only 10 % of the species have a described karyotype, and comparative genomic studies have covered no more than 1 % of the species (Degrandi et al., 2020).

The order Anseriformes (geese, ducks, swans, and others) includes ~180 species (Gill et al., 2023) and together with Galliformes (chickens, guinea fowl, pheasants, and others) forms the Galloanserae clade. Galloanserae taxon diverged from other Neognathae approximately 100–70 million years ago (Van Tuinen, Hedges, 2001; Sun et al., 2017). Karyotypes of 46 Anseriformes species have been described. The diploid number of chromosomes varies from 2*n* = 74 (*Spatula querquedula*) (Ebied et al., 2005) to 2*n* = 98 (*Coscoroba coscoroba*) (Rodrigues et al., 2014). Nonetheless, most karyotypes contain 2*n* = 80–82 ((Bird Chromosome Database and (Degrandi et al., 2020)).

Eight Anseriformes species have been studied by comparative chromosome painting using chicken macrochromosome probes (Degrandi et al., 2020). The chicken (*Gallus gallus*, GGA, 2*n* = 78) genome is widely used as a reference in comparative genomics, and the painting probes have become essential for avian comparative cytogenetic researches (Shetty et al., 1999; Guttenbach et al., 2003; Griffin et al., 2008; Nanda et al., 2011; Islam et al., 2014; Rodrigues et al., 2014; Uno et al., 2019). Based on comparative data obtained using chicken painting probes, a hypothesis about homology of the first 10 macrochromosomes and Z chromosome in the avian ancestral karyotype has been propounded (Shibusawa et al., 2004; Griffin et al., 2007). Integration of microchromosomes into comparative cytogenetics can be implemented using microdissection (Zimmer et al., 1997; Griffin et al., 1999; Guillier-Gencik et al., 1999; Grützner et al., 2001; Masabanda et al., 2004), gene-specific probes (Coullin et al., 2005; Islam et al., 2014), and bacterial artificial chromosomes (BACs) (Shibusawa et al., 2001, 2002; Schmid et al., 2005; Fillon et al., 2007; Damas et al., 2017; Kretschmer et al., 2021). The set of avian probes obtained using chromosome sorting usually does not cover microchromosomes (Habermann et al., 2001).

The use of stone curlew (*Burhinus oedicnemus*; BOE) painting probes helps to include microchromosomes into cytogenetic analysis and to count linkage groups. The stone curlew has a karyotype with a diploid number 2*n* = 42 and contains only four pairs of microchromosomes (Bulatova et al., 1971). BOE painting probes have been applied to fluorescence *in situ* hybridization (FISH) experiments on a wide range of avian karyotypes, including reciprocal chromosome painting between the stone curlew and chicken (Hansmann et al., 2009; Nie et al., 2009, 2015; Wang et al., 2022). On the other hand, this type of analysis has not fully resolved the correspondence between microchromosomes (Nie et al., 2009). These data obtained using cytogenetic approaches and chromosome level genome assembly allow for the inclusion of microchromosomes into reconstructions of the avian ancestral karyotype (Damas et al., 2018).

There are several papers on Anseriformes chromosome evolution. In general, bird karyotypes have a low level of interchromosomal rearrangements (Griffin et al., 2008; Damas et al., 2018) and at the same time a high average rate of intrachromosomal rearrangements (Damas et al., 2019; O'Connor et al., 2024). Chicken BAC probes mapped onto *Anas platyrhynchos* chromosomes have revealed the presence of inversions in all macrochromosomes and in the Z chromosome of the duck relative to the chicken (Fillon et al., 2007; Kiazim et al., 2021). However, conservation of gene order among several Anseriformes species has been detected by means of gene-

specific probes (Islam et al., 2014). Data on chromosome homology and rearrangements can be obtained by chromosome-level assembly analysis (for example (Cabanettes, Klopp, 2018)). Now chromosome-level genome assemblies are available for 26 Anseriformes species in the NCBI database.

In the present work, we analyzed karyotype evolution of Anseriformes species by combining new data and previously published ones. New comprehensive cytogenetic results were obtained for three Anatidae species: the mallard (*A. platyrhynchos*, APL), the northern pintail (*Anas acuta*, AAC), and the common goldeneye (*Bucephala clangula*, BCL). Routine and differential staining (G- and C-banding) and molecular cytogenetic approaches (localization of ribosomal and telomeric painting probes and of the set of *B. oedicnemus* whole-chromosome painting probes) were employed, and the results were integrated with previously obtained findings.

## Materials and methods

**Sampled species and an ethical statement.** Lung necropsy was performed on females of *A. platyrhynchos*, *A. acuta*, and *B. clangula* during the official hunting season in Novosibirsk Oblast (2009, 2008, 2023 years, respectively). For our research, we used one individual of each species. All experiments were approved by the Ethics Committee on Animal and Human Research at the Institute of Molecular and Cellular Biology (IMCB), the Siberian Branch of the Russian Academy of Sciences (SB RAS), Russia (decision No. 1/22 of 29 December 2022), following all relevant guidelines and regulations. This article does not contain any experiments on human subjects by the authors. The study was completed using equipment and materials of large-scale research facilities of the Cryobank of Cell Cultures, IMCB, SB RAS (Novosibirsk, Russia).

**Primary fibroblast culture, metaphase chromosome preparation, and chromosome staining.** Primary cell cultures were derived from lung necropsy as described previously (Romanenko et al., 2015). Metaphase spreads from all species were prepared from primary fibroblast cultures following standard procedures, including colcemid treatment, incubation in a hypotonic solution, and fixation in a methanol:acetic acid (3:1) mixture (Graphodatsky et al., 2001).

Routine Giemsa and/or DAPI staining was performed to identify morphology and count quantity of chromosomes. G-banding was done (Seabright, 1971) before FISH procedures. C-banding was performed for *B. clangula* and *A. acuta* according to a standard protocol (Sumner, 1972). Besides, for *B. clangula*, C-banding was done with modifications: metaphase chromosomes after Ba(OH)2 treatments were stained overnight with Chromomycin A3 (20 μg/ml) and 4′,6-diamidino-2-phenylindole (DAPI; 50 ng/ml) and mounted in an antifade solution containing 1,4-diazabicyclo [2.2.2] octane (DABCO) (20 mg/ml) and 100 mM Tris-HCl pH 7.5 in glycerol.

**Probe preparation, FISH, and microscopy.** A set of whole-chromosome female DNA libraries (created from flow-sorted chromosomes by degenerate-oligonucleotide-primed (DOP)-PCR (Telenius et al., 1992)) of the stone curlew was provided by Cambridge Resource Center for Comparative Genomics (Cambridge University, UK). The characterization of the set was described previously (Nie et al., 2009). All DNA libraries were labeled with biotin-11-dUTP and digoxigenin-11-dUTP (Sigma) during secondary DOP-PCR amplification (Telenius et al., 1992) to create painting probes. A telomeric DNA probe was generated by PCR with oligonucleotides $(TTAGGG)_5$ and $(CCCTAA)_5$ (Ijdo et al., 1991). A ribosomal DNA probe was obtained from plasmid DNA (pHr13) containing a human partial 18S ribosomal RNA (rRNA) gene, a full human 5.8S rRNA gene, a partial 28S rRNA gene, and two internal spacers (Maden et al., 1987). Labeling was performed using the FTP-Display DNA Fragmentation Kit (DNA-Display, Russia) with the addition of 0.05 mM biotin-11-dUTP or digoxigenin-11-dUTP (Sigma).

FISH was carried out by standard protocols (Graphodatsky et al., 2000; Liehr et al., 2017). For suppression of non-specific hybridization of probes in FISH experiments, chicken Cot10 DNA was used. Digoxigenin-labeled probes were detected with the help of anti-digoxigenin-CyTM3 (Jackson Immunoresearch), whereas biotin-labeled probes were identified via sequential use of avidin-FITC (Vector Laboratories) and anti-avidin FITC (Vector Laboratories). Images were captured and processed using a VideoTesT 2.0 Image Analysis System and a Baumer Optronics CCD Camera mounted on an Olympus BX53 microscope (Olympus).

**Bioinformatic analysis.** To perform the analysis of chromosome-level assemblies, we used the available material from GenBank NCBI. Chromosome-level assemblies of different Anseriformies species from different genera were compared with the *A. platyrhynchos* reference genome
(GCA_008746955.3; 1*n* = 40, Z, W):
*A. acuta* (GCA_963932015.1; 1*n* = 37, Z, W);
*Aythya baeri* (GCA_026413565.1; 1*n* = 34, Z);
*Netta rufina* (GCA_964035555.1; 1*n* = 40, Z, W);
*Cairina moschata* (GCA_018104995.1, 29, Z);
*Clangula hyemalis* (GCA_963989345.1; 1*n* = 40, Z);
*Anser cygnoides* (GCA_026259575.1; 1*n* = 39);
*Cygnus olor* (GCF_009769625.2; 1*n* = 34, Z, W);
*Oxyura jamaicensis* (GCF_011077185.1; 1*n* = 33, Z, W).
Dot plot comparative analysis was performed using the web application D-GENIES (Cabanettes, Klopp, 2018) by aligner Minimap2 v2.24.

## Results

### Karyotypes of the three duck species

Karyotypes of the three duck species were available: *A. platyrhynchos* (APL), *A. acuta* (AAC), and *B. clangula* (BCL) (Takagi, Makino, 1966; Beçak et al., 1973; Islam et al., 2014). The diploid chromosome number of all three species is 2*n* = 80. For each individual of the studied duck species, we counted the number of chromosomes in 30 routine stained metaphase plates. The *A. platyrhynchos* (mallard) karyotype was G-, R-, and C-banded before (Wójcik, Smalec, 2007) and the patterns coincided with our study.

Cytogenetic data on *A. acuta* and *B. clangula* have been limited to Giemsa-stained karyotypes (Beçak et al., 1973; Abu-Almaaty et al., 2019). We performed Giemsa staining and C-banding for individuals of both species. The *A. acuta* karyotype was found to contain five pairs of large macroautosomes (submetacentric pairs 1 and 2, and acrocentric pairs 3–5). The Z chromosome is a medium-sized subtelocen-

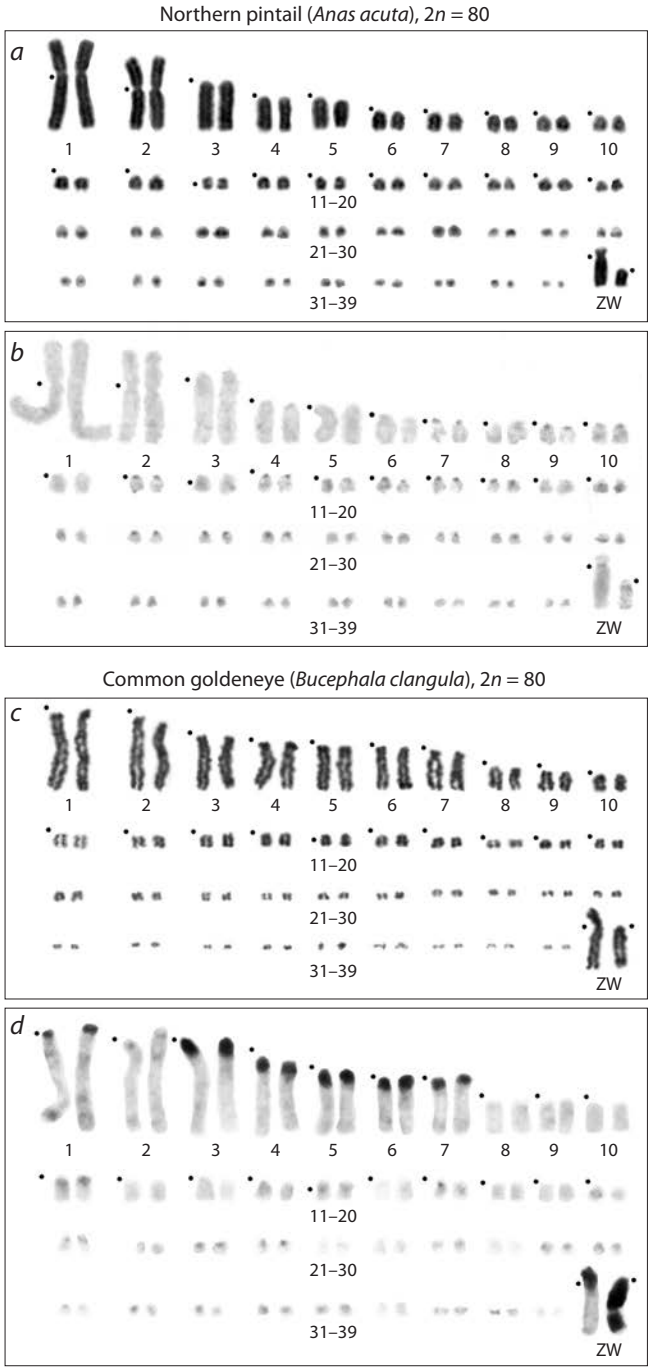



**Fig. 1.** Routinely Giemsa-stained chromosomes of (*a*) *A.acuta* and (*c*) *B. clangula* and C-banded karyotypes of (*b*) *A. acuta* and (*d*) *B. clangula*. Each black dot indicates a centromere position.

tric macrochromosome, and the W chromosome is a small acrocentric chromosome (Fig. 1*a*). C-heterochromatin proved to be dot-like blocks in centromeric regions of almost all chromosomes except for large macroautosomes 1–5 and some microchromosomes (Fig. 1*b*). In the W chromosome, a pericentromeric heterochromatic block was observed.

*B. clangula* large macrochromosomes 1–7 and the W chromosome are acrocentric, whereas the Z chromosome is subtelocentric (Fig. 1*c*). Constitutive heterochromatin in this species was found to be localized in pericentromeric regions of the six macroautosomes and the Z chromosome. In addition, the W chromosome is almost entirely heterochromatic (Fig. 1*d*). Overall, the *B. clangula* karyotype, as compared to the two *Anas* karyotypes, is enriched with repeated DNA organized in constitutive heterochromatin.

### Localization of ribosomal genes and telomere repeats

Sequences carrying ribosomal genes and clusters of repeats homologous to rDNA were located on four small microchromosome pairs in *A. acuta* and *A. platyrhynchos* (Islam et al., 2014) and on two pairs in *B. clangula* (Supplementary Material 1)[1].

In the *A. acuta* and *B. clangula* karyotypes, the probe containing $(5'\text{-TTAGGG-}3')_n$ sequences detected small signals at the ends of all chromosomes except four or five pairs of dot-like microchromosomes with prominent signals. Besides, *A. acuta* was found to have a pair of microchromosomes where half of each chromosome was stained with a telomeric probe, and the other half contained a bright DAPI-positive region (AT rich) (Fig. 3 and Supplementary Material 1). We could not detect interstitial telomeric sites in macrochromosomes of the analyzed species.

### Stone curlew homologies

The complete set of BOE painting probes was hybridized to chromosomes of *A. acuta*, *A. platyrhynchos*, and *B. clangula* (Fig. 2). Owing to G-banding prior to FISH experiments, painted macrochromosomes were easily identified. We obtained specific signals on macrochromosomes and some microchromosomes of the three species. The probe containing the W chromosome also hybridized with Z: in *A. acuta* and *A. platyrhynchos*, weak signals along the chromosome body were detected; in *B. clangula*, a bright signal was detected in the subtelomere region of the q arm (Fig. 3). Also, the W chromosome signals on the *B. clangula* W chromosome covered subtelomere C-pozitive block. Other heterochromatic regions were not hybridized with BOE painting probes on *B. clangula* chromosomes. Signals on some microchromosomes were so weak that we failed to identify them. Quantities of microchromosomes covered by signals are presented in Table. Examples of some FISH results are shown in Fig. 2.

Based on these results, comparative chromosome maps were created showing the homology between BOE and the three Anatidae species (see the Table, Fig. 3). Identification of many microchromosomes was limited by the resolution of the FISH method. Nonetheless, these maps gave an idea of relative sizes of labeled microchromosomes within size groups (pairs 11–20, 21–30, or 31–39). In total, homologies were identified for 66 out of 80 chromosomes of *A. acuta*, 68 out of 80 *A. platyrhynchos* chromosomes, and 76 out of 80 *B. clangula* chromosomes because each species has a different number of painted microchromosomes.

## Discussion

Microchromosomes are typical of most amniotes (snakes, turtles, and lizards) with mammals and crocodilians being exceptions (Srikulnath et al., 2021; O'Connor et al., 2024). The presence of microchromosomes in avian karyotypes may

[1] Supplementary Materials 1–3 are avilable at: https://vavilovj-icg.ru/download/pict-2024-28/appx26.pdf

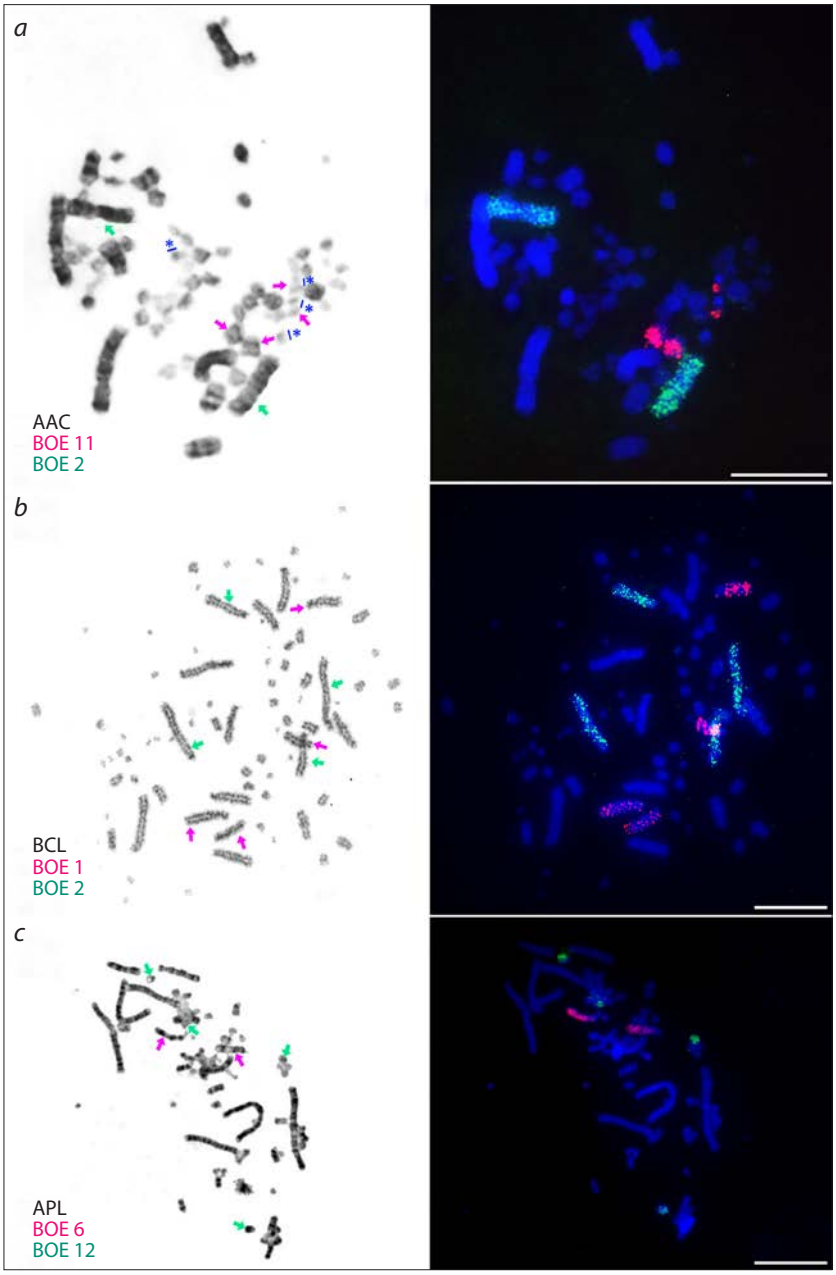



**Fig. 2.** Examples of *Burhinus oedicnemus* (BOE) painting probes' localization on karyotypes of Anatidae representatives (right) with GTG-banding of the same metaphase (left).

*a* – hybridization of BOE 11 (red) and 2 (green) probes to *A. acuta* (AAC) chromosomes; *b* – hybridization of BOE 1 (red) and 2 (green) to *A. platyrhynchos* mallard (APL) chromosomes; *c* – hybridization of BOE 6 (red) and 12 (green) probes to *B.clangula* (BCL) chromosomes. Blue marks and * indicate microchromosomes carrying a large DAPI-positive heterochromatic region in *A. acuta.* Scale bars = 10 μm.

account for the lag of bird cytogenetic studies behind mammalian studies. To date, 131 bird species from 47 families have been analyzed by comparative chromosome painting, i. e., no more than 2 % of extant avian species (Degrandi et al., 2020) (https://sites.unipampa.edu.br/birdchromosomedatabase).

The studied karyotypes of ducks from genera *Anas* and *Bucephala* possess identical diploid chromosome numbers (2*n* = 80). The diploid chromosome number reported for *A. acuta* varies among publications from 80 (Abu-Almaaty et al., 2019) to 82 (Beçak et al., 1973). A similar situation is true for *B. clangula*: diploid chromosome number 2*n* = 80 was shown for *B. clangula americana* (Beçak et al., 1973) and 2*n* = 84 for *B. clangula* (Hammar, 1970). Our results match the data obtained by A.H. Abu-Almaaty with co-authors and M.L. Beçak with co-authors, respectively: diploid chromosome number is 2*n* = 80 for both *A. acuta* and *B. clangula* (Fig. 1).

Previously it was reported that, even at the macrochromosomal level, chromosome banding can be difficult to identify, thereby making necessary a robust analysis of cross-species homology (O'Connor et al., 2024). In this work we provided information about G-banding for three Anatidae species. This allowed reliable identification of all macro- and some microchromosome pairs. We used chromosome painting to gain a better understanding of the transformations that have caused the karyotype differences. Previously, comparative chromosome painting in Anatidae species has been performed only with chicken macrochromosome probes (Guttenbach et al., 2003; Nanda et al., 2011; Islam et al., 2014; Rodrigues et al., 2014; Uno et al., 2019), but Anseriformes karyotypes have not been investigated using BOE painting probes. Applying the set of BOE painting probes, we revealed homologies between *B. oedicnemus*, *A. acuta*, *A. platyrhynchos*, and *B. clangula* chromosomes. According to previously published findings (Nie et al., 2009), we added homologies of *Gallus gallus* (GGA) to our comparative analysis (see the Table).

Due to availability of chromosome-level genome assemblies of *A. platyrhynchos* and *A. acuta*, we prepared a dot plot analysis for genomes of Anseriformies species from different genera (Supplementary Material 3). *A. platyrhynchos* is a domestic species and its genome has been studied in detail and is usually used as the reference for this taxon. We made pairwise comparative analysis with *A. platyrhynchos* and eight genomes: *A. acuta*, *Aythya baeri*, *Netta rufina*, *Cairina moschata*, *Clangula hyemalis*, *Anser cygnoides*, *Cygnus olor*, and *Oxyura jamaicensis*. Karyotypes of most compared species contain 80 chromosomes, except for *A. cygnoides* and *N. rufina* (2*n* = 82) (Degrandi et al., 2020). The karyotype of *C. hyemalis* was not described. In almost all assemblies, the number of chromosome scaffolds is lower than haploid chromosome number and does not cover the smallest microchromosomes. The homologies between *A. platyrhynchos* and *A. acuta*, identified using chromosomal painting and dot plot analysis are consistent with each other (Supplementary Material 3*a*). The chromosome correspondence can also be traced at the G-banding level.

## Macroautosome evolution in Anseriformes

The most notable difference in macrochromosome structure among the three species is the presence of massive heterochromatic blocks in *B. clangula* macrochromosomes 1 and 3–7 (Fig. 1*d*). FISH with the stone curlew painting probes and dot plot analysis revealed conserved macrochromosome homologies between *A. acuta* and *A. platyrhynchos*. The diploid number (2*n* = 80) is typical for most Anatidae species, but in the *B. clangula* karyotype, we found an increase in the number of large chromosomes without a change in the diploid chromosome number. The reason is fission of ancestral chromosomes corresponding to GGA 1 and GGA 2 (Fig. 3, 4). A break in the centromere of ancestral chromosomes homologous to GGA 2 has been documented in other representatives of Galloanseres, for example, in the turkey (Shibusawa et al., 2004). In other avian lineages, there are independent fissions of orthologs of GGA 2 around the centromere (Degrandi et al., 2020). The ancestral homolog of GGA chromosome 1 was split in two within Passeriformes' ancestral karyotype (Damas et al., 2018).

Orthologous elements of GGA 4 are represented by one macrochromosome and one microchromosome in the three examined ducks (see the Table). By contrast, the GGA 4 probe hybridized to one macrochromosome in *Anser anser* (Fig. 4) (Guttenbach et al., 2003). Ancestral chromosome 4 (corresponding to GGA 4q) is the most ancient of all the avian chromosomes, appearing intact even in humans (Chowdhary, Raudsepp, 2000). In the GGA genome, we can see the result of fusion of avian ancestral chromosomes 4 and 10 (4A) (GGA 4q and 4p) (Shibusawa et al., 2002; Guttenbach et al., 2003; Derjusheva et al., 2004; Itoh, Arnold, 2005; Griffin et al., 2007; Damas et al., 2018), and 4p still retains properties of a microchromosome (high gene density and recombination rate, CpG island distribution). The observed interchromosomal changes have also taken place in different lineages (Fig. 4) and led to genome reorganization in the extant species. These observations may be best explained by rearrangements with common breakpoint reuse.

The data of dot plot analysis also identified the presence of interchromosomal rearrangements affecting macrochromosomes in *C. moschata* and *O. jamaicensis* (Supplementary Material 3*d*, *h*). In *C. moschata*, fission of chromosomes orthologous to *A. platyrhynchos* 6 and 7 and fusion parts of these chromosomes with chromosomes 5 and 8 respectively are observed (Supplementary Material 3*d*). Identified rearrangements are not confirmed using chromosome painting (Islam et al., 2014). This may be due to either a weak signal that would not be detected by chicken chromosome painting probes, or an artifact of *de novo* genome assembly. Also, in *O. jamaicensis*, genome fusion of chromosomes orthologous to *A. platyrhynchos* 4 and 11 was detected (Supplementary Material 3*h*). Minor interchromosomal rearrangements in pericentromeric regions of almost all macrochromosomes were observed.

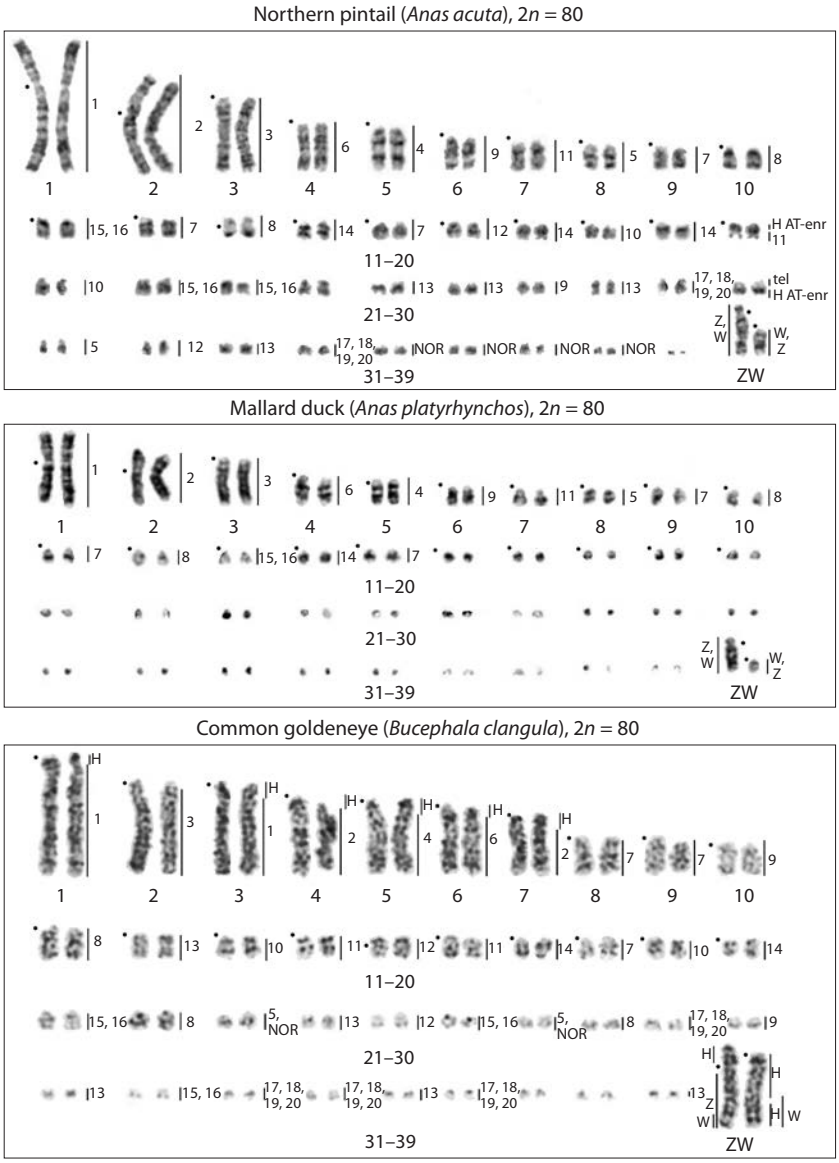



**Fig. 3.** G-banded karyotypes of *A. acuta*, *A. platyrhynchos*, and *B. clangula* with assignment of homologies to *B. oedicnemus* chromosomes.

ID numbers on the right indicate names of homologous stone curlew chromosomes; H – blocks of heterochromatin; NOR – positions of ribosomal DNA clusters; black circles indicate centromere positions; tel – a cluster of telomeric repeated sequences on a microchromosome containing a large DAPI-positive heterochromatic region in *A. acuta*.

## Evolution of microchromosomes in Anseriformes

Bird microchromosomes have attracted research attention because although for most bird species these chromosomes represent 25 % of the genome, they are known to encode 50 % of all genes (Smith et al., 2000). It has been hypo-

Chromosome correspondence between *B. oedicnemus* (BOE) and four Galloanseres species obtained using chromosome painting: *Gallus gallus* (GGA), *A. acuta* (AAC), *A. platyrhynchos* (APL), and *B. clangula* (BCL)

| BOE, 2*n* = 42 | GGA, 2*n* = 78 | AAC, 2*n* = 80 | APL, 2*n* = 80 | BCL, 2*n* = 80 |
|---|---|---|---|---|
| Chr1 | Chr1 | Chr1 | Chr1 | Chr1, Chr7 |
| Chr2 | Chr2 | Chr2 | Chr2 | Chr3, Chr4 |
| Chr3 | Chr3 | Chr3 | Chr3 | Chr2 |
| Chr4 | Chr4q | Chr5 | Chr5 | Chr5 |
| Chr5 | Chr7, Chr8 | Chr8 + 1 MIC | Chr8 + 1 MIC | 2 MICs |
| Chr6 | Chr5 | Chr4 | Chr4 | Chr6 |
| Chr7 | Chr9 + 2 MICs | 3 MICs | 3 MICs | Chr8, Chr9 + 1 MIC |
| Chr8 | Chr4p + 1 MIC | 2 MICs | 3 MICs | 3 MICs |
| Chr9 | Chr6 + 1 MIC | Chr6 + 1 MIC | Chr6 + 1 MIC | Chr10 + 1 MIC |
| Chr10 | 2 MICs | 2 MICs | 2 MICs | 2 MICs |
| Chr11 | 2 MICs | Chr7 + 1 MIC | Chr7 + 1 MICs | 2 MICs |
| Chr12 | 2 MICs | 2 MICs | 2 MICs | 2 MICs |
| Chr13 | 2 MICs | 4 MICs | 4 MICs | 5 MICs |
| Chr14 | 2 MICs | 2 MICs | 1 MIC | 2 MICs |
| Chr15+16 | 3 MICs | 3 MICs | 2–3 MICs | 3 MICs |
| Chr17+18+19+20 | 1 MIC | 2 MICs | 4 MICs | 4 MICs |
| Z | Z | Z, W | Z, W | Z |
| W | W, Zqdist | W | W | W, Zqdist |
| TOT | 56 | 66 | 68 | 76 |

Data on homology between GGA and BOE chromosomes are from (Nie et al., 2009). MICs are quantities of microchromosome pairs; Chr is the chromosome; TOT – the total quantities of chromosomes labeled by BOE painting probes; dist – the distal part of the chromosome.

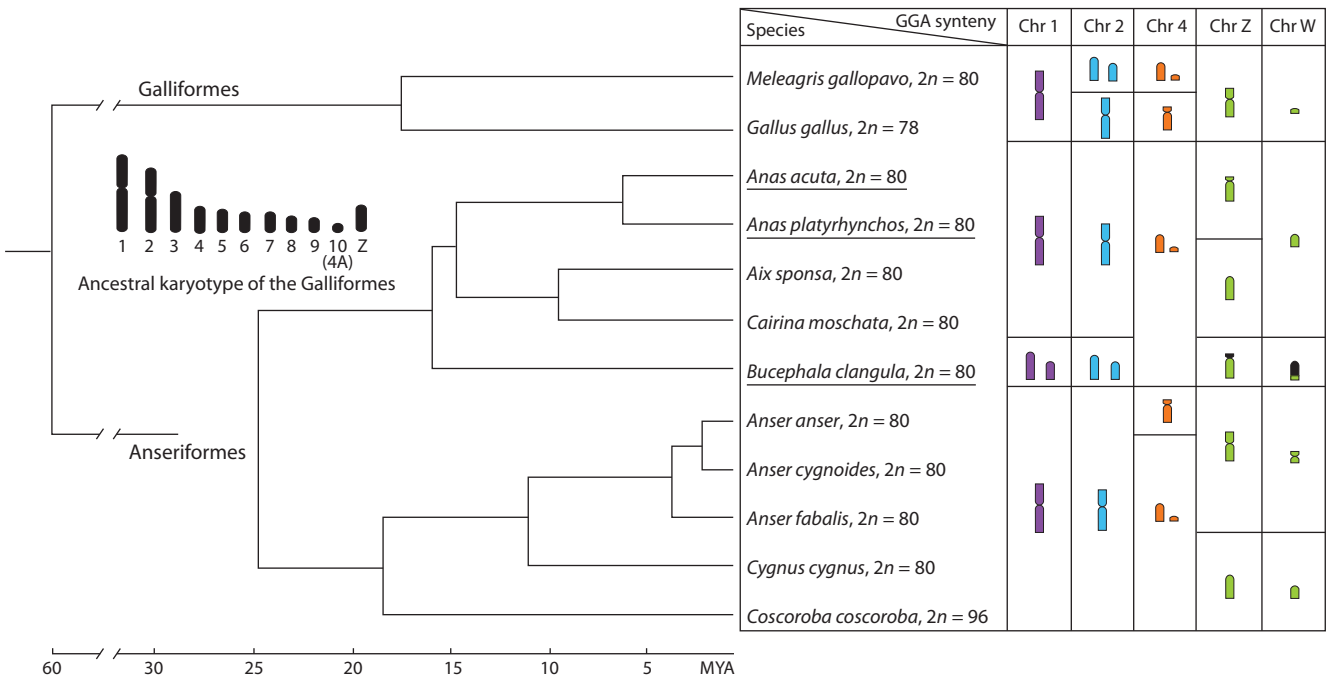



**Fig. 4.** Schematic representation of chromosome rearrangements that occurred in macro- and sex chromosomes during the evolution of Galloanseres according to molecular phylogenies (Prum et al., 2015; Sun et al., 2017).

Rearrangements of chromosomes that were identified by comparative painting with chicken probes (Guttenbach et al., 2003; Nanda et al., 2011; Islam et al., 2014; Rodrigues et al., 2014; Uno et al., 2019) or stone curlew probes (data obtained in this study, species names are underlined) are presented to the right of the molecular phylogeny. The ancestral karyotype of Galliformes and data about Galliformes species are from (Shibusawa et al., 2004; Griffin et al., 2007; Damas et al., 2019).

thesized that avian microchromosomes represent archaic linkage groups of ancestral vertebrates and have been preserved in a conserved state throughout the evolution of birds (Burt, 2002; Nakatani et al., 2007). Among Anatidae species, only *A. platyrhynchos* microchromosomes have been examined by FISH with chicken BAC probes, and the obtained data have confirmed the idea of high conservatism of avian genomes (Fillon et al., 2007; Kiazim et al., 2021).

The set of BOE painting probes covered not all duck microchromosomes (see the Table, Fig. 3). Enrichment of some microchromosomes with ribosomal DNA sequences and telomeric repeats may also influence the detection of homologous sequences on microchromosomes. Sizes of homologous segments on microchromosomes may be below the resolution level of molecular cytogenetics. This problem may be exemplified by BOE 5 and orthologous elements in chicken and duck karyotypes. This probe delineated GGA 7 and GGA 8 but only two pairs of small microchromosomes in *B. clangula*. Most likely, BOE 5 labels more microchromosomes in ducks, and some signals were overlooked when the FISH data were inspected.

Although all three ducks examined here have the same diploid number, in the *B. clangula* chromosomal set, we suggest there are at least two ancestral macrochromosome fissions. It is possible that some microchromosomes have merged or fused with macrochromosomes during the formation of the *B. clangula* karyotype. For Galloanserae, interchromosomal rearrangements between microchromosomes have not yet been recorded, but for other orders of Neognathae, data on interchromosomal rearrangements involving macro- and microchromosomes and the fusion of microchromosomes are gradually accumulating (reviewed in (O'Connor et al., 2024)). In addition, many researchers believe that the ancestral element 10 (4A) is a microchromosome according to its GC content and gene density (Griffin et al., 2007; Damas et al., 2018; O'Connor et al., 2024). Fusion of ancestral elements 10 (4A) with 4 is often observed in Anseriformes (Fig. 4).

Another example of the fusion of micro- and macrochromosomes was revealed by dot plot analysis in Galloanserae species *O. jamaicensis* (Supplementary Material 3*h*). Fusions of microchromosomes occurred in birds; for example, numerous fusions of microchromosomes took place during formation of the stone curlew karyotype (Nie et al., 2009). The fusion of microchromosomes was revealed in the genome of *C. moschata* (Supplementary Material 3*d*). The use of BAC clones to analyze the chromosome set of *B. clangula* and chromosome-level genome assembly will help to identify rearrangements that occurred during the formation of the karyotype of this species.

Given the low number of repeated sequences in bird genomes, the large number of telomeric sequences (2 to 4 %) is a point of interest (Delany et al., 2000). In addition to the canonical localization of $(5'\text{-TTAGGG-}3')_n$ sequences at the ends of chromosomes, these highly conserved sequences may be situated at an interstitial position: around a centromere or inside a chromosome arm. Three types of telomere arrays differing in length, position on a chromosome, age-related stability, and linked either to macro- or to microchromosomes have been described in the chicken (Delany et al., 2000, 2007; Rodrigue et al., 2005). It has been demonstrated that some avian microchromosomes tend to accumulate telomeric repeats. So-called mega-telomeres (200 kbp to 3 Mbp) at one of the ends have been detected in three chicken chromosomes from highly inbred White Leghorn line (9 and two microchromosomes) (Rodrigue et al., 2005; Delany et al., 2007).

Similar features of the distribution and amplification of telomeric sequences have been described for representatives of Passeriformes (Dos Santos et al., 2015, 2017) and three anserid species (Islam et al., 2014). In the *C. moschata* karyotype, 14 pairs of microchromosomes bear extended telomeres (Nanda et al., 2002). We also located extended telomeric sequences at the ends of four microchromosome pairs in *B. clangula* and 4–5 dot-like microchromosome pairs in *A. acuta* (Supplementary Material 1). Further studies involving larger number of individuals, mapping BAC probes to identify individual chromosomes, are needed to determine whether microchromosomes with large clusters of telomeric repeats are common to all Galloanserae, or whether different microchromosomes tend to accumulate telomeric sequences in different bird species.

Genes of 18S–28S rDNA are believed to have been located on a single pair of microchromosomes in the ancestral avian karyotype, as observed in palaeognathous birds and in the chicken (Nishida-Umehara et al., 2007; Delany et al., 2009). We found clusters of rRNA genes on four microchromosome pairs in *A. acuta* and two pairs in *B. clangula*. Localization of rDNA sequences on several pairs of microchromosomes was also described for *A. platyrhynchos*, *C. moschata*, and *Anser cygnoides* (Islam et al., 2014).

Some bird microchromosomes have accumulated different types of repeated sequences during karyotype evolution (Nanda et al., 2002; Dos Santos et al., 2015, 2017; Zlotina et al., 2019; de Oliveira et al., 2024). We revealed the presence of microchromosomes enriched with telomeric sequences and/or constitutive heterochromatin in *A. acuta* (Fig. 3 and Supplementary Material 1). The data presented here are based on the analysis of one individual from each investigated species. Perhaps accumulation of repetitive sequences is subject to intraspecies variation, and we need more information to address this issue.

### Sex chromosome evolution

There is variation in morphology of Z chromosomes of Galloanseres in terms of centromere position and size of heterochromatin regions (Fig. 3). For example, it was reported previously that the centromere position on the Z chromosome is different between the chicken and turkey, but most likely this situation is a consequence of the accumulation of heterochromatin on the chicken Z chromosome (Shibusawa et al., 2004). In Anseriformes, submetacentric, subtelocentric, and acrocentric Z chromosomes were observed (Fig. 4). Recent research indicates that the bird Z chromosome is involved in inter- and intrachromosomal rearrangements (Damas et al., 2019). Dot plot analysis of eight Anseriformies species identified the same gene order on the Z chromosome in all species except for *Anser cygnoides* (Supplementary Material 3*f*). Different chromosome morphology (Fig. 4) with gene order conservation (Supplementary Material 3) indicates the presence

of centromere repositions in Anseriformes Z chromosomes. Amplification of heterochromatin (AT- and GC-rich) could occur in the pericentromeric region of the Z chromosome of *B. clangula* (Fig. 1 and Supplementary Material 2).

Physical size of the W chromosome is one of highly variable characteristics of the avian genome (Stiglec et al., 2007). Usually, the W chromosome of Neognathae is smaller than the Z chromosome and has lost most genes. Nonetheless, large W chromosomes in many avian species have been described (Rutkowska et al., 2012; De Oliveira et al., 2017). The size of the degenerated homologous sex chromosomes may increase due to the enrichment with repetitive elements or to autosomal translocations (Pala et al., 2012; Schartl et al., 2016). Among the three species under study, a large W chromosome has been previously described only for *B. clangula* (Hammar, 1970; Beçak et al., 1973). C-banding of *B. clangula* metaphase chromosomes revealed that its W chromosome carries extended heterochromatic regions with a prominent AT-rich interstitial region (DAPI-positive) and GC-rich subtelomeric and pericentromeric regions (Chromomycin A3–positive) (Supplementary Material 2).

It is interesting that AT- and GC- heterochromatic regions overlapped. It may be explained by alternation of AT- and GC-rich repeats or the presence of repeats with a uniform distribution of AT/GC nucleotides in these regions. The painting probe containing the W chromosome covered the euchromatic region and subtelomere C-positive block, which may indicate the presence of some W-specific repeated sequences in the distal part of W. Remarkably, this probe hybridized also with the distal region of the q arm of the Z chromosome (Fig. 3). This may indicate a pseudoautosomal region or localization of amplified W-specific repeated sequences. Further investigation of *B. clangula* genomic architecture, repeatome, and whole-genome assembly may shed light on the evolutionary rearrangements of sex chromosomes.

## Conclusion

Here, for the first time, we used stone curlew painting probes to analyze karyotypes of three Anatidae representatives. The detailed comparative cytogenetic maps cover all macrochromosomes and some microchromosomes and contain valuable information about the composition of the karyotypes, chromosome morphology, and localization of some repeated DNA sequences. Data of comparative chromosome painting together with dot plot (D-GENIES) analysis confirm the established view about the high level of syntenic group conservation among duck genomes.

Some microchromosomes in the examined duck karyotypes were found to bear amplified clusters of ribosomal DNA, telomeric repeats, or fractions of repeated DNA. The rDNA clusters were located on more than one pair of microchromosomes in Anatidae karyotypes – on four microchromosome pairs in *A. acuta* and two pairs in *B. clangula*. Also, microchromosomes with extended telomeres have been found in this work. This state of affairs may be a contributing factor to the difficulties with assembling avian genomes by conventional bioinformatic methods. Accumulation of heterochromatin in the *B. clangula* karyotype needs further research: karyotyping of more individuals and repeatome analysis.

## References

Abu-Almaaty A.H., Hassan M.K., El Bakary N.E.R., Ahmed S.H. Chromosomal evolution and molecular genetic analysis of four species of genus *Anas* (Aves: Anatidae). *Genetika*. 2019;51(1):104-119. DOI 10.2298/GENSR1901103A

Beçak M.L., Beçak W., Roberts F.L., Shoffner R.N., Volp E.P. Aves. In: Chromosome Atlas: Fish, Amphibians, Reptiles, and Birds. Berlin: Springer, 1973;129-207

Bulatova N.S., Panov E.N., Radzhabli S.I. Description of the karyotypes of several bird species of the USSR fauna. *Dokl. Akad. Nauk SSSR*. 1971;199(6):1420-1423

Burt D.W. Origin and evolution of avian microchromosomes. *Cytogenet. Genome Res.* 2002;96(1-4):97-112

Cabanettes F., Klopp C. D-GENIES: Dot plot large genomes in an interactive, efficient and simple way. *PeerJ*. 2018;6:e4958. DOI 10.7717/PEERJ.4958/TABLE-2

Campagna L., Toews D.P.L. The genomics of adaptation in birds. *Curr. Biol.* 2022;32(20):R1173-R1186. DOI 10.1016/j.cub.2022.07.076

Chowdhary B., Raudsepp T. Cytogenetics and physical gene maps. In: Bowling A.T., Ruvinsky A. (Eds.). The Genetics of the Horse. Oxon, UK: CABI Publishing, CAB International, 2000;171-241

Christidis L. Chordata 3: Aves. In: John B., Kayano H., Levan A. (Eds.). Animal Cytogenetics. Vol. 4. Berlin: Gebrueder Borntraeger, 1990

Coullin P., Bed'Hom B., Candelier J.J., Vettese D., Maucolin S., Moulin S., Galkina S.A., Bernheim A., Volobouev V. Cytogenetic repartition of chicken CR1 sequences evidenced by PRINS in Galliformes and some other birds. *Chromosom. Res.* 2005;13(7):665-673. DOI 10.1007/s10577-005-1004-7

Damas J., O'Connor R., Farré M., Lenis V.P.E., Martell H.J., Mandawala A., Fowler K., Joseph S., Swain M.T., Griffin D.K., Larkin D.M. Upgrading short-read animal genome assemblies to chromosome level using comparative genomics and a universal probe set. *Genome Res.* 2017;27(5):875-884. DOI 10.1101/gr.213660.116

Damas J., Kim J., Farré M., Griffin D.K., Larkin D.M. Reconstruction of avian ancestral karyotypes reveals differences in the evolutionary history of macro- and microchromosomes. *Genome Biol.* 2018;19(1):155. DOI 10.1186/s13059-018-1544-8

Damas J., O'Connor R.E., Griffin D.K., Larkin D.M. Avian chromosomal evolution. In: Kraus R.H.S. (Ed.). Avian Genomics in Ecology and Evolution: From the Lab into the Wild. Springer, 2019;69-92

de Oliveira A.M., Souza G.M., Toma G.A., Dos Santos N., Dos Santos R.Z., Goes C.A.G., Deon G.A., Setti P.G., Porto-Foresti F., Utsunomia R., Gunski R.J., Del Valle Garnero A., Herculano Correa de Oliveira E., Kretschmer R., Cioffi M.B. Satellite DNAs, heterochromatin, and sex chromosomes of the wattled jacana (Charadriiformes; Jacanidae): a species with highly rearranged karyotype. *Genome*. 2024;67(4):109-118. DOI 10.1139/gen-2023-0082

De Oliveira T.D., Kretschmer R., Bertocchi N.A., Degrandi T.M., De Oliveira E.H.C., De Cioffi M.B., Garnero A.D.V., Gunski R.J. Genomic organization of repetitive DNA in woodpeckers (aves, piciformes): Implications for karyotype and ZW sex chromosome differentiation. *PLoS One*. 2017;12(1):e0169987. DOI 10.1371/journal.pone.0169987

Degrandi T.M., Barcellos S.A., Costa A.L., Garnero A.D.V., Hass I., Gunski R.J. Introducing the bird chromosome database: An overview of cytogenetic studies in birds. *Cytogenet. Genome Res.* 2020;160(4):199-205. DOI 10.1159/000507768

Delany M.E., Krupkin A.B., Miller M.M. Organization of telomere sequences in birds: evidence for arrays of extreme length and for in vivo shortening. *Cytogenet. Cell Genet.* 2000;90(1-2):139-145. DOI 10.1159/000015649

Delany M.E., Gessaro T.M., Rodrigue K.L., Daniels L.M. Chromosomal mapping of chicken mega-telomere arrays to GGA9, 16, 28 and W using a cytogenomic approach. *Cytogenet. Genome Res.* 2007;117(1-4):54-63. DOI 10.1159/000103165

Delany M.E., Robinson C.M., Goto R.M., Miller M.M. Architecture and organization of chicken microchromosome 16: Order of the NOR, MHC-Y, and MHC-B subregions. *J. Hered.* 2009;100(5):507-514. DOI 10.1093/jhered/esp044

Derjusheva S., Kurganova A., Habermann F., Gaginskaya E. High chromosome conservation detected by comparative chromosome painting in chicken, pigeon and passerine birds. *Chromosom. Res.* 2004; 12(7):715-723. DOI 10.1023/B:CHRO.0000045779.50641.00

Dos Santos M.S., Kretschmer R., Silva F.A.O., Ledesma M.A., O'Brien P.C.M., Ferguson-Smith M.A., Del Valle Garnero A., de Oliveira E.H.C., Gunski R.J. Intrachromosomal rearrangements in two representatives of the genus *Saltator* (Thraupidae, Passeriformes) and the occurrence of heteromorphic Z chromosomes. *Genetica*. 2015;143(5):535-543. DOI 10.1007/s10709-015-9851-4

Dos Santos M. da S., Kretschmer R., Frankl-Vilches C., Bakker A., Gahr M., O'Brien P.C.M., Ferguson-Smith M.A., De Oliveira E.H.C. Comparative cytogenetics between two important songbird, models: The zebra finch and the canary. *PLoS One*. 2017;12(1):e0170997. DOI 10.1371/journal.pone.0170997

Ebied A.M., Hassan H.A., Abu Almaaty A.H., Yaseen A.E. Karyotypic characterization of ten species of birds. *Cytologia* (*Tokyo*). 2005; 70(2):181-194. DOI 10.1508/cytologia.70.181

Ellegren H. The evolutionary genomics of birds. *Annu. Rev. Ecol. Evol. Syst.* 2013;44:239-259. DOI 10.1146/annurev-ecolsys-110411-160327

Fillon V., Vignoles M., Crooijmans R.P.M.A., Groenen M.A.M., Zoorob R., Vignal A. FISH mapping of 57 BAC clones reveals strong conservation of synteny between Galliformes and Anseriformes. *Anim. Genet.* 2007;38(3):303-307. DOI 10.1111/j.1365-2052.2007.01578.x

Gill F., Donsker D., Rasmussen P. (Eds.). IOC World Bird List (v13.2). 2023. DOI 10.14344/IOC.ML.13.2

Graphodatsky A.S., Yang F., O'Brien P.C.M., Serdukova N., Milne B.S., Trifonov V., Ferguson-Smith M.A. A comparative chromosome map of the Arctic fox, red fox and dog defined by chromosome painting and high resolution G-banding. *Chromosome Res.* 2000;8(3):253-263. DOI 10.1023/A:1009217400140

Graphodatsky A.S., Yang F., O'Brien P.C.M., Perelman P., Milne B.S., Serdukova N., Kawada S.I., Ferguson-Smith M.A. Phylogenetic implications of the 38 putative ancestral chromosome segments for four canid species. *Cytogenet. Cell Genet.* 2001;92(3-4):243-247. DOI 10.1159/000056911

Gregory T.R. The Animal Genome Size Database, 2023

Griffin D.K., Haberman F., Masabanda J., O'Brien P., Bagga M., Sazanov A., Smith J., Burt D.W., Ferguson-Smith M., Wienberg J. Micro- and macrochromosome paints generated by flow cytometry and microdissection: Tools for mapping the chicken genome. *Cytogenet. Cell Genet.* 1999;87(3-4):278-281. DOI 10.1159/000015449

Griffin D.K., Robertson L.B.W., Tempest H.G., Skinner B.M. The evolution of the avian genome as revealed by comparative molecular cytogenetics. *Cytogenet. Genome Res.* 2007;117(1-4):64-77. DOI 10.1159/000103166

Griffin D.K., Robertson L.B., Tempest H.G., Vignal A., Fillon V., Crooijmans R.P.M.A., Groenen M.A.M., Deryusheva S., Gaginskaya E., Carré W., Waddington D., Talbot R., Völker M., Masabanda J.S., Burt D.W. Whole genome comparative studies between chicken and turkey and their implications for avian genome evolution. *BMC Genomics*. 2008;9:168. DOI 10.1186/1471-2164-9-168

Grützner F., Zend-Ajusch E., Stout K., Munsche S., Niveleau A., Nanda I., Schmid M., Haaf T. Chicken microchromosomes are hypermethylated and can be identified by specific painting probes. *Cytogenet. Cell Genet*. 2001;93(3-4):265-269. DOI 10.1159/000056996

Guillier-Gencik Z., Bernheim A., Coullin P. Generation of whole-chromosome painting probes specific to each chicken macrochromosome. *Cytogenet. Cell Genet.* 1999;87(3-4):282-285. DOI 10.1159/000015450

Guttenbach M., Nanda I., Feichtinger W., Masabanda J.S., Griffin D.K., Schmid M. Comparative chromosome painting of chicken autosomal paints 1-9 in nine different bird species. *Cytogenet. Genome Res.* 2003;103(1-2):173-184. DOI 10.1159/000076309

Habermann F.A., Cremer M., Walter J., Kreth G., Von Hase J., Bauer K., Wienberg J., Cremer C., Cremer T., Solovei I. Arrangements of macro- and microchromosomes in chicken cells. *Chromosom. Res.* 2001;9(7):569-584. DOI 10.1023/A:1012447318535

Hammar B.O. The karyotypes of thirty-one birds. *Hereditas*. 1970; 65(1):29-58

Hansmann T., Nanda I., Volobouev V., Yang F., Schartl M., Haaf T., Schmid M. Cross-species chromosome painting corroborates microchromosome fusion during karyotype evolution of birds. *Cytogenet. Genome Res.* 2009;126(3):281-304. DOI 10.1159/000251965

Ijdo J.W., Wells R.A., Baldini A., Reeders S.T. Improved telomere detection using a telomere repeat probe (TTAGGG)n generated by PCR. *Nucleic Acids Res.* 1991;19(17):4780. DOI 10.1093/nar/19.17.4780

Islam F.B., Uno Y., Nunome M., Nishimura O., Tarui H., Agata K., Matsuda Y. Comparison of the chromosome structures between the chickn and three anserid species, the domestic duck (*Anas platyrhynchos*), muscovy duck (*Cairina moschata*), and chinese goose (*Anser cygnoides*), and the delineation of their karyotype evolution by compara. *J. Poult. Sci.* 2014;51(1):1-13. DOI 10.2141/jpsa.0130090

Itoh Y., Arnold A.P. Chromosomal polymorphism and comparative painting analysis in the zebra finch. *Chromosom. Res.* 2005;13(1): 47-56. DOI 10.1007/s10577-005-6602-x

Kiazim L.G., O'Connor R.E., Larkin D.M., Romanov M.N., Narushin V.G., Brazhnik E.A., Griffin D.K. Comparative mapping of the macrochromosomes of eight avian species provides further insight into their phylogenetic relationships and avian karyotype evolution. *Cells*. 2021;10(2):362. DOI 10.3390/cells10020362

Kretschmer R., de Souza M.S., Furo I.O., Romanov M.N., Gunski R.J., Garnero A.D.V., de Freitas T.R.O., de Oliveira E.H.C., O'Connor R.E., Griffin D.K. Interspecies chromosome mapping in caprimulgiformes, piciformes, suliformes, and trogoniformes (Aves): Cytogenomic insight into microchromosome organization and karyotype evolution in birds. *Cells*. 2021;10(4):826. DOI 10.3390/cells10040826

Liehr T., Kreskowski K., Ziegler M., Piaszinski K., Rittscher K. The Standard FISH Procedure. In: Fluorescence In Situ Hybridisation (FISH). Springer Protocols Handbooks. Berlin: Springer, 2017;109-118. DOI 10.1007/978-3-662-52959-1_9

Maden B.E.H., Dent C.L., Farrell T.E., Garde J., McCallum F.S., Wakeman J.A. Clones of human ribosomal DNA containing the complete 18 S-rRNA and 28 S-rRNA genes. Characterization, a detailed map of the human ribosomal transcription unit and diversity among clones. *Biochem. J.* 1987;246(2):519-527. DOI 10.1042/bj2460519

Masabanda J.S., Burt D.W., O'Brien P.C.M., Vignal A., Fillon V., Walsh P.S., Cox H., Tempest H.G., Smith J., Habermann F., Schmid M., Matsuda Y., Ferguson-Smith M.A., Crooijmans R.P.M.A., Groenen M.A.M., Griffin D.K. Molecular cytogenetic definition of the chicken genome: The first complete avian karyotype. *Genetics*. 2004;166(3):1367-1373. DOI 10.1534/genetics.166.3.1367

Nakatani Y., Takeda H., Kohara Y., Morishita S. Reconstruction of the vertebrate ancestral genome reveals dynamic genome reorganization in early vertebrates. *Genome Res.* 2007;17(9):1254-1265. DOI 10.1101/gr.6316407

Nanda I., Schrama D., Feichtinger W., Haaf T., Schartl M., Schmid M. Distribution of telomeric (TTAGGG)n sequences in avian chromosomes. *Chromosoma*. 2002;111(4):215-227. DOI 10.1007/s00412-002-0206-4

Nanda I., Benisch P., Fetting D., Haaf T., Schmid M. Synteny conservation of chicken macrochromosomes 1–10 in different avian lineages revealed by cross-species chromosome painting. *Cytogenet. Genome Res.* 2011;132(3):165-181. DOI 10.1159/000322358

Nie W., O'Brien P.C.M., Ng B.L., Fu B., Volobouev V., Carter N.P., Ferguson-Smith M.A., Yang F. Avian comparative genomics: Reciprocal chromosome painting between domestic chicken (*Gallus gallus*) and the stone curlew (*Burhinus oedicnemus*, Charadriiformes) – An atypical species with low diploid number. *Chromosom. Res.* 2009;17(1):99-113. DOI 10.1007/s10577-009-9021-6

Nie W., O'Brien P.C.M., Fu B., Wang J., Su W., He K., Bed'Hom B., Volobouev V., Ferguson-Smith M.A., Dobigny G., Yang F. Multidi-

rectional chromosome painting substantiates the occurrence of extensive genomic reshuffling within Accipitriformes. *BMC Evol. Biol.* 2015;15(1):205. DOI 10.1186/s12862-015-0484-0

Nishida-Umehara C., Tsuda Y., Ishijima J., Ando J., Fujiwara A., Matsuda Y., Griffin D.K. The molecular basis of chromosome orthologies and sex chromosomal differentiation in palaeognathous birds. *Chromosom. Res.* 2007;15(6):721-734. DOI 10.1007/s10577-007-1157-7

O'Connor R.E., Kretschmer R., Romanov M.N., Griffin D.K. A bird's-eye view of chromosomic evolution in the class aves. *Cells*. 2024; 13(4):310. DOI 10.3390/cells13040310

Pala I., Naurin S., Stervander M., Hasselquist D., Bensch S., Hansson B. Evidence of a neo-sex chromosome in birds. *Heredity* (*Edinb.*). 2012;108(3):264-272. DOI 10.1038/hdy.2011.70

Prum R.O., Berv J.S., Dornburg A., Field D.J., Townsend J.P., Lemmon E.M., Lemmon A.R. A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. *Nature*. 2015;526(7574):569-573. DOI 10.1038/nature15697

Rodrigue K.L., May B.P., Famula T.R., Delany M.E. Meiotic instability of chicken ultra-long telomeres and mapping of a 2.8 megabase array to the W-sex chromosome. *Chromosom. Res.* 2005;13(6):581-591. DOI 10.1007/s10577-005-0984-7

Rodrigues B.S., de Assis M.D.F.L., O'Brien P.C.M., Ferguson-Smith M.A., De Oliveira E.H.C. Chromosomal studies on *Coscoroba coscoroba* (Aves: Anseriformes) reinforce the Coscoroba-Cereopsis clade. *Biol. J. Linn. Soc.* 2014;111(2):274-279. DOI 10.1111/bij.12202

Romanenko S.A., Biltueva L.S., Serdyukova N.A., Kulemzina A.I., Beklemisheva V.R., Gladkikh O.L., Lemskaya N.A., Interesova E.A., Korentovich M.A., Vorobieva N.V., Graphodatsky A.S., Trifonov V.A. Segmental paleotetraploidy revealed in sterlet (*Acipenser ruthenus*) genome by chromosome painting. *Mol. Cytogenet.* 2015;8(1):90. DOI 10.1186/s13039-015-0194-8

Rutkowska J., Lagisz M., Nakagawa S. The long and the short of avian W chromosomes: No evidence for gradual W shortening. *Biol. Lett.* 2012;8(4):636-638. DOI 10.1098/rsbl.2012.0083

Schartl M., Schmid M., Nanda I. Dynamics of vertebrate sex chromosome evolution: from equal size to giants and dwarfs. *Chromosoma*. 2016;125(3):553-571. DOI 10.1007/s00412-015-0569-y

Schmid M., Nanda I., Hoehn H., Schartl M., Haaf T., Buerstedde J.M., Arakawa H., Caldwell R.B., Weigend S., Burt D.W., Smith J., Griffin D.K., Masabanda J.S., Groenen M.A.M., Crooijmans R.P.M.A., Vignal A., Fillon V., Morisson M., Pitel F., Vignoles M., Garrigues A., Gellin J., Rodionov A.V., Galkina S.A., Lukina N.A., Ben-Ari G., Blum S., Hillel J., Twito T., Lavi U., David L., Feldman M.W., Delany M.E., Conley C.A., Fowler V.M., Hedges S.B., Godbout R., Katyal S., Smith C., Hudson Q., Sinclair A., Mizuno S. Second report on chicken genes and chromosomes 2005. *Cytogenet. Genome Res.* 2005;109(4):415-479. DOI 10.1159/000084205

Seabright M. A rapid banding technique for human chromosomes. *Lancet*. 1971;298(7731):971-972. DOI 10.1016/s0140-6736(71)90287-x

Shetty S., Griffin D.K., Graves J.A.M. Comparative painting reveals strong chromosome homology over 80 million years of bird evolution. *Chromosom. Res.* 1999;7(4):289-295. DOI 10.1023/A:1009278914829

Shibusawa M., Minai S., Nishida-Umehara C., Suzuki T., Mano T., Yamada K., Namikawa T., Matsuda Y. A comparative cytogenetic study of chromosome homology between chicken and Japanese quail. *Cytogenet. Cell Genet.* 2001;95(1-2):103-109. DOI 10.1159/000057026

Shibusawa M., Nishida-Umehara C., Masabanda J., Griffin D.K., Isobe T., Matsuda Y. Chromosome rearrangements between chicken and guinea fowl defined by comparative chromosome painting and FISH mapping of DNA clones. *Cytogenet. Genome Res.* 2002; 98(2-3):225-230. DOI 10.1159/000069813

Shibusawa M., Nishibori M., Nishida-Umehara C., Tsudzuki M., Masabanda J., Griffin D.K., Matsuda Y. Karyotypic evolution in the Galliformes: An examination of the process of karyotypic evolution by comparison of the molecular cytogenetic findings with the molecular phylogeny. *Cytogenet. Genome Res.* 2004;106(1):111-119. DOI 10.1159/00007857

Smith J., Bruley C.K., Paton I.R., Dunn I., Jones C.T., Windsor D., Morrice D.R., Law A.S., Masabanda J., Sazanov A., Waddington D., Fries R., Burt D.W. Differences in gene density on chicken macrochromosomes and microchromosomes. *Anim. Genet.* 2000;31(2): 96-103. DOI 10.1046/j.1365-2052.2000.00565.x

Srikulnath K., Ahmad S.F., Singchat W., Panthum T. Why do some vertebrates have microchromosomes? *Cells*. 2021;10(9):2182. DOI 10.3390/cells10092182

Stiglec R., Ezaz T., Graves J.A.M. A new look at the evolution of avian sex chromosomes. *Cytogenet. Genome Res.* 2007;117(1-4):103-109. DOI 10.1159/000103170

Sumner A.T. A simple technique for demonstrating centromeric heterochromatin. *Exp. Cell Res.* 1972;75(1):304-306. DOI 10.1016/0014-4827(72)90558-7

Sun Z., Pan T., Hu C., Sun L., Ding H., Wang H., Zhang C., Jin H., Chang Q., Kan X., Zhang B. Rapid and recent diversification patterns in Anseriformes birds: Inferred from molecular phylogeny and diversification analyses. *PLoS One*. 2017;12(9):e0184529. DOI 10.1371/journal.pone.0184529

Takagi N., Makino S. A revised study on the chromosomes of three species of birds. *Caryologia*. 1966;19(4):443-455. DOI 10.1080/00087114.1966.10796235

Telenius H., Pelmear A.H., Tunnacliffe A., Carter N.P., Behmel A., Ferguson-Smith M.A., Nordenrkjold M., Pfragner R., Ponder B.A.J. Cytogenetic analysis by chromosome painting using DOP-PCR amplified flow-sorted chromosomes. *Genes Chromosom. Cancer*. 1992;4(3):251-263. DOI 10.1002/gcc.2870040311

Uno Y., Nishida C., Hata A., Ishishita S., Matsuda Y. Molecular cytogenetic characterization of repetitive sequences comprising centromeric heterochromatin in three Anseriformes species. *PLoS One*. 2019;14(3):e0214028. DOI 10.1371/journal.pone.0214028

Van Tuinen M., Hedges S.B. Calibration of avian molecular clocks. *Mol. Biol. Evol.* 2001;18(2):206-213. DOI 10.1093/oxfordjournals.molbev.a003794

Wang J., Su W., Hu Y., Li S., O'Brien P.C.M., Ferguson-Smith M.A., Yang F., Nie W. Comparative chromosome maps between the stone curlew and three ciconiiform species (the grey heron, little egret and crested ibis). *BMC Ecol. Evol.* 2022;22(1):23. DOI 10.1186/s12862-022-01979-x

Wójcik E., Smalec E. Description of the mallard duck (*Anas platyrhynchos*) karyotype. *Folia Biol.* 2007;55(3-4):115-120

Zimmer R., King W.A., Verrinder Gibbins A.M. Generation of chicken Z-chromosome painting probes by microdissection for screening large-insert genomic libraries. *Cytogenet. Cell Genet.* 1997;78(2): 124-130. DOI 10.1159/000134643

Zlotina A., Maslova A., Kosyakova N., Al-Rikabi A.B.H., Liehr T., Krasikova A. Heterochromatic regions in Japanese quail chromosomes: comprehensive molecular-cytogenetic characterization and 3D mapping in interphase nucleus. *Chromosom. Res.* 2019;27(3):253-270. DOI 10.1007/s10577-018-9597-9

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

Received February 21, 2024. Revised June 13, 2024. Accepted August 8, 2024.

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2024;28(7):770-779

DOI 10.18699/vjgb-24-85

# Transcription factor TCF4: structure, function, and associated diseases

R.R. Savchenko, N.A. Skryabin

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

**Abstract.** Our understanding of human genes – particularly their structure, functions, and regulatory mechanisms – is still limited. The biological role of approximately 20 % of human proteins has not been established yet, and the molecular functions of the known part of the proteome remain poorly understood. This hinders progress in basic and applied biological and medical sciences, especially in treating hereditary diseases, which are caused by mutations and polymorphic variants in individual genes. Therefore, it is crucial to comprehend the mechanisms of protein functioning to address this problem. This further emphasizes the importance of investigating gene functions and molecular pathogenetic pathways associated with single-gene inherited diseases. This review focuses on the *TCF4* gene that encodes a transcription factor crucial for nervous system development and functioning. Pathogenic variants in this gene have been linked to a rare genetic disorder, Pitt–Hopkins syndrome, and *TCF4* polymorphic variants are associated with several socially significant diseases, including various psychiatric disorders. The pathogenetic mechanisms of these conditions remain unexplored, and the knowledge about *TCF4* upregulation and its target genes is limited. TCF4 can be expressed in various isoforms due to the complex structure and regulation of its gene, which complicates the investigation of the protein's functions. Here, we consider the structure and functions of the TCF4 transcription factor. We discuss its potential target genes and the possible loss-of-function pathogenetic mechanisms identified in animal and cellular models of Pitt–Hopkins syndrome. The review also examines the advantages and limitations of potential therapies for Pitt–Hopkins syndrome that are based on TCF4 dosage compensation or altering the activity of TCF4 target genes.

**Key words:** TCF4; Pitt–Hopkins syndrome; bHLH; mental disorders; autism spectrum disorders; Pitt–Hopkins syndrome therapy.

**For citation:** Savchenko R.R., Skryabin N.A. Transcription factor TCF4: structure, function, and associated diseases. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2024;28(7):770-779. DOI 10.18699/vjgb-24-85

**Funding.** The study was supported by the Russian Science Foundation (project No. 23-75-01138).

# Транскрипционный фактор TCF4: структура, функции и ассоциированные заболевания

Р.Р. Савченко, Н.А. Скрябин

Научно-исследовательский институт медицинской генетики, Томский национальный исследовательский медицинский центр Российской академии наук, Томск, Россия
renata.savchenko@medgenetics.ru

**Аннотация.** На сегодняшний день имеются ограниченные знания об основных характеристиках генов человека, их структуре, функции и механизмах регуляции экспрессии. Биологическая роль около 20 % белковых продуктов генов до сих пор не установлена, а молекулярные функции известной части протеома остаются недостаточно изученными. Данное обстоятельство ограничивает прогресс как фундаментальных, так и прикладных биологических и медицинских наук, в особенности в случае терапии наследственных болезней, патогенез которых обусловлен наличием вариантов в нуклеотидной последовательности отдельных генов. В связи с этим возрастает необходимость проведения исследований, направленных на изучение функций генов, а также молекулярных патогенетических путей, связанных с развитием моногенных заболеваний. Наша статья посвящена гену *TCF4*, кодирующему широко экспрессируемый фактор транскрипции, важный для развития и функционирования нервной системы. К настоящему времени установлено, что патогенные варианты в этом гене приводят к развитию редкого генетического заболевания, известного как синдром Питта–Хопкинса, а полиморфные варианты в *TCF4* ассоциированы с рядом социально значимых заболеваний, представленных различными психическими расстройствами. Молекулярные механизмы патогенеза подобных состояний по-прежнему остаются неизученными, а знания о вышестоящей регуляции TCF4 и его нижестоящих генах-мишенях ограничены. Сложность структурной организации и особенности регуляции экспрессии гена обеспечивают многообразие изоформ TCF4, что затрудняет понимание молекулярных функций белка. В обзоре рассмотрены известные данные

о структуре и функциях фактора транскрипции TCF4. Обсуждаются потенциальные гены-мишени и возможные патогенетические механизмы, обусловленные потерей функции этого белка, выявленные в исследованиях на животных и клеточных моделях синдрома Питта–Хопкинса. Рассмотрены преимущества и ограничения потенциальных стратегий терапии указанного синдрома, основанные на компенсации дозы TCF4 или воздействии на молекулярные мишени изучаемого транскрипционного фактора.

**Ключевые слова:** TCF4; синдром Питта–Хопкинса; bHLH; психические расстройства; расстройства аутистического спектра; терапия синдрома Питта–Хопкинса.

## Introduction

One of the most important problems in medical genetics today is the limited understanding of the role of proteins involved in the molecular pathways underlying the development of several inherited disorders. This problem is especially urgent regarding transcription factors, since these proteins regulate the expression of many genes, have pleiotropic effects, and are critical for various biological processes. One such transcription factor is encoded by the *TCF4* gene. Pathogenic variants in this gene are responsible for Pitt–Hopkins syndrome development (Amiel et al., 2007; Brockschmidt et al., 2007; Zweier et al., 2007), and its polymorphic variants have been associated with various psychiatric disorders including schizophrenia, bipolar disorder, major depressive disorder, and post-traumatic stress disorder (Stefansson et al., 2009; Smoller et al., 2013; Wray et al., 2018).

According to the literature, TCF4 is critical for brain development and function, as it participates in nerve cell differentiation and migration, regulation of neuronal excitability, neuronal plasticity, etc. (Imayoshi, Kageyama, 2014; Kennedy et al., 2016; Li H. et al., 2019; Mesman et al., 2020; Phan et al., 2020). Despite the large amount of data pointing to the important role of TCF4, the molecular mechanisms of its pathogenic variants leading to impaired development and function of cells in the nervous system remain largely unexplored. This review aims to summarize the literature data on TCF4 structure and function, its known molecular targets, diseases associated with variants in *TCF4*, and potential approaches to their therapy.

## TCF4 structure, expression pattern and known functions

The *TCF4* gene, also known as *ITF2* or *PTHS*, is located in the 18q21.2 region on chromosome 18 and contains 41 exons. Twenty of these exons are alternative 5′-exons (non-protein-coding exons that are included or excluded from mature mRNA to regulate the expression of longer or shorter protein isoforms), 20 exons are internal protein-coding, and one is a 3′-non-coding exon (Sepp et al., 2011). *TCF4* encodes a transcription factor containing a basic helix-loop-helix structural motif (bHLH). The proteins of this group contain a DNA-binding domain and can regulate gene expression, forming homo- and heterodimers. The bHLH-containing proteins are categorized into six groups depending on the types of dimers they form, expression pattern, and DNA binding specificity. The transcription factor TCF4 belongs to the E-proteins group (or class I bHLH proteins) that recognize E-box sequences (CANNTG) located in the promoter and enhancer regions of target genes (Schoof et al., 2020). The bHLH domain is a highly conserved motif consisting of a basic amino acid region followed by two amphipathic α-helixes connected by a loop. The basic amino acid region binds to the E-box sequence directly, while the α-helices provide dimerization.

The alternative transcription initiation sites located upstream of non-coding exons 1, 3, 4, 5, 7, 8, and 10 define at least 18 TCF4 isoforms. The isoforms contain relatively conserved C-terminal domains of the basic helix-loop-helix structural motif but differ in N-terminal regions responsible for transcription regulation. However, it should be noted that the diversity of *TCF4* transcripts is even higher due to alternative splicing of internal coding exons (Teixeira et al., 2021).

Full-length *TCF4* transcripts include the following structural elements: the bHLH domain, activation domains (AD1, AD2 and AD3), CE and Rep intramolecular regulatory domains, NLS-1 and NLS-2 nuclear localization signals, NES-1 and NES-2 nuclear export signals, and the RSRS sequence of four amino-acid residues (Fig. 1).

In addition to the bHLH domain, the AD1, AD2, and AD3 activation domains can cooperatively or independently regulate gene expression in a cell type-dependent manner. It has been shown that the AD1 domain can bind transcription coactivators and corepressors. The AD2 domain can bind coactivators of transcription, but no data are currently available on its interactions with corepressors. Transcription coactivators and corepressors compete for binding to the AD1 domain, enabling TCF4 to both activate and repress gene expression (Teixeira et al., 2021).

The AD3 domain interacts directly with the TAF4 subunit of the general transcription factor II D, resulting in enhanced RNA polymerase II preinitiation on target genes, but exactly how AD3 is involved in the regulation of gene expression by TCF4 is currently unclear (Teixeira et al., 2021).

The *TCF4* transcriptional activity is also regulated by the previously mentioned CE and Rep domains. The conserved element (CE) located between the activation domains of AD1 and AD3 can inhibit AD1 activity. The repression domain (Rep), located between AD2 and bHLH, is able to repress the activity of AD1 and AD2. Both of these domains can likely prevent the recruitment of transcriptional cofactors, and as a result, suppress AD1-mediated activation or repression of transcription (Teixeira et al., 2021).

Finally, the presence of the RSRS four amino-acid residue motif (Arg-Ser-Arg-Ser) located between the Rep and bHLH domains may also result in reduced transcriptional activity.

The complex structural organization of *TCF4*, along with the peculiar regulation of its expression, results in a variety of TCF4 isoforms containing different structural domains. Since all *TCF4* transcripts include exons 10–20, all isoforms

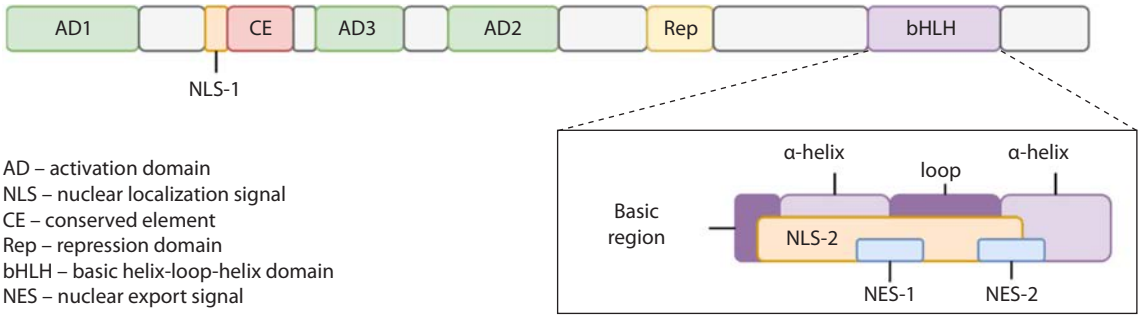



**Fig. 1.** Schematic representation of the full-length TCF4 protein structure.
The figure is modified from J.R. Teixeira and colleagues (Teixeira et al., 2021).

of the encoded protein contain the AD2 and AD3 activation domains, as well as the bHLH, Rep, NLS-2, NES-1, and NES-2 domains. Only the four longer protein isoforms contain the full AD1 activation domain, while the other isoforms contain either only a part of it or none at all (Sepp et al., 2011). In addition, the literature describes "Δ-isoforms", which are characterized by the absence of NLS-1 and CE domains. Finally, the presence of alternative splicing sites in exon 18 leads to the inclusion or exclusion of the segment encoding the RSRS sequence present in positive (+) isoforms and absent in negative (–) isoforms of the protein (Sepp et al., 2011). How various isoforms differ from each other in terms of transcription regulation remains an open question.

Similar to the majority of genes encoding E-proteins, *TCF4* is expressed in almost all tissues of the organism, being the most abundant in the brain (The Human Protein Atlas, https://www.proteinatlas.org). Some *TCF4* transcripts are tissue-specific, while others have a broad spatial expression pattern. Moreover, the quantitative ratios of the same transcripts can vary in different tissues. The expression analysis using RT-PCR showed that most *TCF4* transcripts were expressed in the brain, except for the five found in the testes (Sepp et al., 2011). *TCF4* expression is also upregulated during ontogenesis, with the highest activity during prenatal development (Sepp et al., 2011). It has been shown that *TCF4* expression in the brain increases toward the end of the prenatal period and then decreases to baseline in newborns, persisting throughout life (Li M. et al., 2018).

In humans, TCF4 is expressed in the forebrain and brain ventricular system during fetal development and persists in the forebrain and cerebellum in adults. Besides, TCF4 is found in oligodendrocytes of the spinal cord (Chen H.Y. et al., 2021).

The variety of TCF4 isoforms complicates the understanding of its molecular functions. The functions of specific TCF4 isoforms depend on which 5′-exon and internal exons are included in the translated transcript. Depending on the isoform structure, both subcellular localization and transcription are differentially regulated. For example, isoforms containing NLS are localized in the nucleus, whereas isoforms lacking NLS require a heterodimerization partner to access the nucleus (Chen H.Y. et al., 2021).

As a transcription factor, TCF4 has been associated with the regulation of hematopoiesis, myogenesis, neurogenesis, melanogenesis, osteogenesis, and the differentiation of endothelial, mammary, and Sertoli cells (Teixeira et al., 2021). Additionally, TCF4 appears critical for normal nervous system development and function. This protein forms heterodimers with the transcription factors ATOH1, ASCL1, NEUROD1, and NEUROD2, which play an important role in nervous system development (Wittmann, Häberle, 2018). TCF4 is known to be important for brain development and functioning: it participates in such processes as the differentiation of neuronal progenitor cells into neurons, oligodendrocyte and astrocyte (Imayoshi, Kageyama, 2014), maturation, neuronal migration and function, oligodendrocyte myelination, synaptic plasticity, etc. (Kennedy et al., 2016; Li H. et al., 2019; Mesman et al., 2020; Phan et al., 2020).

## *TCF4*-associated diseases

To date, a number of studies have indicated a possible role for *TCF4* in the pathogenesis of various socially important diseases. Genome-wide association studies show that polymorphic variants in *TCF4*, predominantly localized in non-coding regions of the gene, are associated with various psychiatric disorders, including schizophrenia (Stefansson et al., 2009; Ripke et al., 2011; Steinberg et al., 2011; Smoller et al., 2013; Bocharova et al., 2017), bipolar disorder and autism spectrum disorders (Smoller et al., 2013), major depressive disorder (Wray et al., 2018), and post-traumatic stress disorder (Gelernter et al., 2019). In addition, variants in *TCF4* are associated with Fuchs' corneal endothelial dystrophy (Afshari et al., 2017; Fautsch et al., 2021) and sclerosing cholangitis (Ellinghaus et al., 2013). However, it is currently unknown whether these polymorphic variants are responsible for the development of these diseases. The exception is Fuchs' endothelial corneal dystrophy: most patients with this diagnosis carry an expansion of the trinucleotide repeat (CTG)n in intron 3 of the *TCF4* gene, leading to splicing errors (Du et al., 2015; Papanyan et al., 2019) (see the Table).

In 2007, several independent studies showed that heterozygous carriage of pathogenic variants in the *TCF4* gene leads to the development of a rare inherited disease, Pitt–Hopkins syndrome (PTHS) (Amiel et al., 2007; Brockschmidt et al., 2007; Zweier et al., 2007). Despite phenotypic differences, most patients with this syndrome are characterized by a specific set of dysmorphic facial features combined with in-

Diseases associated with polymorphic variants in *TCF4*

| Variants | Frequency | OR (95 % CI) | Disease | Reference |
|---|---|---|---|---|
| rs9960767[C] | 0.06 | 1.3 (1.11–1.51), $p = 0.001$ | Schizophrenia | Stefansson et al., 2009 |
| rs17512836[C] | 0.02 | 1.23 (1.14–1.31), $p = 1.05 \times 10^{-6}$ | Schizophrenia, BD, ASD | Ripke et al., 2011; Smoller et al., 2013 |
| rs4309482[A] | 0.58 | 1.09 (1.06–1.12), $p = 7.8 \times 10^{-9}$ | Schizophrenia | Steinberg et al., 2011 |
| rs9960767[C] | 0.03 | 0.68 (0.41–1.13), $p = 0.134$ | | Bocharova et al., 2017 |
| rs17594526[T] | 0.01 | 0.60 (0.25–1.42), $p = 0.238$ | | Bocharova et al., 2017 |
| rs12958048 [A] | 0.33 | 1.03 (nd), $p < 1 \times 10^{-5}$ | Major depressive disorder | Wray et al., 2018 |
| rs2123392 [C] | nd | nd | PTSD | Gelernter et al., 2019 |
| rs613872 [G] | 0.37 | 5.47 (3.75–7.99), $p = 1 \times 10^{-18}$ | Fuchs' endothelial dystrophy | Baratz et al., 2010 |
| rs784257[G] | 0.48 | 4.94 (4.45–5.58), $p = 2.5 \times 10^{-200}$ | | Afshari et al., 2017 |
| rs1452787[G] | 0.23 | 0.75 (0.68–0.83), $p = 2.61 \times 10^{-8}$ | Sclerosing cholangitis | Ellinghaus et al., 2013 |

Note. OR – odds ratio; CI – confidence interval; BD – bipolar disorder; ASD – autism spectrum disorders; PTSD – post-traumatic stress disorder; nd – no data.

tellectual disability, sensorimotor impairment, speech delay, and generalized muscular hypotonia. About 78 % of patients frequently perform stereotypical and intense repetitive movements, which allows PTHS to be classified as an autism spectrum disorder. Approximately half of patients with PTHS have abnormal breathing patterns and about one-third develop epileptic seizures. In addition, magnetic resonance imaging has identified several brain anomalies in patients with PTHS, including agenesis of the corpus callosum, large ventricles, and an abnormal shape of the posterior cranial fossa (Teixeira et al., 2021).

The spectrum of *TCF4* mutations identified in patients with PTHS includes missense (~15 % of cases), nonsense (~15 %), splicing site mutations (~10 %), small insertions or deletions resulting in frame shifts (~30 %), translocations and large deletions encompassing *TCF4* partially or fully (~30 %) (Teixeira et al., 2021). Some estimates put the worldwide prevalence of PTHS caused by chromosomal deletions at 1/34,000–1/41,000 (Rosenfeld et al., 2009).

Depending on the localization and mutation type, TCF4 isoforms are affected differently. The majority of missense mutations affect exon 19 encoding the bHLH domain. Certain missense mutations affect exons 15 and 18 encoding the AD2 activation domain and the Rep regulatory domain, respectively. Since all *TCF4* transcripts contain these exons, the pathogenic variants described lead to the disruption of all protein isoforms. Most nonsense, frame-shift, and splice-site mutations also result in damage to all isoforms of the protein. However, if these mutations occur in exons 8 and 9 or are localized upstream of exons 10a-c, the Δ-isoforms and shorter TCF4 isoforms are unchanged. Several translocations and deletions span only initial exons (1 through 4) or inner exons (5 through 9), retaining intermediate and shorter isoforms, respectively (Sepp et al., 2012).

The effects of the structural diversity and cell-specific TCF4 expression pattern on physiologic processes remain poorly understood. However, it is hypothesized that different types of *TCF4* mutations in individuals with PTHS may impair the functions of the encoded protein by diverse mechanisms and to a varying extent, thus leading to the phenotypic variability observed among patients (Bedeschi et al., 2017). For example, missense mutations in the bHLH motif or insertions elongating the reading frame can damage DNA-binding or transactivation functions in a manner dependent on dimer context (Sepp et al., 2012). Pathogenic variants encompassing the bHLH domain responsible for dimerization destabilize the protein, whereas missense mutations outside of the bHLH domain cause no major functional deficiencies (Chen H.Y. et al., 2021).

The majority of the pathogenic variants in *TCF4* found in patients with PTHS lead to haploinsufficiency because they restrict the expression of certain or all transcripts to a single copy of the allele. In addition, certain missense mutations cause the attenuation or loss of TCF4 function as a transcription regulator without affecting its ability to dimerize *in vitro*, which seems to indicate a dominant-negative effect (Forrest et al., 2013). Whether the observed effect occurs *in vivo* is currently unclear. Presumably, it would be weak due to dimer instability with mutant TCF4 (Teixeira et al., 2021). Thus, it is evident that PTHS results from the dysregulation of TCF4-mediated gene expression. How such disturbances can trigger a pathophysiologic process remains unclear. J.R. Teixeira et al. (2021) suggest this process may be related to the general functions of E-proteins during the regulation of the cell cycle and to the specific role of TCF4 in cell differentiation.

## Molecular pathways and potential target genes regulated by TCF4

To date, there has been progress in identifying upstream regulators and target genes of the TCF4 transcription factor. K.M. Henning et al. showed that the pharmacological activation of the WNT/β-catenin signaling pathway in induced pluripotent stem cells (iPSCs), derived from neural progenitor

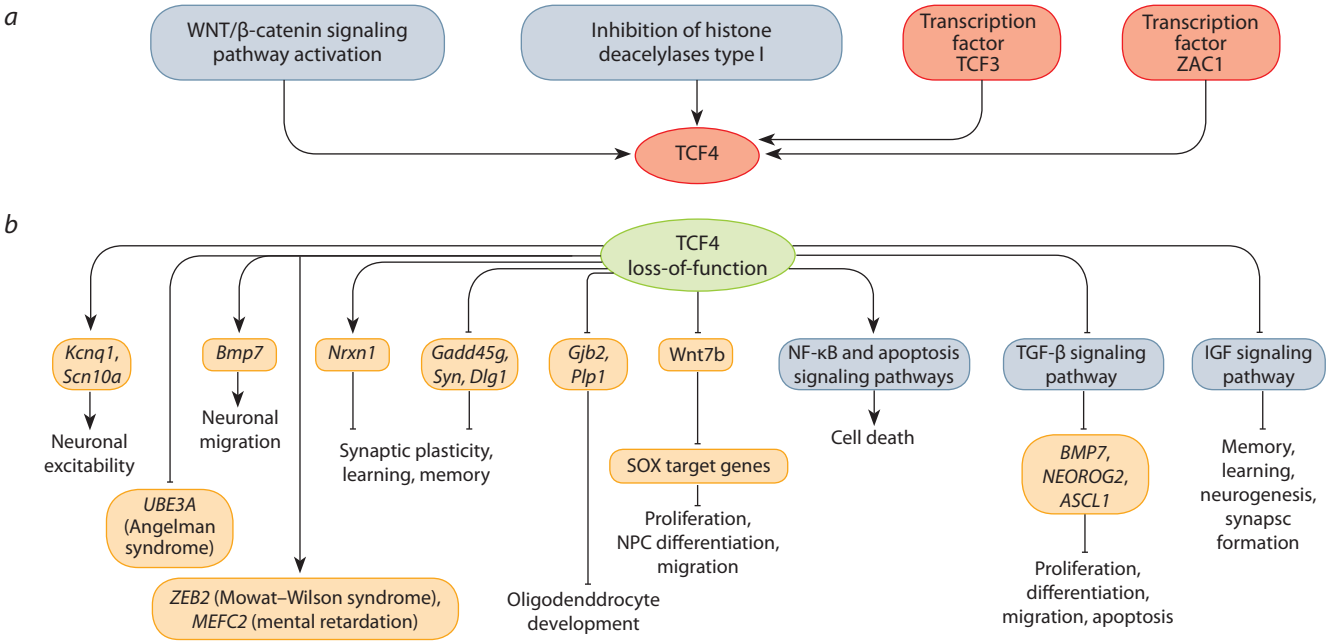



**Fig. 2.** Upregulation of TCF4 and its potential molecular targets.

*a* – upregulation of TCF4; *b* – molecular pathways and potential target genes regulated by TCF4. NPC – neural progenitor cells. Pointed-end arrows indicate activation and blunt-end arrows indicate inhibition.

cells and neurons from patients with PTHS, leads to increased *TCF4* expression (Hennig et al., 2017). Chromatin modification mediated by the inhibition of class I histone deacetylases has a similar effect (Kennedy et al., 2016; Hennig et al., 2017). TCF3, a member of the E-protein subgroup, and the ZAC1 transcription factor also upregulate *Tcf4* expression (Schmidt-Edelkraut et al., 2014; Li H. et al., 2019). The authors suggest that *Tcf4* regulation by TCF3 and other unidentified transcription factors is crucial for normal cortical development (Li H. et al., 2019) (Fig. 2).

Using ChIP-Seq technology, several studies have identified direct TCF4 targets, including Bmp7 (Chen T. et al., 2016), Nrxn1 (D'Rozario et al., 2016), Gadd45g (Sepp et al., 2017), Gjb2, and Plp1 (Wedel et al., 2020). Cellular and animal models have helped identify the following molecular targets of TCF4: *Scn10a* (Nav1.8) and *Kcnq1* (Kv7.1) (Rannals et al., 2016; Martinowich et al., 2022), *Wnt7b* (Wang et al., 2020), *Gadd45g* (Tamberg et al., 2020), *Syn* and *Dlg1* (Tamberg et al., 2020). Collectively, these studies indicate that TCF4 regulates genes involved in brain development, nerve cell differentiation, neuronal excitability, synapse function, and survival (Fig. 2).

However, according to the literature, TCF4 has more than ten thousand binding sites in the genome, potentially connected to more than five thousand genes (Forrest et al., 2018; Xia et al., 2018). In light of these data, it is clear that the vast majority of molecular targets of this transcription factor remain unidentified. Therefore, identifying molecular pathways and target genes regulated by TCF4 is essential for fundamental research of the processes controlled by transcription factors. Moreover, understanding of the mechanisms of the gene network function and identification of the key molecular targets of TCF4 may significantly influence the development of therapeutic strategies for *TCF4*-associated diseases.

To date, several studies using animal models have assessed changes in the transcriptional profile caused by *TCF4* mutations. The genes encoding potassium and sodium ion channels, *Kcnq1* and *Scn10a*, are identified as downstream targets of TCF4 in rodent models (Rannals et al., 2016; Martinowich et al., 2022). Both studies demonstrate overexpression of these genes coupled with the loss of TCF4 function, which allows to consider this transcription factor as a regulator of neuronal excitability. Other studies have reported downregulation of the *Arc* gene, which is important for synaptic plasticity, information processing, and memory (Kennedy et al., 2016), and the *Wnt7b* gene, which is considered a key TCF4 target in the regulation of neuronal progenitor cell migration during dentate gyrus development (Wang et al., 2020) (Fig. 2).

Several studies demonstrate that mice with *Tcf4* mutations are characterized by an increased expression of genes associated with neuronal progenitor cell proliferation and a suppressed expression of genes involved in neuronal differentiation and migration (Li H. et al., 2019), neurogenesis, and neuronal maturation (Mesman et al., 2020). B.D.N. Phan et al. (2020) also reported an abnormal gene expression pattern in oligodendrocytes, in particular the genes involved in myelination, which is critical for the normal function of these cells (Fig. 2).

Thus, animal research suggests that TCF4 is crucial for brain development and function and identifies potential targets of this transcription factor. However, animal model systems have significant limitations when it comes to extrapolating their results to humans. For instance, TCF4 haploinsufficiency is known to result in clinical manifestations

of PTHS in patients, whereas heterozygous mice carrying *Tcf4* mutations (*wt/Tcf4*$^{-}$) tend to exhibit milder phenotypes (Thaxton et al., 2018; Li H. et al., 2019; Mesman et al., 2020; Wang et al., 2020). These differences appear to be due to significant differences between the structure and development of rodent and human brains, which should be taken into account. This fact highlights the need for research to be conducted on human nerve cells.

In the study by F. Papes et al. (2022), fibroblasts derived from patients with PTHS were reprogrammed into iPSCs with subsequent differentiation into neural progenitor cells, neurons, and brain organoids. The authors showed that neuronal progenitor cells with *TCF4* mutations were characterized by reduced proliferation and impaired neuron differentiation, while brain organoids were characterized by abnormal size and cellular composition (Papes et al., 2022). Based on the RNA sequencing results, the authors suggest that the loss of TCF4 function leads to disruptions of the Wnt signaling pathway, and, as a result, a decreased expression of SOX target genes, ultimately leading to the reduced proliferation of progenitor cells (Papes et al., 2022). The rescue of *TCF4* expression or pharmacological correction of the Wnt signaling pathway resulted in a partial recovery of aberrant phenotypes. These data indicate possible therapeutic strategies for *TCF4*-associated genetic disorders.

Several studies using SH-SY5Y to model TCF4 dysfunction are found in the literature and provide valuable insight into the molecular mechanisms regulated by this transcription factor. The microarray analysis of the transcriptional profile performed in the SH-SY5Y cells with *TCF4* knockdown revealed differentially expressed genes (DEGs) involved in the TGFβ signaling pathway, epithelial-mesenchymal transition, neuronal differentiation, and apoptosis (Forrest et al., 2013). The genes encoding EMT, SNAI2, and DEC1 transcription factors, as well as *NEUROG2* and *ASCL1* proneural genes, and genes associated with intellectual disability, such as *UBE3A* (Angelman syndrome), *ZEB2* (Mowat–Wilson syndrome), were characterized by the most pronounced differential expression. The findings suggest that TCF4 regulates several molecular pathways associated with nerve cell differentiation and survival, as well as genes clinically significant to the pathogenesis of intellectual disability.

In another study, H. Xia et al. applied ChIP-seq technology in SH-SY5Y cells to analyze DNA-protein interactions and detect TCF4-binding sites (Xia et al., 2018). This approach has identified more than 10,000 binding sites that can be attributed to more than 5,500 genes. The gene set enrichment analysis (GSEA) of potential target genes revealed the pathways associated with neuronal development and identified genes that overlap with those underexpressed *postmortem* in the brains of patients with schizophrenia. These data further support the importance of TCF4 for brain development and function and indicate the existence of pathogenetic molecular pathways common to PTHS and schizophrenia (Xia et al., 2018).

Thus, studies conducted in animal models of PTHS have identified variability in the phenotypes that provide important biological information about this disorder. The phenotypes described above are observed throughout life, ranging from abnormalities in cortical development and nerve cell differentiation and maturation to impairments in neuronal excitability, synaptic plasticity, and behavior in adult animals. Although the analyses of transcriptional profiles using microarray and RNA sequencing do not point directly to TCF4 target genes, they emphasize the important role of this transcription factor in neurogenesis and demonstrate a large-scale gene network potentially regulated by TCF4. Understanding of the molecular pathways and identification of TCF4 target genes are crucial for comprehending the pathogenesis of TCF4-associated disorders and identifying potential therapeutic targets.

## Potential therapeutic strategies for Pitt–Hopkins syndrome

Pitt–Hopkins syndrome patients require lifelong medical care, but current therapeutic approaches focus on symptomatic treatment. Although there is currently no effective treatment for PTHS, research is ongoing to understand the molecular mechanisms behind the disease and to identify potential therapeutic targets. Several potential therapeutic approaches have been tested in preclinical mouse models of PTHS. The first approach corrects gene transcriptional activity using histone deacetylase inhibitors, which have been associated with improved memory and learning ability. The administration of histone deacetylase inhibitor SAHA improved cognitive function and memory in mice with heterozygous *Tcf4* mutations (a deletion of exons encoding the bHLH domain) (Kennedy et al., 2016). Other studies have selected the sodium potential-dependent NaV1.8 channel encoded by the *SCN10A* gene as a therapeutic target. TCF4 loss-of-function leads to ectopic overexpression of *Scn10a*, and the pharmacological inhibition of NaV1.8 in murine models of PTHS is effective for the restoration of several physiological functions and behavior (Ekins et al., 2020; Cleary et al., 2021; Martinowich et al., 2022). Specifically, S. Ekins and colleagues used Nicardepine, a drug approved by the Food and Drug Administration (FDA) and used in cardiology, as a NaV1.8 inhibitor (Ekins et al., 2020). Other selective NaV1.8 inhibitors have also been proven safe for humans in clinical trials (Hijma et al., 2021, 2022). Given these facts, testing NaV1.8 antagonists for PTHS therapy has significant potential.

The strategies discussed employ either upstream regulators of TCF4 activity or downstream target genes as therapeutic targets. Despite the success of these approaches in animal models, they have some limitations. The effects on upstream regulators of TCF4 are likely to lack specificity and entail undesirable adverse reactions arising from off-target transcriptional effects. The limitations of the second approach stem from the fact that TCF4 regulates the expression of hundreds or thousands of other genes (Forrest et al., 2013; Hill et al., 2017; Xia et al., 2018; Torshizi et al., 2019), which greatly complicates the identification of the key transcription modifier genes and the correction of their expression levels.

Since *TCF4* loss-of-function underlies the disease, it can be hypothesized that rescuing gene expression using antisense oligonucleotides or gene therapy may prove to be the most effective treatment approach. However, given that *TCF4* ex-

pression in humans peaks in the prenatal period and then decreases to the baseline level maintained throughout life (Rannals et al., 2016; Phan et al., 2020), the question arises about the possibility of restoring physiological and behavioral functions of patients with PTHS by normalizing *TCF4* expression in the postnatal period. Moreover, it remains unclear to what extent *TCF4* expression should be upregulated. The regulation of *TCF4* dosage is extremely important because the disease can develop from either too low or too high expression levels. Pathogenic variants in *TCF4* leading to haploinsufficiency may cause neurodevelopmental disorders, whereas polymorphic variants localized in non-coding regions of the gene lead to its overexpression and appear to be associated with schizophrenia.

A recent study by H. Kim et al. (2022) using a mouse model of PTHS showed that the development of the phenotypes characteristic of this syndrome can be prevented or partially corrected by normalizing *Tcf4* expression, with the success of therapeutic intervention depending on the timing of exposure and cell type specificity. Pancellular rescue of *Tcf4* expression in the prenatal period completely prevented the development of PTHS phenotypes. Selective restoration of gene expression in excitatory or inhibitory neurons during embryogenesis resulted in the rescue of a number of behavioral functions. Finally, postnatal restoration of *Tcf4* expression using adeno-associated viral vectors in neurons reduced anxiety-like behavior, stimulated activity, and improved innate behaviors and memory. In addition, this approach led to a partial recovery of EEG parameters and correction of the expression levels of several *Tcf4* target genes (Kim et al., 2022).

Gene therapy based on viral vectors holds great promise for the treatment of diseases previously considered incurable. According to the Gene Therapy Clinical Trials Worldwide database as of March 2023, the vectors based on adenoviruses, retroviruses, lentiviruses, and adeno-associated viruses were the most frequently used in clinical trials (https://a873679.fmphost.com/fmi/webd/GTCT; accessed 29.06.2024). Viral-based vectors have their advantages as well as undesirable effects. The latter include immune response, cytotoxicity, risks of genomic integration, and risks associated with the emergence of *de novo* replicative-competent viruses (Ertl, 2022; Leikas et al., 2023; Lundstrom, 2023).

Thus, gene therapy approaches to rescuing *TCF4* expression developed in animal models may be effective for patients with PTHS. In this regard, further studies are needed to determine whether restoration of *Tcf4* expression at different periods of ontogenesis can help correct behavioral and physiological dysfunction. The results of such studies may help us evaluate the effectiveness of therapy for different age groups of PTHS patients. In addition, potential therapy strategies using *TCF4* expression level correction will have to ensure appropriate biodistribution of the encoded protein, as studies show that restoration of gene activity only in certain cells and brain structures can lead to the normalization of some behavioral and physiological functions in laboratory animals (Kim et al., 2022). One of the main advantages of gene therapy for PTHS is that it does not require an understanding of the molecular pathogenesis mechanisms, as this approach targets the underlying cause of the disorder – the impaired TCF4 and its loss-of-function mutations. However, should gene therapy prove ineffective for humans in the postnatal period or be unfeasible *in utero*, the main focus of research would likely shift towards developing treatment strategies that target TCF4-regulated molecular pathways and downstream target genes.

## Conclusion

To date, substantial experimental data have accumulated, demonstrating the important role of the TCF4 transcription factor in the development and functioning of the nervous system. *TCF4* structure and function anomalies are shown to drive the development of Pitt–Hopkins syndrome, and variants in the gene are associated with a number of psychiatric disorders. However, the molecular mechanisms behind these conditions remain unexplored, and our knowledge of the TCF4 upregulation and its downstream target genes is limited. Moreover, there is insufficient information on the dynamic expression and function of *TCF4* during ontogenesis. It is also unclear how the activity of the encoded transcription factor changes depending on dimerization partners.

Given the broad expression pattern of *TFC4*, as well as its involvement in the development of the nervous system, it can be assumed that pathogenic variants affecting this gene may also be associated with other pathological conditions. This idea is supported by transcriptomic dynamics during TCF4 loss of function, as well as by studies of DNA-protein interactions using ChIP-Seq technology, indicating that common pathogenetic pathways seem to be involved in the pathogenesis of PTHS and some psychiatric disorders (Xia et al., 2018; Phan et al., 2020).

Although many aspects of TCF4 function remain to be explored, this transcription factor is evidently one of the key proteins responsible for learning, memory, verbal contact, and communicative functions in the context of psychiatric disorders. Further study of TCF4 and the identification of molecular pathways and target genes it regulates is crucial for understanding the pathogenesis of *TCF4*-associated diseases. This research direction is also important for finding potential therapeutic strategies for PTHS and possibly other socially significant diseases such as schizophrenia and bipolar disorder.

## References

Afshari N.A., Igo R.P., Morris N.J., Stambolian D., Sharma S., Pulagam V.L., Dunn S., Stamler J.F., Truitt B.J., Rimmler J., Kuot A., Croasdale C.R., Qin X., Burdon K.P., Riazuddin S.A., Mills R., Klebe S., Minear M.A., Zhao J., Balajonda E., Rosenwasser G.O., Baratz K.H., Mootha V.V., Patel S.V., Gregory S.G., Bailey-Wilson J.E., Price M.O., Price F.W., Craig J.E., Fingert J.H., Gottsch J.D., Aldave A.J., Klintworth G.K., Lass J.H., Li Y.J., Iyengar S.K. Genome-wide association study identifies three novel loci in Fuchs endothelial corneal dystrophy. *Nat. Commun.* 2017;8:14898. DOI 10.1038/NCOMMS14898

Amiel J., Rio M., De Pontual L., Redon R., Malan V., Boddaert N., Plouin P., Carter N.P., Lyonnet S., Munnich A., Colleaux L. Mutations in *TCF4*, encoding a class I basic helix-loop-helix transcription factor, are responsible for Pitt-Hopkins syndrome, a severe epileptic encephalopathy associated with autonomic dysfunction. *Am. J. Hum. Genet.* 2007;80(5):988-993. DOI 10.1086/515582

Baratz K.H., Tosakulwong N., Ryu E., Brown W.L., Branham K., Chen W., Tran K.D., Schmid-Kubista K.E., Heckenlively J.R., Swaroop A., Abecasis G., Bailey K.R., Edwards A.O. E2-2 protein and Fuchs's corneal dystrophy. *N. Engl. J. Med.* 2010;363(11):1016-1024. DOI 10.1056/NEJMoa1007064

Bedeschi M.F., Marangi G., Calvello M.R., Ricciardi S., Leone F.P.C., Baccarin M., Guerneri S., Orteschi D., Murdolo M., Lattante S., Frangella S., Keena B., Harr M.H., Zackai E., Zollino M. Impairment of different protein domains causes variable clinical presentation within Pitt-Hopkins syndrome and suggests intragenic molecular syndromology of TCF4. *Eur. J. Med. Genet.* 2017;60(11):565-571. DOI 10.1016/J.EJMG.2017.08.004

Bocharova A.V., Stepanov V.A., Marusin A.V., Kharkov V.N., Vagaitseva K.V., Fedorenko O.Y., Bokhan N.A., Semke A.V., Ivanova S.A. Association study of genetic markers of schizophrenia and its cognitive endophenotypes. *Russ. J. Genet.* 2017;53(1):139-146. DOI 10.1134/S1022795417010033

Brockschmidt A., Todt U., Ryu S., Hoischen A., Landwehr C., Birnbaum S., Frenck W., Radlwimmer B., Lichter P., Engels H., Driever W., Kubisch C., Weber R.G. Severe mental retardation with breathing abnormalities (Pitt–Hopkins syndrome) is caused by haploinsufficiency of the neuronal bHLH transcription factor *TCF4*. *Hum. Mol. Genet.* 2007;16(12):1488-1494. DOI 10.1093/HMG/DDM099

Chen H.Y., Bohlen J.F., Maher B.J. Molecular and cellular function of transcription factor 4 in Pitt-Hopkins syndrome. *Dev. Neurosci.* 2021;43(3-4):159-167. DOI 10.1159/000516666

Chen T., Wu Q., Zhang Y., Lu T., Yue W., Zhang D. Tcf4 controls neuronal migration of the cerebral cortex through regulation of Bmp7. *Front. Mol. Neurosci.* 2016;9:94. DOI 10.3389/FNMOL.2016.00094

Cleary C.M., James S., Maher B.J., Mulkey D.K. Disordered breathing in a Pitt-Hopkins syndrome model involves Phox2b-expressing parafacial neurons and aberrant Nav1.8 expression. *Nat. Commun.* 2021;12(1):1-15. DOI 10.1038/s41467-021-26263-2

D'Rozario M., Zhang T., Waddell E.A., Zhang Y., Sahin C., Sharoni M., Hu T., Nayal M., Kutty K., Liebl F., Hu W., Marenda D.R. Type I bHLH proteins Daughterless and TCF4 restrict neurite branching and synapse formation by repressing Neurexin in postmitotic neurons. *Cell Rep.* 2016;15(2):386. DOI 10.1016/J.CELREP.2016.03.034

Du J., Aleff R.A., Soragni E., Kalari K., Nie J., Tang X., Davila J., Kocher J.P., Patel S.V., Gottesfeld J.M., Baratz K.H., Wieben E.D. RNA toxicity and missplicing in the common eye disease fuchs endothelial corneal dystrophy. *J. Biol. Chem.* 2015;290(10):5979-5990. DOI 10.1074/JBC.M114.621607

Ekins S., Puhl A.C., Davidow A. Repurposing the dihydropyridine calcium channel inhibitor nicardipine as a $Na_V1.8$ inhibitor *in vivo* for Pitt Hopkins syndrome. *Pharm. Res.* 2020;37(7):127. DOI 10.1007/S11095-020-02853-5

Ellinghaus D., Folseraas T., Holm K., Ellinghaus E., Melum E., Balschun T., Laerdahl J.K., Shiryaev A., Gotthardt D.N., Weismüller T.J., Schramm C., Wittig M., Bergquist A., Björnsson E., Marschall H.U., Vatn M., Teufel A., Rust C., Gieger C., Wichmann H.E., Runz H., Sterneck M., Rupp C., Braun F., Weersma R.K., Wijmenga C., Ponsioen C.Y., Mathew C.G., Rutgeerts P., Vermeire S., Schrumpf E., Hov J.R., Manns M.P., Boberg K.M., Schreiber S., Franke A., Karlsen T.H. Genome-wide association analysis in primary sclerosing cholangitis and ulcerative colitis identifies risk loci at GPR35 and TCF4. *Hepatology*. 2013;58(3):1074-1083. DOI 10.1002/HEP.25977

Ertl H.C.J. Immunogenicity and toxicity of AAV gene therapy. *Front. Immunol.* 2022;13:975803. DOI 10.3389/FIMMU.2022.975803

Fautsch M.P., Wieben E.D., Baratz K.H., Bhattacharyya N., Sadan A.N., Hafford-Tear N.J., Tuft S.J., Davidson A.E. TCF4-mediated Fuchs endothelial corneal dystrophy: insights into a common trinucleotide repeat-associated disease. *Prog. Retin. Eye Res.* 2021;81:100883. DOI 10.1016/J.PRETEYERES.2020.100883

Forrest M.P., Waite A.J., Martin-Rendon E., Blake D.J. Knockdown of human TCF4 affects multiple signaling pathways involved in cell survival, epithelial to mesenchymal transition and neuronal differentiation. *PLoS One*. 2013;8(8):e73169. DOI 10.1371/JOURNAL.PONE.0073169

Forrest M.P., Hill M.J., Kavanagh D.H., Tansey K.E., Waite A.J., Blake D.J. The psychiatric risk gene transcription factor 4 (TCF4) regulates neurodevelopmental pathways associated with schizophrenia, autism, and intellectual disability. *Schizophr. Bull.* 2018;44(5):1100-1110. DOI 10.1093/SCHBUL/SBX164

Gelernter J., Sun N., Polimanti R., Pietrzak R., Levey D.F., Bryois J., Lu Q., Hu Y., Li B., Radhakrishnan K., Aslan M., Cheung K.H., Li Y., Rajeevan N., Sayward F., Harrington K., Chen Q., Cho K., Pyarajan S., Sullivan P.F., Quaden R., Shi Y., Hunter-Zinck H., Gaziano J.M., Concato J., Zhao H., Stein M.B. Genome-wide association study of post-traumatic stress disorder reexperiencing symptoms in >165,000 US veterans. *Nat. Neurosci.* 2019;22(9):1394-1401. DOI 10.1038/s41593-019-0447-7

Hennig K.M., Fass D.M., Zhao W.-N., Sheridan S.D., Fu T., Erdin S., Stortchevoi A., Lucente D., Cody J.D., Sweetser D., Gusella J.F., Talkowski M.E., Haggarty S.J. WNT/β-catenin pathway and epigenetic mechanisms regulate the Pitt-Hopkins syndrome and schizophrenia risk gene *TCF4*. *Mol. Neuropsychiatry*. 2017;3(1):53-71. DOI 10.1159/000475666

Hijma H.J., Siebenga P.S., De Kam M.L., Groeneveld G.J. A phase 1, randomized, double-blind, placebo-controlled, crossover study to evaluate the pharmacodynamic effects of VX-150, a highly selective $Na_V1.8$ inhibitor, in healthy male adults. *Pain Med.* 2021;22(8):1814-1826. DOI 10.1093/PM/PNAB032

Hijma H.J., van Brummelen E.M.J., Siebenga P.S., Groeneveld G.J. A phase I, randomized, double-blind, placebo-controlled, single- and multiple dose escalation study evaluating the safety, pharmacokinetics and pharmacodynamics of VX-128, a highly selective $Na_V1.8$ inhibitor, in healthy adults. *Clin. Transl. Sci.* 2022;15(4):981-993. DOI 10.1111/CTS.13215

Hill M.J., Killick R., Navarrete K., Maruszak A., McLaughlin G.M., Williams B.P., Bray N.J. Knockdown of the schizophrenia susceptibility gene *TCF4* alters gene expression and proliferation of progenitor cells from the developing human neocortex. *J. Psychiatry Neurosci.* 2017;42(3):181-188. DOI 10.1503/JPN.160073

Imayoshi I., Kageyama R. bHLH factors in self-renewal, multipotency, and fate choice of neural progenitor cells. *Neuron*. 2014;82(1):9-23. DOI 10.1016/J.NEURON.2014.03.018

Kennedy A.J., Rahn E.J., Paulukaitis B.S., Savell K.E., Kordasiewicz H.B., Wang J., Lewis J.W., Posey J., Strange S.K., Guzman-Karlsson M.C., Phillips S.E., Decker K., Motley S.T., Swayze E.E., Ecker D.J., Michael T.P., Day J.J., Sweatt J.D. *Tcf4* regulates synaptic plasticity, DNA methylation, and memory function. *Cell Rep.* 2016;16(10):2666-2685. DOI 10.1016/J.CELREP.2016.08.004

Kim H., Gao E.B., Draper A., Berens N.C., Vihma H., Zhang X., Higashi-Howard A., Ritola K.D., Simon J.M., Kennedy A.J., Philpot B.D. Rescue of behavioral and electrophysiological phenotypes in a Pitt-Hopkins syndrome mouse model by genetic restoration of *Tcf4* expression. *eLife*. 2022;11:e72290. DOI 10.7554/ELIFE.72290

Leikas A.J., Ylä-Herttuala S., Hartikainen J.E.K. Adenoviral gene therapy vectors in clinical use – basic aspects with a special reference to replication-competent adenovirus formation and its impact on clinical safety. *Int. J. Mol. Sci.* 2023;24(22):16519. DOI 10.3390/IJMS242216519

Li H., Zhu Y., Morozov Y.M., Chen X., Page S.C., Rannals M.D., Maher B.J., Rakic P. Disruption of TCF4 regulatory networks leads to

abnormal cortical development and mental disabilities. *Mol. Psychiatry*. 2019;24(8):1235-1246. DOI 10.1038/S41380-019-0353-0

Li M., Santpere G., Kawasawa Y.I., Evgrafov O.V., Gulden F.O., Pochareddy S., Sunkin S.M., Li Z., Shin Y., Zhu Y., … State M.W., Sanders S.J., Sullivan P.F., Gerstein M.B., Lein E.S., Knowles J.A., Sestan N. Integrative functional genomic analysis of human brain development and neuropsychiatric risks. *Science*. 2018;362(6420): eaat7615. DOI 10.1126/SCIENCE.AAT7615

Lundstrom K. Viral vectors in gene therapy: where do we stand in 2023? *Viruses*. 2023;15(3):698. DOI 10.3390/V15030698

Martinowich K., Das D., Sripathy S.R., Mai Y., Kenney R.F., Maher B.J. Evaluation of $Na_v1.8$ as a therapeutic target for Pitt Hopkins Syndrome. *Mol. Psychiatry*. 2022;28(1):76-82. DOI 10.1038/s41380-022-01811-4

Mesman S., Bakker R., Smidt M.P. *Tcf4* is required for correct brain development during embryogenesis. *Mol. Cell. Neurosci.* 2020;106: 103502. DOI 10.1016/J.MCN.2020.103502

Papanyan S.S., Astakhov S.Yu., Nazarov V.D., Lapin S.V., Novikov S.A., Riks I.A., Anikina L.K., Dovydenko K.S. Expansion of trinucleotide CTG repeats in the *TCF4* gene as a marker of Fuchs' endothelial corneal dystrophy. *Ophthalmology Journal.* 2019;12(2): 11-18. DOI 10.17816/OV2019211-18

Papes F., Camargo A.P., de Souza J.S., Carvalho V.M.A., Szeto R.A., LaMontagne E., Teixeira J.R., Avansini S.H., Sánchez-Sánchez S.M., Nakahara T.S., Santo C.N., Wu W., Yao H., Araújo B.M.P., Velho P.E.N.F., Haddad G.G., Muotri A.R. Transcription factor 4 loss-of-function is associated with deficits in progenitor proliferation and cortical neuron content. *Nat. Commun.* 2022;13(1):2387. DOI 10.1038/s41467-022-29942-w

Phan B.D.N., Bohlen J.F., Davis B.A., Ye Z., Chen H.Y., Mayfield B., Sripathy S.R., Cerceo Page S., Campbell M.N., Smith H.L., Gallop D., Kim H., Thaxton C.L., Simon J.M., Burke E.E., Shin J.H., Kennedy A.J., Sweatt J.D., Philpot B.D., Jaffe A.E., Maher B.J. A myelin-related transcriptomic profile is shared by Pitt-Hopkins syndrome models and human autism spectrum disorder. *Nat. Neurosci.* 2020;23(3):375-385. DOI 10.1038/S41593-019-0578-X

Rannals M.D.D., Hamersky G.R.R., Page S.C.C., Campbell M.N.N., Briley A., Gallo R.A.A., Phan B.D.N., Hyde T.M.M., Kleinman J.E.E., Shin J.H.H., Jaffe A.E.E., Weinberger D.R.R., Maher B.J.J. Psychiatric Risk Gene Transcription Factor 4 regulates intrinsic excitability of prefrontal neurons via repression of SCN10a and KCNQ1. *Neuron*. 2016;90(1):43-55. DOI 10.1016/J.NEURON.2016.02.021

Ripke S., Sanders A.R., Kendler K.S., Levinson D.F., Sklar P., Holmans P.A., Lin D.Y., Duan J., Ophoff R.A., Andreassen O.A., … Williams N.M., Wormley B., Zammit S., Sullivan P.F., O'Donovan M.C., Daly M.J., Gejman P.V. Genome-wide association study identifies five new schizophrenia loci. *Nat. Genet.* 2011;43(10):969-976. DOI 10.1038/ng.940

Rosenfeld J.A., Leppig K., Ballif B.C., Thiese H., Erdie-Lalena C., Bawle E., Sastry S., Spence J.E., Bandholz A., Surti U., Zonana J., Keller K., Meschino W., Bejjani B.A., Torchia B.S., Shaffer L.G. Genotype-phenotype analysis of *TCF4* mutations causing Pitt-Hopkins syndrome shows increased seizure activity with missense mutations. *Genet. Med.* 2009;11(11):797-805. DOI 10.1097/GIM.0B013E3181BD38A9

Schmidt-Edelkraut U., Daniel G., Hoffmann A., Spengler D. Zac1 regulates cell cycle arrest in neuronal progenitors via Tcf4. *Mol. Cell. Biol.* 2014;34(6):1020. DOI 10.1128/MCB.01195-13

Schoof M., Hellwig M., Harrison L., Holdhof D., Lauffer M.C., Niesen J., Virdi S., Indenbirken D., Schüller U. The basic helix-loop-helix transcription factor TCF4 impacts brain architecture as well as neuronal morphology and differentiation. *Eur. J. Neurosci.* 2020; 51(11):2219-2235. DOI 10.1111/EJN.14674

Sepp M., Kannike K., Eesmaa A., Urb M., Timmusk T. Functional diversity of human basic helix-loop-helix transcription factor TCF4 isoforms generated by alternative 5′ exon usage and splicing. *PLoS One*. 2011;6(7):e22138. DOI 10.1371/JOURNAL.PONE.0022138

Sepp M., Pruunsild P., Timmusk T. Pitt-Hopkins syndrome-associated mutations in *TCF4* lead to variable impairment of the transcription factor function ranging from hypomorphic to dominant-negative effects. *Hum. Mol. Genet.* 2012;21(13):2873-2888. DOI 10.1093/HMG/DDS112

Sepp M., Vihma H., Nurm K., Urb M., Page S.C., Roots K., Hark A., Maher B.J., Pruunsild P., Timmusk T. The intellectual disability and schizophrenia associated transcription factor TCF4 is regulated by neuronal activity and protein kinase A. *J. Neurosci.* 2017;37(43): 10516-10527. DOI 10.1523/JNEUROSCI. 1151-17. 2017

Smoller J.W., Kendler K.K., Craddock N., Lee P.H., Neale B.M., Nurnberger J.N., Ripke S., Santangelo S., Sullivan P.S., Neale B.N., Purcell S., Anney R., Buitelaar J., Fanous A., Faraone S.F., Hoogendijk W., Lesch K.P., Levinson D.L., Perlis R.P., Rietschel M., Riley B., Sonuga-Barke E., Schachar R., Schulze T.G., Thapar A., Smoller J.S., Neale M., Perlis R., Bender P., Cichon S., Daly M.D., Kelsoe J., Lehner T., Levinson D., O'Donovan Mick, Gejman P., Sebat J., Sklar P., Devlin B., Sullivan P., O'Donovan Michael. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet*. 2013;381(9875):1371-1379. DOI 10.1016/S0140-6736(12)62129-1

Stefansson H., Ophoff R.A., Steinberg S., Andreassen O.A., Cichon S., Rujescu D., Werge T., Pietiläinen O.P.H., Mors O., Mortensen P.B., … Van Os J., Wiersma D., Bruggeman R., Cahn W., De Haan L., Krabbendam L., Myin-Germeys I. Common variants conferring risk of schizophrenia. *Nature*. 2009;460(7256):744-747. DOI 10.1038/NATURE08186

Steinberg S., de Jong S., Andreassen O.A., Werge T., Børglum A.D., Mors O., Mortensen P.B., Gustafsson O., Costas J., Pietiläinen O.P.H., … Collier D.A., St Clair D., Rietschel M., Cichon S., Stefansson H., Rujescu D., Stefansson K. Common variants at *VRK2* and *TCF4* conferring risk of schizophrenia. *Hum. Mol. Genet.* 2011; 20(20):4076-4081. DOI 10.1093/HMG/DDR325

Tamberg L., Jaago M., Säälik K., Sirp A., Tuvikene J., Shubina A., Kiir C.S., Nurm K., Sepp M., Timmusk T., Palgi M. Daughterless, the *Drosophila* orthologue of TCF4, is required for associative learning and maintenance of the synaptic proteome. *DMM: Dis. Model. Mech.* 2020;13(7):dmm042747. DOI 10.1242/dmm.042747

Teixeira J.R., Szeto R.A., Carvalho V.M.A., Muotri A.R., Papes F. Transcription factor 4 and its association with psychiatric disorders. *Transl. Psychiatry*. 2021;11(1):19. DOI 10.1038/s41398-020-01138-0

Thaxton C., Kloth A.D., Clark E.P., Moy S.S., Chitwood R.A., Philpot B.D. Common pathophysiology in multiple mouse models of Pitt-Hopkins syndrome. *J. Neurosci.* 2018;38(4):918-936. DOI 10.1523/JNEUROSCI.1305-17.2017

Torshizi A.D., Armoskus C., Zhang H., Forrest M.P., Zhang S., Souaiaia T., Evgrafov O.V., Knowles J.A., Duan J., Wang K. Deconvolution of transcriptional networks identifies TCF4 as a master regulator in schizophrenia. *Sci. Adv.* 2019;5(9):eaau4139. DOI 10.1126/SCIADV.AAU4139

Wang Y., Lu Z., Zhang Yilan, Cai Y., Yun D., Tang T., Cai Z., Wang C., Zhang Yandong, Fang F., Yang Z., Behnisch T., Xie Y. Transcription factor 4 safeguards hippocampal dentate gyrus development by regulating neural progenitor migration. *Cereb. Cortex*. 2020;30(5): 3102-3115. DOI 10.1093/CERCOR/BHZ297

Wedel M., Fröb F., Elsesser O., Wittmann M.T., Lie D.C., Reis A., Wegner M. Transcription factor Tcf4 is the preferred heterodimerization partner for Olig2 in oligodendrocytes and required for differentiation. *Nucleic Acids Res.* 2020;48(9):4839-4857. DOI 10.1093/NAR/GKAA218

Wittmann M.T., Häberle B.M. Linking the neuropsychiatric disease gene *TCF4* to neuronal activity-dependent regulatory networks. *J. Neurosci.* 2018;38(11):2653. DOI 10.1523/JNEUROSCI.3475-17.2018

Wray N.R., Ripke S., Mattheisen M., Trzaskowski M., Byrne E.M., Abdellaoui A., Adams M.J., Agerbo E., Air T.M., Andlauer T.M.F., … Werge T., Winslow A.R., Lewis C.M., Levinson D.F., Breen G., Børglum A.D., Sullivan P.F. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat. Genet.* 2018;50(5):668-681. DOI 10.1038/s41588-018-0090-3

Xia H., Jahr F.M., Kim N.K., Xie L., Shabalin A.A., Bryois J., Sweet D.H., Kronfol M.M., Palasuberniam P., McRae M.P., Riley B.P., Sullivan P.F., Van Den Oord E.J., McClay J.L. Building a schizophrenia genetic network: transcription factor 4 regulates genes involved in neuronal development and schizophrenia risk. *Hum. Mol. Genet.* 2018;27(18):3246-3256. DOI 10.1093/HMG/DDY222

Zweier C., Peippo M.M., Hoyer J., Sousa S., Bottani A., Clayton-Smith J., Reardon W., Saraiva J., Cabral A., Göhring I., Devriendt K., De Ravel T., Bijlsma E.K., Hennekam R.C.M., Orrico A., Cohen M., Dreweke A., Reis A., Nurnberg P., Rauch A. Haploinsufficiency of *TCF4* causes syndromal mental retardation with intermittent hyperventilation (Pitt-Hopkins syndrome). *Am. J. Hum. Genet.* 2007;80(5):994-1001. DOI 10.1086/515583

**Conflict of interest.** The authors declare no conflict of interest.
Received January 16, 2024. Revised June 30, 2024. Accepted July 23, 2024.

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2024;28(7):780-791

DOI 10.18699/vjgb-24-86

# Prevalence of AZFc Y chromosome microdeletions and association with spermatogenesis in Russian men from the general population

L.V. Osadchuk [1], G.V. Vasiliev [1], M.K. Ivanov[2], M.A. Prasolova[1, 2], M.A. Kleshchev [1], A.V. Osadchuk [1]

[1] Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
[2] Joint Stock Company Vector Best, Novosibirsk region, Russia
losadch@bionet.nsc.ru

**Abstract.** The Y chromosome contains a set of genes with testis-specific expression that are responsible for the development of testes and spermatogenesis, and it is the most important target in the search for genetic causes of male infertility. Most of these genes are located in the "azoospermia factor" AZF locus (regions AZFa, AZFb, and AZFc) on the long arm of the Y chromosome. Microdeletions of the Y chromosome, leading to the removal of the entire AZF locus as well as one or more regions (complete deletions), are one of the leading causes of spermatogenesis impairment and infertility. However, the role of partial AZFc deletions (gr/gr, b2/b3, b1/b3) in spermatogenesis failure is unclear, and their impact on spermatogenesis varies between populations. The aim of the present study was to assess the frequency of various types of AZFc microdeletions and to search for associations with spermatogenesis parameters in men of Slavic ethnicity from the general Russian population ($n$ = 700, average age 25.8 years). To identify AZF microdeletions, the presence/absence of 15 STS markers was analyzed using multiplex real-time polymerase chain reaction. Age, weight, height, and the volume, concentration, total count, proportion of motile and morphologically normal spermatozoa in the ejaculate were recorded for all participants. In the studied sample, 19.9 % (139/700) of men were found to have AZFc microdeletions, of which 16.7 % (117/700) were carriers of a partial b2/b3 deletion, 3.0 % (21/700) had a partial gr/gr deletion, and 0.14 % (1/700) had a complete b2/b4 deletion. Neither AZFa nor AZFb microdeletions nor other types of AZF deletions were detected. The overall frequency of all types of AZFc deletions, as well as each type of partial microdeletion, b2/b3 and gr/gr, did not differ in the groups of azoospermia, severe oligozoospermia (≤5.0 mill/ml), oligozoospermia (5.0 < SC < 16.0 mill/ml), and normal sperm concentration (≥16.0 mill/ml). Comparison of semen parameters in groups with different types of partial AZFc deletions and the control group (without deletions) also did not reveal significant differences. Thus, partial AZFc microdeletions b2/b3 and gr/gr do not have a significant impact on spermatogenesis in Slavic men. It is suggested that in Slavs, partial AZFc microdeletions b2/b3 and gr/gr are fixed in Y haplogroups N3 and R1a, respectively, and their negative impact on spermatogenesis is balanced by other genetic factors. The higher frequency of partial AZFc deletions (19.7 %) in Slavs compared to European populations (7.3 %) established in our study may be explained by the widespread distribution of these Y haplogroups in the Slavic population of Russia.
**Key words:** AZFc deletions of the Y chromosome; spermatogenesis; male fertility; general population.

**For citation:** Osadchuk L.V., Vasiliev G.V., Ivanov M.K., Prasolova M.A., Kleshchev M.A., Osadchuk A.V. Prevalence of AZFc Y chromosome microdeletions and association with spermatogenesis in Russian men from the general population. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2024;28(7):780-791. DOI 10.18699/vjgb-24-86

**Funding.** This study was supported by the state assignment FWNR-2022-0021.

This article does not contain any studies using animals as an object.

# Распространенность микроделеций AZFc региона Y-хромосомы и влияние на сперматогенез у российских мужчин из общей популяции

Л.В. Осадчук [1], Г.В. Васильев [1], М.К. Иванов[2], М.А. Прасолова[1, 2], М.А. Клещев [1], А.В. Осадчук [1]

[1] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
[2] АО «Вектор-Бест», р. п. Кольцово, Новосибирская область, Россия
losadch@bionet.nsc.ru

**Аннотация.** Y-хромосома содержит набор генов, имеющих тестис-специфическую экспрессию, ответственных за развитие яичек и сперматогенез, и является наиболее важной мишенью в поиске генетических причин мужского бесплодия. Большинство из этих генов расположены в локусе «фактора азооспермии» AZF (регионы AZFa, AZFb и

AZFc) на длинном плече Y-хромосомы. Микроделеции Y-хромосомы, приводящие к удалению всего локуса AZF, а также одного или нескольких регионов (полные делеции), являются одной из ведущих причин нарушения сперматогенеза и бесплодия, однако роль частичных AZFc-делеций (gr/gr, b2/b3, b1/b3) в нарушении сперматогенеза не ясна, а влияние на сперматогенез варьирует между популяциями. Цель настоящего исследования состояла в оценке частоты различных типов AZFc-микроделеций и поиске ассоциаций с параметрами сперматогенеза у мужчин славянской этнической группы из общей российской популяции ($n$ = 700, средний возраст 25.8 года). Для выявления AZF-микроделеций анализировали наличие/отсутствие 15 STS-маркеров методом мультиплексной полимеразной цепной реакции в режиме реального времени. У всех участников записывали возраст, вес, рост, оценивали объем, концентрацию, общее количество, долю подвижных и морфологически нормальных сперматозоидов в эякуляте. В исследуемой выборке выявлены 19.9 % (139/700) мужчин с микроделециями AZFc региона, из них 16.7 % (117/700) являлись носителями частичной делеции b2/b3, 3.0 % (21/700) – частичной делеции gr/gr, 0.14 % (1/700) – полной делеции b2/b4. Не обнаружены AZFa и AZFb микроделеции и другие типы AZF-делеций. Суммарная частота всех типов AZFc-делеций, а также каждого типа частичных микроделеций b2/b3 и gr/gr не различалась в группах азооспермии, тяжелой олигозооспермии (≤5.0 млн/мл), олигозооспермии (5.0 < КС < 16.0 млн/мл) и нормальной концентрации сперматозоидов (≥16.0 млн/мл). Сравнение спермиологических показателей в группах с различными типами частичных AZFc-делеций и контролем (без делеций) тоже не выявило достоверных различий. Таким образом, частичные AZFc-микроделеции b2/b3 и gr/gr не оказывают существенного влияния на сперматогенез у славянских мужчин. Предполагается, что у славян частичные AZFc-микроделеции b2/b3 и gr/gr фиксированы в Y-гаплогруппе N3 и R1a соответственно, а их негативное влияние на сперматогенез уравновешивается другими генетическими факторами. Установленная в нашей работе более высокая частота частичных AZFc-делеций (19.7 %) у славян по сравнению с европейскими популяциями (7.3 %) также может объясняться широким распространением этих Y-гаплогрупп в славянской популяции России.
**Ключевые слова:** AZFc-микроделеции Y-хромосомы; сперматогенез; мужская фертильность; общая популяция.

## Introduction

The prevalence of male infertility in the general population is 7–12 % (Krausz et al., 2018; Cioppi et al., 2021), and in the Russian Federation 10–15 % of married couples suffer from infertility depending on the region (Lebedev et al., 2019). A number of genetic variants that negatively affect male fertility are known, and this list is continuously expanding (Cioppi et al., 2021). The Y chromosome is the most important molecular genetic target in the search for genetic causes of male infertility and subfertility (Krausz, Casamonti, 2017; Colaco, Modi, 2018). The Y chromosome carries genes necessary for the normal development of the testes and testicular functions, such as sex determination and the regulation of spermatogenesis. On the Y chromosome, the AZF locus and its three regions AZFa, AZFb and AZFc are located, and Y chromosome microdeletions leading to the removal of whole AZF regions (complete microdeletions) are the second main cause of spermatogenesis impairment and infertility after Klinefelter syndrome (Krausz, Casamonti, 2017; Krausz et al., 2024). The AZF microdeletions are usually *de novo* mutations, the complete microdeletion rate in the general population is 1:4,000, but in men with oligozoospermia and azoospermia it is significantly higher and can be as high as 14 % (Colaco, Modi, 2018; Cioppi et al., 2021; Deng et al., 2023). Adequate diagnostic methods have been developed for testing microdeletions in the AZF locus of the Y chromosome, and screening for complete microdeletions of AZFa and AZFb has become a mandatory part of routine diagnostic examinations for men with azoospermia and severe oligozoospermia. However, the clinical and diagnostic significance of the AZFc region remains a subject of discussion (Krausz et al., 2018). An indication for testing for Y chromosome microdeletions is a sperm concentration of less than 5 mill/mL or azoospermia, which is often observed in patients with infertility (Krausz et al., 2018).

A feature of the AZF locus of the Y chromosome is the ampliconic structure and multiple copies of genes. Ampliconic sequences are more than 99 % identical and organized into eight massive palindromes. Because palindrome sequences exhibit near-complete symmetry, they tend to form hairpin-like structures and generate homologous recombination (Kuroda et al., 2020). The most common type of AZF deletion is AZFc (70–80 %), followed by AZFa (0.5–9 %), AZFb (1–7 %), and AZFb+c (1–20 %) (Krausz, Casamonti, 2017; Cioppi et al., 2021; Krausz et al., 2024). Complete AZF deletions, which entirely remove one or more AZF regions, are associated with severe spermatogenesis failure, leading to infertility, and are never found in men with normozoospermia.

The AZFa region contains two single-copy genes *USP9Y* and *DDX3Y* and retroviral sequences *HERVyq1* and *HERVyq2*, which are flanking AZFa. Between these directional retroviral sequences, homologous recombination could occur resulting in the deletion of AZFa, azoospermia, and Sertoli cell-only syndrome. The AZFb region contains 32 gene copies and transcription units. With a complete deletion of AZFb, the DNA segment including all 32 copies of genes and transcription units is removed, leading to maturation arrest and azoospermia. The AZFb and AZFc regions are partly overlapping, and complete AZFb or AZFb+c deletions are associated with Sertoli cell-only syndrome and azoospermia (Kuroda et al., 2020; Cioppi et al., 2021).

The AZFc region contains 12 genes in a variable number of copies for a total of 32 transcription units, which are expressed only in the testis and most often undergo deletions (Colaco, Modi, 2018; Cioppi et al., 2021; Krausz et al., 2024). A complete AZFc deletion (b2/b4) occurs as a result of homologous recombination between amplicon b2 and b4 and is characterized by spermatogenic impairment, ranging from severe oligozoospermia to azoospermia. However, in a significant number of cases, it is accompanied by residual

spermatogenesis (Krausz et al., 2024). The AZFc locus contains the *DAZ* gene family that is a key determinant of spermatogenesis and consists of four copies (*DAZ1–4*). The *DAZ* gene contains an RNA-binding protein, which indicates the participation of genes of this family in mRNA translation and, apparently, in the differentiation of spermatogenic cells and meiotic division. Copies of the *DAZ* gene are distributed across two different clusters (*DAZ1/2* and *DAZ3/4*), and their expression is observed at all stages of germ cell development (Colaco, Modi, 2018). The AZFc region is rich in amplicons, therefore, it is predisposed to a number of rearrangements, including partial deletions or duplications, as well as deletions with subsequent duplication. However, their effects on spermatogenesis are not yet clear and are actively discussed (Krausz, Casamonti, 2017). If partial deletions of AZFa and AZFb are extremely rare and are associated with reduced sperm production, then the role of partial deletions of AZFc (the most common are gr/gr, b2/b3, b1/b3) in spermatogenesis is controversial and the association with spermatogenesis varies greatly (from normozoospermia to azoospermia), but they may be compatible with natural conception or successfully overcome by assisted reproductive technologies (Bansal et al., 2016a, b).

The gr/gr partial AZFc deletion removes almost half of the AZFc gene content, including two copies of the *DAZ* gene (*DAZ1/DAZ2* or *DAZ3/DAZ4*) and one copy of the *BPY2* and *CDY1* gene, representing a risk factor of spermatogenic impairment (Bansal et al., 2016b; Krausz et al., 2024). The phenotypic expression of the gr/gr deletion varies from azoospermia to normal sperm concentration, the cause of which is not yet clear. Since some gr/gr deletions are followed by duplications restoring the gene dosage, it is the gene copy number that may be the causal factor modulating sperm production. Geographic and ethnic differences in the frequency and clinical implications of the gr/gr deletion have been found, suggesting that Y chromosomal background can affect the testicular phenotype (Krausz, Casamonti, 2017). Certain Y haplogroups carrying a fixed gr/gr deletion may be present at high frequency in some ethnic populations and may influence the phenotypic manifestation of deletions through as yet unknown genetic factors (Sin et al., 2010; Rozen et al., 2012; Lo Giacco et al., 2014; Mokánszki et al., 2018). In the population of Northern India, the gr/gr deletion is a risk factor for impaired spermatogenesis if this deletion is not fixed in haplogroups R and H, the most common in this region (Bansal et al., 2016b).

Partial AZFc deletions of b1/b3 or b2/b3 remove more than half of the AZFc region and 12 gene copies and transcription units each. The mechanism of b1/b3 deletion formation involves a homologous recombination between sister chromatids or within a chromatid. Due to its low frequency, the effect of b1/b3 deletion on spermatogenesis remains unclear, but some authors find an increased risk of severe spermatogenic failure in men with a b1/b3 deletion (Krausz, Casamonti, 2017). The b2/b3 deletion removes over half of the AZFc, including two copies of *DAZ* and one copy of *CDY1*. The molecular mechanism of the b2/b3 deletion is complex, since it is preceded by an inversion and results in the retention of two *DAZ* gene copies, one *BPY2* gene and one *CDY1* gene. A high frequency of the b2/b3 deletion is observed in populations of Northern Eurasia, where it is fixed in Y haplogroup N and it is not a risk factor for impaired spermatogenesis and infertility. However, it may increase the risk of spermatogenic loss and infertility when occurring outside this haplogroup, for example in Mongoloid, East Asian and African populations (Rozen et al., 2012; Bansal et al., 2016a; Colaco, Modi, 2018; Hallast et al., 2021).

In clinical practice, testing for the presence of AZF deletions on the Y chromosome is recommended for infertile men with azoospermia and severe oligozoospermia for diagnostic purposes. This can, in some cases, help to identify the genetic cause of impaired spermatogenesis. The diagnosis of Y-chromosome deletions also has prognostic value and allows to resolve the issue of the possibility of surgical production of sperm (micro-TESE) for subsequent IVF/ICSI. For patients with a complete AZFc deletion and azoospermia, the prognosis for obtaining sperm is favorable. In contrast, testicular biopsy using the micro-TESE method is generally ineffective for carriers of complete microdeletions in the AZFa or AZFb regions (Krausz, Casamonti, 2017; Kuroda et al., 2020).

Fertile men carrying partial AZFc deletions or infertile men with partial AZFc deletions whose partners give birth to children through assisted reproductive technologies (micro-TESE or TESA/ICSI) can transfer these AZF deletions to the progeny (Pan et al., 2018; Deng et al., 2023). Moreover, in the paper (Pan et al., 2018), in fertile fathers who were carriers of a b2/b3 deletion or a b2/b3 duplication, the sons suffered from infertility and were carriers of a complete AZFc, AZFb+c or AZFa+b+c deletion. Thus, partial deletions of the AZFc region increase the likelihood of other microstructural rearrangements within the AZFc region, which can be a risk factor for complete AZFc deletion and infertility in male offspring.

Although the relationship between spermatogenic failure, testicular phenotype and ART outcomes in men with different types of AZFc microdeletions of the Y chromosome has been extensively studied, the population frequencies of these deletions and therefore their exact contribution to male infertility and subfertility are still insufficiently understood. Genetic testing of male populations is considered a useful approach for obtaining adequate genetic information about the prevalence of genetically determined impairment of spermatogenesis, infertility and subfertility in men of a given population. This information can be used to forecast and plan preventive, diagnostic, and clinical work aimed at preserving and improving the reproductive health of the population. Such data can serve as a basis for further genetic studies on the etiology of infertility and the determination of its causes. They can provide information about the mutation spectrum associated with spermatogenesis disorders and help to define the genetic structure of demographic risks within the population.

The aim of the present study was to analyze the spectrum and prevalence of AZF microdeletions on the Y chromosome and to search for associations with semen parameters in Slavic men from the general Russian population.

## Materials and methods

Young Slavic men (Belarusians, Ukrainians, Russians) ($n$ = 700) from five Russian cities participated in the study: Arkhangelsk ($n$ = 77), Novosibirsk ($n$ = 324), Kemerovo

($n$ = 205), Ulan-Ude ($n$ = 69), Yakutsk ($n$ = 25). The study population also included descendants of mixed marriages between Russians and Belarusians, Ukrainians, Poles (7.1 %). The study design and standardized recruitment protocol had been described earlier in more detail (Osadchuk et al., 2021, 2022). Men from the general population, regardless of fertility status, participated in the study. All participants were either born or had lived for at least 3–5 years in the cities where the study was conducted. Most participants were students or employees of higher educational institutions at the time of the survey and had not previously consulted an andrologist. All participants were volunteers and did not receive financial compensation. The men completed questionnaires that included questions about their age, place of birth, nationality, profession, type of work, military service, smoking, alcohol consumption, and past and current diseases. Ethnic background was assessed for up to two generations – for the participant, their parents, and both maternal and paternal grandparents. All men included in the study gave informed consent to participate in the examination. The Ethics Committee of the Federal Research Center Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences approved the study (protocol No. 160 dated 17.09.2020).

During the examination, the men were examined by a urologist-andrologist, medical histories were collected, current disorders of the urogenital system were diagnosed, and the results of the examination were recorded in the examination protocol. Each volunteer was given a preliminary andrological diagnosis. The age of all participants was documented, and their height (cm) and body weight (kg) were measured. The bitesticular volume (BTV) (ml) was estimated by a Prader orchidometer. Exclusion criteria included acute diseases, taking medications or undergoing procedures affecting sperm quality (such as anabolic steroids, antibiotics and others). A preliminary condition for participation was abstinence from sexual intercourse for 2–7 days before the study. Semen samples for further laboratory analysis were collected by the participants in a specialized laboratory room through masturbation into single-use sterile plastic containers. The period of sexual abstinence was 4 days (median).

The semen samples were analyzed according to the WHO laboratory manual (WHO…, 2010, 2021). The sperm concentration was assessed using Goryaev's hemocytometer after staining an ejaculate aliquot with trypan blue. The proportion of motile sperm with progressive straight-line movement and velocity above 25 and 2–25 μm/s (categories A and B, respectively) was assessed using the sperm analyzer SFA–500-2 ("Biola", Russia). The analysis of sperm morphology was conducted according to WHO guidelines (WHO…, 2021). Ejaculate smears were stained using commercially available Diff-Quik kits ("Abris+", Russia). The first 200 spermatozoa were examined for morphology with an optical microscope Axio Skop.A1 (Carl Zeiss, Germany) at ×1000 magnification with oil immersion. Sperm dimensions were measured using an ocular micrometer. Sperm morphology evaluations were done in duplicates in random and blinded order by a trained staff member. To determine the Teratozoospermia Index (TZI), the total number of identified morphological defects was divided by the number of morphologically abnormal spermatozoa.

Genomic DNA was extracted from peripheral blood leukocytes using a widely accepted phenol-chloroform method. Detection of microdeletions in the AZF locus was performed by multiplex polymerase chain reaction (PCR) with hybridization-fluorescence detection of PCR products in real-time using a CFX96 DNA amplifier (Bio-Rad, USA). In the first stage, to identify deletions, the presence or absence of 13 STS markers was analyzed using "RealBest-Genetics AZF-microdeletions" commercial kits (Vector-Best, Novosibirsk). The amplification reaction protocol included: stage 1: 50 °C – 2 min; stage 2: 95 °C – 2 min; stage 3: 50 cycles (94 °C – 10 sec, 60 °C – 20 sec). The following STS markers were investigated: sY86, sY84, sY615 – for the AZFa region; sY127, sY134, sY142 – for the AZFb region; sY1196, sY1191, sY254, sY255, sY1291, sY1206, sY1125 – for the AZFc region. Partial AZFc deletions b2/b3 and gr/gr were indicated by the absence of markers sY1191 and sY1291, respectively; complete AZFc deletion b2/b4, by the absence of markers sY1191, sY1206, sY1291, sY254, and sY255; partial AZFc deletion b1/b3, by the absence of markers sY1191, sY1196, and sY1291. STS marker typing was conducted using five reaction mixtures (RM). RM1 included markers for the *SRY* gene (sex-determining gene), sY134, sY84, sY254; RM2, for the *HMDS* gene (a gene for additional DNA control), sY127, sY86, sY255; RM3, for the *HMDS* gene, sY142, sY615; RM4, for sY1191, sY1196, sY1206, sY1125; RM5, for the *SRY* gene, sY1296. Genotyping of the *SRY* gene and the autosomal *HMBS* gene (quality control of the material collection) was conducted as an internal control.

In the second stage, for the most common partial microdeletions of the AZFc region – b2/b3 (marker sY1191) and gr/gr (marker sY1291) – verification was performed using two additional STS markers, sY1192, sY1189, which are closely linked to the corresponding markers. Detection of STS markers sY1192 and sY1189 was carried out by their amplification. The amplification reaction protocol was two-step: 50 °C for 2 min, 95 °C for 2 min, 40 cycles (95 °C for 10 sec, 66 °C for 20 sec) followed by electrophoresis in a 1.2 % agarose gel, staining with ethidium bromide, and visualization under ultraviolet light.

A statistical analysis of the obtained data was performed using the statistical package STATISTICA (version 8.0). For all studied parameters, the mean (SD) was calculated. The Kolmogorov–Smirnov test was used to confirm the normal distribution of quantitative variables. Since most parameters did not follow a normal distribution, differences in the studied anthropometric and spermiological parameters between groups with different sperm concentrations or between groups with different types of partial AZFc microdeletions were determined using the Kruskal–Wallis one-way analysis of variance (Kruskal–Wallis ANOVA) and analysis of covariance (ANCOVA). In the latter case, spermiological parameters were adjusted for age and abstinence period. For pairwise comparison of groups, Duncan's test was applied. Comparisons of the frequencies of AZFc microdeletions between groups were conducted using the chi-square ($\chi^2$) test. A $p$-value < 0.05 was considered statistically significant.

**Table 1.** Anthropometric and spermiological parameters of men in the entire study population and after stratification into categories by sperm concentration

| Parameter | Entire study population ($n$ = 700) | Categories by sperm concentration | | | |
|---|---|---|---|---|---|
| | | SC = 0 mill/mL, ($n$ = 19) | SC ≤ 5 mill/mL, ($n$ = 33) | 16 > SC > 5 mill/mL, ($n$ = 74) | SC ≥ 16 mill/mL, ($n$ = 574) |
| Age, years | 25.8 (7.6) | 28.5 (8.6) | 23.2 (4.9) | 25.0 (6.2) | 26.0 (7.8) |
| Weight, kg | 78.3 (14.0) | 81.7 (15.5) | 76.9 (14.0) | 76.9 (13.6) | 78.4 (14.0) |
| Height, cm | 179.0 (6.9) | 179.3 (6.32) | 178.6 (7.9) | 178.7 (7.9) | 179.0 (6.7) |
| BTV, mL | 40.6 (8.8) | 35.7 (12.0)[b] | 35.5 (7.8)[b] | 37.2 (7.1)[b] | 41.5 (8.7)[a] |
| Semen volume, mL | 3.7 (1.7) | 3.4 (1.3) | 3.5 (1.8) | 3.7 (1.6) | 3.7 (1.7) |
| TSC, mill/ejaculate | 190.1 (185.1) | 0 | 10.1 (9.1)[b] | 42.6 (26.7)[b] | 225.8 (185.9)[a] |
| SC, mill/mL | 54.21 (43.31) | 0 | 2.76 (1.55)[b] | 11.12 (2.97)[b] | 64.52 (41.12)[a] |
| Motility, % | 43.3 (27.2) | – | 3.3 (2.2)[b] | 9.4 (7.2)[b] | 50.0 (24.2)[a] |
| Morphology, % | 6.96 (3.23) | – | 2.76 (2.58)[b] | 3.91 (1.92)[b] | 7.56 (3.05)[a] |
| TZI | 1.49 (0.12) | – | 1.66 (0.13)[b] | 1.58 (1.88)[b] | 1.47 (0.11)[a] |

Note. Data are presented as mean (SD). BTV – bitesticular volume (paired testicular volume); TSC – total sperm count per ejaculate; SC – sperm concentration; motility – percentage of motile spermatozoa in categories A+B; morphology – percentage of morphologically normal spermatozoa; TZI – teratozoospermia index.
[a, b] Comparisons with different superscripts within variables were significant ($p < 0.05$).

**Characteristics of the Slavic study population.** According to the results of a physical examination and medical history, the study population consist of 16 (2.3 %) individuals with testicular hypoplasia, 50 (7.1 %) with grade II and III varicocele, 8 (1.1 %) who underwent surgery for cryptorchidism, and 43 (6.4 %) who underwent varicocelectomy. Among 700 participants, 2.7 % suffered from azoospermia, 4.7 %, from severe oligozoospermia, 10.6 %, from moderate oligozoospermia, but 82.0 % had normal sperm concentration (SC) according to WHO recommendations (WHO…, 2021).

Based on sperm concentration (SC), participants were stratified into four groups: 1) SC = 0 mill/mL (azoospermia, absence of spermatozoa in the ejaculate); 2) SC ≤ 5.0 mill/mL (severe oligozoospermia); 3) 16.0 > SC > 5.0 mill/mL (moderate oligozoospermia); 4) SC ≥ 16.0 mill/mL (normal sperm concentration). Anthropometric and spermiological indicators of men in groups with varying sperm concentrations are presented in Table 1. No differences were found between the groups in terms of age, anthropometric parameters, and ejaculate volume. The total sperm count, sperm concentration, percentage of motile and morphologically normal sperm, and BTV in the group with normal sperm concentration were significantly ($p < 0.05$) higher, whereas the TZI was significantly lower ($p < 0.05$) compared to both oligozoospermia groups, which did not differ from each other in these parameters.

## Results

### Prevalence of different types of AZFc microdeletions in the Slavic study population

Since the study population included Slavs residing in 5 cities of Russia, a comparison of the prevalence of AZFc microdeletions in each city group was conducted (Table 2). Statistical analysis did not reveal significant regional differences in the frequency of partial deletions b2/b3 and gr/gr ($\chi^2_8 = 6.46$; $p < 0.595$).

**Table 2.** Frequency of partial AZFc microdeletions (b2/b3 and gr/gr) in the Slavic groups from the studied cities of Russia

| City | $n$ | b2/b3, $n$ (%) | gr/gr, $n$ (%) |
|---|---|---|---|
| Arkhangelsk | 77 | 16 (20.8) | 1 (1.3) |
| Novosibirsk | 324 | 46 (14.2) | 13 (4.0) |
| Kemerovo | 204 | 39 (19.1) | 5 (2.5) |
| Ulan-Ude | 69 | 13 (18.8) | 2 (2.9) |
| Yakutsk | 25 | 3 (12.0) | 0 (0) |
| Entire study population | 699 | 117 (16.7) | 21 (3.0) |

From the study population, two groups were formed: one with normal sperm parameters (normozoospermia, $n$ = 417) and the other with impaired sperm parameters (pathozoospermia, $n$ = 282) in accordance with reference values of the WHO (WHO…, 2021). The latter had either a sperm concentration of less than 16 mill/mL, a proportion of motile spermatozoa (categories A+B) less than 30 %, a proportion of morphologically normal spermatozoa less than 4 %, or any combination of these deviations. The groups were compared for the frequency of AZFc microdeletions b2/b3 and gr/gr, with the results presented in Table 3. Statistical analysis did not reveal significant differences in the frequency of b2/b3 and gr/gr deletions between the normozoospermia and pathozoospermia groups ($\chi^2_2 = 0.21$; $p < 0.90$). Consequently, pathozoospermia

**Table 3.** Frequency of partial AZFc microdeletions (b2/b3 and gr/gr) in the entire Slavic study population and in the normo- and pathozoospermia groups

| | *n* | b2/b3, *n* (%) | gr/gr, *n* (%) | No deletions, *n* (%) |
|---|---|---|---|---|
| Normozoospermia | 417 | 68 (16.3) | 12 (2.9) | 337 (80.8) |
| Pathozoospermia | 282 | 49 (17.4) | 9 (3.2) | 224 (79.4) |
| Entire study population | 699 | 117 (16.7) | 21 (3.0) | 561 (80.3) |

Note. Normozoospermia – sperm concentration ≥ 16.0 mill/mL, proportion of motile sperm (categories A+B) ≥ 30 %, proportion of morphologically normal sperm ≥ 4.0 % (WHO…, 2021); pathozoospermia – concentration, proportion of progressively motile and morphologically normal sperm below reference values (either each indicator or any combination thereof). The carrier of the complete AZFc microdeletion b2/b4 was not included in the table.

**Table 4.** Frequency of various types of AZFc deletions in the entire Slavic study population and in the groups stratified by sperm concentration

| | *n* | AZFc, *n* (%) | b2/b3 *n* (%) | gr/gr *n* (%) | b2/b4 *n* (%) | No deletions, *n* (%) |
|---|---|---|---|---|---|---|
| Azoospermia | 19 | 4 (21.1) | 1 (5.3) | 2 (10.5) | 1 (5.3) | 15 (78.9) |
| SC ≤ 5 mill/mL | 33 | 6 (18.2) | 6 (18.2) | 0 | 0 | 27 (81.8) |
| 16 > SC > 5 mill/mL | 74 | 18 (24.3) | 15 (20.3) | 3 (4.1) | 0 | 56 (75.7) |
| SC ≥ 16 mill/mL | 574 | 111 (19.3) | 95 (16.6) | 16 (2.8) | 0 | 463 (80.7) |
| Entire study population | 700 | 139 (19.9) | 117 (16.7) | 21 (3.0) | 1 (0.1) | 561 (80.1) |

Note. Azoospermia – no spermatozoa in the ejaculate; SC – sperm concentration.

is not associated with an increased frequency of either type of partial AZFc deletions.

The prevalence of various types of AZFc deletions in the entire study population, as well as in groups with different sperm concentrations, is presented in Table 4. Y chromosome microdeletions were detected in 139 (19.9 %) out of 700 men; no AZFa, AZFb, and AZFb+c deletions were found among them. A complete AZFc deletion (b2/b4) was found in one man (0.1 %) and it was associated with azoospermia. The following types of partial AZFc deletions were identified: gr/gr in 21 (3.0 %) and b2/b3 in 117 (16.7 %) men. The combined frequency of both types of AZFc deletions did not differ between groups with different sperm concentrations ($\chi^2_3 = 1.10$, $p = 0.78$), neither did the frequencies of specific types of partial AZFc deletions – gr/gr ($\chi^2_3 = 4.73$, $p = 0.19$) and b2/b3 ($\chi^2_3 = 2.14$, $p = 0.54$). Therefore, no differences were found in the frequency of AZFc deletions gr/gr and b2/b3 between groups with different sperm concentrations, indicating the absence of an impact of these deletions on sperm production in Slavic men.

### Analysis of associations of partial AZFc microdeletions and spermiological parameters

A comparison of spermiological indicators between men with partial AZFc microdeletions (b2/b3 and gr/gr) and those without microdeletions was made. The results are shown in Table 5. No significant differences were found in any spermiological parameters between the carriers of b2/b3 and gr/gr deletions and men without deletions. Therefore, our study did not establish any impact of partial AZFc deletions (b2/b3 and gr/gr) on the examined semen parameters in Slavic men.

## Discussion

The global prevalence of complete AZF deletions, i. e., those that fully remove one or more regions, among infertile men is 7.5 %, which is significantly higher than in the general population – 0.025 % (Colaco, Modi, 2018; Cioppi et al., 2021). In a multi-ethnic group of Russian infertile men with azoospermia/oligozoospermia, the prevalence of complete AZF deletions ranged from 7.5 to 12 % (Chernykh et al., 2006; Mikhaylenko et al., 2019), which is close to the rates observed in other European and Asian countries. In our Slavic study population from the general Russian population, only one man was identified with a complete AZFc b2/b4 deletion and azoospermia, confirming the low frequency of this type of AZF microdeletions in the general population. A significant increase in sample size is required to determine the prevalence of complete AZFc microdeletion among Slavs.

Among Slavic men from European countries, the prevalence of complete deletions of various AZF regions is lower than that in Russian men. For example, the frequency of complete AZF microdeletions in Slovakia among men with azoospermia was 3.35 % (Behulova et al., 2011); in Slovenia among subfertile men, 4.4 % (Peterlin et al., 2002); and in Macedonia among infertile men, 4.1 % (Plaseski et al., 2006). In the non-Slavic population of Europe, the frequency of complete AZF microdeletions in infertile men varied within the same range of 2.4–4.0 % (Lo Giacco et al., 2014; Mokánszki et al., 2018; Johnson et al., 2019).

In Asian countries, higher frequencies of complete AZF microdeletions have been identified in infertile patients with azoospermia/oligozoospermia compared to European countries: in China, 10.7–12.9 % (Liu et al., 2019; Fu et al., 2023);

**Table 5.** Spermiological parameters in Slavic men with different types of partial AZFc deletions (b2/b3 and gr/gr)

| Parameter | b2/b3, (*n* = 117, 16.7 %) | gr/gr, (*n* = 21, 3.0 %) | No deletion, (*n* = 561, 80.1 %) |
|---|---|---|---|
| BTV, mL | 40.6 (9.1) | 41.1 (8.7) | 40.6 (8.8) |
| Semen volume, mL | 3.3 (1.5) | 3.5 (1.7) | 3.7 (1.8) |
| TSC, mill/ejaculate | 163.8 (131.5) | 148.4 (121.8) | 197.5 (195.8) |
| SC, mill/mL | 51.84 (39.54) | 40.62 (29.90) | 55.31 (44.40) |
| Motility, % | 44.1 (25.9) | 34.3 (19.1) | 46.0 (26.6) |
| Morphology, % | 6.74 (3.23) | 6.45 (3.41) | 7.26 (3.13) |
| TZI | 1.48 (0.12) | 1.52 (0.16) | 1.48 (0.12) |

Note. Data are presented as mean (SD); BTV – bitesticular volume; TSC – total sperm count per ejaculate; SC – sperm concentration; motility – proportion of motile sperm of category A+B; morphology – proportion of morphologically normal sperm; TZI – teratozoospermia index. Motility, morphology, and TZI parameters are calculated excluding participants with azoospermia and severe oligozoospermia.

in Japan, 7.5 % (Iijima et al., 2020); in Turkey, 9.6–25.0 % (Akbarzadeh Khiavi et al., 2020); in Iran, 12.1–20.6 % (Bahmanimehr et al., 2018); in India, 10.0–16.1 % (Waseem et al., 2020). Despite extensive study on the geographic and ethnic variability in the frequency of complete Y-chromosome microdeletions, the underlying causes of this variability remain unknown but are largely thought to be influenced by the inappropriate selection criteria of patients.

Complete deletions of the AZF regions of the Y chromosome are rare, with most (over 80 %) being partial microdeletions of the AZFc region (Krausz, Casamonti, 2017; Cioppi et al., 2021). In our Slavic study population from the general Russian population, two types of partial AZFc deletions were identified – the gr/gr and b2/b3 deletions. The combined prevalence of these types of deletions was 19.7 %, with the frequency of the gr/gr deletion being 3.0 %, and that of the b2/b3 deletion being 16.7 %. Notably, the frequency of the b2/b3 and gr/gr deletions in the normozoospermic group did not differ from that in the pathozoospermic group. Since these types of deletions (b2/b3 and gr/gr) are found in men with normozoospermia, they are not markers of impaired spermatogenesis. It should be noted that information on the frequency of these types of partial AZFc deletions in men from the general population is sparse, but there are data on the frequency of these deletions in Russian infertile men with azoo-/oligozoospermia and in Russian fertile men with normozoospermia (Table 6). In a multi-ethnic Russian group of fertile men with normozoospermia (Barkov et al., 2014) or in men from infertile married couples with normozoospermia (Zobkova et al., 2017), the frequencies of the b2/b3 and gr/gr deletions are close to our data (Table 6). In both studies, no differences in the frequency of the b2/b3 deletion were found between the normozoospermic and azoo-/oligozoospermic groups, which also aligns with our conclusions. In Russian fertile men (sperm data not specified), the frequencies of the b2/b3 and gr/gr deletions practically coincide with our data (Chernykh et al., 2022). In Russian studies (Barkov et al., 2014; Zobkova et al., 2017), the higher frequency of the gr/gr deletion in men with infertility or from infertile couples with pathozoospermia compared to those with normozoospermia (although differences are not statistically significant) draws attention to itself, which may be due to the multi-ethnic composition of the groups studied and, accordingly, different genetic background of the Y chromosome. Collectively, the data confirm that the most common types of partial AZFc deletions in Russian men are b2/b3 and gr/gr, and the practical lack of differences in the frequency of these deletions between men with normozoospermia and pathozoospermia indicates the absence of negative effects of partial b2/b3 and gr/gr deletions on spermatogenesis.

Interestingly, in Estonia, fertile men with normozoospermia or with infertility and pathozoospermia have a higher frequency of b2/b3 deletions compared to our data and a similar frequency of gr/gr deletions (Hallast et al., 2021) (Table 6). About two-thirds of Estonian men carrying the gr/gr deletion belonged to haplogroup R1, and almost all (99.4 %) of the men carrying the b2/b3 deletion belonged to Y-haplogroup N3. The frequency of the b2/b3 deletion did not differ between the pathozoospermic and normozoospermic groups, which is consistent with the conclusions of our study. However, the frequency of the gr/gr deletion was significantly higher in the pathozoospermic group compared to the normozoospermic group. At the same time, andrological parameters in men with either b2/b3 or gr/gr deletions and without deletions did not differ.

In men from other European countries, a lower prevalence of partial AZFc deletions is observed, with most AZFc deletions represented by the gr/gr deletion (Table 6). In Italy (Ferlin et al., 2005), Germany (Hucklenbroich et al., 2005), Spain (Lo Giacco et al., 2014), and Hungary (Mokánszki et al., 2018), the frequency of the b2/b3 deletion among patients with normozoospermia ranged from 0 to 2.7 %, and among those with azoospermia/oligozoospermia, from 0.3 to 2.6 %, while the frequency of the gr/gr deletion among men with normozoospermia ranged from 0.4 to 1.8 % and among those with azoo-/oligozoospermia, from 3.9 to 4.7 %. The results of these European studies suggest that the gr/gr deletion is a genetic cause of reduced sperm production, although some authors believe that the gr/gr deletion is only a risk factor predisposing to impaired spermatogenesis.

**Table 6.** Frequency of partial AZFc microdeletions in the Y chromosome in men of different ethnic origin and region of residence

| Country | *n* | Spermiological phenotype | Number of STS | Frequency of partial AZFc b2/b3 and gr/gr microdeletions, % | Reference |
|---|---|---|---|---|---|
| Russia, Slavs | 700 | The general population:<br>417 pathozoospermia<br>282 normozoospermia | 15 | pathozoospermia: b2/b3 – 16.3; gr/gr – 2.9<br>normozoospermia: b2/b3 – 17.4; gr/gr – 3.2 | Our data |
| Russia | 272 | 146 pathozoospermia (infertile)<br>126 normozoospermia (fertile) | 14 | pathozoospermia: b2/b3 – 13.7; gr/gr – 8.2;<br>normozoospermia: b2/b3 – 14.3; gr/gr – 2.4 | Barkov et al., 2014 |
| | 205 | Infertile couples:<br>143 pathozoospermia<br>62 normozoospermia | 14 | pathozoospermia: b2/b3 – 11.00; gr/gr – 9.79;<br>normozoospermia: b2/b3 – 8.06; gr/gr – 3.21 | Zobkova et al., 2017 |
| | 436 | 436 fertile without semen analysis | 6 | fertile: b2/b3 – 14.7; gr/gr – 2.3 | Chernykh et al., 2022 |
| Estonia | 2324 | No complete AZF deletions:<br>1190 azoo-/oligozoospermia (infertile)<br>1134 normozoospermia (fertile/young) | 9 | azoo-/oligozoospermia: b2/b3 – 32.6; gr/gr – 2.7*;<br>normozoospermia: b2/b3 – 32.4; gr/gr – 1.2 | Hallast et al., 2021 |
| Italy | 600 | 337 azoo-/oligozoospermia (infertile)<br>263 normozoospermia (fertile) | 10+8 | azoo-/oligozoospermia: b2/b3 – 0.3; gr/gr – 4.7*;<br>normozoospermia: b2/b3 – 0; gr/gr – 0.4 | Ferlin et al., 2005 |
| Germany | 518 | No complete AZF deletions:<br>348 azoo-/oligozoospermia<br>170 normozoospermia | 5+4 | azoo-/oligozoospermia: b2/b3 – 0.6; gr/gr – 4.0<br>normozoospermia: b2/b3 – 2.9; gr/gr – 1.8 | Hucklenbroich et al., 2005 |
| Spain | 715 | 330 azoo-/oligozoospermia (infertile)<br>385 normozoospermia (fertile) | 6 | azoo-/oligozoospermia: b2/b3 – 1.3*; gr/gr – 3.9*;<br>normozoospermia: b2/b3 – 0; gr/gr – 1.4 | Lo Giacco et al., 2014 |
| Hungary | 456 | 357 azoo-/oligozoospermia (infertile)<br>111 normozoospermia | 5 | azoo-/oligozoospermia: b2/b3 – 2.6; gr/gr – 4.2*;<br>normozoospermia: b2/b3 – 2.7; gr/gr – 1.8 | Mokánszki et al., 2018 |
| Iran | 265 | 154 azoo-/oligozoospermia (infertile)<br>111 normozoospermia | 6+6 | azoo-/oligozoospermia: b2/b3 – 1.8; gr/gr – 5.2;<br>normozoospermia: b2/b3 – 0; gr/gr – 1.8 | Alimardanian et al., 2016 |
| Turkey | 420 | 333 NOA/OAT (infertile)<br>87 normozoospermia (65 fertile | 5 | NOA/OAT: b2/b3 – 8.1; gr/gr – 5.1;<br>normozoospermia: b2/b3 – 10.3; gr/gr – 12.6 | Beyaz et al., 2017 |
| China, Han | 1435 | No complete AZF deletions:<br>654 NOA/oligozoospermia (infertile)<br>781 fertile | 8 | NOA/oligozoospermia: b2/b3 – 9.5*; gr/gr – 10.7;<br>fertile: b2/b3 – 6.3; gr/gr – 8.6 | Lu et al., 2014 |
| China, Yi | 377 | 224 NOA/oligozoospermia (infertile)<br>153 fertile without semen analysis | 6 | NOA/oligozoospermia: b2/b3 – 6.3; gr/gr – 7.6;<br>fertile: b2/b3 – 8.5; gr/gr – 8.5 | Ye et al., 2013 |
| India, Indo-Europeans | 1047 | 619 pathozoospermia (infertile)<br>203 normozoospermia (infertile)<br>225 fertile without semen analysis | 6+6 | pathozoospermia: b2/b3 – 0.32; gr/gr – 4.84*;<br>normozoospermia: b2/b3 – 0; gr/gr – 8.87*;<br>fertile: b2/b3 – 0.44; gr/gr – 0.89 | Bansal et al., 2016b |

Note. NOA – non-obstructive azoospermia; OAT – oligoastenotheratozoospermia. * Significant differences between groups with pathozoospermia and normozoospermia (fertile) ($p < 0.05$).

In Asian countries, there is greater geographical heterogeneity in the frequency of partial AZFc deletions compared to Russia (Table 6). For example, in Iran, the frequency of the b2/b3 deletion among men ranged from 0 to 1.8 %, which is significantly lower than the data from Russia, while the gr/gr deletion ranged from 1.8 to 5.2 %, which is close to the Russian data, and no negative impact of these types of partial deletions on spermatogenesis was identified (Alimardanian et al., 2016). In Turkish men, the frequency of partial b2/b3 deletions is significantly lower compared to Russian men, but both types of deletions (b2/b3 and gr/gr) are not associated with infertility and reduced sperm production (Beyaz et al., 2017). Conversely, screening for partial gr/gr and b2/b3 deletions in the Han Chinese population showed that the frequency of the b2/b3 deletion was significantly higher in patients with infertility and azoo-/oligozoospermia compared to fertile men with normozoospermia, indicating an association of this deletion with impaired spermatogenesis (Lu et al., 2014). However, in Chinese men from another ethnic group, Yi, no such association was found (Ye et al., 2013), underscoring the importance of considering the ethnic composition of the population when studying the effects of partial AZFc deletions in the Y chromosome on spermatogenesis. In the Indian population, the frequency of the b2/b3 deletion was 40–45 times lower, and the frequency of the gr/gr deletion was 2–3 times higher compared to the Russian populations, with gr/gr deletions being the most common and significant among partial AZFc deletions, reducing sperm concentration and increasing the risk of infertility (Bansal et al., 2016b) (Table 6).

The provided data support the previously expressed idea (Rozen et al., 2012) that the geographical and ethnic origin of a population may influence the frequency of partial AZFc deletions b2/b3 and gr/gr. In that study, the prevalence of partial gr/gr and b2/b3 microdeletions was evaluated in 20,884 men from five populations (India, Poland, Tunisia, USA, Vietnam). It was found that the frequency of the gr/gr partial deletion varied from 2.1 % (USA) to 15 % (Vietnam), and b2/b3, from 0.5 % (India) to 2.2 % (Poland). The authors suggested that ethnogeographic differences in the frequency of the b2/b3 deletion are likely due to differences in the prevalence of Y haplogroup N1, the high prevalence of the gr/gr deletion might be due to the prevalence of haplogroup D2a chromosomes (containing these deletions), which corresponds to the hypothesis of the relationship between the frequency of partial AZFc deletions and the genetic background of the Y chromosome.

Analysis of spermiological phenotypes in carriers of Y chromosome microdeletions shows that while complete deletions of one or more AZF regions are associated with impaired spermatogenesis and are specific genetic markers of spermatogenesis failure and infertility, partial AZFc deletions exhibit heterogeneity in terms of spermiological phenotype and are often only risk factors predisposing to pathozoospermia and infertility. In our study population of Slavic men from the general population, no negative effects of b2/b3 and gr/gr deletions on spermatogenesis were detected. In another Russian study, among men with infertility who were carriers of gr/gr and b2/b3 microdeletions (3.5 and 7.9 % of the total number examined), sperm concentration was 12.2 ± 7.1 and 30.3 ± 5.3 mill/mL, respectively, i. e., in carriers of the gr/gr deletion, it was below the reference value of the norm (Chernykh et al., 2014). The negative association of the gr/gr deletion with sperm concentration may be due to the multiethnic composition of the group and, consequently, different genetic backgrounds of the Y chromosome. The gr/gr microdeletion did not show a statistically significant association with spermatogenesis failure in other Slavic populations, such as Bulgarians (Levkova et al., 2020) and Macedonians (Kuzmanovska et al., 2019).

In some Mongoloid populations, no negative effects of the gr/gr deletion on spermatogenesis have been established, for example, in Han Chinese (Yang et al., 2010) or Japanese (Sin et al., 2010), if the gr/gr deletion is fixed in the prevalent haplogroups Q1 and D2b, respectively, which plays a role in the clinical manifestation of the deletion. However, in Koreans, the gr/gr deletion causes spermatogenesis impairment if it is not fixed in the prevalent haplogroup YAP (the precursor of haplogroup D2b), and a normal testicular phenotype is observed when it is in haplogroup YAP (Choi et al., 2012). In the African population of Tunisia, partial gr/gr and b2/b3 deletions are not associated with spermatogenesis failure, which is due to the fixation of these deletions in haplogroup E3b2, which is widespread in North Africa (Ghorbel et al., 2012). Recall that in European populations, the gr/gr deletion is associated with spermatogenesis impairment, particularly in Spaniards (Lo Giacco et al., 2014), Italians (Ferlin et al., 2005), and Hungarians (Mokánszki et al., 2018).

Many authors conclude that the influence of the partial AZFc b2/b3 deletion on spermatogenesis and fertility largely depends on the ethnic composition of the studied population, and the frequency and phenotypic effect are determined by the origin of the Y chromosome. The b2/b3 deletion is a risk factor for spermatogenesis impairment in East Asian and African populations but not in European or South Asian populations (Bansal et al., 2016a; Colaco, Modi, 2018). In the Chinese population, the b2/b3 deletion increases the risk of spermatogenesis impairment and predisposes to the formation of a complete AZFc region deletion (Lu et al., 2014). However, in the ethnic Han Chinese population from eastern China, the b2/b3 deletion is not associated with spermatogenesis impairment, which the authors attribute to interpopulation differences in the frequencies of Y haplogroups in China (Zhang et al., 2007).

In Finno-Ugric, Balto-Slavic, and some Turkish-speaking peoples of Northern Eurasia, the partial AZFc b2/b3 deletion is fixed in Y haplogroup N, which has a high frequency (up to 90 % in some populations) (Repping et al., 2004). It has been established that the b2/b3 deletion does not cause spermatogenesis impairment in Germans (Hucklenbroich et al., 2005); Tunisians (Ghorbel et al., 2012); Iranians (Alimardanian et al., 2016); Hungarians (Mokánszki et al., 2018); and Estonians (Hallast et al., 2021). The population of ethnic Russians can also be included in this group due to the lack of association of this deletion with spermatogenesis failure, the high frequency of the b2/b3 deletion, and the wide prevalence of Y haplogroup N3 carrying this deletion, which varies between 10–19 % (Stepanov et al., 2006; Balanovska, Balanovsky, 2007; Derenko et al., 2007). It is hypothesized that the effect

of the b2/b3 deletion in haplogroup N is balanced by other genetic factors, possibly related to the Y chromosome (Repping et al., 2004).

Thus, the effects of partial AZFc deletions on spermatogenesis can depend on the lineage of the Y chromosome (the Y haplogroup carrying the deletion), increasing or decreasing the risk of spermatogenesis impairment in certain populations. Since partial deletions of the AZFc region can be fixed in specific Y haplogroups, the population frequency of partial AZFc deletions depends on the frequency of these Y haplogroups, and the impact of partial AZFc deletions on spermatogenesis may differ in the Y chromosomes of different haplogroups. For example, in Japanese men, the two most common Y haplogroups are D (34.7 %) and O (51.8 %). The gr/gr deletions were found in 33.7 % of Japanese men, but the frequency of the gr/gr deletion varied significantly depending on the Y haplogroup: it was widespread in haplogroup D (86.2 %) and much less so in haplogroup O (5.1 %), with it being phenotypically neutral in haplogroup D, meaning it did not affect spermatogenesis, while in haplogroup O, it reduced sperm concentration (Sin et al., 2010). In men from Northern Italy, a comparison of the distribution of seven Y haplogroups between a group of fertile men and patients with microdeletions did not reveal any differences, but the frequency of Y haplogroup E with the b2/b4 deletion was significantly higher compared to the control. The results suggest that some haplogroups may be more prone to AZFc b2/b4 microdeletions than others (Arredi et al., 2007).

In the population of ethnic Russians in Russia, the dominant Y haplogroup is R1a, which is the most common (over 40 %), followed by N3 (10–19 %), and I1b (13 %) (Stepanov et al., 2006; Balanovska, Balanovsky, 2007; Ilumäe et al., 2016). Although the association between the phenotypic expression of partial AZFc deletions in the Y chromosome and Y haplogroups remains a subject of discussion, considering the aforementioned facts, it appears promising to study the association of the main haplogroups R1a, N3, and I1b with spermatogenesis indicators in Slavic men. This will help to elucidate a modulatory effect of the haplogroup on the phenotypic expression of partial AZFc deletion. Special attention should be given to haplogroups N3 and R1a, which contain the b2/b3 deletion and gr/gr deletion, respectively (Repping et al., 2004; Rozen et al., 2012). The prevalence of the b2/b3 deletion in our study (16.7 %) coincides with the prevalence of haplogroup N3 in the Russian population (10–19 %); in this haplogroup, this deletion does not affect spermatogenesis, although in another haplogroup it may have a negative effect on the spermatogenic phenotype. It is assumed that in haplogroup N3, the negative impact of the deletion on spermatogenesis is balanced by other genetic factors, possibly also associated with the Y chromosome (Repping et al., 2004).

## Conclusion

In a study population of Slavic men recruited from five cities in Russia ($n$ = 700), the spectrum and frequency of AZFc microdeletions in the Y chromosome were determined. It was found that 19.9 % were carriers of AZFc deletions, of which 16.7 % were carriers of a partial b2/b3 deletion, 3.0 % had a partial gr/gr deletion, and 0.14 % had a complete b2/b4 deletion. No AZFa and AZFb microdeletions or other types of AZF deletions were found.

The overall frequency of all types of AZFc deletions, as well as partial b2/b3 and gr/gr deletions, did not differ between the groups with normozoospermia and pathozoospermia, nor between the groups with azoospermia, severe oligozoospermia, oligozoospermia, and normal sperm concentration. Semen parameters did not differ between the groups with different types of partial AZFc deletions and the group without deletions. The data obtained indicate the absence of a pathogenic role of partial AZFc deletions in spermatogenesis of Slavic men.

## References

Akbarzadeh Khiavi M., Jalili A., Safary A., Gharedaghchi Z., Mirinezhad S.K., Mehdizadeh A., Rahmani S.A. Karyotypic abnormalities and molecular analysis of Y chromosome microdeletion in Iranian Azeri Turkish population infertile men. *Syst. Biol. Reprod. Med.* 2020;66(2):140-146. DOI 10.1080/19396368.2019.1682083

Alimardanian L., Saliminejad K., Razi S., Ahani A. Analysis of partial azoospermia factor c deletion and DAZ copy number in azoospermia and severe oligozoospermia. *Andrologia.* 2016;48(9):890-894. DOI 10.1111/and.12527

Arredi B., Ferlin A., Speltra E., Bedin C., Zuccarello D., Ganz F., Marchina E., Stuppia L., Krausz C., Foresta C. Y-chromosome haplogroups and susceptibility to azoospermia factor c microdeletion in an Italian population. *J. Med. Genet.* 2007;44(3):205-208. DOI 10.1136/jmg.2006.046433

Bahmanimehr A., Zeighami S., Namavar Jahromi B., Anvar Z., Parsanezhad M.E., Davari M., Montazeri S. Detection of Y chromosome microdeletions and hormonal profile analysis of infertile men undergoing assisted reproductive technologies. *Int. J. Fertil. Steril.* 2018;12(2):173-177. DOI 10.22074/ijfs.2018.5244

Balanovska E.V., Balanovsky O.P. The Russian Gene Pool on the Russian Plain. Moscow, 2007 (in Russian)

Bansal S.K., Gupta G., Rajender S. Y chromosome b2/b3 deletions and male infertility: a comprehensive meta-analysis, trial sequential analysis and systematic review. *Mutat. Res. Rev. Mutat. Res.* 2016a; 768:78-90. DOI 10.1016/j.mrrev.2016.04.007

Bansal S.K., Jaiswal D., Gupta N., Singh K., Dada R., Sankhwar S.N., Gupta G., Rajender S. Gr/gr deletions on Y-chromosome correlate with male infertility: an original study, meta-analyses, and trial sequential analyses. *Sci. Rep.* 2016b;6:19798. DOI 10.1038/srep 19798

Barkov I.Yu., Soroka N.E., Popova A.Yu., Gamidov S.I., Belyaeva N.A., Glinkina Zh.I., Kalinina E.A., Trofimov D.Yu., Sukhikh G.T. Diagnosis of male infertility associated with microdeletions at the AZF locus of the Y chromosome. *Akusherstvo i Ginekologiya = Obstetrics and Gynecology.* 2014;1:59-64 (in Russian)

Behulova R., Varga I., Strhakova L., Bozikova A., Gabrikova D., Boronova I., Repiska V. Incidence of microdeletions in the AZF region of the Y chromosome in Slovak patients with azoospermia. *Biomed. Pap. Med. Fac. Univ. Palacky Olomouc Czech. Repub.* 2011;155(1): 33-38. DOI 10.5507/bp.2011.006

Beyaz C.C., Gunes S., Onem K., Kulac T., Asci R. Partial deletions of Y-chromosome in infertile men with non-obstructive azoospermia and oligoasthenoteratozoospermia in a Turkish population. *In Vivo.* 2017;31(3):365-371. DOI 10.21873/invivo.11068

Chernykh V.B., Chukhrova A.L., Beskorovainaya T.S., Grishina E.M., Sorokina T.M., Shileiko L.V., Gogolevsky P.A., Kalugina A.S., Morina G.V., Togobetsky A.S., Tanevsky V.E., Zdanovsky V.M., Gogolevskaya I.K., Kramerov D.A., Polyakov A.V., Kurilo L.F. Types of Y chromosome deletions and their frequency in infertile men. *Russ. J. Genet.* 2006;42(8):936-941. DOI 10.1134/S102279540 6080138

Chernykh V.B., Rudneva S.A., Sorokina T.M., Shileyko L.V., Kurilo L.F., Ryzhkova O.P., Chukhrova A.L., Polyakov A.V. Characteristics of spermatogenesis in infertile men with the AZFc region deletions. *Andrologiya i Genitalnaya Khirurgiya = Andrology and Genital Surgery.* 2014;2:48-57 (in Russian)

Chernykh V.B., Ryzhkova O.P., Kuznetsova I.A., Kazaryan M.S., Sorokina T.M., Kurilo L.F., Schagina O.A., Polyakov A.V. Deletions in AZFc region of Y chromosome in Russian fertile men. *Russ. J. Genet.* 2022;58(7):850-856. DOI 10.1134/s1022795422070043

Choi J., Song S.H., Bak C.W., Sung S.R., Yoon T.K., Lee D.R., Shim S.H. Impaired spermatogenesis and gr/gr deletions related to Y chromosome haplogroups in Korean men. *PLoS One.* 2012;7(8): e43550. DOI 10.1371/journal.pone.0043550

Cioppi F., Rosta V., Krausz C. Genetics of azoospermia. *Int. J. Mol. Sci.* 2021;22(6):3264. DOI 10.3390/ijms22063264

Colaco S., Modi D. Genetics of the human Y chromosome and its association with male infertility. *Reprod. Biol. Endocrinol.* 2018; 16(1):14. DOI 10.1186/s12958-018-0330-5

Deng C.Y., Zhang Z., Tang W.H., Jiang H. Microdeletions and vertical transmission of the Y-chromosome azoospermia factor region. *Asian J. Androl.* 2023;25(1):5-12. DOI 10.4103/aja2021130

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

Ferlin A., Tessari A., Ganz F., Marchina E., Barlati S., Garolla A., Engl B., Foresta C. Association of partial AZFc region deletions with spermatogenic impairment and male infertility. *J. Med. Genet.* 2005;42(3):209-213. DOI 10.1136/jmg.2004.025833

Fu M., Chen M., Guo N., Lin M., Li Y., Huang H., Cai M., Xu L. Molecular genetic analysis of 1,980 cases of male infertility. *Exp. Ther. Med.* 2023;26(1):345. DOI 10.3892/etm.2023.12044

Ghorbel M., Gargouri S.B., Zribi N., Abdallah F.B., Cherif M., Keskes R., Chakroun N., Sellami A., McElreavey K., Fakhfakh F., Ammar-Keskes L. Partial microdeletions in the Y-chromosome AZFc region are not a significant risk factor for spermatogenic impairment in Tunisian infertile men. *Genet. Test. Mol. Biomarkers.* 2012;16(7): 775-779. DOI 10.1089/gtmb.2012.0024

Hallast P., Kibena L., Punab M., Arciero E., Rootsi S., Grigorova M., Flores R., Jobling M.A., Poolamets O., Pomm K., Korrovits P., Rull K., Xue Y., Tyler-Smith C., Laan M. A common 1.6 MB Y-chromosomal inversion predisposes to subsequent deletions and severe spermatogenic failure in humans. *eLife.* 2021;10:e65420. DOI 10.7554/eLife.65420

Hucklenbroich K., Gromoll J., Heinrich M., Hohoff C., Nieschlag E., Simoni M. Partial deletions in the AZFc region of the Y chromosome occur in men with impaired as well as normal spermatogenesis. *Hum. Reprod.* 2005;20(1):191-197. DOI 10.1093/humrep/deh558

Iijima M., Shigehara K., Igarashi H., Kyono K., Suzuki Y., Tsuji Y., Kobori Y., Kobayashi H., Mizokami A. Y chromosome microdeletion screening using a new molecular diagnostic method in 1030 Japanese males with infertility. *Asian J. Androl.* 2020;22(4):368-371. DOI 10.4103/aja.aja_97_19

Ilumäe A.M., Reidla M., Chukhryaeva M., Järve M., Post H., Karmin M., Saag L., Agdzhoyan A., Kushniarevich A., Litvinov S., Ekomasova N., Tambets K., Metspalu E., Khusainova R., Yunusbayev B., Khusnutdinova E.K., Osipova L.P., Fedorova S., Utevska O., Koshel S., Balanovska E., Behar D.M., 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

Johnson M., Raheem A., De Luca F., Hallerstrom M., Zainal Y., Poselay S., Mohammadi B., Moubasher A., Johnson T.F., Muneer A., Sangster P., Ralph D.J. An analysis of the frequency of Y-chromosome microdeletions and the determination of a threshold sperm concentration for genetic testing in infertile men. *BJU Int.* 2019; 123(2):367-372. DOI 10.1111/bju.14521

Krausz C., Casamonti E. Spermatogenic failure and the Y chromosome. *Hum. Genet.* 2017;136(5):637-655. DOI 10.1007/s00439-017-1793-8

Krausz C., Cioppi F., Riera-Escamilla A. Testing for genetic contributions to infertility: potential clinical impact. *Expert Rev. Mol. Diagn.* 2018;18(4):331-346. DOI 10.1080/14737159.2018.1453358

Krausz C., Navarro-Costa P., Wilke M., Tüttelmann F. EAA/EMQN best practice guidelines for molecular diagnosis of Y-chromosomal microdeletions: state of the art 2023. *Andrology.* 2024;12(3):487-504. DOI 10.1111/andr.13514

Kuroda S., Usui K., Sanjo H., Takeshima T., Kawahara T., Uemura H., Yumura Y. Genetic disorders and male infertility. *Reprod. Med. Biol.* 2020;19(4):314-322. DOI 10.1002/rmb2.12336

Kuzmanovska M., Noveski P., Terzic M., Plaseski T., Kubelka-Sabit K., Filipovski V., Lazarevski S., Sukarova Stefanovska E., Plaseska-Karanfilska D. Y-chromosome haplogroup architecture confers susceptibility to azoospermia factor c microrearrangements: a retrospective study. *Croat. Med. J.* 2019;60(3):273-283. DOI 10.3325/cmj.2019.60.273

Lebedev G.S., Golubev N.A., Shaderkin I.A., Shaderkina V.A., Apolikhin O.I., Sivkov A.V., Komarova V.A. Male infertility in the Russian Federation: statistical data for 2000–2018. *Eksperimental'naya i Klinicheskaya Urologiya = Experimental and Clinical Urology.* 2019; 4:4-12. DOI 10.29188/2222-8543-2019-11-4-4-12 (in Russian)

Levkova M., Chervenkov T., Angelova L. The association of gr/gr deletion in the Y chromosome and impaired spermatogenesis in Bulgarian males: a pilot study. *Middle East Fertil. Soc. J.* 2020;25:10. DOI 10.1186/s43043-020-00020-9

Liu T., Song Y.X., Jiang Y.M. Early detection of Y chromosome microdeletions in infertile men is helpful to guide clinical reproductive treatments in southwest of China. *Medicine* (*Baltimore*). 2019; 98(5):e14350. DOI 10.1097/MD.0000000000014350

Lo Giacco D., Chianese C., Sánchez-Curbelo J., Bassas L., Ruiz P., Rajmil O., Sarquella J., Vives A., Ruiz-Castañé E., Oliva R., Ars E., Krausz C. Clinical relevance of Y-linked CNV screening in male infertility: new insights based on the 8-year experience of a diagnostic genetic laboratory. *Eur. J. Hum. Genet.* 2014;22(6):754-761. DOI 10.1038/ejhg.2013.253

Lu C., Jiang J., Zhang R., Wang Y., Xu M., Qin Y., Lin Y., Guo X., Ni B., Zhao Y., Diao N., Chen F., Shen H., Sha J., Xia Y., Hu Z., Wang X. Gene copy number alterations in the azoospermia-associated AZFc region and their effect on spermatogenic impairment. *Mol. Hum. Reprod.* 2014;20(9):836-843. DOI 10.1093/molehr/gau043

Mikhaylenko D.S., Sobol I.Y., Safronova N.Y., Simonova O.A., Efremov E.A., Efremov G.D., Alekseev B.Y., Kaprin A.D., Nemtsova M.V. The incidence of AZF deletions, *CFTR* mutations and long alleles of the *AR* CAG repeats during the primary laboratory diagnostics in a heterogeneous group of infertily men. *Urologiia.* 2019; 3:101-107. DOI 10.18565/urology.2019.3.101-107 (in Russian)

Mokánszki A., Ujfalusi A., Gombos É., Balogh I. Examination of Y-chromosomal microdeletions and partial microdeletions in idiopathic infertility in East Hungarian patients. *J. Hum. Reprod. Sci.* 2018;11(4):329-336. DOI 10.4103/jhrs.JHRS_12_18

Osadchuk L.V., Shantanova L.N., Troev I.V., Kleshchev M.A., Osadchuk A.V. Regional and ethnic differences in semen quality and reproductive hormones in Russia: a Siberian population-based cohort study of young men. *Andrology.* 2021;9:1512-1525. DOI 10.1111/andr.13024

Osadchuk L., Vasiliev G., Kleshchev M., Osadchuk A. Androgen receptor gene CAG repeat length varies and affects semen quality in an ethnic-specific fashion in young men from Russia. *Int. J. Mol. Sci.* 2022;23(18):10594. DOI 10.3390/ijms231810594

Pan Y., Li L.L., Yu Y., Jiang Y.T., Yang X., Zhang H.G., Liu R.Z., Wang R.X. Natural transmission of b2/b3 subdeletion or duplication to expanded Y chromosome microdeletions. *Med. Sci. Monit.* 2018;24:6559-6563. DOI 10.12659/MSM.911644

Peterlin B., Kunej T., Sinkovec J., Gligorievska N., Zorn B. Screening for Y chromosome microdeletions in 226 Slovenian subfertile men. *Hum. Reprod.* 2002;17(1):17-24. DOI 10.1093/humrep/17.1.17

Plaseski T., Novevski P., Kocevska B., Dimitrovski C., Efremov G.D., Plaseska-Karanfilska D. AZF deletions in infertile men from the Republic of Macedonia. *Prilozi.* 2006;27(1):5-16

Repping S., van Daalen S.K., Korver C.M., Brown L.G., Marszalek J.D., Gianotten J., Oates R.D., Silber S., van der Veen F., Page D.C., Rozen S. A family of human Y chromosomes has dispersed throughout northern Eurasia despite a 1.8-Mb deletion in the azoospermia factor c region. *Genomics.* 2004;83(6):1046-1052. DOI 10.1016/j.ygeno.2003.12.018

Rozen S.G., Marszalek J.D., Irenze K., Skaletsky H., Brown L.G., Oates R.D., Silber S.J., Ardlie K., Page D.C. AZFc deletions and spermatogenic failure: a population-based survey of 20,000 Y chromosomes. *Am. J. Hum. Genet.* 2012;91(5):890-896. DOI 10.1016/j.ajhg.2012.09.003

Sin H.S., Koh E., Shigehara K., Sugimoto K., Maeda Y., Yoshida A., Kyono K., Namiki M. Features of constitutive gr/gr deletion in a Japanese population. *Hum. Reprod.* 2010;25(9):2396-2403. DOI 10.1093/humrep/deq191

Stepanov V.A., Khar'kov V.N., Puzyrev V.P. Evolution and phylogeography of human Y-chromosomal lineages. *Informatsionnyy Vestnik VOGiS = The Herald of Vavilov Society for Geneticists and Breeders.* 2006;10(1):57-74 (in Russian)

Waseem A.S., Singh V., Makker G.C., Trivedi S., Mishra G., Singh K., Rajender S. AZF deletions in Indian populations: original study and meta-analyses. *J. Assist. Reprod. Genet.* 2020;37(2):459-469. DOI 10.1007/s10815-019-01661-0

WHO Laboratory Manual for the Examination and Processing of Human Semen. 5th ed. Geneva: World Health Organization, 2010

WHO Laboratory Manual for the Examination and Processing of Human Semen. 6th ed. Geneva: World Health Organization, 2021

Yang Y., Ma M., Li L., Su D., Chen P., Ma Y., Liu Y., Tao D., Lin L., Zhang S. Differential effect of specific gr/gr deletion subtypes on spermatogenesis in the Chinese Han population. *Int. J. Androl.* 2010;33(5):745-754. DOI 10.1111/j.1365-2605.2009.01015.x

Ye J.J., Ma L., Yang L.J., Wang J.H., Wang Y.L., Guo H., Gong N., Nie W.H., Zhao S.H. Partial AZF c duplications not deletions are associated with male infertility in the Yi population of Yunnan Province, China. *J. Zhejiang Univ. Sci. B.* 2013;14(9):807-815. DOI 10.1631/jzus.B1200301

Zhang F., Lu C., Li Z., Xie P., Xia Y., Zhu X., Wu B., Cai X., Wang X., Qian J., Wang X., Jin L. Partial deletions are associated with an increased risk of complete deletion in AZFc: a new insight into the role of partial AZFc deletions in male infertility. *J. Med. Genet.* 2007;44(7):437-444. DOI 10.1136/jmg.2007.049056

Zobkova G.Yu., Baranova E.E., Donnikov A.E., Mskhalaya G.J., Zaletova V.V., Koshkina T.E., Trofimov D.Yu. Range of azoospermia factor (AZF) deletions in men with normal and disturbed spermatogenesis. *Problemy Reproduktsii = Russian Journal of Human Reproduction.* 2017;4:109-113. DOI 10.17116/repro2017234109-113 (in Russian)

**Conflict of interest.** The authors declare no conflict of interest.
Received February 8, 2024. Revised July 1, 2024. Accepted July 8, 2024.

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2024;28(7):792-800

DOI 10.18699/vjgb-24-87

# Polymorphic variants of the genes for enzymes of the antioxidant system, apoptosis and inflammation as potential predictors of myocardial infarction

T.R. Nasibullin [1] ✉, V.V. Erdman [1], Y.R. Timasheva [1], I.A. Tuktarova [1], A.A. Petinseva [1, 2], G.F. Korytina [1]

[1] Institute of Biochemistry and Genetics – Subdivision of the Ufa Federal Research Centre of the Russian Academy of Sciences, Ufa, Russia
[2] Ufa University of Science and Technology, Ufa, Russia
✉ nasibullintr@yandex.ru

**Abstract.** Myocardial infarction (MI) is a multifactorial polygenic disease that develops as a result of a complex interaction of numerous genetic factors and the external environment. Accordingly, the contribution of each of them separately is usually not large and may significantly depend on the state of other accompanying factors. The purpose of the study was to search for informative predictors of MI risk based on polygenic analysis of polymorphic variants of (1) the antioxidant defense enzyme genes *PON1* (rs662), *PON2* (rs7493), *CAT* (rs1001179), *MSRA* (rs10098474) and *GSTP1* (rs1695); (2) the apoptosis genes *CASP8* (rs3834129), *TP53* (rs1042522) and *BCL2* (rs12454712); and (3) the inflammation genes *CRP* (rs1205), *CX3CR1* (rs3732378), *IL6* (rs1800795) and *CCL2* (rs1024611). 591 DNA samples were used in the study (280 patients with the onset at 30 to 60 years, with an average age of 46.02 ± 6.17, and 311 control subjects aged 30 to 62, with an average age of 44.65 ± 7.07). All the participants were male and Tatars by ethnicity. The logistic regression analysis with various models demonstrated associations with MI of polymorphic variants of the genes *CX3CR1* (rs3732378) (overdominant model – G/G + A/A vs A/G $P = 0.0002$, OR = 1.9), *MSRA* (rs10098474) (dominant model – T/T vs T/C + C/C $P = 0.015$, OR = 1.51), *CCL2* (rs1024611) (recessive model – $P = 0.0007$ – A/A + A/G vs G/G OR = 2.63), *BCL2* (rs12454712) (log-additive model – *C allele, $P = 0.005$, OR = 1.38). Using the Monte Carlo method and Markov chains (APSampler), combinations of alleles/genotypes of the studied polymorphic loci associated with a high risk of MI were obtained, which, in addition to those identified during single-locus analysis, contained polymorphic variants of the genes *CASP8*, *TP53*, *CAT*, *PON2*, *CRP*, *IL6*, *GSTP1*. Among the combinations obtained, a pairwise analysis of possible non-linear interactions between the identified combinations of alleles/genotypes was carried out, which showed synergistic interactions of the polymorphic variants *CX3CR1**A/G and *CASP8**I/I, *MSRA**C and *CRP**C, *CAT**C/T and *MSRA**C, *CAT**C/T and *CX3CR1**A contributing to the development of MI. Based on the results obtained using multivariate logistic regression analysis, a predictive model was built to assess the risk of developing MI, the predictive ability of which reached the value AUC = 0.71 (AUC – area under the curve in ROC analysis).
**Key words:** myocardial infarction; oxidative stress; apoptosis; inflammation.

**For citation:** Nasibullin T.R., Erdman V.V., Timasheva Y.R., Tuktarova I.A., Petinseva A.A., Korytina G.F. Polymorphic variants of the genes for enzymes of the antioxidant system, apoptosis and inflammation as potential predictors of myocardial infarction. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov Journal of Genetics and Breeding*. 2024;28(7):792-800. DOI 10.18699/vjgb-24-87

**Funding.** The study was performed with the support of the grant of Russian Science Foundation 24-25-00179 (https://rscf.ru/project/24-25-00179/). The work was performed using the equipment of the Centre for Collective Use "Biomika" and the unique KODINK research facility (Institute of Biochemistry and Genetics, Ufa Federal Research Centre, Russian Academy of Sciences).

# Полиморфные варианты генов ферментов антиоксидантной системы, апоптоза и воспаления как потенциальные предикторы инфаркта миокарда

Т.Р. Насибуллин [1] ✉, В.В. Эрдман [1], Я.Р. Тимашева [1], И.А. Туктарова [1], А.А. Петинцева [1, 2], Г.Ф. Корытина [1]

[1] Институт биохимии и генетики – обособленное структурное подразделение Уфимского федерального исследовательского центра Российской академии наук, Уфа, Россия
[2] Уфимский университет науки и технологий, Уфа, Россия
✉ nasibullintr@yandex.ru

**Аннотация.** Инфаркт миокарда (ИМ) – многофакторное полигенное заболевание, развивающееся в результате сложного взаимодействия многочисленных генетических факторов и внешней среды. Соответственно, вклад каждого из них по отдельности, как правило, невелик и может существенно зависеть от состояния других сопут-

ствующих факторов. Цель исследования – поиск информативных предикторов развития ИМ на основе полигенного анализа полиморфных вариантов генов ферментов антиоксидантной защиты – *PON1* (rs662), *PON2* (rs7493), *CAT* (rs1001179), *MSRA* (rs10098474), *GSTP1* (rs1695); апоптоза – *CASP8* (rs3834129), *TP53* (rs1042522), *BCL2* (rs12454712); воспаления *CRP* (rs1205), *CX3CR1* (rs3732378), *IL6* (rs1800795), *CCL2* (rs1024611). В работе использованы образцы: 591 – ДНК (280 больных, перенесших ИМ в возрасте от 30 до 60 лет, средний возраст 46.02 ± 6.17; 311 – контроль, возраст от 30 до 62 лет, средний возраст 44.65 ± 7.07). Все участники исследования – мужчины, татары по этнической принадлежности. С помощью логистического регрессионного анализа с учетом различных моделей выявлены ассоциации с ИМ полиморфных вариантов генов *CX3CR1* (rs3732378) (сверхдоминантная модель – G/G+A/A vs A/G $P = 0.0002$, OR = 1.9), *MSRA* (rs10098474) (доминантная модель – T/T vs T/C+C/C $P = 0.015$, OR = 1.51), *CCL2* (rs1024611) (рецессивная модель – $P = 0.0007$ – A/A+A/G vs G/G OR = 2.63), *BCL2* (rs12454712) (лог-аддитивная модель – аллель *C, $P = 0.005$, OR = 1.38). С применением метода Монте-Карло и цепей Маркова (APSampler) получены сочетания аллелей/генотипов изученных полиморфных локусов, ассоциированных с высоким риском ИМ, в составе которых, помимо обнаруженных в ходе анализа ассоциаций ИМ и отдельных полиморфных вариантов, присутствуют полиморфные варианты генов *CASP8*, *TP53*, *CAT*, *PON2*, *CRP*, *IL6*, *GSTP1*. Среди этих сочетаний проведен попарный анализ возможного нелинейного взаимодействия между выявленными комбинациями аллелей/генотипов, который показал синергетические взаимодействия полиморфных вариантов *CX3CR1**A/G и *CASP8**I/I, *MSRA**C и *CRP**C, *CAT**C/T и *MSRA**C, *CAT**C/T и *CX3CR1**A, способствующие развитию ИМ. На основе полученных результатов с использованием многофакторного логистического регрессионного анализа построена предиктивная модель для оценки риска развития ИМ, предсказательная способность которой достигла значения AUC = 0.71 (AUC (area under curve) – площадь под кривой при ROC-анализе).

**Ключевые слова:** инфаркт миокарда; окислительный стресс; апоптоз; воспаление.

## Introduction

Myocardial infarction (MI), the most severe clinical variant of coronary heart disease (CHD), significantly reduces life expectancy and quality of life (Shalnova et al., 2022; Sabgayda et al, 2023). In this regard, analysis of its development factors is a crucial task for its prevention. An aggravated family history is one of the main independent risk factors, confirmed by large-scale prospective studies (Colditz et al., 1991; Assmann et al., 2002).

Currently, the molecular genetic basis of hereditary predisposition to MI is actively studied. Genome-wide association studies (GWAS) helped identify a significant number of polymorphic variants associated with CHD in general and MI in particular. At the same time, the obtained results have weak reproducibility; in addition, despite significant advances in the search for genetic variants associated with the pathology under study being made, they do not yield progress in disease prediction.

MI is a multifactorial polygenic disease caused by numerous complexly interacting genetic and environmental factors; the role of individual factors is usually small; moreover, it varies significantly depending on the environment (Domingo et al., 2019). In this regard, a promising direction lies in analyzing associations of polymorphic variant combinations with the studied polygenic trait. At the same time, since increasing the number of elements that make up a combination exponentially increases possible combinations and, as a consequence, decreases the frequency of their occurrence, it seems more rational to limit the number of variables based on already known data on the pathogenesis of the disease, or to include in the analysis polymorphic variants that, according to GWAS, are associated with the studied phenotype.

As known, reactive oxygen species formed during various oxidation-reduction reactions can have a damaging effect on cellular structures and initiate the oxidation of lipids, proteins, and nucleic acids (Batty et al., 2022). Depending on the strength of the effect, they can initiate either inflammation or apoptosis which play a significant role in atherosclerosis development.

Based on the above, this study aims to comprehensively analyze polymorphic variants of the genes of antioxidant defense, inflammation, and apoptosis enzymes (Table 1) as potential predictors of the risk of MI.

## Materials and methods

The study material was DNA samples of unrelated patients with onset of large-focal MI at the age of 30 to 60 years ($N = 280$, the mean age was 46.02 ± 6.17). All the patients were treated at the Republic Center of Cardiology, Ufa.

MI was diagnosed on the basis of the 2012 AHA/ESC guidelines using contemporary instrumental and biochemical methods, including 12- lead ECG, echocardiography, radiography of the thoracic organs, clinical and biochemical blood tests, and assessment of myocardial necrosis markers and lipid spectrum. Patients with endocrine pathology and other concomitant severe chronic diseases were excluded. The control group included unrelated individuals at the age of 30 to 62 years ($N = 311$, the mean age was 44.65 ± 7.08) without clinical signs of cardiovascular pathology. All participants were men belonging to the Tatar ethnic origin, living in Ufa, Republic of Bashkortostan. The study was approved by the Ethics Committee of the Institute of Biochemistry and Genetics (Protocol No. 14 dated 22.12.2010). All participants gave their informed consent.

DNA was isolated by phenol-chloroform extraction. The polymorphic variant rs3834129 (*CASP8*) was genotyped using PCR followed by separation of the obtained fragments in a 7 % polyacrylamide gel. The remaining polymorphic markers were analyzed with allele-specific PCR, followed by analysis of the fragments on a 2 % agarose gel. The selection of primers for the PCR was carried out using the National Center for Biotechnology Information (NCBI) data-

**Table 1.** The polymorphic variants included in the study and their location

| Polymorphic marker | Chromosomal localization (GRCh38.p14) | Gene localization | Gene product |
|---|---|---|---|
| rs1205<br>g.159712443C>T,<br>2042C>T | 1:159712443 | *CRP*<br>3′ untranslated region | C-reactive protein |
| rs3834129<br>g.201232809_201232814del,<br>-654(6)I/D | 2:201232809-201232814 | *CASP8*<br>Promotor | Caspase 8 |
| rs3732378<br>g.39265671G>A,<br>c.935C>T, T280M | 3:39265671 | *CX3CR1*<br>2 exon | Chemokine, CX3C motif, receptor 1 fractalkine receptor |
| rs1800795<br>g.22727026C>G,<br>-174C>G | 7:22727026 | *IL6*<br>Promotor | Interleukin 6 |
| rs662<br>g.95308134T>C,<br>c.575A>G, Q192R | 7:95308134 | *PON1*<br>6 exon | Paraoxonase 1 |
| rs7493<br>g.95405463G>C,<br>c.932C>G, S311C | 7:95405463 | *PON2*<br>9 exon | Paraoxonase 2 |
| rs10098474<br>g.10054107C>T,<br>-410C>T | 8:10054107 | *MSRA*<br>Promotor | Methionine sulfoxide reductase |
| rs1001179<br>g.34438684C>T,<br>-262C>T | 11:34438684 | *CAT*<br>Promotor | Catalase |
| rs1695<br>g.67585218A>G<br>c.313A>G I105V | 11:67585218 | *GSTP1*<br>5 exon | Glutathione S-transferase PI |
| rs1042522<br>g.7676154G>C,<br>c.98C>G, P72R | 17:7676154 | *TP53*<br>4 exon | Tumor protein p53 |
| rs1024611<br>g.34252769A>G,<br>-2518A>G | 17:34252769 | *CCL2*<br>5′-end | Chemokine CCL2 |
| rs12454712<br>g.63178651T>C | 18:63178651 | *BCL2*<br>2 intron | BCL2-related protein |

base (https://www.ncbi.nlm.nih.gov/snp/) and an online tool (https://www.ncbi.nlm.nih.gov/tools/primer-blast/). The primer sequences and expected fragment sizes are presented in Table 2.

Statistical analysis of single polymorphic variants was carried out using the tools of the R programming language and the SNPassoc package. The Fisher exact test was used to analyze the deviation of the obtained genotype frequencies from the Hardy–Weinberg equilibrium. When searching for associations with the disease, a logistic regression analysis was used, taking into account five possible inheritance models (codominant, dominant, recessive, overdominant and additive). The best model was chosen according to the Akaike information criterion. The polymorphic marker was considered to be associated with the trait at $P < 0.05$.

The search for combinations of alleles of genotypes associated with the disease was carried out using the Monte Carlo method and Markov chains using the APSampler software (Favorov et al., 2005). The selection criteria for the identified combinations were $P < 0.05$ after the Benjamin–Hochberg (FDR) procedure and OR < 0.3 (OR is odds ratio) for protective markers or OR > 3 for high-risk markers.

To identify possible nonlinear interaction (synergy) between two elements of the found combinations, the SF (Synergy Factor) indicator was calculated (Cortina-Borja al., 2009). The synergy factor was considered significant if at $P < 0.05$ and the value of 95 % CI for SF did not cross 1. When constructing predictive models (using SPSS v. 22), the method of multifactorial logistic regression with step-by-step inclusion of variables was used, as which the studied polymorphic variants were selected taking into account the selected optimal model and paired combinations with a significant SF indicator.

## Results

Table 3 shows the results of analyzing the genotype frequency distribution of the studied polymorphic variants. In the control group, all the obtained genotype frequency distributions of the

**Table 2.** Primer sequences and sizes of amplified fragments

| Marker | Primers | Alleles (fragment size, bp) |
|---|---|---|
| rs1205**CRP* | F 5′-aga aaa cag ctt gga ctc act ca-3′<br>R 5′-tga gag gac gtg aac ctg gg-3′<br>C 5′-cca gtt tgg ctt ctg tcc tga c-3′<br>T 5′-ttg cca cat gga gag aga cta-3′ | BK* (235)<br>*T (195)<br>*C (85) |
| rs3834129**CASP8* | F 5′-ggt ccc cgc tgt taa cat ttt gat-3′<br>R 5′-ccg agg aag gca ctg aga cg-3′ | *D (126)<br>*I (132) |
| rs3732378**CX3CR1* | F 5′-tgc tgc tca gaa cac ttc ca-3′<br>R 5′-cct tct ggt ggt cat cgt gt-3′<br>A 5′-caa caa tgg cta aat gca atc a-3′<br>G 5′-ccc tca gtg tga ctg aga c-3′ | BK (323)<br>A (163)<br>G (201) |
| rs1800795**IL6* | F 5′-ctt cgt gca tga ctt cag ctt-3′<br>R 5′-gag act cat ggg aaa atc cca ca-3′<br>C 5′-ccc cta gtt gtg tct tgt c-3′<br>G 5′-aat gtg acg tcc ttt agt atc-3′ | BK (279)<br>*G (179)<br>*C (139) |
| rs662**PON1* | F 5′-cta gca cga agg ctc cat cc-3′<br>R 5′-cca cta cat ttc aga gag ttc aca-3′<br>G 5′-ccc aaa tac atc tcc cag cat c-3′<br>A 5′-tat ttt ctt gac ccc tac tta ca-3′ | BK (351)<br>G (222)<br>A (173) |
| rs7493**PON2* | F 5′-cat gtc ccc tta atc agt gtg-3′<br>R 5′-tga gca gct tcc cat cat aca-3′<br>C 5′-tag tca ctg tag gct tct gag-3′<br>G 5′-ccg cat cca gaa cat tca atg-3′ | BK (224)<br>C (152)<br>G (113) |
| rs10098474**MSRA* | F 5′-cct tgct ccc gta ttt tgg c-3′<br>R 5′-cct gtc gta cga agt acg tg-3′<br>C 5′-gtc ctc ttc tat ctt act gag c-3′<br>T 5′-cga ctt cgc agt tta gca gta-3′ | BK (337)<br>T (243)<br>C (136) |
| rs1001179**CAT* | F 5′-ata gct atg gag cgc aag gc-3′<br>R 5′-ggc ctg aag acc gga gat ac-3′<br>C 5′-gcc ctg ggt tcg gct atc-3′<br>T 5′-gcc ctg ggt tcg gct att-3′ | BK (236)<br>C,T (117)** |
| rs1695* *GSTP1* | F 5′-tct cat cct tcc acg cac at-3′<br>R 5′-caa gcc acc tgag ggg taa g-3′<br>A 5′-gtt ggt gta gat gag gga gat-3′<br>G 5′-gac ctc cgc tgc aaa tac g-3′ | BK (333)<br>G (132)<br>A (240) |
| rs1042522**TP53* | F 5′-tca ccc atc tac agt ccc cct-3′<br>R 5′-ata cgg cca ggc att gaa gt-3′<br>C 5′-cca gag gct gct ccc gc-3′<br>G 5′-tgg tgc agg ggc ctc cc-3′ | BK (345)<br>G (149)<br>C (229) |
| rs1024611**CCL2* | F 5′-cgg gcc cag tat ctg gaa tg-3′<br>R 5′-ctg gaa agt gac ttg gcc ttt g-3′<br>G 5′-gaa agt ctt ctg gaa agt gac-3′<br>A 5′-agt ggg agg cag aca gct a-3′ | BK (273)<br>G (201)<br>A (111) |
| rs12454712**BCL2* | F 5′-ctt cct ggt ttc ttt gcc agg-3′<br>R 5′-atc act cct caa agg cgc ag-3′<br>T 5′-gcc cca gac tca ctt gcgt-3′<br>C 5′-ggt gtt gca aca tcc atc acg-3′ | BK (306)<br>T (200)<br>C (145) |

*IC – Internal control.
** – First, testing was carried out for the presence of a rare allele *T, then, with a positive test, the sample was tested for the presence of the allele *C.

studied loci correspond to the Hardy–Weinberg equilibrium distribution.

The analysis of polymorphic loci associations with MI revealed statistically significant results for polymorphic variants of the *CX3CR1* (rs3834129), *MSRA* (rs10098474), *CCL2* (rs1024611), and *BCL2* (rs12454712) genes. Notably, after introducing the Benjamini–Hochberg correction (multiple comparisons), only the results for the *CCL2* and *CX3CR1* genes remained significant.

The APSampler software, which employs the dynamic Monte Carlo method, analyzed possible combinations of the studied polygenic variants associated with a high risk of MI. Combinations with higher OR and P values than those of the components of these combinations were identified (Table 4). In this case, the combinations include not only the *CX3CR1*, *MSRA*, *CCL2*, and *BCL2* gene variants obtained during the analysis of individual loci but also the *CRP*, *CASP8*, *PON2*, *CAT*, *IL6*, *GSTP1*, and *TP53* gene variants.

**Table 3**. Distribution of genotype frequencies according to the studied polymorphic variants and the results of the analysis of associations with myocardial infarction

| Gene marker | Genotype | Control | MI | *P** | The results of the logistic analysis | | |
|---|---|---|---|---|---|---|---|
| | | *n* (%) | *n* (%) | | Model** | *P* | OR 95 % CI |
| *CRP* rs1205 | C/C | 98 (32.67) | 81 (28.93) | 0.906 | Recessive – C/C+C/T vs T/T | 0.099 | 1.4 0.94–2.08 |
| | C/T | 146 (48.67) | 131 (46.79) | | | | |
| | T/T | 56 (18.67) | 68 (24.29) | | | | |
| *CASP8* rs3834129 | I/I | 124 (39.87) | 102 (36.43) | 0.47 | Dominant – I/I vs I/D+D/D | 0.39 | 1.16 0.83–1.61 |
| | I/D | 140 (45.02) | 132 (47.14) | | | | |
| | D/D | 47 (15.11) | 46 (16.43) | | | | |
| *CX3CR1* rs3732378 | G/G | 215 (70.49) | 156 (55.71) | 0.099 | Overdominant – G/G+A/A vs A/G | 0.0002 | 1.9 1.35–2.67 |
| | A/G | 77 (25.25) | 112 (40) | | | | |
| | A/A | 13 (4.26) | 12 (4.29) | | | | |
| *IL6* rs1800795 | C/C | 38 (12.54) | 26 (9.29) | 0.326 | Overdominant – G/G+C/C vs C/G | 0.111 | 1.3 0.94–1.81 |
| | C/G | 151 (49.83) | 158 (56.43) | | | | |
| | G/G | 114 (37.62) | 96 (34.29) | | | | |
| *PON1* rs662 | T/T | 143 (46.58) | 139 (50.18) | 0.3 | Log-additive – allele *C (0, 1, 2) | 0.259 | 0.87 0.68–1.11 |
| | T/C | 127 (41.37) | 112 (40.43) | | | | |
| | C/C | 37 (12.05) | 26 (9.39) | | | | |
| *PON2* rs7493 | G/G | 136 (44.88) | 118 (44.36) | 0.52 | Recessive – G/G+G/C vs C/C | 0.326 | 1.27 0.79–2.06 |
| | G/C | 130 (42.9) | 108 (40.6) | | | | |
| | C/C | 37 (12.21) | 40 (15.04) | | | | |
| *MSRA* rs10098474 | T/T | 152 (50) | 106 (39.85) | 0.494 | Dominant – T/T vs T/C+C/C | 0.015 | 1.51 1.08–2.11 |
| | T/C | 122 (40.13) | 133 (50) | | | | |
| | C/C | 30 (9.87) | 27 (10.15) | | | | |
| *CAT* rs1001179 | C/C | 207 (70.65) | 179 (66.54) | 0.138 | Log-additive – allele *T (0, 1, 2) | 0.277 | 1.18 0.88–1.57 |
| | C/T | 74 (25.26) | 76 (28.25) | | | | |
| | T/T | 12 (4.1) | 14 (5.2) | | | | |
| *GSTP1* rs1695 | A/A | 142 (48.46) | 129 (47.78) | 0.343 | Recessive– A/A+A/G vs G/G | 0.209 | 1.38 0.84–2.27 |
| | A/G | 119 (40.61) | 102 (37.78) | | | | |
| | G/G | 32 (10.92) | 39 (14.44) | | | | |
| *TP53* rs1042522 | C/C | 155 (50) | 134 (53.17) | 0.228 | Log-additive – allele*G (0, 1, 2) | 0.405 | 0.9 0.7–1.16 |
| | C/G | 122 (39.35) | 95 (37.7) | | | | |
| | G/G | 33 (10.65) | 23 (9.13) | | | | |
| *CCL2* rs1024611 | A/A | 169 (54.34) | 131 (46.79) | 0.552 | Recessive – A/A+A/G vs G/G | 0.0007 | 2.63 1.47–4.72 |
| | A/G | 124 (39.87) | 110 (39.29) | | | | |
| | G/G | 18 (5.79) | 39 (13.93) | | | | |
| *BCL2* rs12454712 | T/T | 119 (38.26) | 86 (30.71) | 0.72 | Log-additive – allele *C (0, 1, 2) | 0.005 | 1.38 1.1–1.73 |
| | T/C | 144 (46.3) | 125 (44.64) | | | | |
| | C/C | 48 (15.43) | 69 (24.64) | | | | |

* The exact test for compliance with the Hardy-Weinberg equilibrium for the control group.
** The model was selected based on the results of the Akaike information criterion.

To determine possible non-linear interactions in the identified combinations, the SF factor between all possible pairs of loci included in the obtained combinations was calculated. As a result of the analysis, five statistically significant pairs were obtained: *CAT**T+*MSRA**C (SF = 2.57, 95 % $CI_{SF}$ 1.23–5.4, Z = 2.50, *P* = 0.01), *CAT**C/T+*CX3CR1**A (SF = 2.45, 95 % $CI_{SF}$ 1.08–5.56, Z = 2.15, *P* = 0.03), *CX3CR1**A/G+*CASP8**I/I (SF = 4.71,95 % $CI_{SF}$ 2.22–10.01, Z = 4.03, $P = 5.6 \times 10^{-5}$), *CRP**T+*IL6**C/G (SF = 2.42, 95 % $CI_{SF}$ 1.19–4.94, Z = 2.44, *P* = 0.015), *MSRA**C+*CRP**C (SF = 2.56, 95 % $CI_{SF}$ 1.12–5.86, Z = 2.22, *P* = 0.027). Thus, the results suggest that a synergistic effect is observed for these pairs.

Next, to construct a prognostic model for MI, a multifactorial logistic regression analysis was performed with a step-by-step inclusion of the most significant predictors (individual

**Table 4.** Combinations of alleles/genotypes of the studied polymorphic variants associated with the risk of myocardial infarction

| Combination | Control, % | MI, % | *P* | $P_{FDR}$ | OR | 95 % $CI_{OR}$ |
|---|---|---|---|---|---|---|
| *CAT**C/T+ *MSRA**C+*CRP**C+*CX3CR1**A | 0.34 | 7.17 | $5.91\times10^{-6}$ | 0.0117 | 22.48 | 2.99–169.1 |
| *CX3CR1**A/G+*CASP8**I/I | 6.89 | 18.93 | $8.70\times10^{-6}$ | 0.0142 | 3.16 | 1.85–5.39 |
| *GSTP1**G+*MSRA**C+*CRP**T+*CASP8**I+*IL6**C/G | 3.42 | 13.64 | $8.88\times10^{-6}$ | 0.0141 | 4.45 | 2.16–9.17 |
| *PON2**G+*CAT**T+*MSRA**C+*CRP**C+*CX3CR1**A | 0.69 | 7.92 | $9.47\times10^{-6}$ | 0.0146 | 12.39 | 2.88–53.39 |
| *CAT**T+*MSRA**T/C+*CASP8**D+*BCL2**C | 1.71 | 10.19 | $1.10\times10^{-5}$ | 0.0143 | 6.5 | 2.47–17.17 |
| *MSRA**C+*CX3CR1**A+*TP53**C/C+*IL6**C | 2.00 | 10.71 | $1.29\times10^{-5}$ | 0.0132 | 5.88 | 2.39–14.48 |
| *CX3CR1**A+*CASP8**I/I+*TP53**C+*IL6**C | 0.33 | 13.10 | $1.48\times10^{-5}$ | 0.0135 | 4.4 | 2.12–9.12 |
| *PON2**G/G+*MSRA**C+*CX3CR1**A+*IL6**C | 1.68 | 9.77 | $1.76\times10^{-5}$ | 0.014 | 6.32 | 2.39–16.72 |
| *PON2**G+*CCL2**G/G+*CASP8**D+*IL6**C | 1.00 | 7.89 | $3.27\times10^{-5}$ | 0.016 | 8.46 | 2.49–28.69 |
| *CRP**T+*CCL2**G/G+*TP53**C | 1.34 | 8.33 | $6.73\times10^{-5}$ | 0.0143 | 6.7 | 2.27–19.8 |
| *CAT**T+*GSTP1**A/A+*MSRA**C+*CX3CR1**A | 0.34 | 5.70 | $1\times10^{-4}$ | 0.0145 | 17.6 | 2.31–134.2 |

polymorphic variants, as well as significant paired combinations identified during the SF analysis). Table 5 presents the list of predictors included in the final model. Thus, a model for calculating the genetic risk of MI was obtained, which, according to the ROC analysis, has a fairly high prognostic efficiency (AUC = 0.71, $P_{AUC} = 1.7 \times 10^{-16}$, see the Figure).

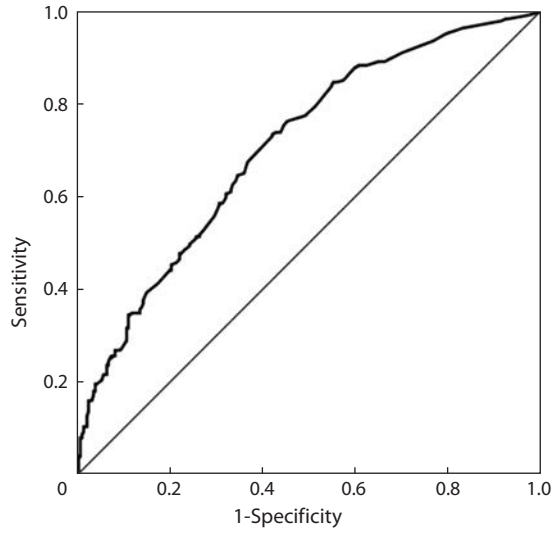
## Discussion

The study aimed to identify informationally significant risk predictors of MI. The results indicate the involvement of genes encoding proteins involved in the inflammatory response, antioxidant protection, and apoptosis in forming a predisposition to MI, which is consistent with modern concepts of CHD etiopathogenesis. Indeed, the consequence of oxidative stress is lipid peroxidation and protein oxidation, which are factors of vascular endothelial damage, leading to the activation of inflammatory processes or apoptosis. Moreover, at the early stages of atherogenesis, apoptosis can be considered a protective factor, while at later stages, it is a factor in atherosclerotic plaque destabilization and activates thrombus formation, which is the direct cause of MI.

Fractalkine (CX3CL1) via its receptor (CX3CR1) triggers chemotaxis and monocyte adhesion in the area of atherosclerotic damage at early stages of atherogenesis (Schulz et al., 2007). Also, this chemokine has an anti-apoptotic effect on smooth muscle cells and monocytes and promotes the proliferation and migration of smooth muscle cells, which contributes to the formation and growth of atherosclerotic plaque (Liu et al., 2010). rs3732378 in the *CX3CR1* gene determines the replacement of the amino acid threonine with methionine. D.H. McDermott (McDermott et al., 2003) showed that the receptor with 280M (allele *A) binds to fractalkine less effectively, i. e. allele *A was considered as a protective factor in relation to atherosclerosis. At the same time, if at early stages of atherogenesis, the decreased activity of the *CX3CL1-CX3CR1* system inhibits the disease development, at later stages, the same effect can lead to apoptosis of monocytes and foam cells, disease progression, and thrombosis (Landsman et al., 2009; Van Vré et al., 2012).

**Table 5.** Coefficients of the logistic regression equation for the multifactorial model for calculating the genetic risk of myocardial infarction

| Predictor | B | *P* | OR | 95 % $CI_{OR}$ |
|---|---|---|---|---|
| *CCL2**G/G | 1.06 | 0.0023 | 2.89 | 1.46–5.71 |
| *BCL2**C (x0. 1. 2) | 0.29 | 0.0300 | 1.33 | 1.03–1.73 |
| *MSRA**C+*CRP**C | 0.71 | 0.0009 | 2.03 | 1.34–3.08 |
| *CRP**T/T | 0.71 | 0.0063 | 2.04 | 1.22–3.39 |
| *CASP8**D | 0.69 | 0.0028 | 1.99 | 1.27–3.12 |
| *CX3CR1**G/A+*CASP8**I/I | 1.25 | 0.0003 | 3.50 | 1.76–6.95 |
| *CRP**T+*IL6**G/C | 0.48 | 0.0172 | 1.61 | 1.09–2.38 |
| *CAT**C/T+*CX3CR1**A | 0.72 | 0.0388 | 2.05 | 1.04–4.03 |
| *CAT**T+*MSRA**C | 0.76 | 0.0020 | 2.15 | 1.32–3.49 |
| Constant | –1.87 | $9\times10^{-11}$ | 0.15 | |




ROC analysis of the effectiveness of a model based on genetic markers of individual risk of myocardial infarction.

*CASP8* belongs to cysteine proteases and triggers a cascade of reactions with the final result of cell apoptosis (Ho, Hawkins, 2005). T. Sun et al. (2007) showed that the deletion of 6 nucleotide pairs in the promoter region of the *CASP8* gene (rs3834129) disrupted the binding site for the stimulating protein (sp1) and reduced the transcriptional activity of the gene. In the same work, *in vivo* experiments demonstrated that the 6N deletion variant was associated with lower apoptotic reactivity of T-lymphocytes when stimulated by cancer cells. Based on this, the identified *CX3CR1**G/A+*CASP8**I/I variant may be associated with increased apoptotic activity and destabilization of the atherosclerotic plaque. At the same time, for carriers of the rs3834129*D allele, a decrease in apoptotic activity at earlier stages of atherogenesis may contribute to disease progression, which is confirmed by the study of this polymorphic variant in a sample of the Russian ethnic group from Novosibirsk, where an association of the *D/D genotype with progressive atherosclerosis was demonstrated (Maksimov et al., 2022).

Catalase belongs to the group of antioxidant enzymes, catalyzes decomposition of hydrogen peroxide formed in biological oxidation into water and molecular oxygen, and protects cells from damage by free radical oxidation products. H. Yang et al. (2004) demonstrated that mice with ApoE-/- and increased expression of catalase had a slowdown in atherosclerosis development.

Information on the association of the polymorphic variant rs1001179**CAT* with enzyme activity is contradictory. Thus, for Americans of European descent, a direct correlation was shown between catalase activity and the allele *C, and the differences in the level of catalase activity varied significantly depending on the level of fruit and vegetable consumption (Ahn et al., 2006); in Italians with chronic lymphocytic leukemia, carriers of the allele *T were characterized by a lower level of methylation and a higher level of the *CAT* gene expression (Galasso et al., 2022); the work on population samples from Russians and Buryats revealed that carriers of the *T/T genotype had lower concentrations of diene conjugates than carriers of the allele *C, which allows assuming greater catalase activity for individuals with the *T/T genotype (Ershova et al., 2016).

Thus, the *CAT**C/T+*CX3CR1**A combination identified in the present study may be associated with higher catalase activity, one of the effects of which is proliferation inhibition and induction of apoptosis of vascular smooth muscle cells (Brown et al., 1999) by catalase and a decrease in the inhibitory effect of smooth muscle cells on apoptosis by fractalkine and its receptor.

Methionine sulfoxide reductase A (MSRA) catalyzes the reduction of methionine sulfoxide to the parent methionine. It is believed that a decrease in MSRA activity decreases cellular resistance to oxidative stress. Y. Xu et al. (2020) demonstrated the ability of MSRA to restore the anti-atherogenic function of oxidized high-density lipoproteins. Previously, the authors of the present study found that the *T/T genotype of the rs10098474 polymorphic locus as part of the polymorphic loci combination of the *CAT* (rs1001179) and *GPX1* (rs1050450) genes was more common among individuals over 90 years of age (Erdman et al., 2021), which is consistent with our results on the negative contribution of the allele *C to forming a hereditary predisposition to MI.

According to a number of studies, the *T/T genotype of the rs1205 (*CRP*) polymorphism is associated with lower plasma CRP levels in Europeans (Kolz et al., 2008), Americans of European descent (Lange et al., 2006), and residents of eastern Mexico (Reynoso-Vilalpando et al., 2021). CRP has pronounced pro-inflammatory effects: according to (Pasceri et al. 2000), it stimulates expression of chemokine intercellular adhesion molecules; H. Fujii et al. (2006) noted that CRP is able to increase the release of reactive oxygen species and induce apoptosis of progenitor endothelial cells, which contributes to endothelial dysfunction.

At the same time, anti-atherogenic properties are also noted – CRP binds modified low-density lipoproteins (Tabuchi et al., 2007); as a result, it can prevent the formation of foam cells and limit complement activation. CRP also inhibits the oxidation of low-density lipoproteins (Badimon et al., 2018). In the model the authors of the present study obtained, the *T/T genotype of the rs1205 polymorphic variant is an MI risk factor, which is consistent with the data on the protective properties of CRP. The authors of the present study also found an unfavorable synergistic interaction in the combination of *MSRA**C+*CRP**C. Probably, the allele *C rs10098474 of the *MSRA* gene is associated with decreased enzyme activity and, as a consequence, decreased cellular resistance to oxidative stress, while *CRP* is known to be able to increase reactive oxygen species release, which can enhance the negative impact of this combination.

The pro-atherogenic role of the chemokine CCL2 (MCP1, a key factor providing chemotaxis of immune competent cells to the site of damage) was demonstrated in the works (Aiello et al., 1999; Öhman et al., 2010). The genotype *G/G rs1024611**CCL2* according to the data (McDermott et al., 2005) is associated with increased CCL2 content in plasma and with MI. The association of the *CCL2**G/G genotype with an increased risk of CHD was confirmed by a meta-analysis of European populations, while statistically significant results were not obtained for Asian populations (Bai et al., 2015).

BCL2 is an inhibitor of apoptosis, an intracellular protein, the main representative of the BCL2 family. The C allele of the polymorphic variant rs12454712 is able to bind to the transcription factor ZNF329, which increases the expression of the *BCL2* gene (Dong et al., 2021). As already noted, in the later stages of atherosclerosis, activation of apoptosis plays a negative role. At the same time, apoptosis can be a significant factor in limiting intimal hyperplasia in atherosclerosis; in addition, macrophage apoptosis can be a factor in excessive limitation of the inflammatory response (Vladimirskaya et al., 2015).

## Conclusion

Notably, the obtained results can be considered intermediate, since the resulting model has limited predictive ability (which can probably be compensated for by introducing additional predictors). In addition, this model needs to be confirmed on alternative samples. Nevertheless, the results give reason to assume that polymorphic variants rs1205**CRP*, rs3732378**CX3CR1*, rs1800795**IL6*, rs1024611**CCL2*,

rs3834129**CASP8*, rs1042522**TP53*, rs12454712**BCL2*, rs1001179**CAT*, and rs10098474**MSRA* significantly contribute to the formation of hereditary predisposition to MI. In addition, it was demonstrated that the identified synergistic interactions between genotypes/alleles in combinations of *CX3CR1**A/G and *CASP8**I/I, *MSRA**C and *CRP**C, *CAT**C/T and *MSRA**C, and *CAT**C/T and *CX3CR1**A can significantly affect the resulting predictive model. The nature of these interactions is subject for further analysis.

## References

Ahn J., Nowell S., McCann S.E., Yu J., Carter L., Lang N.P., Kadlubar F.F., Ratnasinghe L.D., Ambrosone C.B. Associations between catalase phenotype and genotype: modification by epidemiologic factors. *Cancer Epidemiol. Biomarkers Prev*. 2006;15(6):1217-1222. DOI 10.1158/1055-9965.EPI-06-0104

Aiello R.J., Bourassa P.A.K., Lindsey S., Weng W., Natoli E., Rollins B.J., Milos P.M. Monocyte chemoattractant protein-1 accelerates atherosclerosis in apolipoprotein E-deficient mice. *Arterioscler. Thromb. Vasc. Biol*. 1999;19(6):1518-1525. DOI 10.1161/01.ATV.19.6.1518

Assmann G., Cullen P., Schulte H. Simple scoring scheme for calculating the risk of acute coronary events based on the 10-year follow-up of the prospective cardiovascular Munster (PROCAM) study. *Circulation*. 2002;105(3):310-315. DOI 10.1161/hc0302.102575

Badimon L., Peña E., Arderiu G., Padró T., Slevin M., Vilahur G., Chiva-Blanch G. C-reactive protein in atherothrombosis and angiogenesis. *Front. Immunol*. 2018;9:430. DOI 10.3389/fimmu.2018.00430

Bai X.Y., Li S., Wang M., Qu X., Hu G., Xu Z., Chen M., He G.-W., Wu H. Association of monocyte chemoattractant protein-1 (*MCP-1*)-2518A>G polymorphism with susceptibility to coronary artery disease: A meta-analysis. *Ann. Hum. Genet*. 2015;79(3):173-187. DOI 10.1111/ahg.12105

Batty M., Bennett M.R., Yu E. The role of oxidative stress in atherosclerosis. *Cells*. 2022;11(23):3843. DOI 10.3390/cells11233843

Brown M.R., Miller F.J. Jr., Li W.G., Ellingson A.N., Mozena J.D., Chatterjee P., Engelhardt J.F., Zwacka R.M., Oberly L.W., Fang X., Spector A.A., Weintraub N.L. Overexpression of human catalase inhibits proliferation and promotes apoptosis in vascular smooth muscle cells. *Circ. Res*. 1999;85(6):524-533. DOI 10.1161/01.RES.85.6.524

Colditz G.A., Rimm E.B., Giovannucci E., Stampfer M.J., Rosner B., Willett W.C. A prospective study of parental history off myocardial infarction and coronary artery disease in men. *Am. J. Cardiol*. 1991;67(11):933-938. DOI 10.1016/0002-9149(91)90163-F

Cortina-Borja M., Smith A.D., Combarros O., Lehmann D.J. The synergy factor: a statistic to measure interactions in complex diseases. *BMC Res. Notes*. 2009;2(1):105. DOI 10.1186/1756-0500-2-105

Domingo J., Baeza-Centurion P., Lehner B. The causes and consequences of genetic interactions (epistasis). *Annu. Rev. Genomics Hum. Genet*. 2019;20:433-460. DOI 10.1146/annurev-genom-083118-014857

Dong S.S., Zhu D.L., Zhou X.R., Rong Y., Zeng M., Chen J.B., Jiang F., Tuo X.M., Feng Z., Yang T.L., Guo Y. An intronic risk SNP rs12454712 for central obesity acts as an allele-specific enhancer to regulate *BCL2* expression. *Diabetes*. 2021;70(8):1679-1688. DOI 10.2337/db20-1151

Erdman V.V., Tuktarova I.A., Danilko K.V., Timasheva Ya.R., Viktorova T.V., Nasibullin T.R. Polygenic analysis of genetic markers of longevity using the APSampler program. Book of papers of Russian conf. "Modern Problems of Biochemistry, Genetics and Biotechnology", 21-23 Sep. 2021. Ufa, Russia, 2021;245-250. DOI 10.33184/spbgb-2021-09-21.50 (in Russian)

Ershova O.A., Bairova T.A., Kolesnikov S.I., Kalyuzhnaya O.V., Darenskaya M.A., Kolesnikova L.I. Oxidative stress and catalase gene. *Bull. Exp. Biol. Med*. 2016;161(3):400-403. DOI 10.1007/s10517-016-3424-0

Favorov A.V., Andreewski T.V., Sudomoina M.A., Favorova O.O., Parmigiani G., Ochs M.F. A Markov chain Monte Carlo technique for identification of combinations of allelic variants underlying complex diseases in humans. *Genetics*. 2005;171(4):2113-2121. DOI 10.1534/genetics.105.048090

Fujii H., Li S.H., Szmitko P.E., Fedak P.W., Verma S. C-reactive protein alters antioxidant defenses and promotes apoptosis in endothelial progenitor cells. *Arterioscler. Thromb. Vasc. Biol*. 2006;26(11):2476-2482. DOI 10.1161/01.ATV.0000242794.65541.02

Galasso M., Dalla Pozza E., Chignola R., Gambino S., Cavallini C., Quaglia F.M., Lovato O., Dando I., Malpeli G., Krampera M., Donadelli M., Romanelli M.G., Scupoli M.T. The rs1001179 SNP and CpG methylation regulate catalase expression in chronic lymphocytic leukemia. *Cell. Mol. Life Sci*. 2022;79(10):521. DOI 10.1007/s00018-022-04540-7

Ho P.K., Hawkins C.J. Mammalian initiator apoptotic caspases. *FEBS J*. 2005;272(21):5436-5453. DOI 10.1111/j.1742-4658.2005.04966.x

Kolz M., Koenig W., Müller M., Andreani M., Greven S., Illig T., Khuseyinova N., Panagiotakos D., Pershagen G., Salomaa V., Sunyer J., Peters A. DNA variants, plasma levels and variability of C-reactive protein in myocardial infarction survivors: results from the AIRGENE study. *Eur. Heart J*. 2008;29(10):1250-1258. DOI 10.1093/eurheartj/ehm442

Landsman L., Bar-On L., Zernecke A., Kim K.-W., Krauthgamer R., Shagdarsuren E., Lira S.A., Weissman I.L., Weber C., Jung S. $CX_3CR1$ is required for monocyte homeostasis and atherogenesis by promoting cell survival. *Blood*. 2009;113(4):963-972. DOI 10.1182/blood-2008-07-170787

Lange L.A., Carlson C.S., Hindorff L.A., Lange E.M., Walston J., Durda J.P., Cushman M., Bis J.C., Zeng D., Lin D., Kuller L.H., Nickerson D.A., Psaty B.M., Tracy R.P., Reiner A.P. Association of polymorphisms in the CRP gene with circulating C-reactive protein levels and cardiovascular events. *JAMA*. 2006;296(22):2703-2711. DOI 10.1001/jama.296.22.2703

Liu H., Jiang D., Zhang S., Ou B. Aspirin inhibits fractalkine expression in atherosclerotic plaques and reduces atherosclerosis in ApoE gene knockout mice. *Cardiovasc. Drugs Ther*. 2010;24(1):17-24. DOI 10.1007/s10557-009-6210-7

Maksimov V.N., Parkhomenko O.M., Lozhkina N.G., Gurazheva A.A., Maksimova S.V., Ivanova A.A. Some molecular genetic markers of progressive atherosclerosis in patients with coronary heart disease. *Ateroscleroz*. 2022;18(1):6-13. DOI 10.52727/2078-256X-2022-18-1-6-13 (in Russian)

McDermott D.H., Fong A.M., Yang Q., Sechler J.M., Cupples L.A., Merrel M.N., Wilson P.W.F., D'Agostino R.B., O'Donnell Ch.J., Patel D.D., Murphy P.M. Chemokine receptor mutant CX3CR1-M280 has impaired adhesive function and correlates with protection from cardiovascular disease in humans. *J. Clin. Invest*. 2003;111(8):1241-1250. DOI 10.1172/JCI16790

McDermott D.H., Yang Q., Kathiresan S., Cupples L.A., Massaro J.M., Keaney J.F., Larson M.G., Vasan R.S., Hirschhorn J.N., O'Donnell C.J., Murphy Ph.M., Benjamin E.J. CCL2 polymorphisms are associated with serum monocyte chemoattractant protein-1 levels and myocardial infarction in the Framingham Heart Study. *Circulation*. 2005;112(8):1113-1120. DOI 10.1161/CIRCULATIONAHA.105.543579

Öhman M.K., Wright A.P., Wickenheiser K.J., Luo W., Russo H.M., Eitzman D.T. Monocyte chemoattractant protein-1 deficiency protects against visceral fat-induced atherosclerosis. *Arterioscler. Thromb. Vasc. Biol*. 2010;30(6):1151-1158. DOI 10.1161/ATVBAHA.110.205914

Pasceri V., Willerson J.T., Yeh E.T. Direct proinflammatory effect of C-reactive protein on human endothelial cells. *Circulation*. 2000;102:2165-2168. DOI 10.1161/01.CIR.102.18.2165

Reynoso-Villalpando G.L., Casillas-Munoz F.A., Padilla-Gutiérrez J.R., Sevillano-Collantes C., Moreno-Ruiz I., Del Canizo-Gomez F.J., Valdez-Haro A., Martinez-Fernandez E., Valle Y. The genetic variants − 717T>C (rs2794521), 1444G>A (rs1130864), and 1846C>T (rs1205) of *CRP* gene, their haplotypes, and their association with serum CRP levels, acute coronary syndrome, and diabetes in patients from Western Mexico. *Metab. Syndr. Relat. Disord.* 2021; 19(3):127-136. DOI 10.1089/met.2020.0080

Sabgayda T.P., Zubko A.V., Semyonova V.G. Effectiveness of the Federal project "Fight against cardiovascular diseases" in the context of preventable causes of death in the Russian urban and rural settlements. *Voprosy Upravleniya = Management Issues*. 2023;(2):71-85. DOI 10.22394/2304-3369-2023-2-71-85 (in Russian)

Schulz C., Schafer A., Stolla M., Kerstan S., Lorenz M., von Brühl M.L., Schiemann M., Bauersachs J., Gloe T., Busch D.H., Gawaz M., Massberg S. Chemokine fractalkine mediates leukocyte recruitment to inflammatory endothelial cells in flowing whole blood: a critical role for P-selectin expressed on activated platelets. *Circulation*. 2007; 116(7):764-773. DOI 10.1161/CIRCULATIONAHA.107. 695189

Shalnova S.A., Drapkina O.M., Kutsenko V.A., Kapustina A.V., Muromtseva G.A., Yarovaya E.B., Balanova Yu.A., Evstifeeva S.E., Imaeva A.E., Shlyakhto E.V., Boytsov S.A., Astakhova Z.T., Barbarash O.L., Belova O.A., Grinshtein Yu.I., Efanov A.Yu., Kalachikova O.N., Kulakova N.V., Nedogoda S.V., Rotar O.P., Trubacheva I.A., Chernykh T.M. Myocardial infarction in the population of some Russian regions and its prognostic value. *Rossiyskiy Kardiologicheskiy Zhurnal = Russian Journal of Cardiology*. 2022; 27(6):9-19. DOI 10.15829/1560-4071-2022-4952 (in Russian)

Sun T., Gao Y., Tan W., Ma S., Shi Y., Yao J., Guo Y., Yang M., Zhang X., Zhang Q., Zeng Ch., Lin D. A six-nucleotide insertion-deletion polymorphism in the CASP8 promoter is associated with susceptibility to multiple cancers. *Nat. Genet*. 2007;39(5):605-613. DOI 10.1038/ng2030

Tabuchi M., Inoue K., Usui-Kataoka H., Kobayashi K., Teramoto M., Takasugi K., Shikata K., Yamamura M., Ando K., Nishida K., Kasahara J., Kume N., Lopez L.R., Mitsudo K., Nobuyoshi M., Yasuda T., Kita T., Makino H., Matsuura E. The association of C-reactive protein with an oxidative metabolite of LDL and its implication in atherosclerosis. *J. Lipid Res*. 2007;48(4):768-781. DOI 10.1194/jlr.M600414-JLR200

Van Vré E.A., Ait-Oufella H., Tedgui A., Mallat Z. Apoptotic cell death and efferocytosis in atherosclerosis. *Arterioscler. Thromb. Vasc. Biol*. 2012;32:887-893. DOI 10.1161/ATVBAHA.111.224873

Vladimirskaya T.E., Shved I.A., Demidchik Yu.E. importance of the apoptosis of smooth muscle cells and macrophages of the coronary artery wall in the pathogenesis of atherosclerosis. *Vescì Nacyânal'naj Akadèmìì Navuk Belarusì. Seryâ Medycynskìh Navuk = Proceedings of the National Academy of Sciences of Belarus. Medical series*. 2015;(3):69-74 (in Russian)

Xu Y., Li F., Zhao X., Tan C., Wang B., Chen Y., Cao J., Wu D., Yu H. Methionine sulfoxide reductase A attenuates atherosclerosis via repairing dysfunctional HDL in scavenger receptor class B type I deficient mice. *FASEB J*. 2020;34(3):3805-3819. DOI 10.1096/fj. 201902429R

Yang H., Roberts L.J., Shi M.J., Zhou L.C., Ballard B.R., Richardson A., Guo Z.M. Retardation of atherosclerosis by overexpression of catalase or both Cu/Zn-superoxide dismutase and catalase in mice lacking apolipoprotein E. *Circ. Res.* 2004;95:1075-1081. DOI 10.1161/01.RES.0000149564.49410

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

Received April 8, 2024. Revised August 5, 2024. Accepted August 11, 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/



✉ 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 публикации 18.11.2024. Формат 60 × 84 $^{1}/_{8}$. Уч.-изд. л. 17.7.