2025 • 29 • 4

Online ISSN 2500-3259

Сетевое издание

ВАВИЛОВСКИЙ ЖУРНАЛ ГЕНЕТИКИ И СЕЛЕКЦИИ

VAVILOV JOURNAL OF GENETICS AND BREEDING

Основан в 1997 г.
Периодичность 8 выпусков в год
doi 10.18699/vjgb-25-51

Учредители

Сибирское отделение Российской академии наук

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

Межрегиональная общественная организация Вавиловское общество генетиков и селекционеров

Главный редактор

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

Заместители главного редактора

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

Ответственный секретарь

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

Редакционная коллегия

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

2025 • 29 • 4

Online ISSN 2500-3259

Online edition

VAVILOVSKII ZHURNAL GENETIKI I SELEKTSII

VAVILOV JOURNAL OF GENETICS AND BREEDING

Founded in 1997
Publication frequency: 8 issues a year
doi 10.18699/vjgb-25-51

Founders

Siberian Branch of the Russian Academy of Sciences

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

The Vavilov Society of Geneticists and Breeders

Editor-in-Chief

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

Deputy Editor-in-Chief

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

Executive Secretary

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

Editorial board

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

ВАВИЛОВСКИЙ ЖУРНАЛ ГЕНЕТИКИ И СЕЛЕКЦИИ • VAVILOV JOURNAL OF GENETICS AND BREEDING

СОДЕРЖАНИЕ • 2025 • 29 • 4

Молекулярная и клеточная биология

479 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Концепция природной реконструкции генома. Часть 3. Анализ изменения количества теломерной ДНК в клетках колоний как нового амплифицированного признака, возникшего при обработке гемопоэтических стволовых клеток костного мозга. *В.С. Рузанова, С.Г. Ошихмина, Г.С. Риттер, Е.В. Долгова, С.С. Кирикович, Е.В. Левитес, Я.Р. Ефремов, Т.В. Карамышева, А.Г. Богомолов, М.И. Мещанинова, А.Л. Мамаев, О.С. Таранов, С.В. Сидоров, С.Д. Никонов, О.Ю. Леплина, А.А. Останин, Е.Р. Черных, Н.А. Колчанов, А.С. Проскурина, С.С. Богачев*

496 ОБЗОР
Неклассические модели животных для изучения роли теломер в процессах старения и долголетия. *Е.В. Симороз, Е. Василевская, Н.А. Аракелян, А.Д. Манахов, Е.И. Рогаев*

Генетика растений

508 ОБЗОР
Особенности генетического картирования локусов, влияющих на образование эмбриогенного каллуса и регенерацию растений *in vitro* у зерновых и бобовых культур. *Е.К. Потокина, А.С. Сущенко*

517 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Оптимизация этапов технологии получения DH-растений капусты белокочанной. *А.И. Минейкина, К.С. Стебницкая, М.Г. Фомичева, Л.Л. Бондарева, А.С. Домблидес, Е.А. Домблидес*

Иммунитет и продуктивность растений

530 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Устойчивость к засухе фотосинтетического аппарата линий пшеницы *Triticum aestivum* L. с интрогрессиями от *Aegilops tauschii* Coss. в хромосоме 2D. *С.В. Осипова, А.В. Пермяков, А.В. Рудиковский, Е.Г. Рудиковская, Т.А. Пшеничникова*

539 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Цитофизиологические проявления защитных реакций пшеницы от стеблевой ржавчины, индуцируемые биофунгицидом Новохизолем. *А.Б. Щербань, Л.Я. Плотникова, В.В. Кнауб, Е.С. Сколотнева, В.В. Фоменко*

549 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Рецепторподобные киназы с лейцин-богатыми повторами подсемейства III участвуют в распознавании *Pectobacterium* spp. растениями семейства Solanaceae. *Е.В. Шруб, А.В. Колубако, П.В. Вычик, О.А. Бадалян, Е.А. Николайчик*

559 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Влияние отобранного подвоя на параметры роста, накопление ИУК и витаминов в привоях *Cucumis sativus* L. и *Cucumis melo* L. *А.Ж. Шойбекова, С.К. Джантасов, А.С. Джантасова, А.Т. Саматов, Т.С. Сагиндыков, А.Н. Каримова, Г.А. Серикбаева, М.Р. Тойшиманов, Г.Т. Бари (на англ. языке)*

Популяционная генетика

568 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Данные митохондриальной ДНК позволяют выделить субпопуляции широкоареального вида журавлей красавки (*Anthropoides virgo*). *Е.А. Мудрик, Е.И. Ильяшенко, П.А. Казимиров, К.Д. Кондракова, Т.П. Арчимаева, Л.Д. Базаров, О.А. Горошко, Ц.З. Доржиев, А.Н. Куксин, К.А. Постельных, В.В. Шуркина, В.Ю. Ильяшенко, А.В. Шатохина, Д.В. Политов*

578 ОБЗОР
Генетическая изменчивость и филогеография сорок рода *Pica* Голарктики. *А.П. Крюков*

Генетика микроорганизмов

594 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Генетический потенциал к образованию биопленок клинических штаммов *Pseudomonas aeruginosa*. *У.М. Немченко, Н.Л. Белькова, Е.С. Клименко, Н.Е. Смурова, Р.Е. Зугеева, В.В. Синьков, Е.Д. Савилов*

600 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Env-псевдовирусы на основе генетического варианта, циркулирующего на территории Сибири. *Н.Б. Рудометова, А.А. Фандо, Д.Н. Щербаков, Б.Н. Зайцев, А.П. Рудометов, Л.И. Карпенко*

Биоинформатика и системная биология

608 ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ
Определение числа соприкасающихся зерен пшеницы на изображениях на основе эллиптической аппроксимации. *Д.Р. Авзалов, Е.Г. Комышев, Д.А. Афонников*

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

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

CONTENTS • 2025 • 29 • 4

Molecular and cell biology

479 ORIGINAL ARTICLE
Concept of natural genome reconstruction. Part 3. Analysis of changes in the amount of telomeric DNA in colony cells as a new amplified feature that arose during the processing of hematopoietic bone marrow stem cells. *V.S. Ruzanova, S.G. Oshikhmina, G.S. Ritter, E.V. Dolgova, S.S. Kirikovich, E.V. Levites, Y.R. Efremov, T.V. Karamysheva, A.G. Bogomolov, M.I. Meschaninova, A.L. Mamaev, O.S. Taranov, S.V. Sidorov, S.D. Nikonov, O.Y. Leplina, A.A. Ostanin, E.R. Chernykh, N.A. Kolchanov, A.S. Proskurina, S.S. Bogachev*

496 REVIEW
Unconventional animal models to study the role of telomeres in aging and longevity. *E.V. Simoroz, J. Vasilevska, N.A. Arakelyan, A.D. Manakhov, E.I. Rogaev*

Plant genetics

508 REVIEW
Genetic mapping of loci affecting embryogenic callus formation and *in vitro* regeneration in cereals and leguminous crops. *E.K. Potokina, A.S. Sushchenko*

517 ORIGINAL ARTICLE
Optimization of technology steps for obtaining white cabbage DH-plants. *A.I. Mineykina, K.S. Stebnitskaia, M.G. Fomicheva, L.L. Bondareva, A.S. Domblides, E.A. Domblides*

Plant immunity and performance

530 ORIGINAL ARTICLE
Drought tolerance of the photosynthetic apparatus of bread wheat (*Triticum aestivum* L.) lines with introgressions in chromosome 2D from *Aegilops tauschii* Coss. *S.V. Osipova, A.V. Permyakov, A.V. Rudikovskii, E.G. Rudikovskaya, T.A. Pshenichnikova*

539 ORIGINAL ARTICLE
Cytophysiological manifestations of wheat's defense reactions against stem rust induced by the biofungicide Novochizol. *A.B. Shcherban, L.Ya. Plotnikova, V.V. Knaub, E.S. Skolotneva, V.V. Fomenko*

549 ORIGINAL ARTICLE
Receptor-like leucine-rich repeat kinases of subfamily III are involved in the recognition of *Pectobacterium* spp. by Solanaceae plants. *E.V. Shrub, N.V. Kalubaka, P.V. Vychyk, O.A. Badalyan, Y.A. Nikolaichik*

559 ORIGINAL ARTICLE
Influence of selected rootstock on growth parameters, accumulation of IAA and vitamins in scions of *Cucumis sativus* L. and *Cucumis melo* L. *A.Zh. Shoibekova, S.K. Jantassov, A.S. Jantassova, A.T. Samatov, T.S. Sagindykov, A.N. Karimova, G.A. Serikbayeva, M.R. Toishimanov, G.T. Bari*

Population genetics

568 ORIGINAL ARTICLE
Mitochondrial DNA data allow distinguishing the subpopulations in the widespread Demoiselle crane (*Anthropoides virgo*). *E.A. Mudrik, E.I. Ilyashenko, P.A. Kazimirov, K.D. Kondrakova, T.P. Archimaeva, L.D. Bazarov, O.A. Goroshko, Ts.Z. Dorzhiev, A.N. Kuksin, K.A. Postelnykh, V.V. Shurkina, V.Yu. Ilyashenko, A.V. Shatokhina, D.V. Politov*

578 REVIEW
Genetic variation and phylogeography of the magpie's genus *Pica* in the Holarctic. *A.P. Kryukov*

Microbial genetics

594 ORIGINAL ARTICLE
Genetic potential for biofilm formation of clinical strains of *Pseudomonas aeruginosa*. *U.M. Nemchenko, N.L. Belkova, E.S. Klimenko, N.E. Smurova, R.E. Zugeeva, V.V. Sinkov, E.D. Savilov*

600 ORIGINAL ARTICLE
Env-pseudoviruses based on the HIV-1 genetic variant circulating in Siberia. *N.B. Rudometova, A.A. Fando, D.N. Shcherbakov, B.N. Zaitsev, A.P. Rudometov, L.I. Karpenko*

Bioinformatics and systems biology

608 ORIGINAL ARTICLE
Counting touching wheat grains in images based on elliptical approximation. *D.R. Avzalov, E.G. Komyshev, D.A. Afonnikov*

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

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2025;29(4):479-495

doi 10.18699/vjgb-25-52

Concept of natural genome reconstruction.

Part 3. Analysis of changes in the amount of telomeric DNA in colony cells as a new amplified feature that arose during the processing of hematopoietic bone marrow stem cells

V.S. Ruzanova [1#], S.G. Oshikhmina [1, 2#], G.S. Ritter [1], E.V. Dolgova [1], S.S. Kirikovich [1], E.V. Levites [1], Y.R. Efremov [1], T.V. Karamysheva[1], A.G. Bogomolov [1], M.I. Meschaninova [3], A.L. Mamaev[4], O.S. Taranov [5], S.V. Sidorov[2, 6], S.D. Nikonov[7], O.Y. Leplina [8], A.A. Ostanin [8], E.R. Chernykh [8], N.A. Kolchanov [1], A.S. Proskurina [1#], S.S. Bogachev [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
[3] Institute of Chemical Biology and Fundamental Medicine of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
[4] Laboratory Angiopharm LLC, Novosibirsk, Russia
[5] State Scientific Center of Virology and Biotechnology "Vector" of Rospotrebnadzor, Koltsovo, Novosibirsk region, Russia
[6] City Clinical Hospital No. 1, Novosibirsk, Russia
[7] Novosibirsk Tuberculosis Research Institute, Novosibirsk, Russia
[8] Research Institute of Fundamental and Clinical Immunology, Novosibirsk, Russia
✉ labmolbiol@mail.ru

Abstract. The induced "recombinogenic situation" in hematopoietic stem cells and the activation of the cell's reparative systems create the basis for recombination events between fragments of extracellular double-stranded DNA delivered into the cell and chromosomal DNA or other forms of the reparative-recombination process. In mouse and rat model organisms as well as in human bone marrow cells, changes in the amount of telomeric DNA in hematopoietic stem cells were assessed as an indicator of repair and recombination events that have occurred. In all experiments performed, recombinant human angiogenin was used as a comparison factor. Dot blot hybridization showed that in the colony cells obtained from the bone marrow cells of the model organisms as well as from human bone marrow cells treated with a double-stranded DNA preparation, there was a significant increase in the amount of telomeric DNA. Amplification of telomeric DNA in colony cells is not associated with contamination of the original DNA preparation with which the bone marrow cells were treated. Treatment of bone marrow cells with DNA that does not carry telomeric sequences (*AluI* PCR fragment) does not lead to an increase in the amount of telomeric DNA in the cells of grown colonies. This suggests the participation in the amplification of telomeric DNA of an extrachromosomal DNA template carrying telomeric DNA. It has been established that treatment of bone marrow cells with angiogenin also leads to an increase in telomeric DNA in colony cells. A comparison of the type of colonies with the intensity of hybridization (i.e. the amount of telomeric DNA in the sample) suggests that the increase in the amount of detectable telomeric DNA following treatment with angiogenin and hDNA^{gr} has a fundamentally different origin. Western blot analysis and real-time PCR revealed that the increase in the amount of telomeric DNA following treatment of bone marrow cells with a double-stranded DNA preparation does not correlate with the activity of endogenous/exogenous telomerase. For angiogenin, it has been shown that an increase in the amount of telomeric DNA may be the result of activation of endogenous telomerase activity. A principle has been developed for the amplification of a new genetic trait that came into hematopoietic stem cells with extracellular double-stranded DNA material and was fixed in the recipient genome or was transitively present in the cell as new genetic information.

Key words: hematopoietic stem cells; dot blot hybridization; telomeric DNA; angiogenin; recombinogenic situation

For citation: Ruzanova V.S., Oshikhmina S.G., Ritter G.S., Dolgova E.V., Kirikovich S.S., Levites E.V., Efremov Y.R., Karamysheva T.V., Bogomolov A.G., Meschaninova M.I., Mamaev A.L., Taranov O.S., Sidorov S.V., Nikonov S.D., Leplina O.Y., Ostanin A.A., Chernykh E.R., Kolchanov N.A., Proskurina A.S., Bogachev S.S. Concept of natural genome reconstruction. Part 3. Analysis of changes in the amount of telomeric DNA in colony cells as a new amplified feature that arose during the processing of hematopoietic bone marrow stem cells. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov J Genet Breed*. 2025;29(4):479-495. doi 10.18699/vjgb-25-52

Funding. This work was done with the support of the Ministry of Science and Higher Education of the Russian Federation for the Institute of Cytology and Genetics SB RAS (state budget-funded project No. FWNR-2022-0016) and by A.A. Purtov, I.N. Zaitseva and LLC "ES.LAB DIAGNOSTIC".

© Ruzanova V.S., Oshikhmina S.G., Ritter G.S., Dolgova E.V., Kirikovich S.S., Levites E.V., Efremov Y.R., Karamysheva T.V., Bogomolov A.G., Meschaninova M.I., Mamaev A.L., Taranov O.S., Sidorov S.V., Nikonov S.D., Leplina O.Y., Ostanin A.A., Chernykh E.R., Kolchanov N.A., Proskurina A.S., Bogachev S.S., 2025

# These authors contributed equally to this work.

This work is licensed under a Creative Commons Attribution 4.0 License

Концепция природной реконструкции генома.
Часть 3. Анализ изменения количества теломерной ДНК в клетках колоний как нового амплифицированного признака, возникшего при обработке гемопоэтических стволовых клеток костного мозга

В.С. Рузанова [1#], С.Г. Ошихмина [1, 2#], Г.С. Риттер [1], Е.В. Долгова [1], С.С. Кирикович [1], Е.В. Левитес [1], Я.Р. Ефремов [1], Т.В. Карамышева[1], А.Г. Богомолов [1], М.И. Мещанинова [3], А.Л. Мамаев[4], О.С. Таранов [5], С.В. Сидоров[2, 6], С.Д. Никонов[7], О.Ю. Леплина [8], А.А. Останин [8], Е.Р. Черных [8], Н.А. Колчанов [1], А.С. Проскурина [1#], С.С. Богачев [1#] ✉

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

Аннотация. Индуцированная «рекомбиногенная ситуация» в гемопоэтических стволовых клетках и активация репаративных систем клетки создают основу для рекомбинационных событий между доставленными в клетку фрагментами экстраклеточной двуцепочечной ДНК и ДНК хромосом или иных форм репаративно-рекомбинационного процесса. На модельных организмах мыши и крысы, а также с использованием в качестве исходного материала клеток костного мозга человека было оценено изменение количества теломерной ДНК в гемопоэтических стволовых клетках как показатель произошедших репарационно-рекомбинационных событий. Во всех проведенных экспериментах в качестве фактора сравнения использовался ангиогенин рекомбинантный человеческий. Методом дот-блот гибридизации показано, что в клетках колоний, полученных из клеток костного мозга модельных организмов, а также из клеток образцов костного мозга человека, обработанных препаратом двуцепочечной ДНК, произошло достоверное увеличение количества теломерной ДНК. Амплификация теломерной ДНК в клетках колоний не связана с контаминацией препаратом исходной ДНК, которым обрабатывались клетки костного мозга. Обработка клеток костного мозга ДНК, не несущей теломерных последовательностей (*AluI* ПЦР-фрагмент), не приводит к увеличению количества теломерной ДНК в клетках выросших колоний. Это предполагает участие в амплификации теломерной ДНК экстрахромосомальной ДНК-матрицы, несущей ДНК теломер. Установлено, что обработка клеток костного мозга ангиогенином также сопровождается увеличением теломерной ДНК в клетках колоний. Сопоставление типа колоний с интенсивностью гибридизации (т. е. количества теломерной ДНК в образце) предполагало, что увеличение количества детектируемой теломерной ДНК при обработке ангиогенином и hDNA^{gr} имеет принципиально разное происхождение. Вестерн-блот анализом и методом ПЦР в реальном времени установлено, что увеличение количества теломерной ДНК при обработке клеток костного мозга препаратом двуцепочечной ДНК не коррелирует с активностью эндогенной/экзогенной теломеразы. Для ангиогенина показано, что увеличение количества теломерной ДНК может быть результатом активации эндогенной теломеразной активности. Разработан принцип амплификации нового генетического признака, пришедшего в гемопоэтические стволовые клетки с экстраклеточным двуцепочечным ДНК материалом и закрепившимся в реципиентном геноме или транзитно присутствующим в клетке в качестве новой генетической информации.

Ключевые слова: гемопоэтические стволовые клетки; дот-блот гибридизация; теломерная ДНК; ангиогенин; рекомбиногенная ситуация

Introduction

The central idea of this part of the study is to prove that extracellular double-stranded DNA fragments internalized into hematopoietic stem cells (HSCs) (Ruzanova et al., 2024) are involved in recombination repair processes activated by these fragments in undifferentiated precursors. A telomere, which consists of repetitive homogeneous DNA sequences, was used as the model target as changes in its content can be easily detected experimentally. In all the mammals, telomeric repeats have an identical nucleotide sequence. Therefore, human DNA can be used as a substrate for assessing changes in telomeric DNA content caused by recombination repair processes in different experimental model systems. Quantitative dot blot hybridization with a telomeric repeat DNA probe was chosen as the main method for assessing the events that have taken place.

Induction of pangenomic single-strand breaks by double-stranded DNA fragments internalized into HSCs via a natural

mechanism is the underlying phenomenon in a cascade of events defined by us as a "recombinogenic situation" (Likhacheva et al., 2008). This state of the cell drives the recombination repair machinery, resulting in numerous interactions between chromatin and intranuclear DNA fragments.

In essence, it is the same situation as the one when double-strand breaks are formed or disruption of higher-order chromatin structure is induced in the cell. Two key aspects of this process can be differentiated within the recombinogenic situation: the enzymatic molecular machinery activated in the cell and recombinant intermediates of chromatin and DNA fragments involved in recombination repair. Both aspects have been thoroughly analyzed for double-strand breaks and single-stranded DNA, while the data on nick-initiated recombinogenic situation are very sparse as there has been a lack of research community's attention to these processes over the past two decades.

According to what has been said, in Supplementary Materials 1 and 2[1], we briefly describe the molecular events characterizing the emergence of double-strand breaks and single-stranded DNA or disruption of higher-order chromatin structure assuming that many of the described details will also be typical of the nick-initiated recombinogenic situation. Supplementary Material 1 lists brief information about factors involved in the processes described. An analysis reported in Supplementary Material 2 indicates that a comprehensive response to damage and various perturbations in higher-order chromatin structure is induced in the cell. The system of hierarchical kinases (ATM, ATR, DNA-PK, belonging to the family of phosphatidylinositol-3-kinase-dependent kinases) is activated, and molecular systems of either restoring chromatin integrity or normalizing its spatial organization are brought into an active state.

Single-strand DNA breaks (nicks) are among the key factors initiating the disruption of higher-order chromatin structure. This type of chromatin structure disruption was shown to have its own repair pathway and activate the palette of recombination repair factors that is intrinsic to this pathway and differs from the machinery of double-strand DNA break repair. Homologous recombination, the main feature of which is the high precision of correction of genetic information, is activated upon nick-induced recombinogenic situation (Vriend, Krawczyk, 2017; Maizels, Davis, 2018). It is known that ATM and ATR kinases are not absolutely required for the DNA break repair process upon nick emergence and repair via the homologous recombination mechanism. It means that nick-initiated homologous recombination may proceed without involvement of hierarchically organized kinases and, therefore, without activation of the checkpoint mechanism. Furthermore, there are grounds to believe that nick-initiated homologous recombination can be independent of the phase of the cell cycle. Like double-strand break repair, repair of DNA nicks depends on the formation of replication factor A filaments. While double-strand break repair is dependent on activity of the BRCA1, RAD51, and BRCA2 complexes, repair of DNA nicks is related to BRCA1 activity, but is independent of RAD51 (Vriend, Krawczyk, 2017; Maizels, Davis, 2018).

Therefore, DNA fragments internalized into the cell induce nicks. A recombinogenic situation develops, and the homologous recombination mechanism is initiated. In this scenario, intracellular DNA fragments act as an extrachromosomal substrate for recombination repair processes activated by them. As a result of the interaction of fragments and chromatin, changes in the DNA of chromosomes will occur. We suppose that if these alterations are large-scale, they can be detected by modern analytical methods.

Chromosomal loci in which genomic alterations can be detected will obviously be those carrying repetitive DNA sequences, including intercalary heterochromatin domains, centromeres, and telomeres. The DNA content in these chromosome domains is high, so changes in DNA content in a selected locus can be detected using various experimental approaches, including real-time PCR, fluorescent *in situ* hybridization (FISH), and quantitative dot blot hybridization assay, in the case of large-scale alterations. Intercalary heterochromatin and centromeric satellites are species-specific, and allogeneic DNA should be used for analyzing alterations that have occurred in the genome. In all mammals and humans, telomeric satellites are represented by the same hexanucleotide repeat TTAGGG, and any mammalian or human double-stranded DNA can be used for conducting experiments and analyzing telomeric chromatin alterations in different species (Giardini et al., 2014).

Telomeres are specific chromatin structures at the ends of eukaryotic chromosomes. As mentioned above, telomeric DNA in most eukaryotes consists of hexanucleotide repeats (for humans, TTAGGG). Human telomeres are approximately 10 kbp long. The forward and complementary telomeric strands are known as the G-strand and the C-strand, respectively. The 3′-end of the G-strand is single-stranded DNA known as the telomeric G-tail. It has been demonstrated that the G-tail penetrates into the proximal double-stranded repeat and anneals to the C-strand to form a special structure, the t-loop. Each telomere is the region of DNA sequence at the end of a chromosome that is protected by the closed circle of the t-loop structure and specific proteins, shelterin and CST (CTC1-STN1-TEN1) heterotrimeric complex, against degradation causing chromosomal instability. The CST heterotrimer binds to the t-loop intermediate at the site of annealing of the single-stranded G-overhang and the complementary sequence of the 3′–5′ strand (C-strand) to form a protective capping complex (Fig. 1*A*) (Giardini et al., 2014; Soman et al., 2022; Alanazi et al., 2024).

Shelterin, a specialized protein complex, is a functional basis of telomeric chromatin and, in mammalian cells, consists of one (POT1) and two (TRF1 and TRF2) telomeric DNA binding proteins, as well as specific proteins linking these DNA binding proteins (Fig. 1*A*) (Giraud-Panis et al., 2010; Lee et al., 2014; Soman et al., 2022). Together, these structural complexes form telomeric heterochromatin in mammals (Lu W. et al., 2013).

[1] Supplementary Materials 1–5 are available at:
https://vavilov.elpub.ru/jour/manager/files/Suppl_Ruzanova_Engl_29_4.pdf

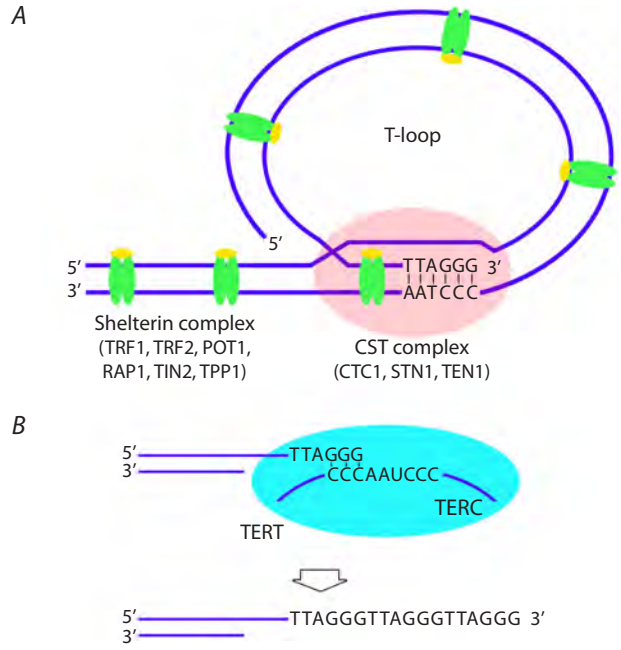


Fig. 1. Telomere structure and the mechanism of its elongation by the telomerase complex.

A – protein complexes in the telomeric region. *B* – the mechanism of telomere lengthening by the telomerase complex. Several nucleotides at the 3′ end of the telomeric G-strand complementarily bind to the TERC template sequence of telomerase RNA. The chromosome end is lengthened by telomerase reverse transcriptase (TERT).

Telomeric DNA in dividing cells is prone to shortening (the so-called end replication problem). Semiconservative replication cannot complete the synthesis of ends of linear DNA. Hence, after several cell division rounds, somatic cells have shortened telomeric DNA, resulting in irreversible cell cycle arrest (the so-called replicative senescence) (Chan, Blackburn, 2003; Doksani, 2019; Jones et al., 2023).

Two mechanisms preventing telomere shortening and preserving the feasibility of infinite division have been described; these mechanisms are active in stem cells, including HSCs, and in immortalized cancer cells. The main mechanism is related to telomerase activity (Fig. 1*B*). Telomerase is a specific reverse transcriptase that elongates the telomeric G-strand. Stem cells and cancer cells in ~90 % of tumors maintain telomere length in a telomerase-dependent manner (Chan, Blackburn, 2003; Nandakumar, Cech, 2013).

The second mechanism, known as alternative lengthening of telomere, has been described for a small number of tumors. This pathway is characterized by a specific mechanism of telomeric DNA metabolism where the key elements are recombination and replication-associated recombination (Lundblad, 2002; Hande, 2004; Pickett et al., 2009; Nabetani, Ishikawa, 2011; Rovatsos et al., 2011; Doksani, 2019; Loe et al., 2020; Lu R., Pickett, 2022; Jones et al., 2023).

In the present study, for assessing large-scale genomic alterations, we chose to analyze the content of telomeric DNA (as a target consisting of non-species-specific repeats in mammals) in cells treated with therapeutic DNA, hDNA^{gr}. The analysis was performed using three approaches: FISH, real-time PCR, and dot blot hybridization assay (Supplementary Material 3). As mentioned above, a simple way to assess changes in the telomeric DNA content in HSCs will be to analyze this parameter in progeny cells after treatment of HSCs within bone marrow cells and their amplification to form colonies on methylcellulose (up to 1,000 cells per colony). Genetically altered HSCs on methylcellulose will produce genetically homogeneous progeny. In other words, a new detectable trait will be amplified (the technology is the property of OJSC "ES.LAB DIAGNOSTIC", Patent Application No. 2023124343 dated September 20, 2023). Telomeres are formed by repeats that are identical for all mammals (the TTAGGG repeat sequence in vertebrates/humans). In principle, this fact made it possible to use hDNA^{gr} in the mouse or rat models to assess alterations in the telomeric DNA content.

Additionally, changes in the telomerase level in colony cells were assessed. Fifteen days after the initial induction within bone marrow cells, colony cells were re-treated with the same factors. Evaluation was performed for colony cell samples collected at time points 0 (untreated samples) and 1, 2, 4, 8, 16, and 32 h after re-treatment. All the evaluations involved comparison between cells treated with three inducers: angiogenin, hDNA^{gr}, and angiogenin+hDNA^{gr}.

Materials and methods

Experimental animals. Young male CBA/Lac mice aged 2–5 months, old male CBA/Lac mice aged 9–12 months, and old male Wistar rats aged 18–22 months, bred at the Conventional Vivarium (Institute of Cytology and Genetics, SB RAS; Novosibirsk, Russia), were used in this study. The animals were housed in groups (6–10 mice and 3–4 rats per cage) with *ad libitum* access to food and water. All animal experiments were approved by the Animal Care and Use Committee of the Institute of Cytology and Genetics, SB RAS. Mice were withdrawn from the experiment by cervical dislocation; rats, by euthanasia using CO_2 or decapitation.

Human bone marrow cells. Cells from cryopreserved bone marrow specimens collected from patients with Hodgkin lymphoma, provided by the cryobank of the Research Institute of Fundamental and Clinical Immunology, were utilized. At the Clinic of Immunopathology of the Research Institute of Fundamental and Clinical Immunology (Hematology Department), having a bone marrow transplant unit, patients with hemoblastosis receive high-dose chemotherapy and transplantation of autologous or allogeneic peripheral HSCs. When harvesting peripheral stem cells, along with the main apheresis product (which is transplanted to the patient), two or three samples (satellite test tubes) of separated cells are also collected to ensure quality control of the apheresis product and for research purposes. Together with the main specimen, these samples were used in the present study. Each bone marrow specimen, including satellite ones, is accompanied by the required documentation package including the Informed Consent Form, Protocol of Bone Marrow Examination, and

Treatment Protocol, which are signed by the patient in accordance with the statutory standards. After the treatment and application of the main product, the satellite samples are either disposed of in compliance with the Sanitary Rules and Regulations or used for research purposes. Documents accompanying each procedure of bone marrow harvesting are stored in the archive of the cryobank of the Research Institute of Fundamental and Clinical Immunology and can be claimed upon first demand.

DNA preparation. Human DNA genome reconstructor ($hDNA^{gr}$) and placenta DNA were isolated from placentas of healthy women. $hDNA^{gr}$ was fragmented to 1–10 nucleosome monomers (200–2,000 bp) by ultrasonic disintegration, deproteinized using proteinase K, and isolated by phenol–chloroform extraction. Placenta DNA was extracted in a similar manner, without fragmentation.

Angiogenin was procured from the Angiopharm Laboratory LLC (Novosibirsk, Russia).

pBSM13-*AluI*-pBSM13 PCR fragment. The human *AluI* repeat (the pBSM13-*AluI*-pBSM13 fragment) was amplified by PCR. *AluI* repeat DNA cloned into pUC19, including the beginning and end of the tandemly repeated AluJ and AluY sequences (NCBI: AC002400.1, 53494–53767), was used as a template. Standard M13 primers were used for amplification (M13 for: 5′ GTAAAACGACGGCCAGT 3′; M13 rev: 5′ CAGGAAACAGCTATGAC 3′). The PCR fragment was resuspended in 0.1 V NaAc 3 M pH 5.2 and 1 V isopropanol for 10 min at –20 °C. The precipitate was washed in 70 % ethanol and dissolved in sterile water.

Isolation of bone marrow cells. After cervical dislocation, femoral and tibial bones were isolated, epiphyses were removed, and the bone marrow cavity was washed with IMDM+2 % FBS. The resulting cell suspension was passed through a 21-gauge needle several times to eliminate bone marrow rosettes and then through a 40-μm filter. Cells were pelleted by centrifuging for 10 min at 400g and resuspended in red blood cell lysis buffer containing 130 mM ammonium chloride for 3–5 min. The buffer was then diluted tenfold with PBS; cells were re-pelleted, resuspended in IMDM medium, and counted in a Goryaev chamber.

Treatment of bone marrow cells with inducers. Bone marrow cells isolated from old animals and bone marrow sections from patients with Hodgkin lymphoma were incubated with inducers for 1 h in the 5 % CO_2 atmosphere with 95 % humidity at 37 °C at the following ratio: 500 μg of $hDNA^{gr}$, or 500 ng of angiogenin, or 500 μg of $hDNA^{gr}$ + 500 ng angiogenin in 1 mL of serum-free IMDM medium per 3×10^6 cells. Control (untreated) bone marrow cells were incubated in serum-free IMDM medium supplemented with PBS volume equal to that of the inducer added to activate bone marrow cells.

Cultivation of bone marrow cells in methylcellulose medium. Bone marrow cells with/without inducer activation were pelleted for 10 min at 400g and resuspended in IMDM+2 % FBS. To quantify and analyze myeloid precursors, the mouse bone marrow cells were placed in the MethoCult M3434 methylcellulose medium; the rat and human bone marrow cells were placed in the MethoCult H4034 methylcellulose medium (Stem Cell Technologies). Colony counting and cell isolation from the methylcellulose medium after cultivation were carried out according to the manufacturer's instructions. Cells were cultivated for 9–15 days depending on the experiment objective.

DNA isolation from colony cells and the liver of young mice. Colony cells were pelleted at 400g for 5–7 min; the precipitate was resuspended in 50 mM EDTA.

After mice had been subjected to cervical dislocation, a liver fragment was dissected and homogenized in a buffer supplemented with 100 mM EDTA pH 8.0 and 20 mM Tris-HCl pH 7.5. SDS was then added to the cells in both cases until a concentration of 1 %, and the homogenate was incubated in the presence of 100 μg/mL proteinase K at 58 °C for 60 min. DNA was isolated by phenol-chloroform extraction and re-pelleting of 1 V isopropanol from 0.3 M NaAc. The pelleted DNA was washed with 70 % ethanol and dissolved in sterile water. DNA was quantified on a Qubit 4 fluorometer (Thermo Fisher Scientific, USA).

Total RNA isolation. Colony cells were pelleted at 400g for 5–7 min. The precipitate was resuspended in TRIzol Reagent (Thermo Fisher Scientific, USA). Total RNA was isolated in accordance with the manufacturer's instructions. RNA content was measured on a Qubit 4 fluorometer (Thermo Fisher Scientific, USA).

Obtaining cDNA. Reverse transcriptase PCR was performed on the poly-A mRNA template using a T100 Thermal Cycler (Bio-Rad Laboratories, USA) and an MMLV RT kit (Evrogen, Russia) according to the manufacturer's protocol.

Dot blot hybridization. DNA samples isolated from mouse and human colony cells were used to quantify telomeric DNA. DNA samples were sonicated to a size of 100–500 bp. DNA was denatured in 0.2 M NaOH at 100 °C for 10 min, and equal quantities of DNA were applied to the Hybond N membrane using specialized equipment, a dot chamber. The samples were annealed to the membrane for 10 min using an ultraviolet lamp and stored until hybridization.

The membrane with attached DNA was transferred to 50 mL of a pre-hybridization buffer containing 0.1 % SDS, 5×SSC, 5× Denhardt's solution, and 100 μg/mL yeast total RNA, and incubated at 37 °C for 1–3 h. The labeled DNA sample 54 bp (P^{32} oligonucleotide G-probe – $(TTAGGG)_9$; C-probe – $(CCCTAA)_9$) was denatured by 10-min boiling and added to 50 mL of the hybridization buffer containing 0.1 % SDS; 5×SSC; 5 % dextran sulfate 500,000; and 100 μg/mL yeast total RNA. The pre-hybridization solution was removed, and the hybridization buffer containing labeled material was added to the membrane after stirring. Hybridization was carried out at 37 °C overnight under constant stirring. After hybridization, the membrane was washed thrice with a solution containing 0.1 % SDS and 0.1×SSC (for 15 min each time) at 37 °C. The hybridization regimen (the buffer system, temperature, and number of washings) of short oligonucleotides was selected empirically in numerous experiments with radioactive isotopes of phosphorus and lies within the range of 37–42 °C (Dolgova et al., 2012).

The membrane with the samples transferred to it was projected onto a K-type screen. Radioisotope-labeled samples were scanned using the PharosFX system. The recorded images were analyzed employing the Quantity One software using the spot density parameter (intensity/mm^2).

Pulsed-field gel electrophoresis. Rat colony cells were used for quantifying telomeric DNA by pulsed-field gel electrophoresis. The cells were pooled, washed to remove the methylcellulose medium, and counted in a Goryaev chamber. The colony cells were embedded into blocks based on 1 % low melting point agarose (5×10^5 cells per block). Before the analysis, the blocks were stored in 0.5 M EDTA at 4 °C. Before pulsed-field gel electrophoresis, the blocks were rinsed in TE buffer and incubated with a lysis buffer (50 mM EDTA, 1 % sarcosyl (Serva, Germany), 1 mg/mL proteinase K (Thermo Fisher Scientific, USA)) for 20 min at 50 °C. Next, the low melting point agarose blocks were fixed in agarose block pockets and subjected to electrophoretic separation in a pulsed-field gel electrophoresis system according to the following regimen: forward – 3 s; reverse – 1 s; RAM-factor – 0.9.

DNA was then transferred to a Hybond N membrane using the capillary method in 20×SSC (Maniatis et al., 1984). DNA samples were attached to the membrane for 10 min using an ultraviolet lamp and stored until hybridization. Hybridization with the P^{32}-labeled oligonucleotide and scanning of radioisotope-labeled samples were then performed in a manner identical to that for dot blot hybridization assay.

Analysis of TERT expression. Bone marrow cells isolated from bone marrow sections from patients with Hodgkin lymphoma were incubated in the presence of inducers (hDNA^{gr}, angiogenin, angiogenin+hDNA^{gr}) and without them (untreated bone marrow cells, control) in IMDM for 1 h in an atmosphere of 5 % CO_2, at 95 % humidity and 37 °C. Next, the bone marrow cells were cultured in methylcellulose medium for 15 days. When isolating cells from the methylcellulose medium, the colonies were pooled and washed to remove the medium according to the manufacturer's instructions. The colony cells were then counted in a Goryaev chamber and incubated again with inducers. After inducer activation or without it, the cells were re-pelleted for 10 min at 400g, resuspended in DMEM/F-12 (1:1) medium (BioloT, Russia) supplemented with 10 % fetal bovine serum (Capricorn Scientific, Germany), 100 µg/mL gentamicin (Dalkhimpharm, Russia) and 1 µg/mL amphotericin B (Sintez, Russia), and inoculated into wells of a 24-well plate. A sample of cells was collected and divided into two parts consisting of equal amounts of cells 1, 2, 4, 8, 16, and 32 h after re-induction. The zero point corresponded to the colony cells before re-treatment with inducers.

One portion of the cells was pelleted; the precipitate was lysed in TRIzol Reagent, and total RNA was isolated. RNA samples were pooled into two groups: 0–4 and 8–32 hrs. RT-qPCR was performed on the poly-A mRNA template using a T100 Thermal Cycler amplifier and an MMLV RT kit according to the manufacturer's protocol. qPCR was conducted in 96-well plates using the BioMaster HS-qPCR SYBR (2×) mix according to the manufacturer's protocol on a QuantStudio 5 Real-Time PCR System (Applied Biosystems, USA). The primer sequences are summarized in the Table. The qPCR analysis of each sample was performed in three replicates. The relative expression level was determined by the $2^{-\Delta\Delta Ct}$ method. The 0–4 h group of samples was used as a control group; the expression level of the target gene in them was taken as unity. Rplp0 was used as the reference gene. The PCR protocol was as follows: 95 °C for 5 min, 40 cycles of 95 °C for 20 s, 57 °C for 30 s, 72 °C for 30 s; the final melting step with slow heating from 60 to 95 °C.

Sequences of the primers used

Name	Structure
Rplp0-for	5′ CGTCCTCGTGGAATGACAT 3′
Rplp0-rev	5′ GCATCATGGTGTTCTTGCCC 3′
TERT-for	5′ GGCACGGCTTTTGTTCAGAT 3′
TERT-rev	5′ ACATGCGTGAAACCTGTACG 3′

The other portion of the cells was pelleted and resuspended in saline solution. Protease inhibitors were added to the cell suspension: PMSF, N-ethylmaleimide, and TPSK to a concentration of 1 mM and aprotinin, to a final concentration of 2 µg/mL. A sample buffer (66 mM Tris-HCl, pH = 6.8; 26.3 % glycerol; 2.1 % SDS; and 0.011 % bromophenol blue) was then added; lysates were boiled at 96 °C for 10 min and centrifuged for 5 min at 12,000 rpm. The lysates were used to conduct electrophoresis, and samples were not pooled over time. The samples were equilibrated according to the number of lysed cells before being applied to the electrophoresis system. Commercially available recombinant human TERT protein (Cloud-Clone-Corp, USA) (2 µg per lane) was used as a control. Western blotting with antibodies was performed after electrophoresis and transfer to the nitrocellulose membrane. Non-specific binding was blocked by incubation in 0.01 M phosphate-buffered saline (PBS) supplemented with 0.02 % Tween 20 overnight at 4 °C. Membranes were then incubated with polyclonal antibodies specific to human TERT (Cloud-Clone-Corp, USA) or monoclonal primary antibodies specific to human TERT (Antibody System, France) and anti-mouse IgG (H+L) secondary antibodies (Affinity Biosciences, USA). Western blotting was performed using an ECL Western blotting detection system (Abcam, UK) and visualized using an iBright imaging system (Thermo Fisher Scientific, USA).

Statistical analysis was carried out using the Statistica 8 software (StatSoft, USA). Statistical significance was assessed using the Mann–Whitney U-test; the differences were considered significant at $p < 0.05$.

Results

Choosing the adequate method for quantifying the telomeric DNA content

A series of analytical experiments was conducted. At the first stage, three methodological approaches for quantifying telomeric DNA were elaborated using the mouse model.

In order to choose the method for quantifying the telomeric DNA, mouse bone marrow cells were treated with hDNA^{gr} and angiogenin activators and inoculated onto methylcellulose. The cells were harvested after nine days. DNA was isolated from a portion of the cells, followed by real-time PCR and dot blot hybridization. Some cells from the same sample were treated with colchicine, and FISH was carried out. Hence, the experiments were conducted using the same cell material at a single time point, "here and now", so we successfully assessed the adequacy of each approach for quantifying telomeric DNA in the analyzed samples (Supplementary Material 3).

The findings indicated that real-time PCR and FISH used for analyzing telomere length under the selected experimental conditions yielded conflicting results that could have mechanistic interpretation. Only dot blot hybridization allows one to detect high statistically significant difference in changes in telomeric DNA content. Therefore, we chose quantitative dot blot hybridization for measuring telomeric DNA content. This approach allows one to directly quantify the content of DNA homologous to the probe being used in the experimental sample regardless of the circumstances summarized in Supplementary Material 3.

Quantification of telomeric DNA content in colony cells by dot blot hybridization

There can be several reasons for the increased telomeric DNA content in HSC progeny that was treated as part of bone marrow cells and gave rise to colonies with a higher telomeric DNA content:

1) integration of telomeric DNA that is initially present in the hDNA^{gr} sample into the HSC genome and its amplification as part of genetically homogeneous colony cells;
2) amplification of cyclic telomeric repeats present in hDNA^{gr} (rolling circle amplification or alternative lengthening of telomeres);
3) induction of endogenous HSC telomerase or a transient telomerase gene incorporated along with extracellular hDNA^{gr} internalized into HSCs, stochastically containing telomerase gene DNA;
4) activation of quiescent HSCs, previously never activated by life events, containing an initially given, maximum possible, number of telomeric repeats (and thus telomeric DNA);
5) the increase in the amount of telomeric DNA is a consequence of the presence of colonies of residual initial hDNA^{gr} in the cells;
6) mixed variants are also possible.

Mouse and human bone marrow cells

The telomeric DNA content in HSC progeny cells treated within bone marrow cells with hDNA^{gr} activators, angiogenin and their combination, in the mouse model and in human bone marrow cells using the quantitative dot blot hybridization was estimated (Fig. 2). Experiments were repeated multiple times (see Figure 2 captions) with DNA from different extractions and using forward and reverse hybridization probe primer. Two approaches to normalizing DNA quantity in the treated cell samples were selected. First, DNA quantities were normalized with respect to intercalator (Qubit), and quantitative dot blot hybridization was performed (mouse model, human bone marrow cells). Second, normalization was performed with respect to the number of colony cells taken into the treatment (rat model).

Figure 2*H* compares the results of hybridization of angiogenin, hDNA^{gr}, and angiogenin+hDNA^{gr}. One can see that the hybridization intensity varies differently in the angiogenin samples in different experiments. For hDNA^{gr}, the hybridization intensity is always higher than that of the control; telomeric DNA content for the hDNA^{gr} inducer exceeded that in the control 1.1–2.5-fold. When using a combination of two inducers, the hybridization signal was insignificantly higher than that in the control. The findings revealed that the considered trait for angiogenin as a monopreparation is unstable.

Finding an association between colony type and hybridization intensity. We compared the dependence of hybridization signal intensity (the telomeric DNA content in the sample) on the type of colonies in four independent experiments, where the indicated parameters were taken into account. Two lineages, BFU-E and CFU-GM, were analyzed (Supplementary Material 4).

These findings indicated that there are fundamentally different reasons for the increase in quantities of detectable telomeric DNA upon treatment with angiogenin and hDNA^{gr}. For angiogenin, the increase in telomeric DNA content can possibly be related to induction of G0 activity of CFU-GM colonies, which had previously remained inactive in the bone marrow and contained the embryologically predetermined quantity of telomeric DNA. For hDNA^{gr}, comparison of all the data obtained for both model systems revealed no correlation between the hybridization intensity and prevalence of a certain colony type.

Assessment of the intensity of hybridization with P^{32}-labeled telomeric probe using the pBSM13-*AluI*-pBSM13 PCR fragment as an inducer. Special mention should be given to the results of comparative hybridization with DNA isolated from colonies derived from HSCs within bone marrow cells treated with the pBSM13-*AluI*-pBSM13 PCR fragment, with placenta DNA and DNA isolated from colonies on day 15 after treatment with inducers. It appeared that the PCR DNA fragment did not stimulate the increase in telomeric DNA content in colony cells (Fig. 2*E–G*). This fact means that extracellular DNA fragments (in this particular experiment) do not induce endogenous telomerase activity.

Justification of the feasibility of changing the hybridization intensity depending on quantity and composition of internalized DNA fragments, as well as the variant of the P^{32}-labeled telomeric (C/G) probe. It is noteworthy that changes in hybridization intensity in the samples exposed to hDNA^{gr} in different experiments can be related to the quantity of telomeric DNA internalized by the cell. Since the cell may contain approximately 0.2 % (mouse) to 0.02 % of extracellular fragments (humans) (Potter et al., 2024; Ruzanova et al.,

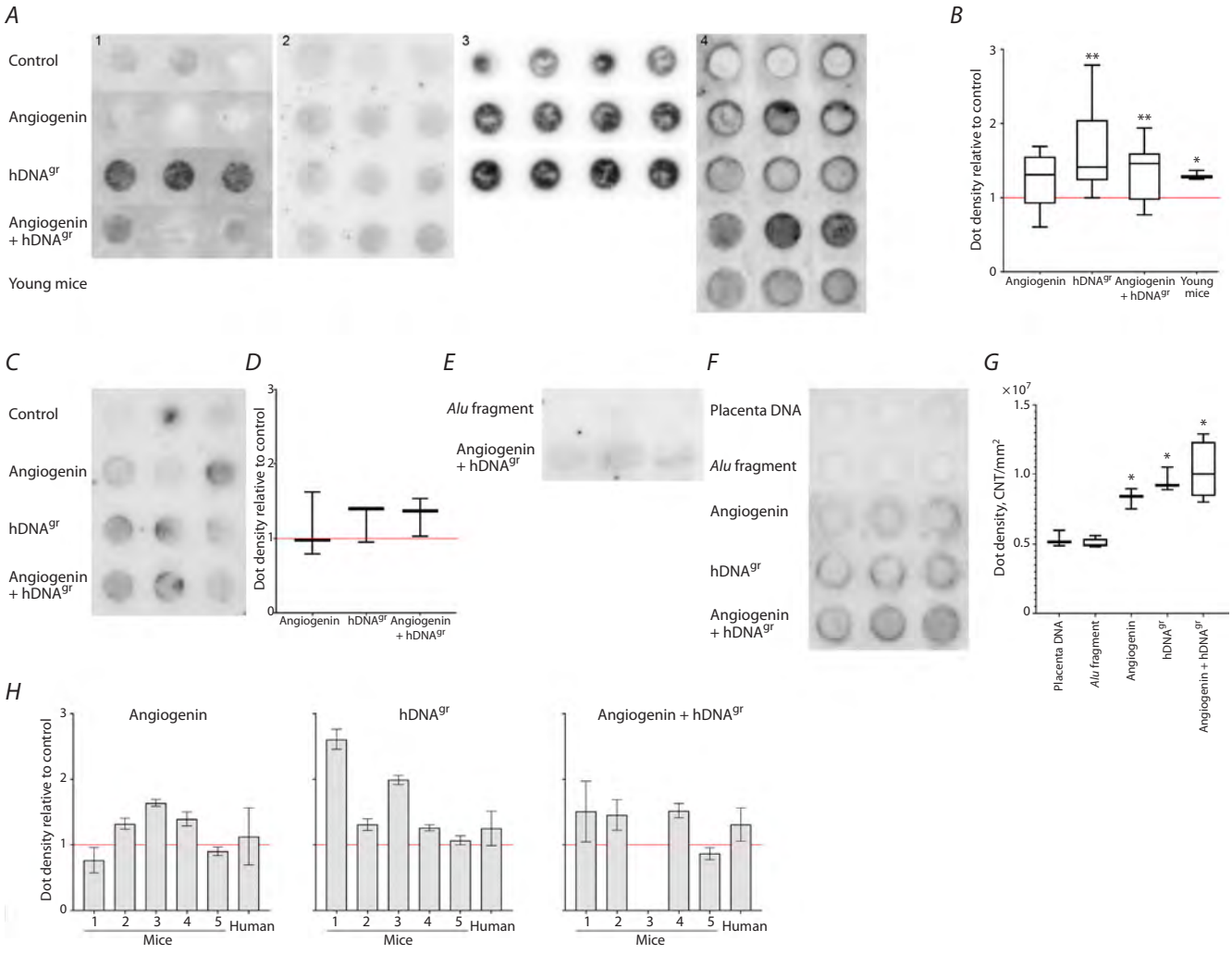


Fig. 2. Quantitative dot blot hybridization of DNA extracted from colonies of hematopoietic stem cells within mouse (*A*, *B*) and human bone marrow (*C*–*G*) without induction (Control) and after induction with angiogenin, hDNA^{gr}, and angiogenin+hDNA^{gr} using a telomeric repeat (54 bp, $n = 9$) as a probe. *A*1, *A*3, *C*, *F* – C-probe; *A*2, *A*4 – G-probe. DNA of young mice (*A*4), human placenta DNA (*F*), and DNA extracted from colonies derived from HSCs within bone marrow cells subjected to treatment with the pBSM13-*AluI*-pBSM13 PCR fragment (*E*, *F*). The membrane was analyzed using a phosphor imaging system. Signal intensity was analyzed in the Quantity1 software. *A*, *C*, *E*, *F* – images of the membranes after hybridization. *B*, *D*, *G* – the diagrams showing spot density (intensity/mm^2) with respect to the control group where spot density was taken as unity (the red line). Significant differences were determined using the Mann–Whitney U-test compared to the control group (*B*) and compared to the group treated with the *AluI* fragment (*G*); * $p < 0.05$, ** $p < 0.01$. *H* –comparative analysis of hybridization intensity (the telomeric DNA content) from individual experiments between colony DNA samples after treatment with angiogenin, hDNA^{gr}, or angiogenin+hDNA^{gr}.

2024), the result of competitive internalization will always be vague when it comes to the qualitative composition of the internalized fragments. It means that the number of telomeric repeats can vary significantly from experiment to experiment. Furthermore, the variation in hybridization signal intensity could be because either forward or reverse primer had been used. An analysis of hybridization intensity using two different probes indicated that DNA homologous to the G-tail (C-probe) underwent maximum amplification. DNA homologous to the C-tail (G-probe) was also amplified, but not significantly.

Comparison of the intensity of hybridization response to P^{32}-labeled telomeric DNA probe between the DNA extracted from colonies of the control sample and DNA extracted from the liver of young animals. The intensities of hybridization response to the P^{32}-labeled telomere DNA probe were compared to DNA extracted from colonies of the control sample and from the liver of young animals (Fig. 2*A*4). One can see an unambiguously interpretable rise in the intensity of hybridization response in the young animal sample. This result supports the known fact that old organisms have a lower telomeric DNA content in HSCs than young individuals. Moreover, this finding indicates that if inducers activated proliferation of quiescent HSCs of embryonic origin, the hybridization pattern would not differ significantly for the samples obtained from young animals and experimental mice.

Assessment of the chances that residual hDNA^{gr} can remain in HSCs after they are treated with this DNA

within bone marrow cells, which may cause an artifact of increased telomeric DNA content. If the DNA internalized by the cell is not integrated into the genome, there is a chance that it is present as extrachromosomal material for a long time, which may produce this artifact of increased telomeric DNA content (Dolgova et al., 2012).

The quantity of foreign DNA in the progeny cells of human bone marrow cells on day 15 of cultivation on methylcellulose after HSCs within bone marrow cells had been treated with TAMRA-labeled *AluI* repeat DNA flanked by pBS sequences with primers M13 was estimated (Supplementary Material 5). The findings indicate that the *AluI* repeat DNA molecules, which had initially been internalized into HSCs during primary processing of bone marrow cells, are not detected in colony cells. In other words, the increased telomeric DNA content detected in dot blot hybridization experiments cannot result from the presence of residual original DNA in colony cells. In addition, it follows from the experiments that *AluI* fragments, together with the nonhomologous ends of pBSM13, are not included in the genome and are not amplified by PCR.

To sum up all the findings obtained, the following conclusion can be drawn. For hDNA^{gr}, the increase in hybridization intensity is not associated with the prevalence of a certain colony type; therefore, it is not associated with HSCs, which have previously been inactive throughout the entire life of the organism. There are several variants for activation of the telomerase gene of exogenous origin, direct integration of telomeric DNA into the HSC genome, or emergence of extra telomeric DNA resulting from replication upon treatment of bone marrow cells with an inducer. The variant that DNA fragments initially internalized by HSCs at an amount sufficient to alter the hybridization response intensity can be persisting in the non-integrated state in the cell throughout the entire time of culturing on methylcellulose is ruled out.

Nevertheless, the cumulative result suggests that there is more likely to be true integration of telomeric DNA internalized by the cell into the genome or emergence of replication-related extra telomere DNA.

For angiogenin, the increase in hybridization intensity can be associated with induction of the CFU-GM-derived cells that previously remained inactive throughout the entire life of the organism. Activation of the endogenous telomerase gene is also possible. Both possibilities are indicated by the results of experiments on internalization of angiogenin protein into primitive murine and human hematopoietic stem cells. It has been demonstrated that angiogenin is internalized by primitive murine Sca1 hematopoietic cells and human CD34+ stem cells (Ruzanova et al., 2024). In this case, the telomerase gene can be activated by angiogenin internalized by active HSCs. Integration is not an option, since there is no necessary substrate.

Hence, this analysis has reliably revealed that changes occur in the cells in response to induction of bone marrow cells by DNA preparation, angiogenin, or their combination, affecting the length of telomeric repeats (the telomeric DNA content) in as many cells as are needed to enable imaging of the observed phenomenon.

Rat bone marrow cells

For the rat model experiments, the number of colony cells was chosen as the normalization criterion. After being washed to remove methylcellulose, colony cells were embedded into blocks of low melting point agarose (500,000 per block, corresponding to about 3 μg of DNA). The blocks were lysed, and electrophoresis was carried out using a pulse controller as described in the Materials and methods section. The electrophoretic data were analyzed, and Southern blot analysis was conducted. The results obtained are summarized in Fig. 3. The electrophoresis images were processed using the GelPro 3.0 software (Fig. 3*A*). The relative ratio between DNA quantities in the lanes was estimated from the luminescence of the intercalary dye. In the sample containing DNA-treated cells, DNA quantity in the bands that were subsequently evaluated by hybridization increased by a total of 10 %. Meanwhile, the increase in the three bands was nonuniform: the top two bands almost did not increase, whereas in the third band, the quantity of DNA increased twofold (Fig. 3*B*).

The hybridization data showed that the number of telomeric repeats in the sample of DNA-treated cells increased by 17–30 % (Fig. 3*C*, *D*). The number of telomeric repeats was increased in all the bands, in contrast to the rise in the total amount of DNA. In angiogenin-treated cells, the amount of DNA decreased compared to the control group. It can be related to the fact that a large number of erythroid colonies grew from HSCs treated with angiogenin within the bone marrow (Fig. 3*E*), which may contain mature red blood cells lacking DNA but which could have been counted during cell selection.

Analysis of the potential mechanisms for the increase in telomeric DNA content

Murine HSCs internalize extracellular DNA fragments (Potter et al., 2024). The model system of cryopreserved human bone marrow was used in the main experiments of further studies. It was found that human CD34+ HSCs also capture extracellular DNA fragments. A total of 0.02 % of the haploid genome of extracellular DNA (in a particular experiment) is internalized by the cell (Ruzanova et al., 2024).

Two inducers were selected to analyze the mechanism for the increase in telomeric DNA content: one of those, hDNA^{gr}, carries telomeric DNA as a potential sensing substrate in HSCs, while the other one, angiogenin, does not carry any DNA, including telomeric DNA.

It means that the increase in quantity of detectable telomeric DNA after angiogenin treatment is associated with induction of either endogenous telomerase or activation of quiescent primary HSCs that previously have never been activated and that have been formed and have occupied the bone marrow niches during embryogenesis. Activation and integration of the exogenous telomerase gene are infeasible, since the necessary substrate is lacking.

The following options are being considered for hDNA^{gr}: the option related to the feasibility of incorporation of hDNA^{gr} *per se* into the telomeric DNA genome or an increase

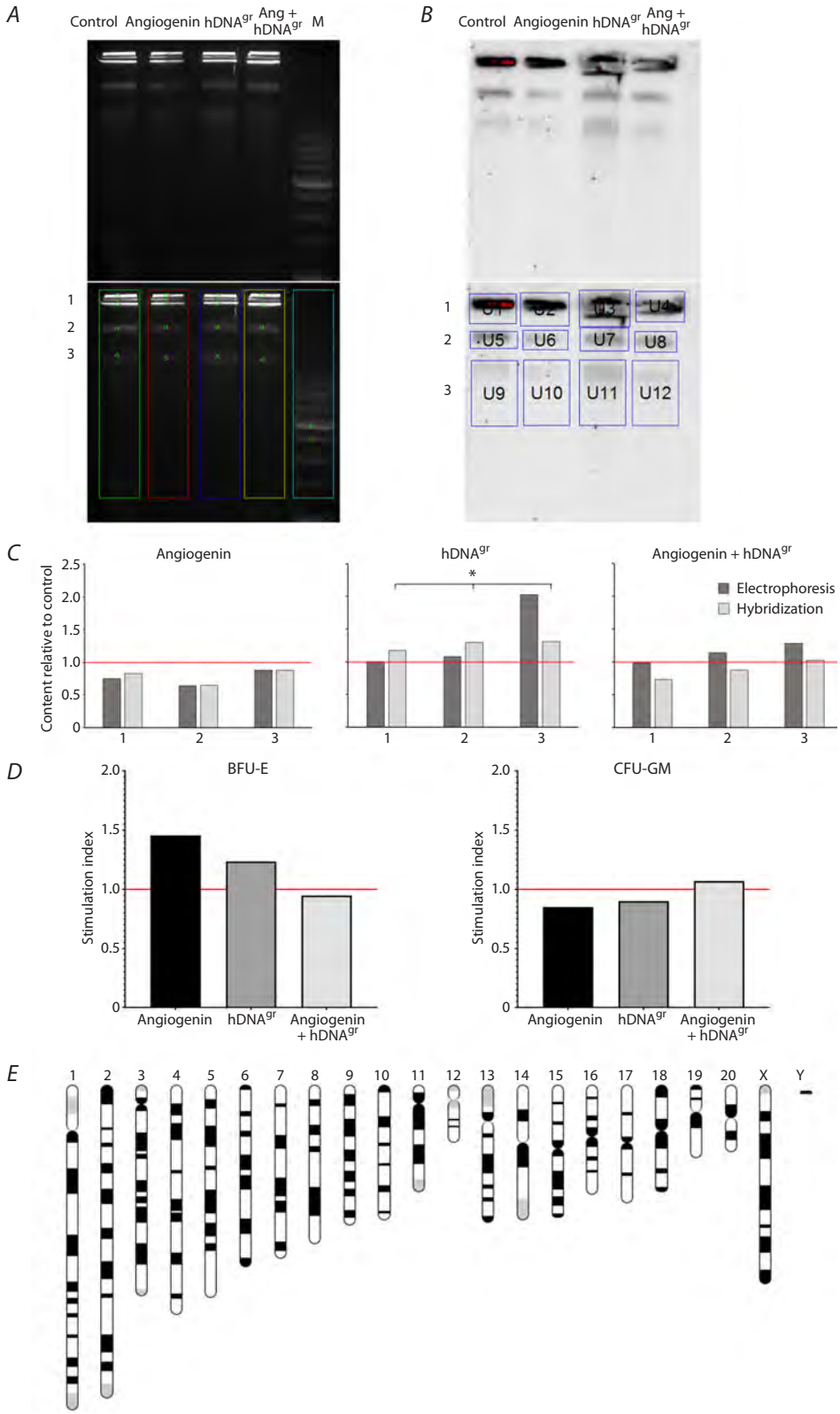


Fig. 3. The results of treating rat bone marrow cells with angiogenin, hDNA^{gr}, and angiogenin+hDNA^{gr}.

A – electrophoresis with DNA isolated from colonies, from low melting point agarose blocks. In the bottom block, numbers 1–3 denote the fragments used for quantitative analysis. *B* – hybridization with telomeric repeats (C-probe) of DNA isolated from colonies. In the lower block, numbers 1–3 denote the regions used for quantitative analysis. *C* – DNA content according to dye luminescence and hybridization intensity for three different fragments compared to the control group (values are taken as unity; shown with a red line). * Differences are significant compared to the control group; $p < 0.05$, Mann–Whitney U-test. *D* – the content of the BFU-E and CFU-GM colonies on methylcellulose after treating rat bone marrow cells with different inducers, expressed as an index with respect to the control group (values taken as unity, shown with a red line). *E* – rat chromosomes (https://rgd.mcw.edu/rgdweb/report/genomeInformation/genomeInformation.html?species = Rat&mapKey = 372&details = true).

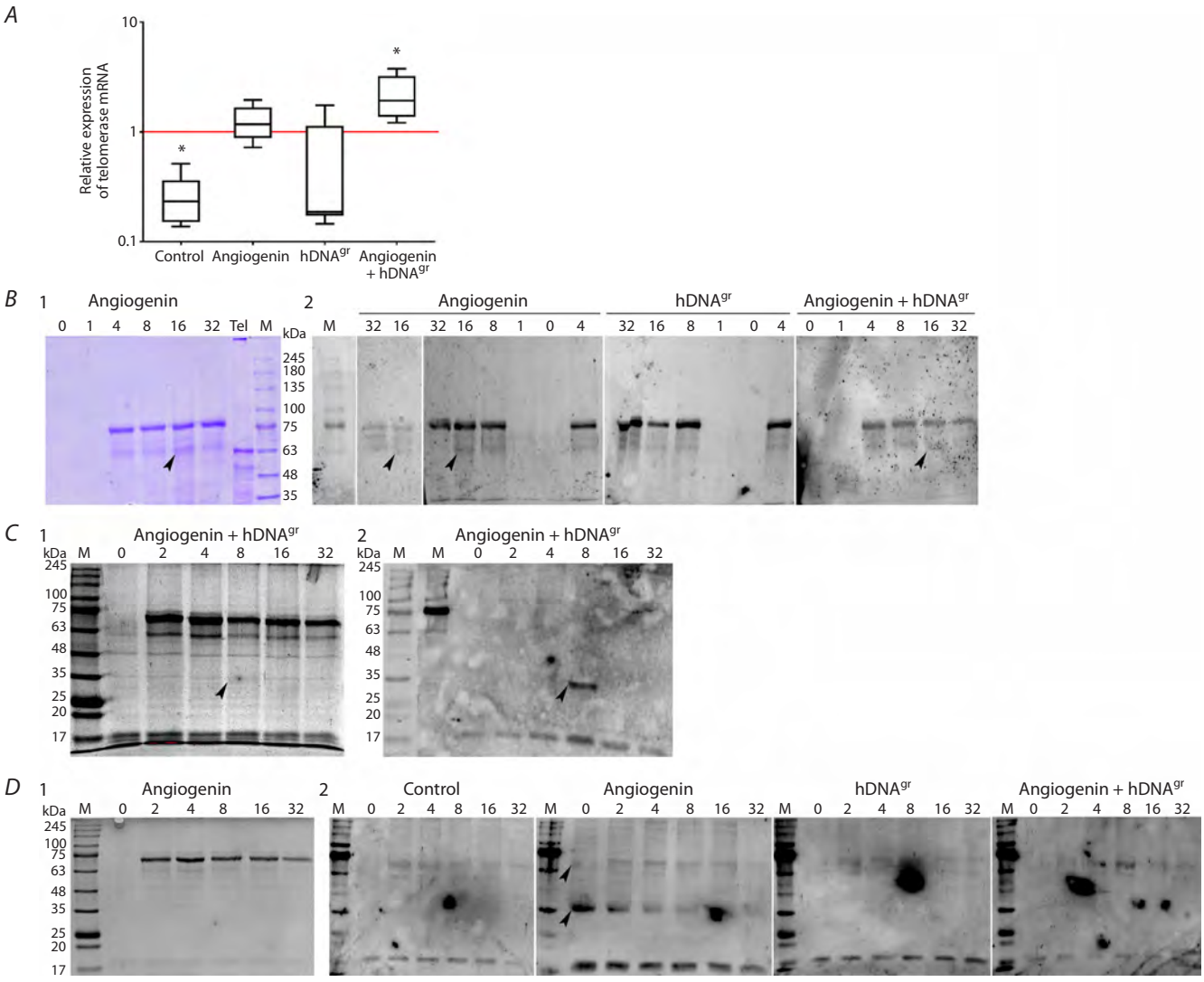


Fig. 4. Real-time PCR and Western blotting of RNA and protein lysates to detect telomerase mRNA and telomerase protein.

A – real-time PCR of the pooled samples (8–32 h). Values are shown relative to the respective 0–4 h groups (values are taken as unity; indicated with a red line). * Differences are significant compared to the 0–4 h group, $p < 0.005$, Mann–Whitney U-test. *B–D* – Western blotting to detect telomerase in the lysates of activator-treated cells; 1 – acrylamide gels with Coomassie staining, 2 – blots with antibodies to telomerase. Incubation time (h) with the corresponding inducers is indicated above the lanes. Arrows indicate specific bands at 63 and 35 kDa. Results from three independent experiments are shown.

in the number of G-telomeric repeats as a result of replication repair of the G-strand (tail), which will be observed as a larger quantity of telomeric DNA in the descendants of the primary bone marrow HSCs. hDNA^{gr} can also possibly activate the transient telomerase gene residing within the extracellular fragments internalized into the HSC compartments. The dot blot hybridization data do not support the option that the endogenous telomerase gene in cells accepting the DNA fragment is activated (Fig. 4*E–G*). An option similar to that described for angiogenin, involving activation of quiescent primary HSCs, which have previously never been activated and which have been formed and have occupied the bone marrow niches during embryogenesis, was not confirmed. Furthermore, the option that the residual amount of hDNA^{gr} is preserved in colony cells after the initial uptake of extracellular fragments by HSCs within bone marrow cells and subsequent contamination of DNA samples in colonies by the residual amount of telomeric DNA, which is initially present in the hDNA^{gr} sample and is sufficient to be sensed by dot blot hybridization, has not been confirmed as well.

The analysis suggested that, at least for angiogenin, the most likely explanation for the increased telomeric DNA content would be induction of telomerase activity by this factor. For hDNA^{gr}, activation of the transient telomerase gene was still possible.

Assessment of the effect of inducers on telomerase activity

It was analyzed in direct experiments whether telomerase activity can be activated. The experiments were carried out using the model of HSCs within human bone marrow cells

after treatment with hDNA^{gr}, angiogenin, and their combination. It has previously been demonstrated that markers of primitive cells are still retained in colonies grown after HSCs had been activated by the hDNA^{gr} inducer to ~15 % for murine HSCs (c-Kit/Sca-1) and up to 3 % for human HSCs (CD34) (Potter et al., 2024). It meant that reactivation of colony-cultured cells would have similar effects on primitive progenitors and induce similar events in them; in particular, the telomerase gene can be activated in the options suggested above. There would be sufficient material for characterizing cell lysates to detect any telomerase present by real-time PCR or (and) Western blotting.

On incubation day 18, colony cells (pangenomic single-strand breaks being completely repaired), after activation of HSCs within bone marrow cells by three inducers, were re-treated using the same substances. Samples were collected and lysed at the zero point, 1, 2, 4, 8, 16, 32 h after re-induction. Samples were prepared for real-time PCR and Western blotting. For real-time PCR, RNA samples were pooled into two groups, 0–4 and 8–32 h. Time-specific samples were used for Western blotting. The results of the experiments are summarized in Figure 4.

The real-time PCR data (three independent replicates) indicated that when a response develops between 8–32 h after re-induction, telomerase mRNA was synthesized in the samples that had been treated with angiogenin and angiogenin+hDNA^{gr}. In the control sample, telomerase mRNA synthesis was blocked. In the hDNA^{gr}-treated sample, no telomerase mRNA was detected under the selected conditions (Fig. 4*A*).

Western blotting was conducted using samples corresponding to all the time points. Three independent experiments were performed. Polyclonal antibodies were used in the first one, while monoclonal anti-telomerase antibodies were used in the second and third experiments. The following results were obtained. In the first experiment, when utilizing polyclonal antibodies in the samples treated with angiogenin and angiogenin+hDNA^{gr}, a 63 kDa band corresponding to a cloned fragment (EcoRI-NotI clone 712562) of human telomerase (Cech et al., 1998) was detected 16 h post-induction, which correlates with the overall pattern of telomerase mRNA synthesis (Fig. 4*B*). Two consecutive experiments utilizing monoclonal antibodies against human telomerase revealed a ~35 kDa band that was undetectable by Coomassie staining (Fig. 4*C*, *D*). In the third experiment, a ~63 kDa band was detected along with the 35 kDa band (Fig. 4*D*). The specific ~35 kDa band was detected in groups of samples treated with angiogenin+hDNA^{gr} (second experiment, Fig. 4*C*) and angiogenin (third experiment, Fig. 4*D*). The modes of emergence of this specific band differ for the two experiments. In the second experiment, the ~35 kDa band clearly appeared at the time point of 8 h after starting the re-induction for the angiogenin+hDNA^{gr} sample. No band was detected in the other samples of the second experiment. For the third experiment, an intense ~35 kDa band along with a 63 kDa band was detected at the zero time point (i. e., before the induction in colony cells of the angiogenin-treated sample after washing to remove methylcellulose). As the incubation proceeds, by 32 h of the experiment, the intensity of the 35 kDa band dropped almost to the background value. The 63-kDa band disappeared within the first hour after re-induction. No bands were detected in other images. We found a single publication mentioning a 35-kDa protein related to telomerase activity. During affinity chromatography used for telomerase isolation, the 35-kDa protein was detected along with the 120-kDa and 43-kDa proteins. This protein was not analyzed in this study because it was not detected in preparations of fully purified telomerase (Lingner, Cech, 1996).

The data obtained using two independent approaches indicate that angiogenin activates the molecular mechanisms inducing telomerase activity in HSCs, while exposure to hDNA^{gr} does not abolish the activity of this mechanism.

The results obtained mean that: 1) hDNA^{gr} does not induce expression of the telomerase gene; therefore, the increased telomeric DNA content cannot be related to telomerase activity; 2) angiogenin induces expression of the telomerase gene, and the increase in the telomeric DNA content quantified by dot blot hybridization can be related to this very activity of angiogenin. Exposure to a combination of angiogenin and hDNA^{gr} also enhances synthesis of telomerase mRNA. It was shown in the hybridizations performed that in some cases, the telomeric DNA content in the sample treated with a combination of two inducers is higher than for the samples treated separately. This fact may mean that in this case, the three mechanisms of increasing the telomeric DNA content overlap due to telomerase activity and either direct integration of extra telomeric DNA into the genome of the recipient cell or via replication of quasi t-rings formed by exogenous telomeric repeats during concatamerization and closure into a ring.

Discussion

The analysis revealed that the two inducers used simultaneously increase telomeric DNA content in colony cells via two independent mechanisms: the conventional telomerase-dependent complementary synthesis in the case of angiogenin and the alternative lengthening mechanism for telomeres or true integration of telomeric DNA into the telomeric heterochromatin domain in the case of hDNA^{gr}.

Telomerase is a heterodimer formed by the non-coding RNA template (telomerase RNA component with a size of 400 bp, carrying the basic telomeric sequence complementary to the G-strand) used for *de novo* synthesis of telomeric DNA sequences and the catalytic subunit of the enzyme (telomerase reverse transcriptase, TERT). The telomerase complex orchestrates telomere length homeostasis by inserting telomeric repeat sequences to the 3′ end of the chromosome using the RNA template (Fig. 1*B*). With a few exceptions (e. g., lymphocytes and endothelial cells), most human somatic cells exhibit no telomerase activity, mostly because of suppression of TERT expression. On the other hand, stem cells, germline cells, and most tumors exhibit it (Chan, Blackburn, 2003; Giraud-Panis et al., 2010; Nandakumar, Cech, 2013; Soman et al., 2022).

As shown previously, angiogenin is internalized into Sca1 (mouse) and CD34 (human) hematopoietic stem cells (Ruzanova et al., 2024). It has also been demonstrated using human bone marrow cells that angiogenin treatment stimulates the GM hematopoietic lineage and induces telomerase activity. Earlier, it was revealed (Goncalves et al., 2016) that recombinant angiogenin stimulates proliferation of myeloid progenitors (like in our experiments), while enhancing the quiescent properties of stem cells. These characteristics are attributed to generation of stress-induced tiRNAs, reduction of the synthetic activity of blood stem cell, enhanced ribosomal RNA synthesis, and stimulation of protein synthesis in myeloid progenitor cells. It is possible that the emergence of telomerase in cells of GM colonies, which has been demonstrated in our study, is a consequence of the first process (stimulation of proliferation of myeloid progenitor cells).

Internalized DNA fragments initiate the formation of nicks required for chromatin rearrangement toward the selected pathway of terminal differentiation of progenitor cells, which actually trigger the mechanism of this differentiation (Ruzanova et al., 2024). A similar concept was discussed in ref. (Sjakste, Riekstiņa, 2021), where it was suggested that chromosomal DNA damage in stem cells can trigger differentiation. Nick-induced chromatin perturbations (chromatin relaxation) will result in an induced recombinogenic situation and activation of the recombination repair machinery comprising numerous active and structural proteins (Nabetani, Ishikawa, 2011). This fact means that the initially extracellular DNA fragments residing in the intranuclear space of HSCs can participate in recombination events that they initiated. The results of the present study suggest that the increase in telomeric DNA content when using $hDNA^{gr}$ as an inducer implies either integration of extracellular fragments carrying telomeric repeats into homologous telomeric regions or activation of the mechanism of alternative lengthening of telomeres using concatamerized cyclic telomeric repeat sequences.

An analysis of the intensity of hybridization response revealed that the telomeric DNA content was significantly increased in the analyzed samples, suggesting that the mechanism of alternative lengthening of telomeres was involved in this process to a greater extent. One of the most characteristic features of cells having an active mechanism of alternative lengthening of telomeres is the presence of extrachromosomal telomeric circles, which are either double-stranded (t-circles) or partially single-stranded (C- or G-circles) Cesare, Griffith, 2004; Wang et al., 2004; Henson et al., 2009). The t-loops are closed double-stranded DNA. C-circles are extrachromosomal telomere DNA with a C-strand forming a circle and a broken G-strand annealed to it. G-circles are extrachromosomal telomere DNA, with the G-strand forming a circle and the broken C-strand annealed to it. The origin of these extrachromosomal structures is usually attributed to nick-initiated replication of telomeric DNA (Fig. 5). Break repair is known to be stimulated by induction of telomeric double-strand breaks (McEachern, Haber, 2006; Dilley et al., 2016) and, as suggested by the present study, single-strand breaks (nicks). Break-induced replication can be initiated by integration of the strand end of the ruptured telomeric agglomerate between the strands of the intact telomere and proceed via the branch migration mechanism. The migrating D-loop copies telomeric repeats from the strand invasion point towards the end of the donor telomere (Saini et al., 2013; Wilson et al., 2013), which is accompanied by restoration of telomere length and structure.

Another mechanism of restoring telomere length is associated with insertion of the 3′ end of the telomeric G-strand between the strands of extrachromosomal t- or C-circles. Rolling circle replication is induced in this case, leading to accumulation of a single-stranded G-rich telomeric strand (Fig. 6*A*) (Nabetani, Ishikawa, 2011; Lu W. et al., 2013; Doksani, 2019).

Two mechanisms of formation of t-, C-, and G-circles have been described. t-circles can be formed by intrachromosomal recombination and t-loop liberation as a result of Holliday recombination of the 3′ end of the t-loop annealed to the complementary sequence of the 3′–5′ strand (Fig. 5*A*) (Wang et al., 2004; Nabetani, Ishikawa, 2011; Claussin, Chang, 2015; Doksani, 2019; Jones et al., 2023). In the other case, both variants of the C- and G-circles are formed (Fig. 5*B*).

It was demonstrated that both variants of DNA circles containing telomeric repeats are bulged upon replicative stress associated with replication fork stalling in difficult-to-replicate telomeric heterochromatin or with chromosomal DNA damage (e. g., double-strand breaks or nicks). In this case, the G-circle can initiate the rolling circle replication starting with the nick and synthesis of the C-rich tail carrying telomeric repeats of the 3′–5′ strand up to 100 kbp long, which can be detected both using a phi29 polymerase kit and *in vivo* (Zhang et al., 2017). The amplified single-stranded C-rich tail can assumedly be annealed to the truncated G-telomeric strand (e. g., after a t-loop is separated from the telomere because of replicative stress) and as a homologous template to stimulate synthesis of the truncated G-rich strand (Zhang et al., 2017). Topo II, the NHEJ mechanism, and DNA-PK activity are involved in circle formation in this case (Fig. 6*B*).

The more than twofold increase in telomeric DNA content in some experiments is attributed to telomeric repeat amplification via the putative rolling circle replication mechanism using circular structures as a template and nick-initiated replication. For this type of amplification, the single-stranded G-strand region can be as long as 70–100 kbp (Doksani, 2019; Jones et al., 2023) (Fig. 6*A*, *B*).

The large number of C-circles in the cell is the basic criterion for the mechanism of alternative lengthening of telomeres. The single-stranded 3′ end of the 5′–3′ G-strand can be paired with both the t- and C-circle to form a D-loop (Fig. 6*A*). The following multiple rounds of rolling circle replication amplify telomeric repeats. The number of telomeric repeats synthesized via this mechanism can be degenerate and will differ for different telomeres in different chromosomes (Lee et al., 2014; Jones et al., 2023).

When extracellular fragments carrying telomeric repeats are internalized into the nucleus, the following events occur. A recombinogenic situation triggered by the emergence of

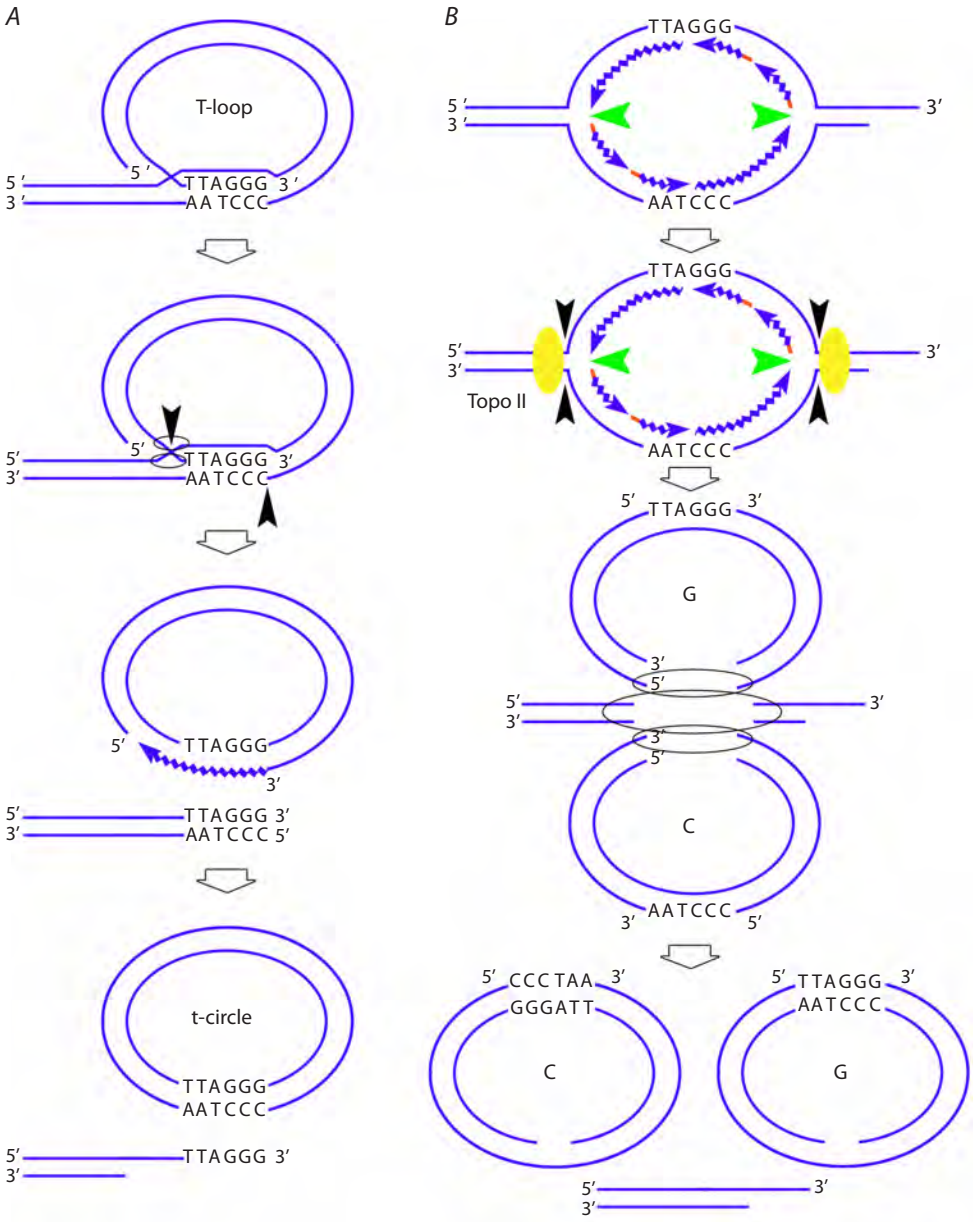


Fig. 5. The mechanisms of formation of an extrachromosomal circle.

A – formation of t-circles as a result of detachment of the telomere end structure of the t-circle. *B* – formation of G- and C-circles as a result of replication fork stalling, induction of break-induced repair, and G- and C-strand bulging with involvement of Topo II, the NHEJ mechanism, and DNA-PK activity (Zhang et al., 2017).

single-strand breaks is initiated. If the factors activated by the nick-initiated recombinogenic situation are similar to those activated by the recombinogenic situation initiated by double-stranded breaks (Dolgova et al., 2013), the internalized double-stranded fragments will promptly form a circle (Dolgova et al., 2013; Potter et al., 2018, 2024). During the time they exist in a linear form, they can be integrated into the genome via the *ends in/ends out* mechanism. Once ligated into a circle, these structures will be virtually indistinguishable from the t- and C-circles formed via the mechanism of alternative lengthening of telomeres. This means that the amplification of telomeric DNA during the internalization of extracellular DNA in HSCs is mostly associated with the elongation of the G-chain of the telomere as a result of activation of replicative synthesis by the rolling ring mechanism, presumably induced by nicks. The significant (more than twofold) increase in telomeric DNA content in some experiments can be attributed to this very fact.

Integration of extrachromosomal fragments carrying telomeric repeats (leading to extension of telomeric DNA) can also be implemented via the *ends in/ends out* homologous recombination mechanism (Fig. 7) (Rubnitz, Subramani, 1984; Hastings et al., 1993; Cromie et al., 2001; Li et al., 2001; Langston, Symington, 2004). Other mechanisms of homologous recombination (single strand annealing or gene conversion) will not increase telomeric DNA content.

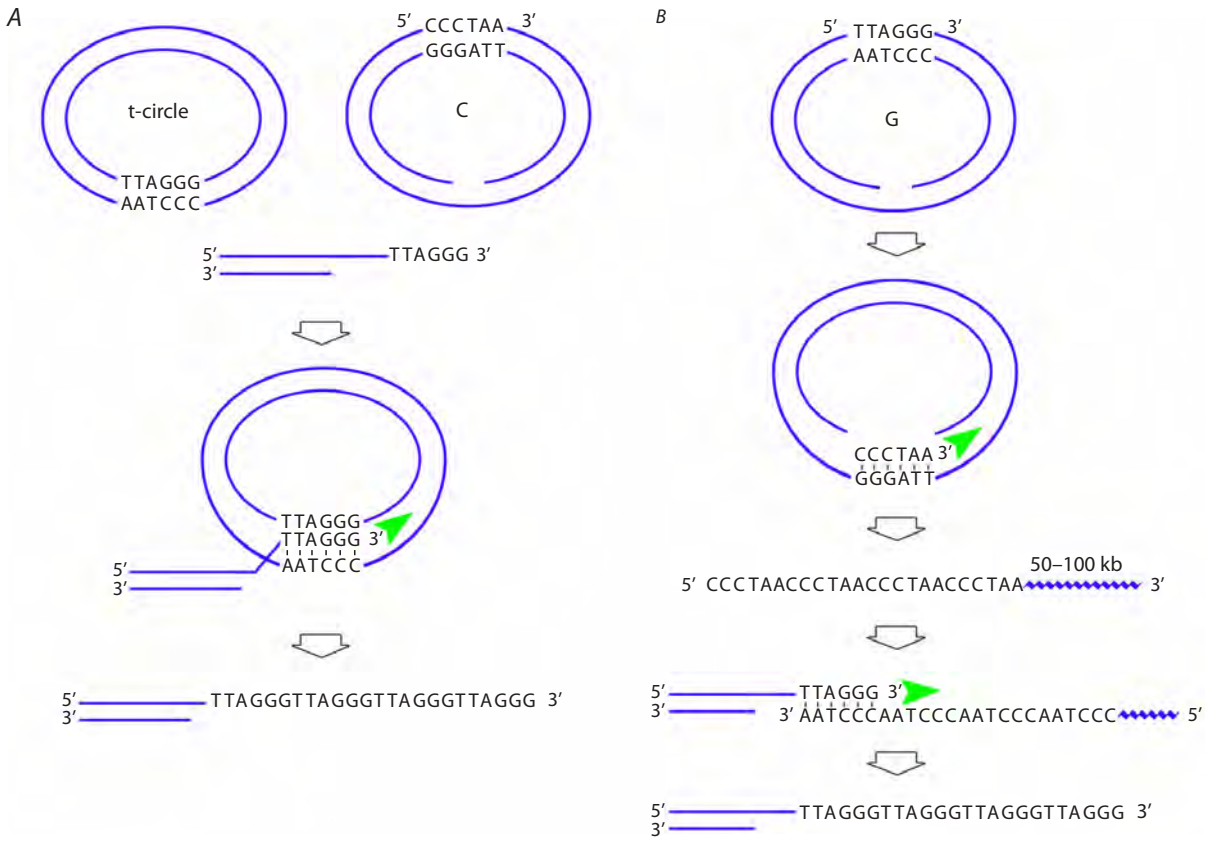


Fig. 6. The mechanisms of telomere lengthening.

A – lengthening of the telomeric tail at the t-circle and C-circle. *B* – lengthening of the telomeric tail at the G-circle (see explanation in the text).

Conclusion

Hence, the conducted studies demonstrate that extracellular DNA fragments internalized by HSCs and carrying telomeric repeats can be either directly integrated into telomeric heterochromatin or become a template for alternative lengthening of telomeres, which is accompanied by a rise in telomeric DNA quantity and, presumably, an increase in telomere length.

References

Alanazi A.F.R., Parkinson G.N., Haider S. Structural motifs at the telomeres and their role in regulatory pathways. *Biochemistry*. 2024; 63(7):827-842. doi 10.1021/acs.biochem.4C00023

Cech T.R., Lingner J., Nakamura T., Chapman K.B., Morin G.B., Harley C.B., Andrews W.H. Telomerase reverse transcriptase. Patent WO1998/14592.1998.

Cesare A.J., Griffith J.D. Telomeric DNA in ALT cells is characterized by free telomeric circles and heterogeneous t-loops. *Mol Cell Biol*. 2004;24(22):9948-9957. doi 10.1128/MCB.24.22.9948-9957.2004

Chan S.W.L., Blackburn E.H. Telomerase and ATM/Tel1p protect telomeres from nonhomologous end joining. *Mol Cell*. 2003;11(5): 1379-1387. doi 10.1016/S1097-2765(03)00174-6

Claussin C., Chang M. The many facets of homologous recombination at telomeres. *Microb Cell*. 2015;2(9):308-321. doi 10.15698/MIC2015.09.224

Cromie G.A., Connelly J.C., Leach D.R.F. Recombination at double-strand breaks and DNA ends: conserved mechanisms from phage to

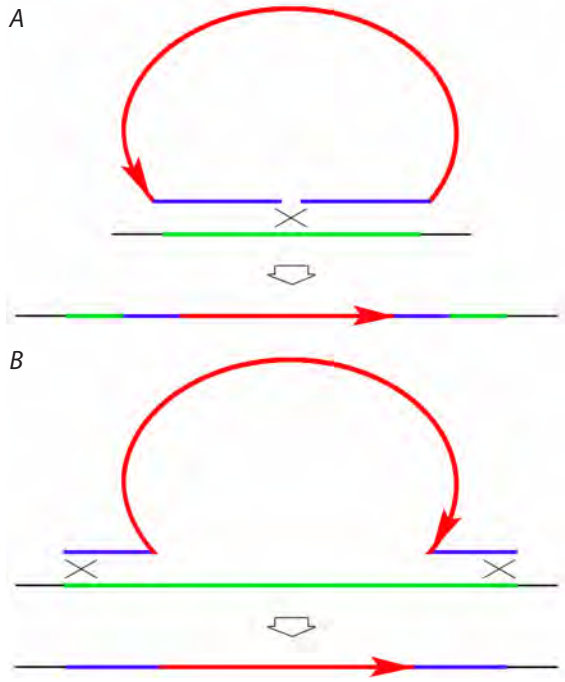


Fig. 7. The mechanisms of *ends in/ends* out integration of extracellular DNA fragments into the recipient genome.

A – ends in integration. *B – ends out* integration.

humans. *Mol Cell*. 2001;8(6):1163-1174. doi 10.1016/S1097-2765(01)00419-1

Dilley R.L., Verma P., Cho N.W., Winters H.D., Wondisford A.R., Greenberg R.A. Break-induced telomere synthesis underlies alternative telomere maintenance. *Nature*. 2016;539(7627):54-58. doi 10.1038/nature20099

Doksani Y. The response to DNA damage at telomeric repeats and its consequences for telomere function. *Genes*. 2019;10(4):318. doi 10.3390/genes10040318

Dolgova E.V., Nikolin V.P., Popova N.A., Proskurina A.S., Orishenko K.E., Alyamkina E.A., Efremov Y.R., … Taranov O.S., Rogachev V.A., Sidorov S.V., Bogachev S.S., Shurdov M.A. Internalization of exogenous DNA into internal compartments of murine bone marrow cells. *Russ J Genet Appl Res*. 2012;2:440-452. doi 10.1134/S2079059712060056

Dolgova E.V., Efremov Y.R., Orishchenko K.E., Andrushkevich O.M., Alyamkina E.A., Proskurina A.S., Bayborodin S.I., … Omigov V.V., Minkevich A.M., Rogachev V.A., Bogachev S.S., Shurdov M.A. Delivery and processing of exogenous double-stranded DNA in mouse CD34+ hematopoietic progenitor cells and their cell cycle changes upon combined treatment with cyclophosphamide and double-stranded DNA. *Gene*. 2013;528(2):74-83. doi 10.1016/j.gene.2013.06.058

Giardini M.A., Segatto M., da Silva M.S., Nunes V.S., Cano M.I.N. Telomere and telomerase biology. *Prog Mol Biol Transl Sci*. 2014;125:1-40. doi 10.1016/B978-0-12-397898-1.00001-3

Giraud-Panis M.J., Pisano S., Poulet A., Le Du M.H., Gilson E. Structural identity of telomeric complexes. *FEBS Lett*. 2010;584(17):3785-3799. doi 10.1016/j.febslet.2010.08.004

Goncalves K.A., Silberstein L., Li S., Severe N., Hu M.G., Yang H., Scadden D.T., Hu G.F. Angiogenin promotes hematopoietic regeneration by dichotomously regulating quiescence of stem and progenitor cells. *Cell*. 2016;166(4):894-906. doi 10.1016/j.cell.2016.06.042

Hande M.P. DNA repair factors and telomere-chromosome integrity in mammalian cells. *Cytogenet Genome Res*. 2004;104:116-122. doi 10.1159/000077475

Hastings P.J., McGill C., Shafer B., Strathern J.N. Ends-in vs. ends-out recombination in yeast. *Genetics*. 1993;135(4):973-980. doi 10.1093/genetics/135.4.973

Henson J.D., Cao Y., Huschtscha L.I., Chang A.C., Au A.Y.M., Pickett H.A., Reddel R.R. DNA C-circles are specific and quantifiable markers of alternative-lengthening-of-telomeres activity. *Nat Biotechnol*. 2009;27(12):1181-1185. doi 10.1038/nbt.1587

Jones C.Y., Williams C.L., Moreno S.P., Morris D.K., Mondello C., Karlseder J., Bertuch A.A. Hyperextended telomeres promote formation of C-circle DNA in telomerase positive human cells. *J Biol Chem*. 2023;299(5):104665. doi 10.1016/j.jbc.2023.104665

Langston L.D., Symington L.S. Gene targeting in yeast is initiated by two independent strand invasions. *Proc Natl Acad Sci USA*. 2004;101(43):15392-15397. doi 10.1073/pnas.0403748101

Lee M., Hills M., Conomos D., Stutz M.D., Dagg R.A., Lau L.M.S., Reddel R.R., Pickett H.A. Telomere extension by telomerase and ALT generates variant repeats by mechanistically distinct processes. *Nucleic Acids Res*. 2014;42(3):1733-1746. doi 10.1093/nar/gkt1117

Li J., Read L.R., Baker M.D. The mechanism of mammalian gene replacement is consistent with the formation of long regions of heteroduplex DNA associated with two crossing-over events. *Mol Cell Biol*. 2001;21(2):501510. doi 10.1128/MCB.21.2.501-510.2001

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)

Lingner J., Cech T.R. Purification of telomerase from *Euplotes aediculatus*: requirement of a primer 3′ overhang. *Proc Natl Acad Sci USA*. 1996;93(20):10712-10717. doi 10.1073/PNAS.93.20.10712

Loe T.K., Zhou Li J.S., Zhang Y., Azeroglu B., Boddy M.N., Denchi E.L. Telomere length heterogeneity in ALT cells is maintained by PML-dependent localization of the BTR complex to telomeres. *Genes Dev*. 2020;34(9-10):650-662. doi 10.1101/gad.333963.119

Lu R., Pickett H.A. Telomeric replication stress: the beginning and the end for alternative lengthening of telomeres cancers. *Open Biol*. 2022;12(3):220011. doi 10.1098/rsob.220011

Lu W., Zhang Y., Liu D., Songyang Z., Wan M. Telomeres-structure, function, and regulation. *Exp Cell Res*. 2013;319(2):133-141. doi 10.1016/j.yexcr.2012.09.005

Lundblad V. Telomere maintenance without telomerase. *Oncogene*. 2002;21(4):522-531. doi 10.1038/sj.onc.1205079

Maizels N., Davis L. Initiation of homologous recombination at DNA nicks. *Nucleic Acids Res*. 2018;46(14):6962-6973. doi 10.1093/nar/gky588

Maniatis T., Fritch E., Sambrook D. Methods of Genetic Engineering. Molecular Cloning. Moscow: Mir Publ., 1984 (in Russian)

McEachern M.J., Haber J.E. Break-induced replication and recombinational telomere elongation in yeast. *Annu Rev Biochem*. 2006;75:111-135. doi 10.1146/annurev.biochem.74.082803.133234

Nabetani A., Ishikawa F. Alternative lengthening of telomeres pathway: recombination-mediated telomere maintenance mechanism in human cells. *J Biochem*. 2011;149(1):5-14. doi 10.1093/jb/mvq119

Nandakumar J., Cech T.R. Finding the end: recruitment of telomerase to telomeres. *Nat Rev Mol Cell Biol*. 2013;14(2):69-82. doi 10.1038/nrm3505

Pickett H.A., Cesare A.J., Johnston R.L., Neumann A.A., Reddel R.R. Control of telomere length by a trimming mechanism that involves generation of t-circles. *EMBO J*. 2009;28(7):799-809. doi 10.1038/emboj.2009.42

Potter E.A., Proskurina A.S., Ritter G.S., Dolgova E.V., Nikolin V.P., Popova N.A., Taranov O.S., Efremov Y.R., Bayborodin S.I., Ostanin A.A., Chernykh E.R., Kolchanov N.A., Bogachev S.S. Efficacy of a new cancer treatment strategy based on eradication of tumor-initiating stem cells in a mouse model of Krebs-2 solid adenocarcinoma. *Oncotarget*. 2018;9(47):28486-28499. doi 10.18632/oncotarget.25503

Potter E.A., Dolgova E.V., Proskurina A.S., Ruzanova V.S., Efremov Y.R., Kirikovich S.S., Oshikhmina S.G., … 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 10.1590/1414-431X2024e13072

Rovatsos M.T., Marchal J.A., Romero-Fernández I., Fernández F.J., Giagia-Athanosopoulou E.B., Sánchez A. Rapid, independent, and extensive amplification of telomeric repeats in pericentromeric regions in karyotypes of arvicoline rodents. *Chromosome Res*. 2011;19(7):869-882. doi 10.1007/S10577-011-9242-3

Rubnitz J., Subramani S. The minimum amount of homology required for homologous recombination in mammalian cells. *Mol Cell Biol*. 1984;4(11):2253-2258. doi 10.1128/mcb.4.11.2253-2258.1984

Ruzanova V.S., Oshikhmina S.G., Proskurina A.S., Ritter G.S., Kirikovich S.S., Levites E.V., Efremov Y.R., Karamysheva T.V., Meschaninova M.I., Mamaev A.L., Taranov O.S., Bogachev A.S., Sidorov S.V., Nikonov S.D., Leplina O.Y., Ostanin A.A., Chernykh E.R., Kolchanov N.A., Dolgova E.V., Bogachev S.S. A concept of natural genome reconstruction. Part 2. Effect of extracellular double-stranded DNA fragments on hematopoietic stem cells. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov J Genet Breed*. 2024;28(8):993-1007. doi 10.18699/vjgb-24-106

Saini N., Ramakrishnan S., Elango R., Ayyar S., Zhang Y., Deem A., Ira G., Haber J.E., Lobachev K.S., Malkova A. Migrating bubble

during break-induced replication drives conservative DNA synthesis. *Nature*. 2013;502(7471):389-392. doi 10.1038/nature12584

Sjakste N., Riekstiņa U. DNA damage and repair in differentiation of stem cells and cells of connective cell lineages: a trigger or a complication? *Eur J Histochem*. 2021;65(2):3236. doi 10.4081/ejh.2021.3236

Soman A., Korolev N., Nordenskiöld L. Telomeric chromatin structure. *Curr Opin Struct Biol*. 2022;77:102492. doi 10.1016/J.SBI.2022.102492

Vriend L.E.M., Krawczyk P.M. Nick-initiated homologous recombination: protecting the genome, one strand at a time. *DNA repair*. 2017;50:1-13. doi 10.1016/j.dnarep.2016.12.005

Wang R.C., Smogorzewska A., De Lange T. Homologous recombination generates t-loop-sized deletions at human telomeres. *Cell*. 2004;119(3):355-368. doi 10.1016/j.cell.2004.10.011

Wilson M.A., Kwon Y., Xu Y., Chung W.H., Chi P., Niu H., Mayle R., Chen X., Malkova A., Sung P., Ira G. Pif1 helicase and Polδ promote recombination-coupled DNA synthesis via bubble migration. *Nature*. 2013;502(7471):393-396. doi 10.1038/nature12585

Zhang T., Zhang Z., Li F., Hu Q., Liu H., Tang M., Ma W., Huang J., Songyang Z., Rong Y., Zhang S., Chen B.P., Zhao Y. Looping-out mechanism for resolution of replicative stress at telomeres. *EMBO Rep*. 2017;18(8):1412-1428. doi 10.15252/embr.201643866

Conflict of interest. The authors declare no conflict of interest.
Received July 4, 2024. Revised February 10, 2025. Accepted February 10, 2025.

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2025;29(4):496-507

doi 10.18699/vjgb-25-53

Unconventional animal models to study the role of telomeres in aging and longevity

E.V. Simoroz[1], J. Vasilevska[1] ✉, N.A. Arakelyan[1], A.D. Manakhov [1, 2], E.I. Rogaev [1, 3]

[1] Research Center for Genetics and Life Sciences, Sirius University of Science and Technology, Sirius Federal Territory, Krasnodar region, Russia
[2] Vavilov Institute of General Genetics of the Russian Academy of Sciences, Moscow, Russia
[3] Department of Psychiatry, UMass Chan Medical School, Worcester, MA, USA

✉ vasilevskaya.e@talantiuspeh.ru

Abstract. The progressive shortening of telomeres is significantly implicated in various cellular processes related to aging, including the limitation of cellular proliferative lifespan through the activation of DNA damage response pathways, ultimately leading to replicative senescence. Telomere shortening is considered an indicator of biological age rather than chronological age. The restoration of telomere length is mediated by the enzyme telomerase; however, it is crucial to maintain a balance in this process, as excessive telomerase activity and overly elongated chromosomes may increase the susceptibility of individuals to cancer. It has been proposed that variations in telomere length among individuals of the same chronological age may be associated with differences in potential lifespan. However, recent studies suggest that telomere length may serve only as a rough estimate of the aging process and is likely not a clinically relevant biomarker for age-related diseases or mortality risk. Furthermore, variations in telomere length are not solely determined by chronological age; rather, they are modulated by a multitude of factors, including genetic predispositions, environmental conditions, and heightened metabolic activities such as reproduction and body weight, which may lead to increased telomere attrition in certain species. It has been argued that traditional animal models, such as the mouse (*Mus musculus*) and the rat (*Rattus norvegicus domestica*), are suboptimal for investigating the relationship between telomere length and aging, as their lifespans and telomere lengths do not adequately reflect those of humans. Consequently, it is recommended to use long-lived species as they would provide a more appropriate framework for such research initiatives. This review aims to examine the correlation between telomere length and longevity in various non-traditional long-lived animal models, evaluating their suitability for investigating the molecular mechanisms underlying telomere attrition in the context of aging. Nevertheless, the question of whether telomere length is a causative factor or a consequence of longevity remains an area that necessitates further investigation.

Key words: telomere length; telomerase; aging; longevity; unconventional animal models

For citation: Simoroz E.V., Vasilevska J., Arakelyan N.A., Manakhov A.D., Rogaev E.I. Unconventional animal models to study the role of telomeres in aging and longevity. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov J Genet Breed*. 2025; 29(4):496-507. doi 10.18699/vjgb-25-53

Funding. The research conducted by V.J. and A.N.A. was funded by the Sirius Federal Territory. Agreement 18-03, project MB-BFT-2403.

Acknowledgements. We express our gratitude to the BioRender.com platform for its contribution to the creation of the schematic illustrations.

Неклассические модели животных для изучения роли теломер в процессах старения и долголетия

Е.В. Симороз[1], Е. Василевская[1] ✉, Н.А. Аракелян[1], А.Д. Манахов [1, 2], Е.И. Рогаев [1, 3]

[1] Научный центр генетики и наук о жизни, Научно-технологический университет «Сириус», федеральная территория «Сириус», Краснодарский край, Россия
[2] Институт общей генетики им. Н.И. Вавилова Российской академии наук, Москва, Россия
[3] Медицинская школа Чан Массачусетского университета, департамент психиатрии, Шрусбери, США

✉ vasilevskaya.e@talantiuspeh.ru

Аннотация. Укорочение теломер играет важную роль в различных клеточных процессах, связанных со старением. Оно ограничивает продолжительность клеточной пролиферации и активирует механизмы, связанные с повреждением ДНК, что, в свою очередь, приводит к репликативному старению. Укорочение теломер отражает биологический, а не хронологический возраст. Механизм восстановления длины теломер осуществляется с помощью фермента теломеразы. Однако в этом процессе необходимо соблюдать баланс, поскольку избыточная активность теломеразы и чрезмерно длинные хромосомы могут привести к развитию онкологических заболеваний. Предполагается, что различия в потенциальной продолжительности жизни могут быть связаны с вариа-

© Simoroz E.V., Vasilevska J., Arakelyan N.A., Manakhov A.D., Rogaev E.I., 2025
This work is licensed under a Creative Commons Attribution 4.0 License

циями в длине теломер у особей одного возраста. Тем не менее последние исследования показывают, что длина теломер может служить лишь приблизительной оценкой скорости старения и, вероятно, не является клинически значимым маркером риска возрастных заболеваний и смертности. Вариации в длине теломер часто обусловлены не только возрастом, но и генетическими изменениями, факторами окружающей среды, а также метаболически затратными процессами, такими как размножение и даже масса тела. Эти факторы могут способствовать ускоренной потере теломер у некоторых видов. Существует мнение, что для изучения роли длины теломер в контексте старения и долголетия классические модели, например мышь (*Mus musculus*) и крыса (*Rattus norvegicus domestica*), не оптимальны, поскольку продолжительность жизни этих животных и длина их теломер не сопоставимы с человеческими. Целесообразно использовать виды с более длительным жизненным циклом. В данном обзоре рассматривается степень корреляции между длиной теломер и долголетием в различных неклассических моделях долгоживущих животных, а также их пригодность для изучения молекулярных механизмов, приводящих к истощению теломер в контексте старения. Важно помнить, что вопрос о причинно-следственной связи длины теломер с продолжительностью жизни по-прежнему требует дальнейшего исследования.
Ключевые слова: длина теломер; теломераза; старение; долголетие; неклассические модели животных

Introduction

Biological aging is characterized by a progressive decline in the functional capacities of an organism following the attainment of maturity, ultimately culminating in its demise. Substantial evidence indicates that telomere length may serve as a significant biomarker for aging and longevity. Telomeres are nucleoprotein structures that protect the ends of linear chromosomes from DNA degradation and are involved in repair processes, thereby playing a crucial role in maintaining chromosomal stability. With each cell division, telomeres undergo shortening due to the phenomenon known as the end-replication problem, particularly exacerbated by oxidative stress, which arises from an imbalance between the production of reactive oxygen species and the organism's antioxidant defense mechanisms (Allsopp et al., 1995; Armstrong, Boonekamp, 2023).

It has been proposed that telomere length may significantly influence the allocation of resources between growth and reproduction, as well as the maintenance of the somatic state of cells (Young, 2018). This hypothesis is grounded in the aging theory articulated by T. Kirkwood (Kirkwood, 1977; Kirkwood, Rose, 1991), who posited that mortality associated with advanced age may result from an energy-conserving strategy designed to minimize the regulation of errors in somatic cells. Consequently, the metabolic costs associated with early reproduction may ultimately result in a depletion of energy resources essential for sustaining a stable somatic state, which could, in turn, expedite the aging process and increase mortality rates. Notably, in certain species, a negative correlation has been observed between telomere length and active reproduction.

The principle of age-related telomere shortening is a complex phenomenon, and the biological mechanisms underlying this process remain poorly understood. Specifically, it is not yet clear whether telomeric aging functions as an analogue of the mitotic clock or serves as a biomarker of cellular stress (Koliada et al., 2015; Lin, Epel, 2022). Telomere attrition can be counteracted by telomerase, a specialized enzyme, the primary function of which is to extend telomeres. Telomerase is a ribonucleoprotein DNA polymerase composed of two subunits: telomerase RNA (TER) and telomerase reverse transcriptase (TERT). This enzyme facilitates the *de novo* addition of TTAGGG repeats to the terminal regions of chromosomes, thereby compensating for telomere loss. Figure 1 illustrates the variations in telomere length in relation to age and telomerase activity. Telomerase initiates its action by binding to the end of a telomere, with TER first interacting with the single-stranded DNA at the terminus of the chromosome. TERT utilizes the incorporated RNA as a template to synthesize new DNA repeats. This synthesis is typically mediated by reverse transcriptase, which generates a new segment of DNA by adding nucleotides to the single-stranded DNA of the telomere. The newly synthesized DNA fragment is complementary to the existing strand, as complementary nucleotides bind to the template, resulting in telomere elongation. Following the synthesis of several repeats, telomerase translocates along the telomere, allowing the enzyme to repeatedly add new nucleotides to the end of the chromosome. This process is repeated multiple times, significantly lengthening the telomere. After synthesizing a new stretch of DNA, the single-stranded DNA of the telomere can form a double helix by pairing with complementary strands. Additional enzymes, such as ligase and DNA polymerases, are also involved in this process to ensure proper telomere joining and completion (Nguyen, 2021).

There are two main telomere protection complexes: CST (Centriole- and Spindle-Associated Telomerase) and shelterin. These complexes can function in parallel in most mammals, including humans. The shelterin complex consists of six key components: TRF1 and TRF2 (Telomeric Repeat Binding Factor 1 and 2), TIN2 (TRF1-interacting protein), TPP1 (Telomeric Protein 1), POT1 (Protection of Telomeres 1), and RAP1 (Repressor-Activator Protein 1). The CST complex includes three components: Ctc1 (Cell Cycle Protein 1), Stn1 (Suppressor of Telomere Lengthening 1), and Ten1 (Telomere Length Maintenance 1). Both complexes play a crucial role in protecting and maintaining telomere structure and are also involved in regulating telomere length (Jenner et al., 2022; Zaug et al., 2022). Telomerase activity is notably elevated during the early stages of human fetal development; however, it is considerably restricted in the majority of normal adult cells. When telomeres attain critically short lengths, they initiate a persistent DNA damage response, which subsequently leads to various cellular processes, including cellular senescence and/or apoptosis. It also diminishes the capacity of stem cells to regenerate tissue. Furthermore, accelerated telomere shortening is a characteristic feature of age-related diseases,

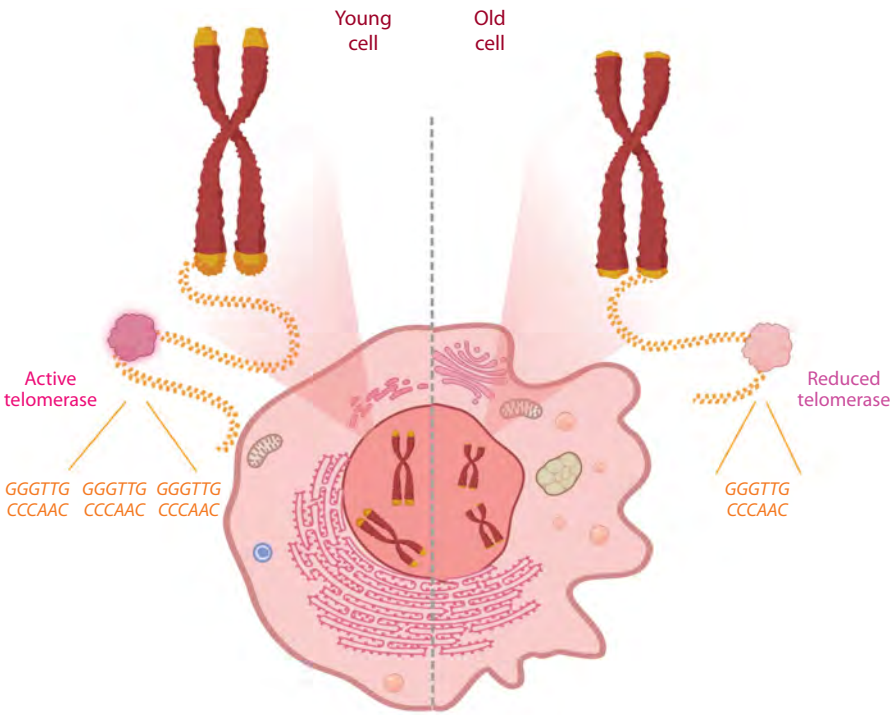


Fig. 1. Variation in telomere length in relation to age and telomerase activity.

which adversely impacts overall health and lifespan (Rossiello et al., 2022).

Research has indicated that telomere length may serve as a positive predictor of lifespan in humans. For instance, centenarians tend to possess longer telomeres, and studies have demonstrated that their offspring inherit this characteristic (Atzmon et al., 2010). Furthermore, healthy centenarians exhibit significantly longer telomeres compared to individuals afflicted with various diseases (Terry et al., 2008). Nevertheless, not all studies support this assertion, which raises questions regarding the reliability of telomere length as a biomarker for longevity and underscores the necessity for further investigation (Arai et al., 2015).

In many studies, traditional animal models, such as the mouse (*Mus musculus*) and the rat (*Rattus norvegicus domestica*), have been used to investigate the molecular mechanisms underlying aging and longevity (Sahm et al., 2018). Research involving genetically modified mice with hyperlong telomeres has demonstrated that these organisms exhibit reduced DNA damage as they age. Furthermore, these mice typically possess a lean body composition, lower cholesterol levels, enhanced glucose and insulin sensitivity, a decreased incidence of cancer, and an extended lifespan (Muñoz-Lorente et al., 2019). However, the applicability of these findings to humans is limited due to significant differences in telomere dynamics between the two species. For instance, the lifespan of mice is approximately 30 times shorter than that of humans, while their telomeres are 5 to 10 times longer and undergo shortening at a rate approximately 100 times faster. Additionally, there are notable disparities in the organization of repetitive elements and the shelterin complex within subtelomeric regions (Vera et al., 2012; Smoom et al., 2023). The complete absence of telomerase in mice results in a relatively weak phenotype over several generations, whereas heterozygosity for telomerase mutations in humans is sufficient to induce defects in organ regeneration and facilitate cancer progression (Calado, Dumitriu, 2013). Moreover, the majority of studies are conducted on specific inbred strains of rodents, such as C57BL/6 and BALB/c mice (Bernardes de Jesus et al., 2012). However, the lifespan observed in natural populations significantly surpasses that achieved in inbred strains, particularly due to anti-aging interventions conducted *in vitro* (Miller et al., 2002). This observation has prompted the consideration of alternative models with exceptionally long lifespans as potentially valuable for elucidating the mechanisms of telomere shortening and their viability as reliable biomarkers for aging and longevity.

This review aims to explore non-traditional long-lived animal models that may offer various advantages in supporting or challenging the role of telomeres as biomarkers that determine age and predict longevity (see the Table).

Birds

Birds serve as a distinctive model for investigating essential cellular mechanisms that are associated with extended lifespan. Despite their significant energetic expenditures throughout life, the majority of bird species can be classified as long-lived homeotherms exhibiting relatively slow aging processes. Notably, birds tend to have longer lifespans compared to mammals of equivalent size. Furthermore, critical aging processes in birds, such as cellular responses to oxidative stress and telomere dynamics, frequently exhibit similarities to those observed in mammals (Harper, Holmes, 2021).

Relationship between telomere length, telomerase activity, and age in non-classical animal models

Animal Model	Lifespan	The trait associated with telomere length	Telomerase activity	Tissues	Technology
Southern giant petrel *Macronectes giganteus*	12–40 years	Telomere length is correlated with life expectancy	No data	Blood	Southern Blotting
Parrots (order: *Psittaciformes*)	33–65 years (long-lived)	Telomere length is correlated with life expectancy	No data	Blood	qRT-PCR
	10–17 years (short-lived)				
53 species of birds	No data	Telomere length is correlated with longevity, smaller clutch sizes, and slower embryonic development	No data	No data	Meta-analysis
19 species of birds	7–50 years	Telomere length is correlated with life expectancy	No data	Blood	TRF
Zebra finch *Taeniopygia guttata*	5 years (short-lived)	No data	In short-lived species, telomerase activity decreases with age, whereas in long-lived species, telomerase activity remains stable	Tissues of killed birds	TRAPeze® XL
Tree swallows *Tachycineta bicolor*	11 years (short-lived)				
Common terns *Sterna hirundo*	27 years (long-lived)				
Leach's storm petrels *Oceanodroma leucorhoa*	36 years (long-lived)				
Greater horseshoe bat *Rhinolophus ferrumequinum*	30 years	Telomere length is correlated with life expectancy	No data	Wing biopsy	qRT-PCR (Cawthon)
Common longwing *Miniopterus schreibersii*	22 years				
Evening bat (genus *Myotis*)	37–40 years	Telomere length does not correlate with chronological age but is linked to various environmental factors	Increased expression of 20 genes associated with telomere maintenance and DNA repair	Wing biopsy and buccal swabs	TRF, qRT-PCR (Cawthon)
Little brown bat *Myotis lucifugus*	34 years (Puechmaille et al., 2011)	Telomere length does not correlate with chronological age, but it is correlated with the incidence of fungal diseases	No data	Wing biopsy and buccal swabs	qRT-PCR (Cawthon)
Rodrigues flying fox *Pteropus rodricensis*	28 years	No correlation between telomere length and chronological age	No data	Wing biopsy and buccal swabs	qRT-PCR (Cawthon)
Black musk deer *Eptesicus fuscus*	19 years[1]				
Brazilian free-tailed bat *Tadarida brasiliensis*	8 years[2]				
Ocean quahog *Arctica islandica*	10–226 years (maximum 507)	Constant telomere length	Constant telomerase activity	Foot tissues and muscles	DNA sequencing
				Gill tissue	Southern Blotting, qRT-PCR
Coral *Stylophora pistillata*	3–6 months (Bythell et al., 2018)	Telomere length is not correlated with chronological age, but it is correlated with exposure to darkness	Decreased expression of Pot2 protein	Tips of the branches	TRF, Southern Blotting

Table (end)

Animal Model	Lifespan	The trait associated with telomere length	Telomerase activity	Tissues	Technology
Species of acroporid corals *Acropora digitifera*	Three stages of development: spermatozoa, planula larvae, and polyps of adult colonies	Telomere length correlates with chronological age and cell proliferation rate	No data	Spermatozoa, planula larvae, and polyps of adult colonies	TRF
Stony corals *Pocillopora* spp. (short-lived)	No data	Telomere length in short-lived colonies is more influenced by seasonal temperature changes than in long-lived colonies	No data	Upper and lower segments of coral branches	TRF
Small polyp stony corals *Porites* spp. (long-lived)	421–438 years (Smith et al., 2021)				
Red sea urchin *Mesocentrotus franciscanus*	More than 100 years	Constant telomere length	Constant telomerase activity and elevated expression of DNA stability genes	Aristotle's lantern, esophagus, radial nerve bundle	RNA sequencing
Sea urchin *Stomopneustes variolaris*	3–4 years	Constant telomere length	Constant telomerase activity	Caviar	DNA sequencing
American lobster *Homarus americanus*	Up to 100 years	No data	Constantly high telomerase activity	Tissues	TRAP-PCR, DNA sequencing
Naked mole rat *Heterocephalus glaber*	37 years	Telomere length is correlated with life expectancy	Low telomerase activity	Bone marrow lymphocytes	qFISH
				Heart, liver, spleen, kidney, skin, lung, testicle	TRAP-PCR
				Skin, kidneys, lungs, cornea	TRF
Asian elephant *Elephas maximus*	60–70 years	Telomere length is not correlated with chronological age, but it is correlated with body weight	No data	Skin fibroblasts, kidney, lung or cornea	TRF
Southern lantern shark *Etmopterus granulosus*	48–57 years[3]	Telomere length is not correlated with chronological age but is correlated with body weight	No data	Muscle tissue	qRT-PCR
Greenland shark *Somniosus microcephalus*	400 years	No data	No data	Muscle tissue, liver	RNA sequencing
28 Humpback whales *Megaptera novaeangliae*	95 years[4]	Telomere length is correlated with life expectancy	No data	Skin	qRT-PCR
Baleen whales Mysticeti	75–100 years[5]	Telomere length is not correlated with chronological age but is correlated with body weight	No data	Skin	qRT-PCR
Toothed whale Odontoceti	58–62 years[6]				

Note. TRF – restriction fragment assay; PCR – polymerase chain reaction; qFISH – quantitative fluorescence *in situ* hybridization; TRAPeze® XL – telomerase activity assay.
Source of animal age data:
[1] https://animaldiversity.org/accounts/Eptesicus_fuscus/#:~:text=Lifespan%2FLongevity,die%20in%20their%20first%20winter
[2] https://animaldiversity.org/accounts/Tadarida_brasiliensis/
[3] https://fish.gov.au/docs/SharkReport/2023_FRDC_Etmopterus_granulosus%20_final.pdf
[4] https://animaldiversity.org/accounts/Megaptera_novaeangliae/
[5] https://marilimitado.com/blog/fin-whale/
[6] https://dlnr.hawaii.gov/dar/whales-and-dolphins/

A study conducted by C.G. Foote et al. examined telomere length as a potential indicator of the relationship between life history and fitness in southern giant petrels (*Macronectes giganteus*). In a cohort of adults aged between 12 and 40 years, no significant association was found between telomere length and age. However, individuals who died within an eight-year period following the measurement of telomere length exhibited significantly shorter telomeres compared to those who survived, irrespective of age or sex, which were not significant predictors of survival. These findings suggest that relatively short telomere length may serve as a biomarker for predicting life expectancy and may also indicate adult health (Foote et al., 2011). In a recent study, the telomere length in red blood cells and the markers of oxidative stress in plasma were examined in both long-lived and short-lived avian species belonging to the order Psittaciformes over a four-year duration. The findings indicated that long-lived birds, with lifespans ranging from 33 to 65 years, exhibited longer telomeres in comparison to their short-lived counterparts, which have lifespans of 10 to 17 years. However, it was noted that the long-lived birds experienced a higher rate of telomere shortening. Notably, the study established a significant correlation between the rate of telomere shortening and the levels of accumulated oxidative stress in short-lived birds. This correlation enhances the understanding of the underlying causes and dynamics associated with changes in telomere length (Domínguez-de-Barros et al., 2023, 2024).

Species exhibiting lower metabolic costs of reproduction at a young age are posited to develop more effective mechanisms for the maintenance and repair of somatic cells. This phenomenon may subsequently contribute to an increased potential lifespan and a deceleration of aging processes. In 2021, a phylogenetic meta-analysis encompassing data from 53 avian species was conducted to investigate the relationships among average telomere length in both chicks and adults, the average rate of change in telomere length throughout the lifespan, and various life history traits. The findings indicated that, irrespective of body size, longer-lived species characterized by smaller clutch sizes and slower embryonic growth rates demonstrate a reduced decline in telomere length over their lifespan (Criscuolo et al., 2021). Comparable results were observed in another study that analyzed telomere length across 19 bird species with lifespans ranging from 7 to 50 years. This study concluded that species with extended lifespans exhibited a slower decline in telomere length in comparison to those with shorter lifespans (Tricola et al., 2018).

Telomerase activity plays a crucial role in determining the rate of telomere attrition and is likely to have a direct influence on lifespan. A study conducted by M.F. Haussmann et al. examined telomerase activity in the bone marrow of two short-lived avian species: zebra finches (*Taeniopygia guttata*) and tree swallows (*Tachycineta bicolor*), which have maximum lifespans of 5 and 11 years, respectively. Additionally, two long-lived avian species were investigated: common terns (*Sterna hirundo*) and Leach's storm petrels (*Oceanodroma leucorhoa*), which can live for 27 and 36 years, respectively. The findings indicated that the short-lived species exhibited high telomerase activity in nestlings; however, this activity declined significantly in both young and older adults. In contrast, the long-lived species maintained relatively high levels of telomerase activity in their bone marrow, which did not diminish with age (Haussmann et al., 2004).

In conclusion, the inherent anti-aging mechanisms observed in avian species render them more appropriate models for the study of longevity compared to short-lived laboratory rodents. Investigations focused on birds may ultimately contribute to the identification of therapeutic interventions for diseases related to human aging.

Bats

Bats represent a distinctive subject for the investigation of aging and longevity. Similar to birds, they exhibit an atypical combination of small body size and extended lifespan relative to other mammals. For instance, individuals of the species *Myotis brandtii* can live for 40 years or more (Garg et al., 2023), whereas *M. myotis* has an average lifespan of approximately 37 years, *Rhinolophus ferrumequinum* about 30 years, and *Miniopterus schreibersii* around 22 years (Foley et al., 2018).

The existing literature on telomere length in bats presents conflicting findings. For instance, research indicates that telomeres shorten with age in species such as *R. ferrumequinum* and *M. schreibersii*, whereas this phenomenon is not observed in the genus *Myotis*, which is characterized by an extended lifespan (Gomes et al., 2011; Ineson et al., 2020). Furthermore, telomere shortening does not exhibit a correlation with age in species including *Myotis lucifugus*, *Pteropus rodricensis*, *Eptesicus fuscus*, and *Tadarida brasiliensis*. Notably, *M. lucifugus* individuals infected with the fungal disease known as white-nose syndrome (WNS) displayed significantly shorter telomeres compared to their uninfected counterparts (Ineson et al., 2020). These findings lend additional support to the hypothesis that environmental factors may exert an influence on telomere length, suggesting that telomere length is not invariably associated with the aging process.

The analysis of telomerase expression in wing fibroblasts and blood cells of *M. myotis* revealed that this enzyme is not expressed, suggesting the existence of alternative mechanisms that maintain telomere length. A more comprehensive investigation demonstrated a significant upregulation of 20 genes associated with telomere maintenance and DNA repair in *Myotis* bats when compared to other mammalian species. Notable among these genes are *Atm* and *SETX*, which have been identified as evolving under divergent selection in *Myotis*, alongside *Abl1*, *Cct4*, *Dclre1a*, *Dot1l*, *Gnl3l*, *Mlh3*, *Mre11a*, *Parp1*, *Rad50*, *Rb1*, *Rfc3*, *Rpa1*, *Sde2*, *Ssb*, *Terf2ip*, *Wrap53*, *Wrn*, and *Xrcc5*, all of which may play a role in maintaining genomic stability in bats. Furthermore, it was observed that variations in mean and minimum temperature, precipitation, and wind speed were significantly correlated with telomere length in bats (Foley et al., 2020). Consequently, telomere length may serve primarily as a biomarker for aging and longevity in specific contexts, although this does not appear to apply to the majority of bat species. Nonetheless, this model could be instrumental in examining the impact of environmental stressors on telomeres. Future research should explore whether variations in telomere length are linked to body size or survival strategies, such as hibernation or aestivation.

Mollusks

Another noteworthy model for longevity studies is the mollusk. The North Atlantic oceanic quahog (*Arctica islandica*) is the most extensively researched long-lived bivalve, possessing the longest documented lifespan of at least 507 years. These organisms exhibit a remarkable tolerance to various environmental factors, including elevated salinity, temperature, and oxygen levels. It is important to highlight that the Icelandic population of *A. islandica* demonstrates an unusually high lifespan, whereas populations in the Baltic and White Seas have a maximum lifespan ranging from 30 to 50 years (Basova et al., 2012; Gruber et al., 2015).

The investigation of telomere length and telomerase activity in the longest-lived non-colonial organism, *A. islandica*, is crucial for comprehending the mechanisms underlying telomere length maintenance and its contribution to the organism's exceptionally prolonged lifespan. An analysis of both short-lived and long-lived populations of young and old specimens (ranging from 10 to 226 years of age) revealed significant heterogeneity in telomere length among *A. islandica*. Notably, consistent telomerase activity and telomere length were observed across all age groups, with no correlation identified between these factors and population age or habitat. It is posited that stable telomere maintenance may play a role in the longevity of *A. islandica*; however, telomere dynamics alone do not account for its extraordinary lifespan (Gruber et al., 2014). Currently, the molecular mechanisms and potential mutations associated with this organism remain largely unknown, as the genome of *A. islandica* has yet to be published. Consequently, the specific factors that contribute to the longevity of this species are yet to be fully understood.

Corals

Due to their extended lifespans, corals represent a compelling yet underexplored model for investigating telomere responses to aging and environmental stressors. Recent radiocarbon dating of the deep-sea proteinaceous corals *Gerardia* sp. and *Leiopathes* sp. has revealed that these species exhibit radial growth rates ranging from 4 to 35 μm per year, with individual colonies possessing lifespans that extend to thousands of years. Notably, the oldest recorded individuals of *Gerardia* sp. and *Leiopathes* sp. were found to be 2,742 and 4,265 years old, respectively (Roark et al., 2009).

Research involving corals predominantly investigates the function of telomeres in response to stress. For instance, A. Rouan et al. (2022) analyzed telomere alterations in the symbiotic coral species *Stylophora pistillata*, which was subjected to a continuous dark environment for a duration of six months. This stressful condition led to a significant loss of symbiotic organisms. The study revealed that prolonged darkness was correlated with a reduction in telomere DNA length and a decrease in the expression of the Pot2 protein. In mammals, Pot2 forms a heterodimeric complex with Tpp1 and is essential for the recruitment of telomerase to telomeres. However, the authors did not establish connections between their findings and the concepts of aging or accelerated aging (Rouan et al., 2022). Additionally, another investigation explored the feasibility of using telomere length as a biomarker for estimating the age of colonial corals, specifically examining *Acropora digitifera* at three developmental stages: sperm, planula larvae, and polyps of adult colonies (Tsuta et al., 2014). The findings indicated that telomere length diminishes throughout coral development, with the highest values observed in spermatozoa, and the lowest in the polyps of adult colonies. It has been established that telomere length is affected not only by the chronological age of polyps but also by the rate of cell proliferation. Consequently, it can be inferred that attempts to accurately determine the age of corals based on telomere length may yield inconclusive results (Tsuta et al., 2014).

The long-term and short-term water temperature regimes are critical factors that influence intercolony variations in the Pacific Ocean. In this regard, telomere length was examined in short-lived, more stress-sensitive colonies of *Pocillopora* spp. and in long-lived, more stress-tolerant colonies of *Porites* spp. The findings indicated that telomere DNA length in short-lived colonies was significantly affected by seasonal temperature fluctuations. Conversely, telomere DNA length in long-lived colonies was not influenced by seasonal variations; rather, it was determined by historical thermal anomalies (Rouan et al., 2023).

In conclusion, the length of telomeres in corals is significantly influenced by environmental factors. The mechanisms involved in telomere maintenance are associated with the productivity of the organism, which is particularly relevant in light of the effects of climate change on overall health. Nevertheless, corals may not serve as optimal models for investigating the mechanisms of aging and lifespan, as their distinct environmental adaptations and slow metabolic rates could hinder the generalizability of findings to more complex multicellular organisms.

Sea urchins

Sea urchins serve as a compelling model organism for the investigation of developmental biology, longevity, and aging. Within this group, one can differentiate between short-lived species, such as *Lytechinus variegatus* and *L. pictus*, which have a lifespan that does not exceed four years, and long-lived species, including *Mesocentrotus franciscanus*, *Strongylocentrotus purpuratus*, *Echinometra mathaei*, and *Stomopneustes variolaris*, which can live for over 100 years and are recognized as some of the longest-lived organisms. The genomes of several species within this group have been sequenced and published, providing valuable resources for the study of genomic characteristics in relation to longevity (Sea Urchin Genome Sequencing Consortium et al., 2006; Sergiev et al., 2016; Polinski et al., 2024). Due to their capacity for indeterminate growth, sustained reproductive activity, and the absence of increased mortality with age, these organisms represent an ideal model for examining the phenomenon of negative aging (Ebert, 2019).

Research conducted within the established frameworks of aging theories, such as alterations in telomere length, has demonstrated that both short-lived and long-lived species of sea urchins exhibit minimal signs of aging. These marine organisms maintain consistent telomere length and telomerase activity. Furthermore, they preserve antioxidant and proteasome enzyme activities, and there is a negligible accumulation of oxidative cellular damage as they age. Their regenerative potential remains robust throughout their lifespan, irrespec-

tive of its length (Francis et al., 2006; Du et al., 2013). To investigate the mechanisms associated with longevity and aging in this model organism, a comparative analysis of gene expression in the radial nerve cord of *M. franciscanus* at various ages was conducted. This analysis identified over 4,000 differentially expressed genes that encompass a broad spectrum of cellular functions and molecular pathways, including neural function, metabolism, and the maintenance of DNA stability. Additionally, two genes, the expression levels of which increase with age, are implicated in the preservation of telomere length (Polinski et al., 2020).

A comparative study examined the amino acid sequence of the telomere-binding protein Pot1, which is crucial for the maintenance of telomere length through the regulation of telomerase-mediated elongation (Aramburu et al., 2020). In cell culture experiments, it was observed that mutations in the coding gene result in various telomere phenotypes, and the absence of this protein contributes to cellular aging (Zade, Khattar, 2023). Notably, the amino acid at position 198 of Pot1 exhibits variability across different species. In long-lived organisms, such as the red sea urchin and the bat (*M. brandtii*), this position is occupied by valine, whereas in short-lived species of sea urchins and bats, it is represented by threonine and serine, respectively. Interestingly, in humans and the long-lived naked mole rat, the amino acid at this position is isoleucine (Sergiev et al., 2016).

Although telomere length does not function as a reliable marker for longevity and aging in these models, it remains of significant interest for investigating the mechanisms underlying longevity and the preservation of DNA stability.

Arthropods

Crustaceans represent a diverse group of arthropods characterized by significant variability in the size and structure of their genomes. For instance, the genus *Homarus*, which includes lobsters, is estimated to have a lifespan of up to 50 years in the wild and potentially up to 100 years in captivity. These organisms exhibit continuous growth throughout their lives, possess the ability to regenerate limbs even in advanced age, and older individuals may demonstrate greater fertility compared to their younger counterparts (Koopman et al., 2015; Bowden et al., 2020). A particularly noteworthy species within this group is the American lobster (*H. americanus*), which is distinguished by its remarkable longevity, potentially reaching 100 years. Lobsters serve as compelling subjects for research on longevity, aging, and telomerase function, as they may provide critical insights into the molecular mechanisms that underlie these exceptional longevity traits. Nevertheless, the available data on this model organism remain limited (Louzon et al., 2019).

Lobsters exhibit a slow aging process, making the analysis of telomerase activity in these organisms particularly intriguing. A study conducted in 1998 identified elevated telomerase activity across all organs of the lobster species *H. americanus*. The authors concluded that the activation of telomerase represents a conserved mechanism that aids in sustaining long-term cellular proliferative capacity and mitigating aging, not only in cellular models or during embryonic development but also in adult multicellular organisms (Klapper et al., 1998). Furthermore, the lobster genome was scrutinized for genes that contribute to DNA stability. Comparative analysis of expanded gene families in lobsters versus short-lived arthropods revealed the presence of *Fancc* and *Ddb2*, both of which play critical roles in maintaining genomic integrity (Polinski et al., 2021). Importantly, there is currently a lack of data regarding the relationship between telomere length and age in lobsters, underscoring the necessity for further investigation in this area.

Naked mole-rat

The naked mole rat (*Heterocephalus glaber*) is a unique species notable for its remarkably extended lifespan, which can exceed 38 years. Traditional markers of aging, such as reduced reproductive capacity, neurodegenerative disorders, and cancer, are observed to manifest only minimally in this species (Yang et al., 2024). A study was conducted to examine telomere length across three age categories of naked mole rats: young, adult, and old. The results indicated that telomere length increased with age when compared to the young cohort; however, the authors acknowledged a limitation due to the small sample size. These findings support the hypothesis that telomeres are preserved in these animals as they age (Leonida et al., 2020). Additionally, naked mole rats exhibit low levels of telomerase activity (Seluanov et al., 2007), and a comparative analysis has revealed a negative correlation between telomerase expression levels and rodent size (Gomes et al., 2011). To date, the precise role of telomeres in the aging process remains unclear, suggesting that naked mole rats may offer valuable insights into the relationship between telomeres and aging.

Large animals (elephants, whales, and sharks)

The lifespan of large mammals is significantly influenced by their habitat, behavioral strategies, and physiological adaptations. For instance, forest elephants (*Loxodonta cyclotis*), which typically weigh between 2,000 and 2,500 kg, have an average lifespan of approximately 50 years. In contrast, African savannah elephants (*L. africana*), which weigh between 4,000 and 7,000 kg, and Asian elephants (*Elephas maximus*), weighing between 2,500 and 5,500 kg, can live for 60 to 70 years. These interspecies variations in longevity present a valuable model for investigating the molecular mechanisms of aging, including telomere dynamics (Crawley et al., 2017; Chusyd et al., 2021).

Comparative analyses of telomere length across species with differing lifespans indicate that initial telomere length may influence longevity. For instance, Asian elephants exhibit relatively longer telomeres at a young age when compared to Chihuahuas, a small dog breed that typically weighs between 1 and 3 kg and has a considerably shorter lifespan. Despite this initial advantage in telomere length, both elephants and Chihuahuas demonstrate similar rates of telomere shortening over time. This observation implies that while initial telomere length may partially contribute to determining maximum lifespan, the process of telomere shortening with age appears to be a relatively conserved trait across species.

An important factor associated with interspecies variations in longevity is the activity of telomerase and its correlation with body size. Research indicates that the suppression of telomerase activity in the somatic cells of larger mammals may represent an evolutionary adaptation aimed at decreasing the

incidence of cancers commonly observed in organisms with greater body mass. Consequently, the restriction of telomerase activity and the presence of shorter telomeres in large, long-lived species seem to function as an additional barrier against tumorigenesis (Buddhachat et al., 2017).

The lifespan of large mammals is intricately linked not only to their terrestrial counterparts but also to those inhabiting marine environments. Among cetaceans (Cetacea), which encompass the parvorders of baleen whales (Mysticeti) and toothed whales (Odontoceti), a considerable variation in longevity is observed. Numerous whale species, characterized by substantial body mass – ranging from several tons in dolphins to tens and even hundreds of tons in the largest species – exhibit remarkable lifespans. For instance, the bowhead whale (*Balaena mysticetus*), belonging to the family Balaenidae, is documented to live for over 200 years, thereby earning its status as one of the longest-lived mammals on Earth (Buddhachat et al., 2021; Lagunas-Rangel, 2021).

In contrast to several terrestrial species, for which a correlation between initial telomere length and potential lifespan has been established, the data concerning cetaceans remain more ambiguous. A study using quantitative PCR on skin samples from 28 humpback whales (*Megaptera novaeangliae*) aged 0 to 26 years revealed a statistically significant correlation between telomere length and age. However, the considerable variability observed among individuals of the same age suggests that telomere length cannot be regarded as a reliable indicator for determining age in free-swimming whales. This variability may be attributed to both methodological factors, such as measurement accuracy, and biological influences, including hereditary traits, adaptations to unpredictable environmental conditions, and stochastic processes related to resource allocation (Olsen et al., 2014).

Additional comparative studies encompassing 23 marine mammal species, which include four Mysticeti and 19 Odontoceti species, revealed no significant association between relative telomere length and maximum lifespan. Statistical analyses indicated that longevity exhibited a stronger correlation with body size; specifically, adult mass and length were identified as robust predictors, whereas relative telomere length did not demonstrate a significant association (Buddhachat et al., 2021).

Another marine species notable for its exceptional lifespan is the shark; however, determining their age can be particularly challenging. For example, the Greenland shark (*Somniosus microcephalus*) has recently been identified as the longest-living vertebrate on Earth. Despite this recognition, many aspects of its biology, physiology, and ecology remain insufficiently understood. This species can live for nearly 400 years and reaches sexual maturity at approximately 150 years of age, with weights ranging from 700 to 1,000 kilograms (Nielsen et al., 2016). Recent studies involving samples from Greenland sharks have enabled researchers to analyze RNA and identify a highly expressed long interspersed nuclear element-like (LINE-like) transcript (Bartas et al., 2023). It has been suggested that this transcript may be linked to an increased lifespan and enhanced resistance to age-related diseases. The authors of the study hypothesize that this factor could contribute to improved telomere maintenance. However, there is currently no scientific evidence to support this hypothesis, nor are there any data regarding telomere length in these animals (Bartas et al., 2023).

A separate study focusing on other shark species, such as the southern lantern shark (*Etmopterus granulosus*), examined the relative telomere length in relation to age, which in this case was assessed based on body mass (Nehmens et al., 2021). The findings indicated that telomeres in this species do shorten in accordance with size; however, it remains uncertain whether age directly influences telomere length in this context (Nehmens et al., 2021).

In conclusion, although the current body of research is insufficient to reach definitive conclusions, it can be suggested that large body size is a significant factor for the organisms discussed. This phenomenon may reflect a combination of evolutionary and physiological strategies that are aimed at maintaining genomic stability, regulating cell division, and reducing the risk of cancer.

Conclusion

In this review, we summarize the distinctive characteristics of various alternative animal models that exhibit delayed and accelerated aging phenotypes. The species examined are notable for their exceptional longevity, significant regenerative capabilities, or resistance to age-related diseases. However, the findings from studies investigating the relationship between telomere length and age-related diseases, as well as lifespan, remain inconsistent and contradictory. In this context, birds and naked mole rats appear to be the most appropriate models for studying the mechanisms of telomere shortening in relation to aging and longevity, while bats and corals are more suitable for analyzing the effects of stressors on telomere length. Large, long-lived animals such as elephants, whales, and sharks demonstrate a correlation between telomere length and body mass. Sea urchins and lobsters are particularly intriguing for exploring alternative aging mechanisms that have yet to be identified (Fig. 2).

It is evident that all the aforementioned models possess unique mechanisms for life extension and telomere length maintenance that lack analogs in traditional model organisms. Nevertheless, the extent to which insights gained from long-lived model organisms regarding aging and longevity can be applied to humans to promote a longer and healthier life remains uncertain. Therefore, optimizing the application of these models in applied research is crucial. Species exhibiting constant telomere length and active telomerase, such as *Arctica islandica*, *Mesocentrotus franciscanus*, and *Homarus americanus*, are particularly well-suited for in-depth investigations into the molecular mechanisms underlying extreme longevity and the identification of potential geroprotective targets. Organisms, in which telomere length varies significantly due to external factors, such as bats of the genus *Myotis* and corals of the genus *Pocillopora*, are appropriate for assessing the impact of environmental stressors, diseases, or habitat quality on aging processes. A combined study of sea urchins (e. g., *Strongylocentrotus variolaris* and *Mesocentrotus franciscanus*) or corals (e. g., *Pocillopora* spp. and *Porites* spp.) will facilitate the identification of factors influencing telomere dynamics across different lifespan scenarios. Large-scale models, including elephants, sharks, and whales, exhibit a correlation between telomere length, body mass, and lifespan, which may

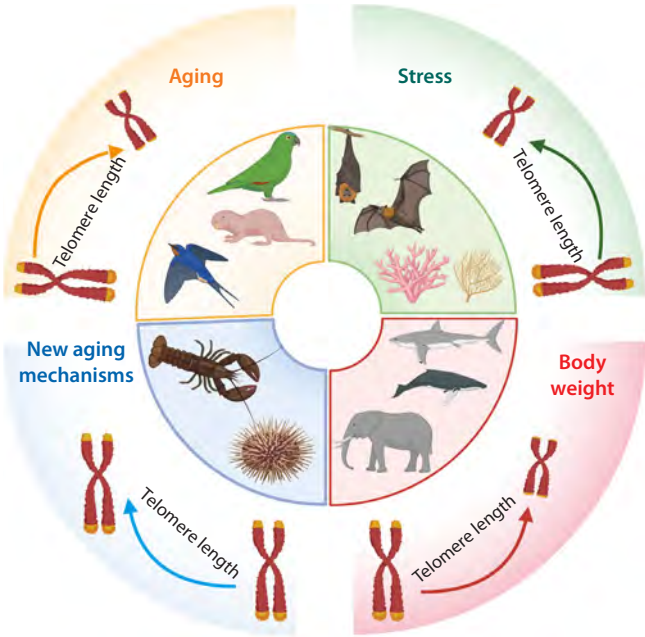


Fig. 2. Factors affecting telomere length in various long-lived organisms.

be beneficial for developing biomarkers of population health in natural settings. Data obtained from species with notable resistance to age-related pathologies, such as naked mole rats and *Myotis* bats, can be extrapolated to identify new therapeutic targets for age-related diseases in humans.

Thus, the use of alternative species with diverse aging strategies and unique telomere maintenance mechanisms not only broadens the scope for testing the hypothesis of telomeres as a universal biomarker of aging but also enables a more objective evaluation of their contributions to longevity and age-related pathologies. Incorporating these extreme models alongside classical organisms enhances our understanding of the fundamental mechanisms of longevity and opens new strategies for applying the knowledge acquired in medicine, ecology, and the conservation of endangered species, potentially aiding in the extension of the active and healthy lifespan of humans.

References

Allsopp R.C., Chang E., Kashefi-Aazam M., Rogaev E.I., Piatyszek M.A., Shay J.W., Harley C.B. Telomere shortening is associated with cell division *in vitro* and *in vivo*. *Exp Cell Res*. 1995; 220(1):194-200. doi 10.1006/excr.1995.1306

Arai Y., Martin-Ruiz C.M., Takayama M., Abe Y., Takebayashi T., Koyasu S., Suematsu M., Hirose N., von Zglinicki T. Inflammation, but not telomere length, predicts successful ageing at extreme old age: a longitudinal study of semi-supercentenarians. *EBioMedicine*. 2015;2(10):1549-1558. doi 10.1016/j.ebiom.2015.07.029

Aramburu T., Plucinsky S., Skordalakes E. POT1-TPP1 telomere length regulation and disease. *Comput Struct Biotechnol J*. 2020;18:1939-1946. doi 10.1016/j.csbj.2020.06.040

Armstrong E., Boonekamp J. Does oxidative stress shorten telomeres *in vivo*? A meta-analysis. *Ageing Res Rev*. 2023;85:101854. doi 10.1016/j.arr.2023.101854

Atzmon G., Cho M., Cawthon R.M., Budagov T., Katz M., Yang X., Siegel G., Bergman A., Huffman D.M., Schechter C.B., Wright W.E., Shay J.W., Barzilai N., Govindaraju D.R., Suh Y. Genetic variation in human telomerase is associated with telomere length in Ashkenazi centenarians. *Proc Natl Acad Sci USA*. 2010;107(1):1710-1717. doi 10.1073/pnas.0906191106

Bartas M., Červeň J., Valková N., Volná A., Dobrovolná M., Šislerová L., Baldvinsson H., Pečinka P., Brázda V. RNA analysis of the longest living vertebrate Greenland shark revealed an abundance of LINE-like elements in its transcriptome. *Czech Polar Rep*. 2023;13(2):210-227. doi 10.5817/CPR2023-2-17

Basova L., Begum S., Strahl J., Sukhotin A., Brey T., Philipp E., Abele D. Age-dependent patterns of antioxidants in Arctica islandica from six regionally separate populations with different lifespans. *Aquat Biol*. 2012;14(2):141-152. doi 10.3354/ab00387

Bernardes De Jesus B., Vera E., Schneeberger K., Tejera A.M., Ayuso E., Bosch F., Blasco M.A. Telomerase gene therapy in adult and old mice delays aging and increases longevity without increasing cancer. *EMBO Mol Med*. 2012;4(8):691-704. doi 10.1002/emmm.201200245

Bowden T.J., Kraev I., Lange S. Extracellular vesicles and post-translational protein deimination signatures in haemolymph of the American lobster (*Homarus americanus*). *Fish Shellfish Immunol*. 2020; 106:79-102. doi 10.1016/j.fsi.2020.06.053

Buddhachat K., Kriangwanich W., Kumoun I., Brown J.L., Chailangkarn S., Somgird C., Thitaram C., Prasitwattanaseree S., Nganvongpanit K. Telomeric attrition with increasing age in short- (Chihuahua dog) and long- (Asian elephant) life span animals. *Kafkas Univ Vet Fak Derg*. 2017;23(4):643-649. doi 10.9775/kvfd.2017.17504

Buddhachat K., Brown J.L., Kaewkool M., Poommouang A., Kaewmong P., Kittiwattanawong K., Nganvongpanit K. Life expectancy in marine mammals is unrelated to telomere length but is associated with body size. *Front Genet*. 2021;12:737860. doi 10.3389/fgene.2021.737860

Bythell J.C., Brown B.E., Kirkwood T.B.L. Do reef corals age? *Biol Rev Camb Philos Soc*. 2018;93(2):1192-1202. doi 10.1111/brv.12391

Calado R.T., Dumitriu B. Telomere dynamics in mice and humans. *Semin Hematol*. 2013;50(2):165-174. doi 10.1053/j.seminhematol.2013.03.030

Chusyd D.E., Ackermans N.L., Austad S.N., Hof P.R., Mielke M.M., Sherwood C.C., Allison D.B. Aging: what we can learn from elephants. *Front Aging*. 2021;2:726714. doi 10.3389/fragi.2021.726714

Crawley J.A.H., Mumby H.S., Chapman S.N., Lahdenperä M., Mar K.U., Htut W., Thura Soe A., Aung H.H., Lummaa V. Is bigger better? The relationship between size and reproduction in female Asian elephants. *J Evol Biol*. 2017;30(10):1836-1845. doi 10.1111/jeb.13143

Criscuolo F., Dobson F.S., Schull Q. The influence of phylogeny and life history on telomere lengths and telomere rate of change among bird species: a meta-analysis. *Ecol Evol*. 2021;11(19):12908-12922. doi 10.1002/ece3.7931

Domínguez-de-Barros A., Sifaoui I., Borecka Z., Dorta-Guerra R., Lorenzo-Morales J., Castro-Fuentes R., Córdoba-Lanús E. An approach to the effects of longevity, sexual maturity, and reproduction on telomere length and oxidative stress in different Psittacidae species. *Front Genet*. 2023;14:1156730. doi 10.3389/fgene.2023.1156730

Domínguez-de-Barros A., Sifaoui I., Dorta-Guerra R., Lorenzo-Morales J., Castro-Fuentes R., Córdoba-Lanús E. Telomere- and oxidative stress dynamics in Psittacidae species with different longevity trajectories. *GeroScience*. 2024. doi 10.1007/s11357-024-01397-5

Du C., Anderson A., Lortie M., Parsons R., Bodnar A. Oxidative damage and cellular defense mechanisms in sea urchin models of aging. *Free Radic Biol Med*. 2013;63:254-263. doi 10.1016/j.freeradbiomed.2013.05.023

Ebert T.A. Negative senescence in sea urchins. *Exp Gerontol*. 2019; 122:92-98. doi 10.1016/j.exger.2019.04.018

Foley N.M., Hughes G.M., Huang Z., Clarke M., Jebb D., Whelan C.V., Petit E.J., … Kerth G., Rebelo H., Rodrigues L., Puechmaille S.J., Teeling E.C. Growing old, yet staying young: the role of telomeres in bats' exceptional longevity. *Sci Adv*. 2018;4(2):eaao0926. doi 10.1126/sciadv.aao0926

Foley N.M., Petit E.J., Brazier T., Finarelli J.A., Hughes G.M., Touzalin F., Puechmaille S.J., Teeling E.C. Drivers of longitudinal telomere dynamics in a long-lived bat species, *Myotis myotis*. *Mol Ecol*. 2020;29(16):2963-2977. doi 10.1111/mec.15395

Foote C.G., Daunt F., González-Solís J., Nasir L., Phillips R.A., Monaghan P. Individual state and survival prospects: age, sex, and telomere length in a long-lived seabird. *Behav Ecol*. 2011;22(1):156-161. doi 10.1093/beheco/arq178

Francis N., Gregg T., Owen R., Ebert T., Bodnar A. Lack of age-associated telomere shortening in long- and short-lived species of sea urchins. *FEBS Lett*. 2006;580(19):4713-4717. doi 10.1016/j.febslet.2006.07.049

Garg K.M., Lamba V., Sanyal A., Dovih P., Chattopadhyay B. Next generation sequencing revolutionizes organismal biology research in bats. *J Mol Evol*. 2023;91(4):391-404. doi 10.1007/s00239-023-10107-2

Gomes N.M.V., Ryder O.A., Houck M.L., Charter S.J., Walker W., Forsyth N.R., Austad S.N., Venditti C., Pagel M., Shay J.W., Wright W.E. Comparative biology of mammalian telomeres: hypotheses on ancestral states and the roles of telomeres in longevity determination. *Aging Cell*. 2011;10(5):761-768. doi 10.1111/j.1474-9726.2011.00718.x

Gruber H., Schaible R., Ridgway I.D., Chow T.T., Held C., Philipp E.E.R. Telomere-independent ageing in the longest-lived non-colonial animal, *Arctica islandica*. *Exp Gerontol*. 2014;51:38-45. doi 10.1016/j.exger.2013.12.014

Gruber H., Wessels W., Boynton P., Xu J., Wohlgemuth S., Leeuwenburgh C., Qi W., Austad S.N., Schaible R., Philipp E.E.R. Age-related cellular changes in the long-lived bivalve *A. islandica*. *Age*. 2015;37(5):90. doi 10.1007/s11357-015-9831-8

Harper J.M., Holmes D.J. New perspectives on avian models for studies of basic aging processes. *Biomedicines*. 2021;9(6):649. doi 10.3390/biomedicines9060649

Haussmann M.F., Winkler D.W., Huntington C.E., Nisbet I.C.T., Vleck C.M. Telomerase expression is differentially regulated in birds of differing life span. *Ann NY Acad Sci*. 2004;1019(1):186-190. doi 10.1196/annals.1297.029

Ineson K.M., O'Shea T.J., Kilpatrick C.W., Parise K.L., Foster J.T. Ambiguities in using telomere length for age determination in two North American bat species. *J Mammal*. 2020;101(4):958-969. doi 10.1093/jmammal/gyaa064

Jenner L.P., Peska V., Fulnečková J., Sýkorová E. Telomeres and their neighbors. *Genes*. 2022;13(9):1663. doi 10.3390/genes13091663

Kirkwood T.B.L. Evolution of ageing. *Nature*. 1977;270(5635):301-304. doi 10.1038/270301a0

Kirkwood T.B.L., Rose M.R. Evolution of senescence: late survival sacrificed for reproduction. *Philos Trans R Soc Lond B Biol Sci*. 1991;332(1262):15-24. doi 10.1098/rstb.1991.0028

Klapper W., Kühne K., Singh K.K., Heidorn K., Parwaresch R., Krupp G. Longevity of lobsters is linked to ubiquitous telomerase expression. *FEBS Lett*. 1998;439(1-2):143-146. doi 10.1016/S0014-5793(98)01357-X

Koliada A.K., Krasnenkov D.S., Vaiserman A.M. Telomeric aging: mitotic clock or stress indicator? *Front Genet*. 2015;6:82. doi 10.3389/fgene.2015.00082

Koopman H.N., Westgate A.J., Siders Z.A. Declining fecundity and factors affecting embryo quality in the American lobster (*Homarus americanus*) from the Bay of Fundy. *Can J Fish Aquat Sci*. 2015;72(3):352-363. doi 10.1139/cjfas-2014-0277

Lagunas-Rangel F.A. Deciphering the whale's secrets to have a long life. *Exp Gerontol*. 2021;151:111425. doi 10.1016/j.exger.2021.111425

Leonida S.R.L., Bennett N.C., Leitch A.R., Faulkes C.G. Patterns of telomere length with age in African mole-rats: new insights from quantitative fluorescence in situ hybridisation (qFISH). *PeerJ*. 2020;8:e10498. doi 10.7717/peerj.10498

Lin J., Epel E. Stress and telomere shortening: insights from cellular mechanisms. *Ageing Res Rev*. 2022;73:101507. doi 10.1016/j.arr.2021.101507

Louzon M., Coeurdassier M., Gimbert F., Pauget B., De Vaufleury A. Telomere dynamic in humans and animals: review and perspectives in environmental toxicology. *Environ Int*. 2019;131:105025. doi 10.1016/j.envint.2019.105025

Miller R.A., Harper J.M., Dysko R.C., Durkee S.J., Austad S.N. Longer life spans and delayed maturation in wild-derived mice. *Exp Biol Med* (*Maywood*). 2002;227(7):500-508. doi 10.1177/153537020222700715

Muñoz-Lorente M.A., Cano-Martin A.C., Blasco M.A. Mice with hyper-long telomeres show less metabolic aging and longer lifespans. *Nat Commun*. 2019;10(1):4723. doi 10.1038/s41467-019-12664-x

Nehmens M.C., Varney R.M., Janosik A.M., Ebert D.A. An exploratory study of telomere length in the deep-sea shark, *Etmopterus granulosus*. *Front Mar Sci*. 2021;8:642872. doi 10.3389/fmars.2021.642872

Nguyen T.H.D. Structural biology of human telomerase: progress and prospects. *Biochem Soc Trans*. 2021;49(5):1927-1939. doi 10.1042/BST20200042

Nielsen J., Hedeholm R.B., Heinemeier J., Bushnell P.G., Christiansen J.S., Olsen J., Ramsey C.B., Brill R.W., Simon M., Steffensen K.F., Steffensen J.F. Eye lens radiocarbon reveals centuries of longevity in the Greenland shark (*Somniosus microcephalus*). *Science*. 2016;353(6300):702-704. doi 10.1126/science.aaf1703

Olsen M.T., Robbins J., Bérubé M., Rew M.B., Palsbøll P.J. Utility of telomere length measurements for age determination of humpback whales. *NAMMCO Sci Publ*. 2014;10. doi 10.7557/3.3194

Polinski J.M., Kron N., Smith D.R., Bodnar A.G. Unique age-related transcriptional signature in the nervous system of the long-lived red sea urchin *Mesocentrotus franciscanus*. *Sci Rep*. 2020;10(1):9182. doi 10.1038/s41598-020-66052-3

Polinski J.M., Zimin A.V., Clark K.F., Kohn A.B., Sadowski N., Timp W., Ptitsyn A., Khanna P., Romanova D.Y., Williams P., Greenwood S.J., Moroz L.L., Walt D.R., Bodnar A.G. The American lobster genome reveals insights on longevity, neural, and immune adaptations. *Sci Adv*. 2021;7(26):eabe8290. doi 10.1126/sciadv.abe8290

Polinski J.M., Castellano K.R., Buckley K.M., Bodnar A.G. Genomic signatures of exceptional longevity and negligible aging in the long-lived red sea urchin. *Cell Rep*. 2024;43(4):114021. doi 10.1016/j.celrep.2024.114021

Puechmaille S.J., Frick W.F., Kunz T.H., Racey P.A., Voigt C.C., Wibbelt G., Teeling E.C. White-nose syndrome: is this emerging disease a threat to European bats? *Trends Ecol Evol*. 2011;26(11):570-576. doi 10.1016/j.tree.2011.06.013

Roark E.B., Guilderson T.P., Dunbar R.B., Fallon S.J., Mucciarone D.A. Extreme longevity in proteinaceous deep-sea corals. *Proc Natl Acad Sci USA*. 2009;106(13):5204-5208. doi 10.1073/pnas.0810875106

Rossiello F., Jurk D., Passos J.F., d'Adda Di Fagagna F. Telomere dysfunction in ageing and age-related diseases. *Nat Cell Biol*. 2022;24(2):135-147. doi 10.1038/s41556-022-00842-x

Rouan A., Pousse M., Tambutté E., Djerbi N., Zozaya W., Capasso L., Zoccola D., Tambutté S., Gilson E. Telomere dysfunction is associated with dark-induced bleaching in the reef coral *Stylophora pistillata*. *Mol Ecol*. 2022;31(23):6087-6099. doi 10.1111/mec.16199

Rouan A., Pousse M., Djerbi N., Porro B., Bourdin G., Carradec Q., Hume B.C., … Furla P., Voolstra C.R., Forcioli D., Lombard F., Gilson E. Telomere DNA length regulation is influenced by seasonal temperature differences in short-lived but not in long-lived reef-building corals. *Nat Commun*. 2023;14(1):3038. doi 10.1038/s41467-023-38499-1

Sahm A., Bens M., Henning Y., Vole C., Groth M., Schwab M., Hoffmann S., Platzer M., Szafranski K., Dammann P. Higher gene expression stability during aging in long-lived giant mole-rats than in short-lived rats. *Aging*. 2018;10(12):3938-3956. doi 10.18632/aging.101683

Sea Urchin Genome Sequencing Consortium; Sodergren E., Weinstock G.M., Davidson E.H., Cameron R.A., Gibbs R.A., Angerer R.C., … Okwuonu G., Parker D., Pu L.-L., Thorn R., Wright R. The genome of the sea urchin *Strongylocentrotus purpuratus*. *Science*. 2006;314(5801):941-952. doi 10.1126/science.1133609

Seluanov A., Chen Z., Hine C., Sasahara T.H.C., Ribeiro A.A.C.M., Catania K.C., Presgraves D.C., Gorbunova V. Telomerase activity coevolves with body mass, not lifespan. *Aging Cell*. 2007;6(1):45-52. doi 10.1111/j.1474-9726.2006.00262.x

Sergiev P.V., Artemov A.A., Prokhortchouk E.B., Dontsova O.A., Berezkin G.V. Genomes of *Strongylocentrotus franciscanus* and *Lytechinus variegatus*: are there any genomic explanations for the two order of magnitude difference in the lifespan of sea urchins? *Aging*. 2016;8(2):260-271. doi 10.18632/aging.100889

Smith A., Cook N., Cook K., Brown R., Woodgett R., Veron J., Saylor V. Field measurements of a massive *Porites* coral at Goolboodi (Orpheus Island), Great Barrier Reef. *Sci Rep*. 2021;11(1):15334. doi 10.1038/s41598-021-94818-w

Smoom R., May C.L., Ortiz V., Tigue M., Kolev H.M., Rowe M., Reizel Y., Morgan A., Egyes N., Lichtental D., Skordalakes E., Kaestner K.H., Tzfati Y. Telomouse – a mouse model with human-length telomeres generated by a single amino acid change in RTEL1. *Nat Commun*. 2023;14(1):6708. doi 10.1038/s41467-023-42534-6

Terry D.F., Nolan V.G., Andersen S.L., Perls T.T., Cawthon R. Association of longer telomeres with better health in centenarians. *J Gerontol A Biol Sci Med Sci*. 2008;63(8):809-812. doi 10.1093/gerona/63.8.809

Tricola G.M., Simons M.J.P., Atema E., Boughton R.K., Brown J.L., Dearborn D.C., Divoky G., … Wheelwright N.T., Winkler D.W., Young R., Vleck C.M., Haussmann M.F. The rate of telomere loss is related to maximum lifespan in birds. *Phil Trans R Soc B*. 2018;373(1741):20160445. doi 10.1098/rstb.2016.0445

Tsuta H., Shinzato C., Satoh N., Hidaka M. Telomere shortening in the colonial coral *Acropora digitifera* during development. *Zoolog Sci*. 2014;31(3):129-134. doi 10.2108/zsj.31.129

Vera E., Bernardes de Jesus B., Foronda M., Flores J.M., Blasco M.A. The rate of increase of short telomeres predicts longevity in mammals. *Cell Rep*. 2012;2(4):732-737. doi 10.1016/j.celrep.2012.08.023

Yang W., Hu Y., Cui S. Fighting with aging: the secret for keeping health and longevity of naked mole rats. *Aging Dis*. 2024;16(1):137-145. doi 10.14336/AD.2024.0109

Young A.J. The role of telomeres in the mechanisms and evolution of life-history trade-offs and ageing. *Phil Trans R Soc B*. 2018;373(1741):20160452. doi 10.1098/rstb.2016.0452

Zade N.H., Khattar E. POT1 mutations cause differential effects on telomere length leading to opposing disease phenotypes. *J Cell Physiol*. 2023;238(6):1237-1255. doi 10.1002/jcp.31034

Zaug A.J., Goodrich K.J., Song J.J., Sullivan A.E., Cech T.R. Reconstitution of a telomeric replicon organized by CST. *Nature*. 2022;608(7924):819-825. doi 10.1038/s41586-022-04930-8

Conflict of interest. The authors declare no conflict of interest.
Received November 19, 2024. Revised December 26, 2024. Accepted December 28, 2024.

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2025;29(4):508-516

doi 10.18699/vjgb-25-54

Genetic mapping of loci affecting embryogenic callus formation and *in vitro* regeneration in cereals and leguminous crops

E.K. Potokina, A.S. Sushchenko

Skolkovo Institute of Science and Technology (Skoltech), Moscow, Russia
e.potokina@skoltech.ru

Abstract. Recalcitrance is defined as the inability of plant species or individual genotypes to effectively regenerate and/or to be transformed in *in vitro* culture, and is the most significant limitation for genome editing of agricultural crops. To develop protocols for genotype-independent transformation and regeneration of cultivated plants, knowledge of the genetic factors that determine recalcitrance in various plant species under *in vitro* conditions is required. Their search by classical QTL mapping in populations segregating for callus formation efficiency, regeneration, and transformation is considered a complex and labor-intensive process due to a specific nature of the analyzed phenotypes and a strong genotype-environment relationship. The article provides an overview of the methodology, prospects, and most outstanding achievements of "forward" genetics in identifying genetic determinants of recalcitrance in the most popular and at the same time most difficult to work with *in vitro* cereal and legume crops. Examples of genetic mapping and successful cloning of genes responsible for various aspects of recalcitrance in cereals are discussed. Thus, it was found that the formation of rapidly proliferating type II embryogenic callus in maize is determined by increased expression of the *Wox2a* gene. The Koshihikari rice variety, popular in Japan, poorly regenerates *in vitro* due to impaired nitrate metabolism, since it has a low expression level of nitrite reductase (*NiR*), which converts nitrite into ammonia. Callus browning, which occurs among many plant species and leads to a decrease in regenerative capacity and even to plant death, in rice varieties (*Oryza sativa* ssp. *indica*) depends on the expression level of the *Browning of Callus1* (*BOC1*) gene, which encodes the SRO protein (Similar to RCD One), regulating the plant response to oxidative stress. Similar studies on mapping loci for somatic embryogenesis traits in soybean have revealed major QTLs explaining 45 and 26 % of phenotypic variation. Studies on genetic mapping of loci affecting the efficiency of regeneration and embryogenesis in recalcitrant plant species have obvious prospects due to the emergence of annotated reference genomes, high-throughput genotyping and high-resolution genetic maps.

Key words: plants; *in vitro*; genotype-dependent regeneration; recalcitrance; genetic control; QTLs of morphogenetic traits

For citation: Potokina E.K., Sushchenko A.S. Genetic mapping of loci affecting embryogenic callus formation and *in vitro* regeneration in cereals and leguminous crops. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov J Genet Breed*. 2025;29(4):508-516. doi 10.18699/vjgb-25-54

Acknowledgements. The work was supported by the Russian Science Foundation, grant No. 24-26-00073.

Особенности генетического картирования локусов, влияющих на образование эмбриогенного каллуса и регенерацию растений *in vitro* у зерновых и бобовых культур

Е.К. Потокина, А.С. Сущенко

Сколковский институт науки и технологий (Сколтех), Москва, Россия
e.potokina@skoltech.ru

Аннотация. Рекальцитрантность определяется как неспособность видов или отдельных генотипов растений к эффективной регенерации и/или трансформации в культуре *in vitro* и представляет собой самое существенное ограничение для геномного редактирования сельскохозяйственных культур. Для разработки протоколов генотип-независимой трансформации и регенерации культурных растений необходимы знания о генетических факторах, детерминирующих рекальцитрантность у различных видов растений в условиях *in vitro*. Поиск их путем классического картирования QTL для признаков эффективности каллусообразования, регенерации, трансформации в расщепляющихся популяциях считается сложным и трудоемким процессом из-за специфичной природы анализируемых фенотипов и сильной взаимосвязи «генотип – среда».

© Potokina E.K., Sushchenko A.S., 2025
This work is licensed under a Creative Commons Attribution 4.0 License

В статье приводится обзор методологии, перспектив и наиболее ярких достижений «прямой» генетики в идентификации генетических детерминант рекальцитрантности у самых востребованных и одновременно наиболее трудных для работы *in vitro* зерновых и бобовых культур. Приведены примеры генетического картирования и успешного клонирования генов, отвечающих за разные аспекты рекальцитрантности у злаков. Так, установлено, что формирование быстро пролиферирующего эмбриогенного каллуса II типа у кукурузы определяется повышенной экспрессией гена *Wox2a*. Популярный в Японии сорт риса Koshihikari плохо регенерирует в культуре *in vitro* из-за нарушенного метаболизма нитратов, так как отличается низким уровнем экспрессии нитритредуктазы (*NiR*), преобразующей нитрит в аммиак. Побурение каллуса, встречающееся среди многих видов растений и приводящее к снижению регенерационной способности, у сортов риса (*Oryza sativa* ssp. *indica*) зависит от уровня экспрессии гена *Browning of Callus1* (*BOC1*), который кодирует белок SRO (Similar to RCD One), регулирующий реакцию растения на окислительный стресс. Аналогичные работы по картированию локусов для признаков соматического эмбриогенеза у сои позволили обнаружить мажорные (major) QTL, объясняющие 45 и 26 % изменчивости признака. Исследования по генетическому картированию локусов, влияющих на эффективность регенерации и эмбриогенеза у рекальцитрантных видов растений, имеют очевидные перспективы в связи с появлением аннотированных референсных геномов, высокопроизводительного генотипирования и генетических карт с высоким разрешением.
Ключевые слова: растения; *in vitro*; генотип-зависимая регенерация; рекальцитрантность; генетический контроль; QTL морфогенетических признаков

Introduction

In recent years, significant achievements in the field of plant genome editing have contributed to the rising number of new varieties and clones with introduced mutations that are of interest for agricultural practice. In 2017, 20 crops were reported to be improved using CRISPR/Cas9 technology (Ricroch et al., 2017); by 2020, genome editing had been applied to 40 crops in 25 countries to improve their yield and resistance to biotic and abiotic stresses (Menz et al., 2020). However, if we classify the current status of crop genome editing projects into five sequential stages of development and implementation: (1) discovery; (2) proof of concept; (3) early development; (4) advanced development; (5) commercialization, then by 2022, most such developments were in the "early development" stage, and only rice genome editing was categorized as "advanced development" (Pixley et al., 2022). The lack of new commercial varieties improved using CRISPR/Cas9 is explained not only by legal restrictions, but also by the fact that in most cultivated plant species, only a small number of tested genotypes are capable of regular and efficient development of embryogenic and regenerative tissues under standard *in vitro* conditions (Nam et al., 1997; Salvo et al., 2018; Nivya, Shah, 2023; Nagle et al., 2024).

Recalcitrance *in vitro* is defined as the inability of plant cells, tissues and organs to respond to manipulations in tissue culture (Benson, 2000). Recalcitrance concerns not only regeneration but also transformation efficiency: sometimes successfully regenerating cells fail to be transformed using *Agrobacterium*, and vice versa, successfully transformed cells fail to regenerate. The failure of plants to effectively regenerate and/or transform represents the most significant limitation for transgenesis and genome editing in crops (Altpeter et al., 2016).

The traditional approach to overcoming plant recalcitrance *in vitro* is to work on optimizing external factors, including the composition of the basal medium, pH, lighting conditions, types of explants, etc. In most cases, the plant cell development program is changed by adding growth regulators (auxins and cytokinins) to the medium. In this case, the choice of growth regulators, their arrangement and time of exposure are usually determined empirically for each species and are often adjusted for each genotype (Altpeter et al., 2016). At the same time, research aimed at identifying the genetic and epigenetic mechanisms controlling somatic embryogenesis and callus formation has made it possible to manipulate these processes more finely using hormonal signals (Maren et al., 2022).

Significant progress in the technology of transformation of monocots and recalcitrant dicot species has been achieved by manipulating so-called "morphogenic genes" to reprogram somatic cells to initiate embryogenesis. Such morphogenic genes include, in particular, key regulators of the development and determination of meristematic cells, such as *Baby Boom* (*BBM*), *Wuschel* (*WUS*) and *Wuschel-Related Homeobox* (*WOX*) (Chen Z. et al., 2022).

The development of "reverse" genetics methods has led to the fact that today several dozen such morphogenic genes that regulate the growth and development of plants *in vitro* are known. The term "fine-tuning" has appeared in the scientific literature, meaning precise adjustment of the expression level of key morphogenic genes, ensuring successful transformation and regeneration of plants (Maren et al., 2022). For example, applying such adjustments to the expression of the *BBM* and *WUS2* morphogenic genes, it was possible to induce somatic embryogenesis and regenerate fertile transgenic plants of corn, sorghum and sugarcane from calli of an immature embryo (Lowe K. et al., 2016). In this particular study, low expression of the *WUS2* gene under the low-efficiency monocot nopaline synthase promoter (Nos:ZmWUS2) was combined with increased expression of the *BBM* gene under the "strong" maize ubiquitin promoter (ZmUbi:ZmBBM). As a result of such tuning of the expression of two morphogenic genes in

maize, it was possible to obtain transformed fertile plants from 40 % of the calli of the inbred line Pioneer PHH5G, which had previously not been transformed using bioballistics or agrobacterium. For maize, 53 potential morphogens have been described to date that affect the efficiency of regeneration and transformation, and manipulation of the most effective of them – transcription factors ZmWIND1 and ERF/AP2 – allows increasing the frequency of callus formation by 60.22–47.85 % and transformation by 16.56–37.2 %, depending on the genotype (Jiang et al., 2024).

Dozens of similar examples of successful manipulation of morphogenic genes expression for efficient transformation of agricultural crops (corn, rice, wheat, triticale, barley, sorghum, soybean, beet, rapeseed, tomato, pepper, potato, turnip, grapes) (Chen Z. et al., 2022) indicate that the development of protocols for genotype-independent transformation and regeneration of crop plants may eventually become not so much an art as a technology. However, this requires knowledge of the genetic factors influencing somatic embryogenesis, the formation of embryogenic callus and the regeneration of various types of cultivated plants *in vitro*. The search for them using the classical QTL mapping for traits of callus formation efficiency, regeneration, transformation in segregating populations is considered a complex and labor-intensive process due to the specific nature of the analyzed phenotypes and the strong genotype-environment relationship affecting the plant's responsiveness to manipulations *in vitro* (McFarland et al., 2023).

The purpose of this article is to review the methodology, prospects and most impressive achievements of forward genetics in identifying genetic determinants of recalcitrance in the most popular and at the same time most difficult to work with *in vitro* cereal and legume crops.

Mapping of loci that negatively affect the regeneration of cereal crops

The issues of low regenerative capacity of explants *in vitro* and genotype-dependent transformation of cereals have received the most attention in the literature because these crops provide the majority of calories consumed by humanity (Chen Z. et al., 2022). In many important cereals such as rice, wheat, barley and maize, embryogenic regenerating callus cultures have been limited to a few genotypes for several decades, restricting the possibilities of breeding these crops using biotechnology (Kausch et al., 2021).

A good example is the search for loci that determine the genotype-specific regenerative capacity of inbred maize lines (McFarland et al., 2023). Several maize genotypes, namely H99 (Duncan et al., 1985), B104 (Frame et al., 2011) and LH244 (Altschul et al., 1990), are capable of forming slow-growing, compact, highly heterogeneous embryogenic type I callus *in vitro*. From a biotechnological point of view, it is much preferable to work with embryogenic callus type II, which is looser, proliferates rapidly, has high embryogenic and regenerative capacity. This type II callus was identified in a single inbred line A188 more than 40 years ago (Green, Phillips, 1975). Since then, despite an active search for new maize lines suitable for manipulation *in vitro*, highly embryogenic type II callus has remained a specific attribute of a single genotype A188 and its derivatives (McFarland et al., 2023). Numerous attempts to optimize the composition of the culture medium have allowed some increase in the efficiency of callus formation and regeneration of fertile transgenic plants (Gordon-Kamm et al., 1990). However, effective regeneration of transgenic maize plants was achieved only for a few genotypes that were of little interest from an agronomic point of view (McFarland et al., 2023).

In response to the challenge from practical selection, a breeding program was initiated in the 1970s to obtain "culturable" lines of maize by crossing the unique line A188 with the inbred line B73, which was valuable from a selection point of view, but recalcitrant in *in vitro* culture (Russell, 1972). As a result of a series of recurrent backcrosses, lines were obtained with an introgressive fragment A188 on chromosome 3, which determines regeneration ability (Armstrong et al., 1992). Another decade later, through additional crosses, it was possible to obtain "culturable" lines that inherited only 15 % of their genome from A188 (Lowe B.A. et al., 2006). At this stage, it was still not possible to identify the causative gene, but molecular markers linked to it were identified. With the publication in 2009 of the reference genome of the B73 maize line, as well as the advent of high-throughput genotyping tools (Illumina 55k Maize SNP Chip), it became possible to conduct more accurate QTL mapping for the trait "ability to form embryogenic callus in *in vitro* culture", using the same material from crossing contrasting parents A188 and B73, converted into almost isogenic and double haploid lines. As a result, the desired interval on chromosome 3 was narrowed to 3,035 Gb (Salvo et al., 2018).

In 2023, following a series of additional backcrosses and with the help of the annotated reference genome of the parental line B73, 93 potential candidate genes were identified. Based on the results of their transcription analysis using the RNAseq method, three most likely candidates were identified, the increased expression of which in explants was achieved using vectors with a "strong" maize ubiquitin promoter (ZmUbi1), and this made it possible to assess the effect of the expression level of potential candidate genes on the development of embryogenic callus. As a result, the *Wox2a* gene was identified for the first time, underlying the QTL for the ability to form embryogenic callus, which was mapped in the population of offspring from crossing the inbred maize lines A188 and B73. The differences in the structural part of the *Wox2a* gene in the contrasting parental genotypes were minimal, but the promoter region coincided only by 69 %. It was concluded that the increased expression of the *Wox2a* gene could be the cause of the formation of type II embryogenic callus in the A188 line (McFarland et al., 2023).

Another example of successful cloning of a gene affecting regeneration and embryogenesis has been described in rice, for which an efficient *in vitro* system had previously been developed in model varieties such as Nipponbare (*Oryza sativa* ssp. *japonica*) and Kasalath (*O. sativa* ssp. *indica*). However, many leading rice varieties used for food production in Japan, such as the Koshihikari variety, had low regeneration capacity of mature embryo *in vitro,* which was a serious obstacle to the efficient production of transgenic plants (Nishimura et al., 2005). The recalcitrant genotype Koshihikari formed calli that invariably turned brown in tissue culture and never gave rise to green shoots. The contrasting genotype Kasalath, used for crossing with Koshihikari, on the contrary, was distinguished by the ability to form viable calli from which new shoots were successfully regenerated.

The trait "regeneration ability" (the number of shoots regenerated from the callus) was mapped using a population of 99 progeny of the BC_1F_1 generation using 262 PCR markers regularly distributed over 12 rice chromosomes. Four significant QTLs were mapped on chromosomes 1, 2, 3 and 6, and at all these loci, Kasalath alleles had a positive effect on the regeneration ability of plants. The QTL on chromosome 1 showed the greatest effect, was designated as *PSR1* (*Promoter of Shoot Regeneration 1*) and was subjected to fine map-based cloning using 3,800 recombinants of the BC_3F_2 generation. The desired chromosomal interval was narrowed to 50.8 kb, but to identify the *PSR1* gene, it was necessary to construct a BAC (Bacterial Artificial Chromosome) library from the genomic DNA of the Kasalath variety, in which the BHAL15 clone covering the desired region of the genome was identified. Next, several sequences covering possible candidate genes were subcloned from BHAL15 and used to transform calli of the recalcitrant Koshihikari genotype. One of these subclones, 12.2 kb in size, overlapping the sequence of the candidate gene *NiR* encoding ferredoxin-nitrite reductase, restored the regenerative capacity of Koshihikari calli and, on this basis, was identified as the causative gene for the trait in question.

Comparison of the *NiR* gene sequences in the Koshihikari and Kasalath varieties revealed several SNPs and InDels, especially in the promoter region of the gene. The mutations found in the structural part of the gene led to only two conservative amino acid substitutions in the encoded protein; on the other hand, the expression level of this gene in the recalcitrant Koshihikari variety was 2.5 times lower than in the Kasalath variety. It is also interesting that in the Koshihikari variety, in addition to the full-length *NiR* transcript, a transcript with a retained (third) intron was also detected. Reduced *NiR* expression in Koshihikari apparently led to a disruption of nitrate metabolism in this rice variety, since in this metabolic pathway nitrate reductase catalyzes the reduction of nitrates to nitrites, and nitrite reductase (NirBD) converts nitrite to ammonia. Disrupted nitrate metabolism probably caused the low embryogenic capacity of Koshihikari rice calli.

Another example of positional mapping of recalcitrance loci in rice concerns the callus browning effect *in vitro*, which is typical to the widespread varieties of *O. sativa* ssp. *indica* (Zhang K. et al., 2020). Callus browning occurs in many plant species and results in reduced regenerative capacity, poor *in vitro* growth, and plant death (He et al., 2009). The use of antioxidants, adsorbents, low salt concentrations and growth regulators can reduce the effects of callus browning to some extent, but there is no universal solution to this problem (Zhang K. et al., 2020).

To search for loci responsible for callus browning in rice, a population of offspring was created by crossing the YJCWR genotype of the wild species *O. rufipogon* Griff., which is relatively resistant to callus browning (donor), and the elite variety Teqing (*O. sativa* ssp. *indica*) (recipient). In the population of hybrids, the YIL25 progeny line was isolated, in which the callus browning frequency were significantly lower than in Teqing, while in the YIL25 line, introgressions from the donor parent YJCWR were found on chromosomes 2, 3, and 5.

Backcrossing of the YIL25 line with the recipient parent Teqing yielded a population of 198 BC_1F_2 lines, which were genotyped using microsatellite markers and used to map the QTL for the callus browning trait. The QTL mapped on chromosome 3 explained 14 % of the observed variability. To map this QTL at a higher resolution, a fraction of BC_1F_2 lines heterozygous at the QTL interval were selfed, resulting in 6,377 recombinants. Their genotyping using SNP markers allowed to narrow the QTL interval to 18.6 kb, in which only one coding sequence, LOC_Os03g12820, annotated with the reference genome (The Rice Genome Annotation Project Database) was identified. Comparison of the sequence of this gene in the parental lines Teqing and YIL25, which were contrast in the analyzed trait, did not reveal polymorphism in the structural part of the gene, but a deletion of 337 bp, 1 bp and three SNPs were found in the promoter region of the Teqing variety. The sequence LOC_Os03g12820 was thus identified as a candidate gene *Browning of Callus1* (*BOC1*), which affects callus browning.

A comparative analysis of the *BOC1* expression level in the parental genotypes YIL25 and Teqing showed an almost twofold difference that reached its maximum values starting from the 21st day of callus cultivation and did not appear at earlier stages. Additional experiments on protoplasts with constructs representing various variants of mutations in the *BOC1* promoter integrated into the pGreenII 0800-LUC vector made it possible to evaluate the effect of these mutations on the expression level of the luciferase reporter gene (*LUC*), and to establish that it is the insertion of 337 bp in the *BOC1* promoter in the YIL25 genotype, resistant to callus browning, that significantly increases the expression level of this gene in callus culture. It was also found that the 337 bp insertion in the *BOC1* gene promoter is a transposon (Tourist MITE), and the presence of this insertion not only reduces callus browning in rice

varieties, but also increases the transformation efficiency by 2.5 times.

At least 16 scientific publications devoted to genetic mapping of QTL for recalcitrance traits in cereals were published between 2000 and 2020, but only 6 of them reported identified candidate genes (QTGs) as final results (see review by Lardon, Geelen, 2020). The identification of QTGs has not always involved the long and thorough process of positional mapping of genetic loci with multiple crosses and obtaining thousands of recombinants. For example, in barley, the only variety that can be transformed by agrobacterium is Golden Promise (Hisano, Sato, 2016).

To identify the loci that ensure the "cultivability" of this variety, Golden Promise was crossed with the recalcitrant genotype Haruna Nijo. A total of 3,013 immature embryos were isolated from F_2 caryopses that were inoculated with *Agrobacterium tumefaciens* carrying a plasmid with a reporter gene for resistance to hygromycin *HPT* (*Hygromycin PhosphoTransferase*). Of the 3,013 inoculated explants, 293 formed calli on a selective medium, and 60 of these calli regenerated into full-sized plants (Hisano, Sato, 2016). DNA analysis of these 60 plants showed the presence of the *HPT* transgene, and the fact that these plants regenerated from callus indicated that they inherited alleles from Golden Promise at loci critical for transformation and regeneration processes that the parent Haruna Nijo did not possess. Such *TFA* (transformation amenability) loci were identified on chromosome 2 (*TFA2, TFA3*) and chromosome 3 (*TFA1*), and the confidence interval of each of these QTLs was, on average, 40 centiMorgan (cM).

The next step was to answer the question of whether the genes of the transcription factors *BBM* and *WUS2*, already known for cereals, were localized in these large chromosomal intervals. For this purpose, the sequences of these genes in maize were used to search for homologous sequences in the barley genome, and it turned out that the barley *BBM* homolog falls into the *TFA2* interval, and the *WUS2* homolog falls into the *TFA1* interval (Hisano et al., 2017). Although this study did not provide direct evidence on the effect of the *BBM* and *WUS2* morphogenic genes on barley transformation efficiency, it was shown that introgression of chromosome regions 2 and 3, where these genes are localized, from the Golden Promise variety into the desired barley genotype helps to achieve a transformation level of 15.5–23.7 %, which can be considered a high result, since for the Golden Promise variety itself, the transformation efficiency is approximately 30 % (Hisano et al., 2017).

By using segregating populations from biparental crosses, it is possible to map loci that have different alleles only in a particular parent pair. With the advent of high-throughput genotyping, it has become possible to analyze the variability of traits associated with callus formation, regeneration, and somatic embryogenesis in large samples of unrelated genotypes using association analysis (GWAS, Genome Wide Association Mapping). So, 510 rice varieties were genotyped using several thousand SNPs identified by sequencing of this global sample with Illumina HiSeq 2000 (Zhang Z. et al., 2019).

An association study between SNPs and variability of three callus formation traits was performed: callus induction rate (CIR), callus induction speed (CIS), and time of the first callus appearance (T0). The first two traits (CIR and CIS) were correlated with each other ($r^2 = 0.881$), while the correlation between T0 and the other two traits was low (–0.337 and 0.286). As a result, 88 significantly linked loci were identified: 33 loci for CIR, 31 for CIS, and 24 for T0, with the identified loci for the three traits not overlapping. Of the total 88 loci identified, 21 were detected within QTL intervals previously mapped for rice in other studies. Among others, candidate genes *CRL1*, *OsBMM1*, and *OsSET1*, which are orthologs of the *LBD17/LBD29, BBM*, and *SWN* genes in *Arabidopsis*, were proposed for callus induction frequency, where the role of these genes in callus formation has been previously demonstrated (Boutilier et al., 2002; Chanvivattana et al., 2004; Fan et al., 2012).

A similar genome-wide association study was conducted for the callus induction frequency trait on 110 rice (ssp. *indica*) accessions genotyped with 2,385,475 SNP markers (Kamolsukyeunyong et al., 2024). A unique feature of this study was that the trait was tested on three culture media: B5 (Gamborg), MS (Murashige–Skoog), and N6 (CHU). Notably, callus induction was affected by different loci on different media: for B5, such a QTL was mapped to chromosome 6, for MS, to chromosomes 2 and 6, on N6, callus induction was affected by four QTLs, two on chromosome 6, and two more QTLs on chromosomes 7 and 11. As in the previous study, the intervals of mapped QTLs did not overlap. This suggested that different genes may influence successful callus induction on different culture media. This noticeable example partly explains why QTL mapping for traits associated with callus induction and subsequent production of fertile transgenic plants is not a popular area of research today – there are too many factors that can affect the reproducibility of the results.

Another difficulty with association analysis is that GWAS allows one to identify interesting patterns related to the physiological mechanisms of the studied traits, but rarely results in the identification of causative genes. More often, genes in close proximity to a reliably associated SNP, or haplotypes in an identified region of a chromosome that differ in the manifestation of a trait, are proposed for subsequent detailed study.

Mapping of QTLs for transformation efficiency and callus formation ability in legume crops

Widespread legumes of the tribe Phaseoleae (soybean, beans, cowpea), as well as pea and guar, are recalcitrant plants for *in vitro* culture, in contrast to some other legumes like Medicago and Lotus (Nivya, Shah, 2023). In the most

popular crop, soybean, the efficiency of regeneration and transformation depends on a specific genotype and is acceptable for a few varieties, such as, for example, cv. Jack (Yang et al., 2009) or Williams, Williams79 and Williams82 varieties (Xu et al., 2022).

Two features of the behavior of legumes *in vitro* have been reported (Nivya, Shah, 2023). First, the regeneration efficiency can be quite high, but only in the absence of any transformation attempts that imply selective pressure. The reasons for this phenomenon are unknown, although optimization of transformation protocols may improve the situation (Bekalu et al., 2023). Second, most published experiments on legumes failed to demonstrate the inheritance of transgenes or edited genes in the T1 generation (Nivya, Shah, 2023). The reason for this low heritability of transgenes is most likely the chimerism of regenerants, in which the floral meristem cells giving rise to gametes remain untransformed, which ultimately also explains the low transformation efficiency.

Despite the obvious difficulties in overcoming the recalcitrance of legumes *in vitro*, studies on mapping morphogenic genes for this group of crop plants are very rare and not comparable in scale with similar studies in cereals.

For example, soybean is a popular object of reverse genetics of morphogenic genes (e. g., Chen F. et al., 2019; Hao et al., 2019), but only two studies on QTL mapping in biparental populations are known: for traits of somatic embryogenesis efficiency (Song et al., 2010) and callus induction (Yang et al., 2011). In the first study, a population of 126 recombinant inbred lines (RILs) from a cross between Peking (higher somatic embryogenesis capacity) and Keburi (low capacity) was generated. The population was genotyped with microsatellite markers and highly significant QTLs were mapped to chromosome C2(6) for the somatic embryogenesis frequency trait, explaining a very high percentage of the observed variability – 45.2 % (Satt307) and 25.97 % (Satt286). Such a significant effect may indicate the presence of so-called major QTLs in these intervals of chromosome 6 in soybean. Additional QTLs with less pronounced effects (6–7 %) were identified on chromosomes "H" and "G", which correspond to chromosomes 12 and 18 according to the current nomenclature (https://www.soybase.org/about/lgs_and_chromosomes/). The second QTL mapping study for the callus induction frequency (CIF) trait was performed on a population of RILs from a cross between Kefeng (CIF = 0.69) and Nannong (CIF = 0.86). The most significant QTL for this trait was mapped to chromosome 14 (B2) and explained 16.6 % of the observed variability (Song et al., 2010).

An example of a genome-wide association study (GWAS) for *in vitro* culture-related traits in legumes has been published for peanut (Luo et al., 2024). To identify accessions with potential for regeneration, the authors compared the genotyping results of 353 peanut accessions from 26 countries with their ability to form embryogenic callus *in vitro*. Embryos isolated from sterilized seeds were placed on MS medium with vitamins, and explants were subcultured onto fresh medium every 4 weeks. It is reported that after the sixth passage, the physiological state of the callus began to stabilize, and the number of calli was recorded in the seventh, eighth, and ninth subculture each (T7, T8, and T9, respectively).

The analyzed trait, callus formation frequency, was designed as the ratio of the number of formed calli to the initial number of explants for each passage separately. 864,179 SNPs and 71,052 InDels were used for population genotyping. The correlation coefficient between the callus formation frequency in the T7, T8 and T9 subcultures varied from 0.56 to 0.61. As a result of the GWAS, 23 significantly associated SNPs were identified for the T7 subculture, 30 SNPs for T8, and 8 SNPs for T9. An important fact is that in this study, the same interval on chromosome 13 containing several SNPs associated with the trait was identified for all three passages. This fact may indicate the presence of a major QTL on chromosome 13 in peanut. The most reliable SNP in this region of the chromosome was identified in the gene encoding peroxisomal ABC transporter 1, which affects plant growth and development processes (Baker et al., 2015).

Another SNP from the same interval introduced an amino acid substitution in the *Arahy.MIX90M* gene encoding auxin response factor 19. The confidence interval on chromosome 13 also included SNPs in close proximity to the gene encoding the MYB transcription factor. In maize, genes of this family are involved in the formation of embryogenic callus via gibberellin signaling (Ge et al., 2016).

Problems and prospects in searching for genetic determinants of recalcitrance in plants using QTL mapping

Mapping of QTLs controlling regeneration and transformation capacity is currently not a widely used research approach for overcoming *in vitro* recalcitrance in plants. The main reason is that mapped QTLs are often specific to particular experimental conditions, thus the results depend on the specific culture medium in which the explants are grown or on the specific stage of explant development at which the trait variability begins to manifest itself. Often, the identified QTLs reflect polymorphisms inherent only to a particular parental pair, and QTL mapping does not always result in the identification of a candidate gene.

Nevertheless, it is clear that the low regeneration and transformation efficiency of many crop species severely limits the potential of CRISPR-Cas technology to improve the agronomic performance of agricultural crops. Experience shows that knowledge of key genes encoding "global" transcription factors, the expression of which is capable of stimulating cell proliferation, makes it possible to solve this problem using biotechnological methods. An example of such an approach is the work of J.M. Debernardi et al. (2020), who created a construct expressing a chimeric pro-

tein combining the transcription factor Growth-Regulating Factor 4 (GRF4) of wheat and its cofactor GRF-Interacting Factor 1 (GIF1). GRF factors mediate interactions between proteins and between proteins and DNA, and *GRF* genes are highly conserved in angiosperms, gymnosperms, and mosses, indicating their fundamental importance for growth and development processes (Omidbakhshfard et al., 2015).

Expression of the chimeric GRF4–GIF1 protein in tetraploid wheat calli increased regeneration by 7.8 times and significantly reduced the time required to obtain regenerants. The same effect was observed when transforming triticale and rice calli with the same *GRF4–GIF1* construct (Debernardi et al., 2020), as well as in experiments with barley (Timonova et al., 2023). J.M. Debernardi et al. (2020) also showed that homologs of the wheat *GRF4–GIF1* genes expressed in the epicotyl of Carrizo citrus (a hybrid of *Citrus triptera* × *C. sinensis*) also increased regeneration by 4.7 times compared to explants transformed with a vector without the *GRF–GIF* insert. This shows that this approach can also be used to overcome recalcitrance in dicotyledonous species, in particular, in legumes.

In soybean, for example, 22 genes of the *GmGRFs* (*Glycine max* GRFs) family have been identified to date, localized on 14 chromosomes (Chen F. et al., 2019). Another family of transcription factors, WUSCHEL-related homeobox (WOX), is represented in soybean by 33 genes, and out of 19 soybean chromosomes, these genes are absent only on one chromosome, 16 (Hao et al., 2019). In this regard, experiments on QTL mapping of regeneration and transformation efficiency would help to find out which genes of these families of "global" transcription factors have the greatest effect on plant regeneration in culture. For example, the above-mentioned study on mapping a QTL in soybean that explains 26 % of the variability in somatic embryogenesis frequency (Song et al., 2010) indicates the presence of possible candidate genes on chromosome 6 in the region of the microsatellite marker Satt286 (physical position ~16,171,913 bp). One of the *GRF* family genes, the *GmGRF5* gene (Glyma.06G134600), is located at a physical distance of ~5 Mb from the Satt286 marker, at a position of ~11,067,587 bp. Considering that the average genetic distance between markers on the used map was 28.4 cM, linkage between the Satt286 marker and the *GmGRF5* gene can be assumed.

Genetic mapping is by no means the only way to identify morphogenetic regulators; today, multi-omics approaches are also used to search for them. For example, X. Liu et al. (2023) identified 446 key transcription factors regulating callus induction in wheat by combining three omics approaches at once: RNA-seq, ATAC-seq (Assay for Transposase-Accessible Chromatin using sequencing) and CUT&Tag (Cleavage Under Targets and Tagmentation). Based on the results of transcriptome profiling and analysis of the dynamics of epigenetic changes accompanying the process of regeneration from immature wheat embryo of Fielder variety, the authors identified two new genes, *TaDOF5.6* and *TaDOF3.4*, the overexpression of which significantly increased the induction of callus and the efficiency of transformation in the wheat varieties Fielder, JM22 and Kenong 199.

Today, based on available resources, researchers have the opportunity to choose between a multi-omics approach to find factors influencing the efficiency of regeneration and transformation of the plants they work with, and the classical method of mapping these factors in segregating populations. The latter still seems less expensive, so studies on genetic mapping of regeneration and embryogenesis efficiency QTLs in recalcitrant species have obvious prospects.

Conclusion

Annotated reference genomes available for many crop species, as well as modern genotyping and high-resolution genetic mapping capabilities, can significantly simplify the search for genes, the expression level or allelic polymorphism of which influences plant behavior *in vitro*.

References

Altpeter F., Springer N.M., Bartley L.E., Blechl A.E., Brutnell T.P., Citovsky V., Conrad L.J., Gelvin S.B., Jackson D.P., Kausch A.P., Lemaux P.G., Medford J.I., Orozco-Cárdenas M.L., Tricoli D.M., Van Eck J., Voytas D.F., Walbot V., Wang K., Zhang Z.J., Stewart C.N. Advancing crop transformation in the era of genome editing. *Plant Cell*. 2016;28(7):1510-1520. doi 10.1105/tpc.16.00196

Altschul S.F., Gish W., Miller W., Myers E.W., Lipman D.J. Basic local alignment search tool. *J Mol Biol*. 1990;215(3):403-410. doi 10.1016/S0022-2836(05)80360-2

Armstrong C.L., Romero-Severson J., Hodges T.K. Improved tissue culture response of an elite maize inbred through backcross breeding, and identification of chromosomal regions important for regeneration by RFLP analysis. *Theor Appl Genet*. 1992;84(5-6):755-762. doi 10.1007/BF00224181

Baker A., Carrier D.J., Schaedler T., Waterham H.R., van Roermund C.W., Theodoulou F.L. Peroxisomal ABC transporters: functions and mechanism. *Biochem Soc Trans*. 2015;43(5):959-965. doi 10.1042/BST20150127

Bekalu Z.E., Panting M., Bæksted Holme I., Brinch-Pedersen H. Opportunities and challenges of *in vitro* tissue culture systems in the era of crop genome editing. *Int J Mol Sci*. 2023;24(15):11920. doi 10.3390/ijms241511920

Benson E.E. Special symposium: *In vitro* plant recalcitrance: an introduction. *In Vitro Cell Dev Biol Plant*. 2000;36:141-148. doi 10.1007/s11627-000-0029-z

Boutilier K., Offringa R., Sharma V.K., Kieft H., Ouellet T., Zhang L., Hattori J., Liu C.M., van Lammeren A.A., Miki B.L., Custers J.B., van Lookeren Campagne M.M. Ectopic expression of BABY BOOM triggers a conversion from vegetative to embryonic growth. *Plant Cell*. 2002;14:1737-1749. doi 10.1105/tpc.001941

Chanvivattana Y., Bishopp A., Schubert D., Stock C., Moon Y.H., Sung Z.R., Goodrich J. Interaction of Polycomb-group proteins controlling flowering in *Arabidopsis*. *Development*. 2004;131(21):5263-5276. doi 10.1242/dev.01400

Chen F., Yang Y., Luo X., Zhou W., Dai Y., Zheng C., Liu W., Yang W., Shu K. Genome-wide identification of GRF transcription factors in soybean and expression analysis of *GmGRF* family under shade stress. *BMC Plant Biol*. 2019;19(1):269. doi 10.1186/s12870-019-1861-4

Chen Z., Debernardi J.M., Dubcovsky J., Gallavotti A. Recent advances in crop transformation technologies. *Nat Plants*. 2022;8(12): 1343-1351. doi 10.1038/s41477-022-01295-8

Debernardi J.M., Tricoli D.M., Ercoli M.F., Hayta S., Ronald P., Palatnik J.F., Dubcovsky J. A GRF–GIF chimeric protein improves the regeneration efficiency of transgenic plants. *Nature Biotechnol*. 2020;38(11):1274-1279. doi 10.1038/s41587-020-0703-0

Duncan D.R., Williams M.E., Zehr B.E., Widholm J.M. The production of callus capable of plant regeneration from immature embryos of numerous *Zea mays* genotypes. *Planta*. 1985;165(3):322-332. doi 10.1007/BF00392228

Fan M., Xu C., Xu K., Hu Y. LATERAL ORGAN BOUNDARIES DOMAIN transcription factors direct callus formation in *Arabidopsis* regeneration. *Cell Res*. 2012;22(7):1169-1180. doi 10.1038/cr.2012.63

Frame B., Main M., Schick R., Wang K. Genetic transformation using maize immature zygotic embryos. In: Thorpe T., Yeung E. (Eds) Plant Embryo Culture. Methods in Molecular Biology. Vol. 710. Humana Press, 2011;327-341. https://doi 10.1007/978-1-61737-988-8_22

Ge F., Luo X., Huang X., Zhang Y., He X., Liu M., Lin H., Peng H., Li L., Zhang Z., Pan G., Shen Y. Genome-wide analysis of transcription factors involved in maize embryonic callus formation. *Physiol Plant*. 2016;158(4):452-462. doi 10.1111/ppl.12470

Gordon-Kamm W., Spencer T.M., Mangano M.L., Adams T.R., Daines R.J., Start W.G., O'Brien J.V., Chambers S.A., Adams W.R. Jr., Willetts N.G., Rice T.B., Mackey C.J., Krueger R.W., Kausch A.P., Lemaux P.G. Transformation of maize cells and regeneration of fertile transgenic plants. *Plant Cell*. 1990;2(7):603-618. doi 10.1105/tpc.2.7.603

Green C.E., Phillips R.L. Plant regeneration from tissue cultures of maize. *Crop Sci*. 1975;15(3):417-421. doi 10.2135/cropsci1975.0011183X001500030040x

Hao Q., Zhang L., Yang Y., Shan Z., Zhou X.A. Genome-wide analysis of the WOX gene family and function exploration of GmWOX18 in soybean. *Plants*. 2019;8(7):215. doi 10.3390/plants8070215

He Y., Guo X., Lu R., Niu B., Pasapula V., Hou P., Cai F., Xu Y., Chen F. Changes in morphology and biochemical indices in browning callus derived from *Jatropha curcas* hypocotyls. *Plant Cell Tiss Organ Cult*. 2009;98:11-17. doi 10.1007/s11240-009-9533-y

Hisano H., Sato K. Genomic regions responsible for amenability to *Agrobacterium*-mediated transformation in barley. *Sci Rep*. 2016; 6(1):37505. doi 10.1038/srep37505

Hisano H., Meints B., Moscou M.J., Cistue L., Echávarri B., Sato K., Hayes P.M. Selection of transformation-efficient barley genotypes based on *TFA* (transformation amenability) haplotype and higher resolution mapping of the *TFA* loci. *Plant Cell Rep*. 2017;36(4):611-620. doi 10.1007/s00299-017-2107-2

Jiang Y., Wei X., Zhu M., Zhang X., Jiang Q., Wang Z., Cao Y., An X., Wan X. Developmental regulators in promoting genetic transformation efficiency in maize and other plants. *Curr Plant Biol*. 2024; 40:100383. doi 10.1016/j.cpb.2024.100383

Kamolsukyeunyong W., Dabbhadatta Y., Jaiprasert A., Thunnom B., Poncheewin W., Wanchana S., Ruanjaichon V., Toojinda T., Burns P. Genome-wide association analysis identifies candidate loci for callus induction in rice (*Oryza sativa* L.). *Plants*. 2024;13(15):2112. doi 10.3390/plants13152112

Kausch A.P., Wang K., Kaeppler H.F., Gordon-Kamm W. Maize transformation: history, progress, and perspectives. *Mol Breed*. 2021; 41(6):38. doi 10.1007/s11032-021-01225-0

Lardon R., Geelen D. Natural variation in plant pluripotency and regeneration. *Plants*. 2020;9(10):1261. doi 10.3390/plants9101261

Lowe B.A., Way M.M., Kumpf J.M., Rout J., Warner D., Johnson R., Armstrong C.L., Spencer M.T., Chomet P.S. Marker assisted breeding for transformability in maize. *Mol Breed*. 2006;18:229-239. doi 10.1007/s11032-006-9031-4

Lowe K., Wu E., Wang N., Hoerster G., Hastings C., Cho M.J., Scelonge C., Lenderts B., Chamberlin M., Cushatt J., Wang L., Ryan L., Khan T., Chow-Yiu J., Hua W., Yu M., Banh J., Bao Z., Brink K., Igo E., Rudrappa B., Shamseer P.M., Bruce W., Newman L., Shen B., Zheng P., Bidney D., Falco C., Register J., Zhao Z.Y., Xu D., Jones T., Gordon-Kamm W. Morphogenic regulators *Baby boom* and *Wuschel* improve monocot transformation. *Plant Cell*. 2016;28(9):1998-2015. doi 10.1105/tpc.16.00124

Liu X., Bie X.M., Lin X., Li M., Wang H., Zhang X., Yang Y., Zhang C., Zhang X.S., Xiao J. Uncovering the transcriptional regulatory network involved in boosting wheat regeneration and transformation. *Nat Plants*. 2023;9(6):908-925. doi 10.1038/s41477-023-01406-z

Luo D., Shi L., Sun Z., Qi F., Liu H., Xue L., Li X., Liu H., Qu P., Zhao H., Dai X., Dong W., Zheng Z., Huang B., Fu L., Zhang X. Genome-wide association studies of embryogenic callus induction rate in peanut (*Arachis hypogaea* L.). *Genes*. 2024;15(2):160. doi 10.3390/genes15020160

Maren N.A., Duan H., Da K., Yencho G.C., Ranney T.G., Liu W. Genotype-independent plant transformation. *Hortic Res*. 2022;9: uhac047. doi 10.1093/hr/uhac047

McFarland F.L., Collier R., Walter N., Martinell B., Kaeppler S.M., Kaeppler H.F. A key to totipotency: *Wuschel-like homeobox 2a* unlocks embryogenic culture response in maize (*Zea mays* L.). *Plant Biotechnol J*. 2023;21(9):1860-1872. doi 10.1111/pbi.14098

Menz J., Modrzejewski D., Hartung F., Wilhelm R., Sprink T. Genome edited crops touch the market: a view on the global development and regulatory environment. *Front Plant Sci*. 2020;11:586027. doi 10.3389/fpls.2020.586027

Nagle M.F., Yuan J., Kaur D., Ma C., Peremyslova E., Jiang Y., Niño de Rivera A., Jawdy S., Chen J.G., Feng K., Yates T.B., Tuskan G.A., Muchero W., Fuxin L., Strauss S.H. GWAS supported by computer vision identifies large numbers of candidate regulators of *in planta* regeneration in *Populus trichocarpa*. *G3*. 2024;14(4):jkae026. doi 10.1093/g3journal/jkae026

Nam J., Matthysse A.G., Gelvin S.B. Differences in susceptibility of Arabidopsis ecotypes to crown gall disease may result from a deficiency in T-DNA integration. *Plant Cell*. 1997;9:317-333. doi 10.1105/tpc.9.3.317

Nishimura A., Ashikari M., Lin S., Takashi T., Angeles E.R., Yamamoto T., Matsuoka M., Khush G.S. Isolation of a rice regeneration quantitative trait loci gene and its application to transformation systems. *Proc Natl Acad Sci USA*. 2005;102(33):11940-11944. doi 10.1073/pnas.0504220102

Nivya V.M., Shah J.M. Recalcitrance to transformation, a hindrance for genome editing of legumes. *Front Genome Ed*. 2023;5:1247815. doi 10.3389/fgeed.2023.1247815

Omidbakhshfard M.A., Proost S., Fujikura U., Mueller-Roeber B. Growth-regulating factors (GRFs): a small transcription factor family with important functions in plant biology. *Mol Plant*. 2015;8(7): 998-1010. doi 10.1016/j.molp.2015.01.013

Pixley K.V., Falck-Zepeda J.B., Paarlberg R.L., Phillips P.W., Slamet-Loedin I.H., Dhugga K.S., Campos H., Gutterson N. Genome-edited crops for improved food security of smallholder farmers. *Nat Genet*. 2022;54(4):364-367. doi 10.1038/s41588-022-01046-7

Ricroch A., Clairand P., Harwood W. Use of CRISPR systems in plant genome editing: toward new opportunities in agriculture. *Emerg Top Life Sci*. 2017;1(2):169-182. doi 10.1042/etls20170085

Russell W.A. Registration of B70 and B73 parental lines of maize (Reg. Nos. PL16 and PL17). *Crop Sci*. 1972;12:721. doi 10.2135/cropsci1972.0011183X001200050085x

Salvo S., Cook J., Carlson A.R., Hirsch C.N., Kaeppler S.M., Kaeppler H.F. Genetic fine-mapping of a quantitative trait locus (QTL) associated with embryogenic tissue culture response and plant regeneration ability in maize (*Zea mays* L.). *Plant Genome*. 2018; 11(2):170111. doi 10.3835/plantgenome2017.12.0111

Song X., Han Y., Teng W., Sun G., Li W. Identification of QTL underlying somatic embryogenesis capacity of immature embryos in soybean (*Glycine max* (L.) Merr.). *Plant Cell Rep*. 2010;29(2):125-131. doi 10.1007/s00299-009-0804-1

Timonova E.M., Kiseleva A.A., Berezhnaia A.A., Nesterov M.A., Adonina I.G., Kochetov A.V., Salina E.A. Modification of agricultural traits in cultivated varieties of barley and wheat. *Ecol Genet*. 2023; 21:24-25. doi 10.17816/ecogen568184

Xu H., Guo Y., Qiu L., Ran Y. Progress in soybean genetic transformation over the last decade. *Front Plant Sci*. 2022;13:900318. doi 10.3389/fpls.2022.900318

Yang C., Zhao T., Yu D., Gai J. Somatic embryogenesis and plant regeneration in Chinese soybean (*Glycine max* (L.) Merr.) – impacts of mannitol, abscisic acid, and explant age. *In Vitro Cell Dev Biol Plant*. 2009;45:180-188. doi 10.1007/s11627-009-9205-y

Yang C., Zhao T., Yu D., Gai J. Mapping QTLs for tissue culture response in soybean (*Glycine max* (L.) Merr.). *Mol Cells*. 2011; 32(4):337-342. doi 10.1007/s10059-011-0063-1

Zhang K., Su J., Xu M., Zhou Z., Zhu X., Ma X., Hou J., Tan L., Zhu Z., Cai H., Liu F., Sun H., Gu P., Li C., Liang Y., Zhao W., Sun C., Fu Y. A common wild rice-derived BOC1 allele reduces callus browning in indica rice transformation. *Nat Commun*. 2020;11(1):443. doi 10.1038/s41467-019-14265-0

Zhang Z., Zhao H., Li W., Wu J., Zhou Z., Zhou F., Chen H., Lin Y. Genome-wide association study of callus induction variation to explore the callus formation mechanism of rice. *J Integr Plant Biol*. 2019;61(11):1134-1150. doi 10.1111/jipb.12759

Conflict of interest. The authors declare no conflict of interest.
Received October 31, 2024. Revised November 26, 2024. Accepted December 5, 2024.

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2025;29(4):517-529

doi 10.18699/vjgb-25-55

Optimization of technology steps for obtaining white cabbage DH-plants

A.I. Mineykina, K.S. Stebnitskaia, M.G. Fomicheva, L.L. Bondareva, A.S. Domblides, E.A. Domblides

Federal Scientific Vegetable Center, VNIISSOK, Odintsovo district, Moscow region, Russia
edomblides@mail.ru

Abstract. White cabbage is one of the economically important crops among the representatives of the genus *Brassica* L. To create highly productive F_1 hybrids with improved characteristics, the breeders need genetically diverse breeding material, which takes a long time to produce. It is possible to significantly accelerate this stage of breeding by obtaining doubled haploids (DH-plants). The lack of standardized, efficient and reproducible protocols for *in vitro* cultivation of different plant species, covering several factors and their interactions, often hinders the practical implementation of the method. Plant material, cultivation conditions and composition of nutrient media are determinants of embryogenesis efficiency. As a result of this study, the protocol for obtaining doubled haploids in *in vitro* culture of isolated microspores was optimized for late maturing white cabbage. The optimal bud size for introduction into *in vitro* culture varied from 3.5 to 5.0 mm. For the studied genotypes, the combined effect of high-temperature stress at 32 °C for 48 h and pH 5.8 stimulated the highest embryoid yield. The use of 3.5 g/L phytogel as a gelling agent was not effective. The use of flow cytometry allowed for separation of doubled haploids (69.8 %) from haploids (8.4 %), triploids (1.5 %) and tetraploids (20.3 %) at an early stage of development. Molecular genetic analysis with polymorphic microsatellite loci (SSR-analysis) confirmed the haploid origin of the diploid regenerant plants.

Key words: white cabbage; *Brassica oleracea* L.; DH-plants; *in vitro* microspore culture; androgenesis; doubled haploids; acidity of nutrient medium; shaker platform; flow cytometry; SSR-analysis

For citation: Mineykina A.I., Stebnitskaia K.S., Fomicheva M.G., Bondareva L.L., Domblides A.S., Domblides E.A. Optimization of technology steps for obtaining white cabbage DH-plants. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov J Genet Breed*. 2025;29(4):517-529. doi 10.18699/vjgb-25-55

Acknowledgements. The authors express their sincere gratitude to Cand. Sci. (Agriculture) G.A. Kostenko for providing white cabbage accessions for this study.

Оптимизация этапов технологии получения DH-растений капусты белокочанной

А.И. Минейкина, К.С. Стебницкая, М.Г. Фомичева, Л.Л. Бондарева, А.С. Домблидес, Е.А. Домблидес

Федеральный научный центр овощеводства, пос. ВНИИССОК, Одинцовский район, Московская область, Россия
edomblides@mail.ru

Аннотация. Одна из экономически важных сельскохозяйственных культур среди представителей рода *Brassica* L. – капуста белокочанная. Для создания высокопродуктивных F_1 гибридов с улучшенными показателями необходима генетически разнообразная база выровненного материала, получение которого занимает продолжительную часть селекционного процесса. Ускорить этот этап селекции можно за счет получения удвоенных гаплоидов (DH-растений). Отсутствие стандартизированных, эффективных и воспроизводимых протоколов для культивирования *in vitro* разных видов растений, охватывающих несколько факторов и их взаимодействие, часто препятствует практической реализации метода. Растительный материал, условия культивирования и состав питательных сред являются определяющими факторами эффективности эмбриогенеза. В результате проведенного исследования оптимизирован протокол получения удвоенных гаплоидов в культуре изолированных микроспор *in vitro* для капусты белокочанной позднеспелого срока созревания. Оптимальный размер бутонов позднеспелой капусты белокочанной для введения в культуру *in vitro* варьировал от 3.5 до 5.0 мм. Для изученных генотипов капусты белокочанной совместное влияние высокотемпературного стресса при 32 °С в течение 48 ч и pH питательной среды 5.8 способствовало наибольшему выходу эмбриоидов. За счет использования платформы-шейкера при режиме 40 об./мин достигнуто ускорение развития эмбриоидов до семядольной стадии на 7–10 суток. Благодаря подобранному комплексу условий для успешного эмбриогенеза из микроспор капусты белокочанной достигнут выход эмбриоидов до 273.6 ± 32.2 шт./чашку

© Mineykina A.I., Stebnitskaia K.S., Fomicheva M.G., Bondareva L.L., Domblides A.S., Domblides E.A., 2025
This work is licensed under a Creative Commons Attribution 4.0 License

Петри. Применение проточной цитометрии позволило отделить удвоенные гаплоиды (69.8 %) от гаплоидов (8.4 %), триплоидов (1.5 %) и тетраплоидов (20.3 %) на ранней стадии развития. Молекулярно-генетический анализ с использованием микросателлитных маркеров (SSR-анализ) подтвердил гаплоидное происхождение диплоидных растений-регенерантов.

Ключевые слова: капуста белокочанная; *Brassica oleracea* L.; DH-растения; культура микроспор *in vitro*; андрогенез; удвоенные гаплоиды; кислотность питательной среды; платформа-шейкер; метод проточной цитометрии клеточных ядер; SSR-анализ

Introduction

Among the variety of vegetable crops belonging to the genus *Brassica* L., the most popular among consumers is white cabbage (*Brassica oleracea* L. var. *capitata*). In the Russian Federation this crop accounts for 14.3 % of the area occupied by vegetable crops in the open field. In the recent years, import substitution is a pressing issue in the Russian Federation. The Doctrine of Food Security of the Russian Federation (the document was approved by Presidential Decree No. 20 of January 21, 2020) strategic plan includes the task of expanding export potential in the vegetable growing industry. In this regard, vegetable producers have a great need for F_1 hybrids, as they are economically advantageous for cultivation in terms of quality and resistance to adverse environmental factors.

Genetic homogeneity of parental lines is required for the development of hybrids. It can be achieved by inbreeding over several generations. Since the white cabbage is a cross-pollinated crop with a two-year development cycle, it takes 12–14 years to obtain a homogenous line by conventional breeding.

In vitro culture of isolated microspores is one of the advanced biotechnological tools for obtaining homozygous lines – doubled haploids. This technology enables to achieve homozygosity in one generation. Due to genetic homogeneity, doubled haploids can be used not only in practical breeding, but also for basic research, for instance, for genetic transformation and induced mutagenesis.

The first successful experiments on the cabbage microspore culture were carried out in the early 1980s (Lichter, 1982). Then, a basic protocol for rape microspore culture was developed, which serves as a basis of DH technology for *Brassica* plants (Pechan, Keller, 1988). At present, this technology is being actively developed, but a high level of efficiency is required for its full-fledged inclusion in the breeding process. Plant genotype, cultivation conditions and nutrient medium ingredients are determinants of the quality and quantity of plant material obtained in any *in vitro* cell culture protocol. Many strategies have been implemented worldwide to improve cell culture protocols for plants of the Brassicaceae family. Important advances have been observed in all major members of the genus *Brassica*. Due to the high responsiveness of some species, the main factors that have impact on the induction of embryogenesis in *B. napus* (Weber et al., 2005), *B. rapa* (Ferrie et al., 1995; Gu et al., 2003; Shumilina et al., 2020), *B. carinata* (Barro et al., 2003), *B. juncea* (Prem et al., 2008) have been studied. For white cabbage, such studies are sparse due to lower responsiveness in different genotypes and low embryo yield (Cao et al., 1990; Rudolf et al., 1999; Bhatia et al., 2018).

It is known that the *in vitro* culture of isolated microspores is based on the ability of microspores to switch from the gametophytic to the sporophytic developmental pathway under the influence of certain factors. For plants of the genus *Brassica*, this ability is observed in microspores at the late one-cell stage and in pollen grains at the early two-cell stage (Pechan, Keller, 1988; Kott, 1998). In more recent studies, the viability of microspores and the embryoid yield have been shown to be dependent on the bud size, which helps to reduce the selection time when a large amount of plant material is being processed (Takahata, Keller, 1991; Bhatia et al., 2018; Kozar et al., 2022). The proper selection of the bud size ensures the homogeneity of the microspore population in terms of developmental stage, which is key to the success of this technology (Cristea et al., 2020).

Other key factors affecting the induction of embryogenesis are bud cold pretreatment and microspore heat shock during the first days of cultivation (Takahata, Keller, 1991; Bhatia et al., 2018). In all tested protocols, temperature shock initiated microspore division. According to studies (Pechan, Smykal, 2001), temperature treatment at 32 °C for 1–4 days is a required condition for microspore induction in *B. napus*. Thus, the combination of cold pretreatment (4 °C) for 1 or 2 days and heat shock (32.5 °C) for 1 day significantly enhanced microspore embryogenesis in broccoli (Yuan et al., 2012), and 32.5 °C for 1 or 2 days was optimal for white cabbage (Yang et al., 2013).

The liquid nutrient NLN medium with 13 % sucrose was developed in 1982 (Lichter, 1982) and has been used in microspore cultivation protocols for multiple cultures. Studies have shown that the acidity of the medium has a significant influence on embryogenesis. pH varies between 5.6 and 6.6 for different genotypes of cabbage cultures (Yuan et al., 2012). To increase the viability of induced microspores and developing embryoids, it is also beneficial to add activated charcoal to the medium (Prem et al., 2008).

An unambiguously positive response was observed when a shaker platform at 40 to 50 revolutions per minute was used for culturing microspores on the liquid medium. This increased the formation of *Brassica rapa* L. ssp. *chinensis* embryoids by 11.6–69.37 %, as well as shortened the culturing time by 1–4 days and accelerated plant regeneration (Yang et al., 2013). It is also important to support further embryoid and plantlet development. MS solid medium (Murashige, Skoog, 1962) is most commonly used for propagation, while S. Yuan and colleagues (Yuan et al., 2012) use 1/2 MS medium with 50 % salt content for rooting (Yang et al., 2013).

The procedure for obtaining DH lines includes two main steps: induction of embryogenesis and chromosome doubling.

The second step is necessary for the practical application of the obtained regenerants, since plants are sterile before genome doubling. At present, the mechanism underlying spontaneous chromosome doubling is unclear in many cases, and its efficiency varies greatly among species and cultivars of the same species (Kasha, 2005). A study by J.C. da Silva Dias and coauthors found 43–88 % spontaneous diploidization in broccoli and 7–91 % spontaneous diploidization in other cabbage species (da Silva Dias, 2003). Since plants with different ploidy can occur among regenerants, it is necessary to analyze the ploidy level of all obtained plants.

Early studies have shown that embryogenic ability in cabbage crops is generally a quantitatively inherited trait controlled by several genes that varies among cultivars and genetic groups (Zhang et al., 2003; Kitashiba et al., 2016; Ji et al., 2023). Thus, the selected optimal factors for successful *in vitro* cultivation of microspores are suitable for a particular genotype. However, there is no universal cultivation protocol. Identification and modification of potentially interacting factors would help to improve embryoid yield, which is particularly important for low-responsive genotypes.

The aim of this study is to improve the basic protocol of isolated microspore culture for late-maturing genotypes of white cabbage by identifying optimal cultivation conditions.

Materials and methods

Material and growing conditions. Eight varieties of late-maturing white cabbage (*Brassica oleracea* var. *capitata* L.) from the collection of the LLC "Agrofirma Poisk" (No. 2403, 2404, 2405, 2406, 2407) and FSBSI FSVC (No. 127, 303, 360) were used. All genotypes had a maturity period of 160 to 180 days from sprouting and represented valuable breeding samples of different genetic origin, selected for economically valuable traits.

Donor plants were grown in a climatic chamber at 19 °C, illumination of 65 μmol · m^{-2} · s^{-1} and photoperiod of 16 h – day, 8 h – night. Plant vernalization was carried out at 6 °C in the dark for three months. At the end of vernalization, at the acclimatization stage, the growing vessels with plants were placed in a climatic chamber to obtain inflorescences. During 15–35 days, the temperature was gradually increased from +8 °C to 16±2 °C under the regime of 16 h – day, 8 h – night and illumination 65 μmol · m^{-2} · s^{-1}.

Culture of isolated microspores *in vitro*. At the first stage, buds containing the maximum number of microspores at the stages potentially capable of embryogenic development from late uninucleate to early binucleate were selected by linear size. Stages were identified by staining anthers with a differential dye (Alexander, 1969) and observing them under an Axio Imager A2 microscope (Zeiss, Germany). A microspore suspension was prepared from the selected buds so that microspores from 1 bud were placed in 1 ml of NLN medium (Lichter, 1982) with 13 % sucrose and pH 5.8; 6.0; 6.1; 6.2; 6.4, depending on the experiment. For this purpose, buds were collected at the beginning of donor plant flowering and sterilized for 30 s in 96 % ethanol. Then, they were sterilized for 15 min in a sodium hypochlorite commercial solution ("Belizna", Russia) diluted by sterile distilled water in a 1:1 ratio with the addition of Tween 20 (Panreac, Spain) (1 drop per 100 ml of solution), followed by three 7 min washes in sterile distilled water.

Sterile buds were placed in the glass flasks with the NLN medium and magnets on a magnetic stirrer (BioSan, Latvia). The resulting microspore suspension was passed through a 40 μm nylon filter and then centrifuged at 920g for 5 min using an Eppendorf 5804R centrifuge (Germany). Microspores were washed twice. After isolation and washing, microspores were placed in 60 mm Petri dishes. The sterile solution of activated charcoal and agarose (Sigma-Aldrich, USA) (1 g of activated charcoal per 100 ml of 0.5 % agarose solution) was melted in a microwave oven. Then 3–4 drops of this solution were added to each Petri dish. Petri dishes were then incubated in the dark at 32 °C shock temperature for 1, 2, 3 days, followed by 25 °C incubation in the dark at 40 rpm in a shaker incubator (New Brunswick Innova® 44/44R Eppendorf, Germany) until embryoids reached the cotyledonary stage. All experiments were performed in triplicate.

Plant regeneration. Embryoids that reached the cotyledonary stage were transferred to the solid MS medium with 2 % sucrose and 0.7 % agar, pH 5.8, supplemented with 1 mg/L 6-benzylaminopurine (6-BAP), 0.1 mg/L 1-naphthylacetic acid (NAA), or 0.1 mg/L gibberellic acid (GA). Shoots were subcultured every 4 weeks on the same medium without the addition of growth regulators. The cultivation was performed on racks under mixed illumination of two types of fluorescent lamps: OSRAM Fluora L36W/77 (predominantly blue and red spectrum) and Philips 36W/54-765 (predominantly white spectrum), at a total illumination of 24 μmol · m^{-2} · s^{-1} at 16 h day/8 h night and 24 ± 2 °C.

Adaptation of plants to *in vivo* conditions. The regenerant plants with developed leaves, stems and roots were transplanted into 8-cm-diameter pots with peat and perlite (7:3) for adaptation to *in vivo* conditions. To improve adaptation and maintain high humidity during the first week after transplanting into the soil, the regenerant plants were covered with perforated plastic transparent cups tightly adhering to the substrate. Then the cups were gradually lifted and removed. Adaptation took place in a climate room with the same parameters as for donor plants.

DNA extraction for PCR analysis from regenerant plants of white cabbage. Young leaves of each plant were ground in 200 μL of CTAB buffer using tungsten carbide beads (3 mm in diameter) and a TissueLyser II homogenizer (Qiagen, Germany) (1560 oscillations/min, duration 1.7 min) to a suspension. After grinding, 15 μL of proteinase K was added to each sample. Further DNA extraction was performed by CTAB method using Sorb-GMO-B reagent kit (Syntol, Russia) according to the manufacturer's protocol. The final purity and concentration of total DNA were determined using a spectrophotometer (Smart Spec Plus, Bio-Rad, USA). The obtained ratio OD260/280 = 1.6–1.8 corresponded to the pure DNA solution. Preparations of isolated DNA were stored in a freezer at –70 °C.

PCR analysis of white cabbage regenerant plants. Nine microsatellite loci (Table 1) with known primer sequences for amplification showing a primer PIC of at least 0.5 (Tonguç, Griffiths, 2004; Louarn et al., 2007) were used for microsatellite analysis.

Table 1. Characteristics of microsatellite loci

Marker name in GenBank	Primer sequence	Motive	Annealing temperature (Ta), °C	Fragment size/ range	Known allele number
BoDCTD1/ AF458409	AGAAAGCAGACGGGA TGGTTAAAGCGAAAGTGTGC	$(aga)_6$	55.7	166/ 133–168	5
BoPLD2/ AF113918	GACCACCGACTCCGATCTC AGACAAGCAAAATGCAAGGAA	$(ct)_7(at)_{7-1}$	56.4	267/ 258–273	11
BoAP1/ U67451	GGAGGAACGACCTTGATT GCCAAAATATACTATGCGTCT	$(at)_{9-1}$	56	138/ 136–150	6
BoCAM1/ AJ427337	GCTGATGTTGATGGTGATGG GCCGAAGCAGACAAATAAAAC	$(ga)_5$	56	206/ 187–212	3
BoABI1/ AF180355	TATCAGGGTTTCCTGGGTTG GTGAACAAGAAGAAAAGAGAGCC	$(tc)_{16}$	55	172/ 164–190	2
BoKAH45TR/ BZ523957	ATTATGACGCCTGGTTTTA ATTGGTTAGAAGTTATGGGAAC	$(ttg)_6$	50	269/ 231–272	5
BoIAB19TF/ CC969431	AAGCCACCTCACCTTAGCC GAAATCCCAGAGACTGAAAACC	$(ga)_6$	55	258/ 237–272	13
BoCALa/ AF241115 (1)	TTGTAAATGTAAACAAAGGGG CAAAATGAAACAATTCTCAGGG	$(at)_5(ta)_6$	53	204/ 202–237	6
BoCALb/ AF241115 (2)	GTAATTCCTTGATAATTGC TCTGATTGGTTTTGATGTGTCC	$(ta)_{6-1}$	46	236/ 237–242	2

Basic PCR was performed in a volume of 25 μL, including 1x PCR buffer B, 2.5 mM $MgCl_2$, 0.25 mM of each dNTP, 0.3 μM of each primer, 1.5 units of Taq DNA polymerase (Syntol, Russia) and 3 μL of DNA from each plant sample tested. The C1000 Touch instrument (BioRad, USA) was used for amplification. The basic amplification protocol consisted of a denaturation step of 2–5 min at 92–95 °C; an annealing step of 30 sec at temperatures from 52 to 58 °C; and an elongation step of 30 sec to 1 min at 72 °C. The program was designed for 35 cycles of amplification.

PCR products were separated by vertical electrophoresis using the Mini-PROTEAN Tetra Cell system (BioRad, USA) in a 6 % polyacrylamide gel. After electrophoresis, gels were stained with SYBR™ Safe DNA Gel Stain (Invitrogen, USA) according to the manufacturer's instructions and documented using the ChemiDoc XRS+ system (BioRad, USA).

The sizes of the amplified fragments were determined by comparison with Thermo Scientific GeneRuler 100 bp Plus DNA Ladder molecular mass marker (Thermo Fisher Scientific Baltics UAB, Lithuania). The obtained digital photographs of amplification products were analyzed using ImageLab 3.0 software (BioRad, USA).

Ploidy determination by flow cytometry. Young healthy leaves were chopped with a razor blade in 300 μL of Galbraith buffer (45 mM $MgCl_2$, 20 mM MOPS, 30 mM sodium citrate, 0.1 % Triton X-100, pH 7.0) on ice supplemented with 50 μg/mL RNase I (Syntol, Russia). The sample was then filtered through a 30 μm nylon filter. Then, propidium iodide (Sigma, USA) was added to a final concentration of 50 μg/mL. DNA content was determined by the fluorescence intensity of propidium iodide staining on a Beckman Coulter CytoFLEX flow cytometer with the B2-RO-V2 kit (Beckman Coulter, USA) with a 532 nm laser light source.

Histogram visualization and data processing were performed using CytExpert 2.4 software (Beckman Coulter, USA).

Donor plant diploid samples were used as an external standard to determine ploidy and DNA content. The ploidy was determined by the index of the difference between the diploid standard and the sample peaks:

$$\text{Index} = \frac{\text{the average peak of the sample}}{\text{the average peak of the standard}}.$$

DNA content (2C, pg) was calculated according to the formula:

$$2C = \frac{\text{the average peak of the sample}}{\text{the average peak of the standard}} \times 2C \text{ of standard.}$$

Statistical analysis. Statistical analysis was performed using analysis of variance (one-way, two-way ANOVA) and mean values were compared using Duncan's multiple range test (DMRT) with 95 % probability. Statistical analysis was performed using Statistica 8.0 (Statsoft, www.statsoft.com).

Results

Stages of white cabbage embryogenesis

After isolation, white cabbage microspores were subjected to high-temperature stress at 32 °C for 1–2 days, after which microspore division was observed (Fig. 1*a*, *b*). For all studied samples of late-maturing white cabbage, multicellular structures could be observed on day 15 of cultivation. Their further

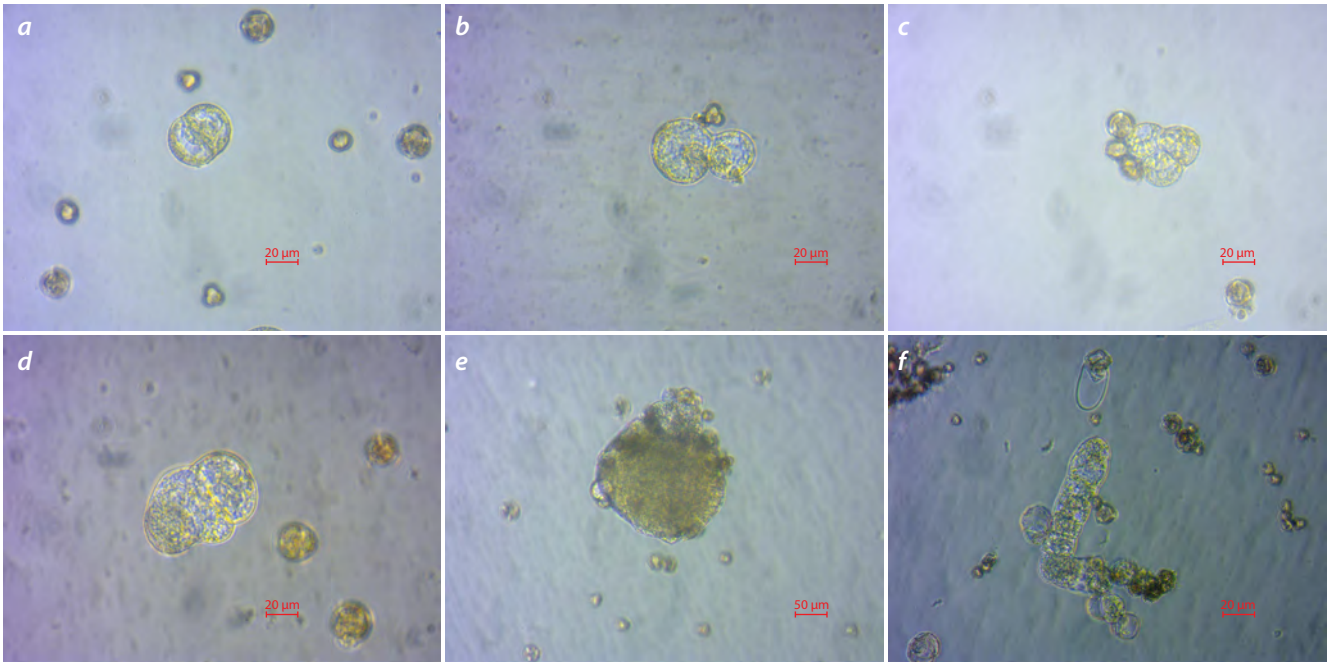


Fig. 1. Microspore division and embryoid formation in white cabbage.
a – the first microspore division (day 1); *b* – day 2; *c* – day 3; *d* – day 6; *e* – globule formation (14 days); *f* – a suspensor-like structure.

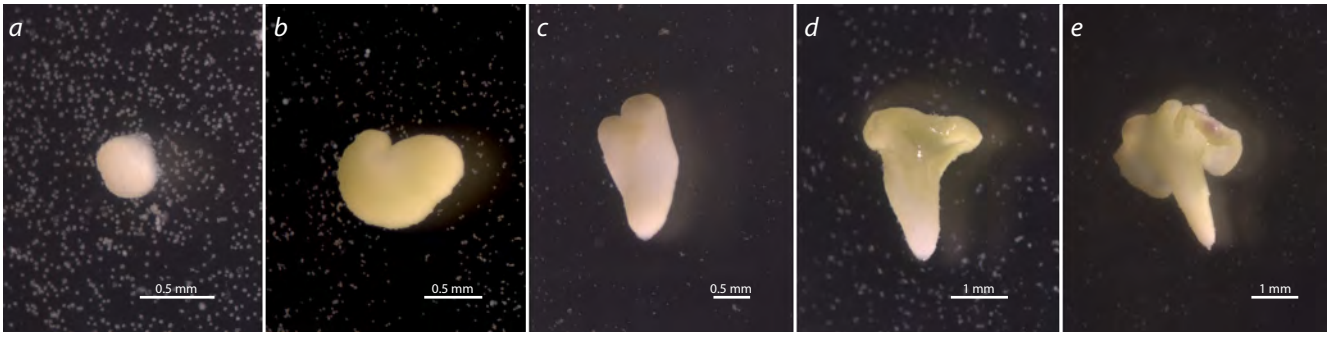


Fig. 2. Growth and development of white cabbage embryoids in the microspore culture *in vitro*.
a – a globular stage after 16 days of cultivation; *b* – heart-shaped stage after 20 days of cultivation; *c* – heart-shaped stage on day 25; *d* – an embryoid at the cotyledonary stage after 30 days; *e* – an embryoid with expanded cotyledons after 30 days of cultivation.

development could occur by different embryoid formation pathways. The most common is the formation of globules by dividing cells. Further, similar to zygotic embryoids *in vivo*, during the transition to the heart-shaped stage, radial symmetry changes to bipolar symmetry with the development of two future cotyledons (Fig. 2). Another way of embryoid formation is through suspensor-like structures. In such structures, we observed transverse division of daughter cells, resulting in the formation of a suspensor shaped as a long filament (Fig. 1*f*). The embryoid formation pathways have been investigated in more detail in other *Brassicaceae* crops in different studies (Tang et al., 2013; Kozar et al., 2021).

Along with normal embryoids, in all genotypes we also observed abnormally developed embryoids characterized by the absence of cotyledons and hypocotyl (Fig. 3*a*, *b*), as well as various twin forms (Fig. 3*c*–*f*). The percentage of abnormal embryoids ranged from 1 % to 15 % depending on the number of embryoids formed per Petri dish (embryoid yield). The percentage of abnormal embryoids increased in highly responsive genotypes.

The bud size affects embryogenesis efficiency

The predominant stage of microspore development in culture, which correlates with the bud size, was a determinant of embryogenesis efficiency in white cabbage. It was noted that it is impossible to determine a single optimal bud size for all genotypes. Before introduction into *in vitro* culture, it is necessary to determine for each genotype, which bud linear size has the maximum percentage of microspores at the optimal developmental stages for androgenesis induction,

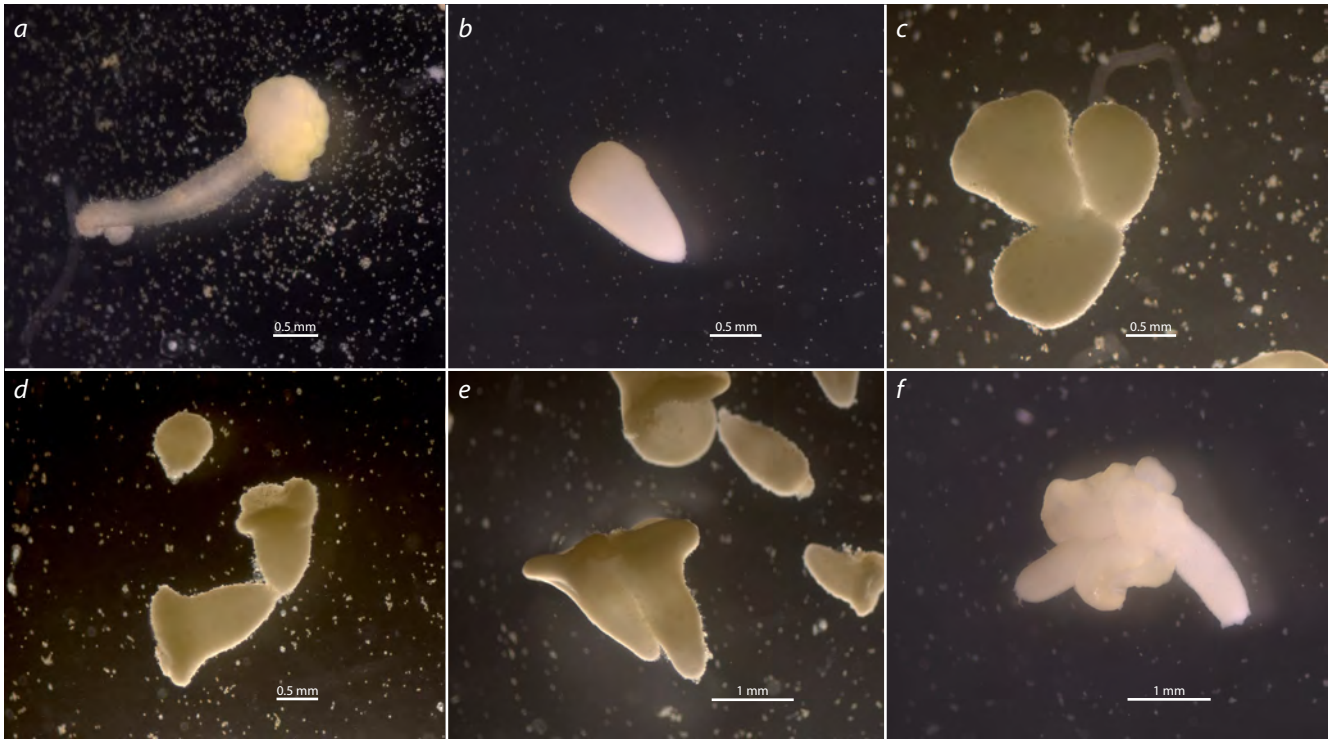


Fig. 3. Abnormal development of white cabbage embryoids in *in vitro* microspore culture.
a – an embryoid with no pronounced hypocotyl; *b* – an embryoid without formed cotyledons; *c–f* – conjoined twin embryoids.

Table 2. White cabbage embryoids yield in *in vitro* microspore culture depending on bud size

Genotype	The number of embryoids, number/Petri dish (mean ± SE)*			
	3.5–3.9	3.9–4.5	4.5–5.0	5.0–5.5
2403	0a	10.85 ± 0.69c	0.50 ± 0.29b	–
2404	16.00 ± 4.00c	2.00 ± 1.08b	0a	–
2405	–	0a	2.43 ± 1.04b	0.11 ± 0.11a
2406	–	22.25 ± 4.91b	273.56 ± 32.21c	0.25 ± 0.25a
2407	–	0a	1.33 ± 0.88b	0a
127	–	136.50 ± 17.97b	265.00 ± 5.11c	2.50 ± 0.50a
360	–	16.75 ± 0.75b	136.25 ± 3.84c	1.00 ± 0.71a
303	0a	38.50 ± 0.65c	4.00 ± 1.58b	–

Note. "–" – microspores at the optimal developmental stage were absent in the study variants.
* Here and in the Table 3: values in a row within a genotype with the same lowercase letter (a–c) are not significantly different with 95 % probability, according to Duncan's multiple range test.

since genotypes have different phenotypic features, including bud form. A round or elongated form of the bud will have a significant effect on the range of bud lengths suitable for microspores isolation.

In preliminary experiments, it was determined that embryoid induction per Petri dish was significantly reduced or inhibited, if the bud length range exceeded 1 mm in a bud sample. This was due to the fact that many microspores/pollen grains were introduced into the *in vitro* culture, which died by day 10–14 of cultivation and had a toxic effect on the culture. The male gametophyte stages optimal for embryogenesis were in bud samples varying by 0.5 mm in length.

In eight genotypes, we isolated several bud groups differing in length and studied the embryoid yield (Table 2). For all genotypes, the maximum embryoid yield was obtained only in one of the groups, which confirms the need to select buds by size with a variation of no more than 0.5 mm in the sample. Only for genotype No. 2404 the optimum bud size

was 3.5–3.9 mm. For two genotypes (No. 2403 and 303), the optimum size was 3.9–4.5 mm, and for five genotypes, (No. 2405, 2406, 2407, 127, 360) it was in the range of 4.5–5.0 mm. The highest embryoid yield was achieved for genotype No. 2406 from the 4.5–5.0 mm buds and averaged 273.56 ± 32.21 embryos/Petri dish (Table 2).

The effect of temperature treatment, pH of the medium and their combined effect on microspores embryogenesis

No embryoids were formed when microspores were cultured at a constant temperature of 25 °C in all genotypes used in the study. Short-term 32 °C shock temperature treatment proved to be a key factor in reprogramming microspores to the sporophytic path of development with the formation of embryoids for all studied white cabbage genotypes. The duration of high temperature treatment had a significant effect on embryoids yield. The maximum embryoid yield for genotype No. 2406 was achieved after 48 h treatment. One-day high temperature treatment also initiated embryogenesis, but it was insufficient for the majority of potentially embryogenic microspores in the studied genotype (Supplementary Materials, Fig. S1)[1]. For genotypes No. 127 and 360, the difference between embryoid yield under high-temperature treatment for 24 and 48 h was insignificant. When the treatment time was increased up to three days, there was a significant decrease in embryoids yield for all studied white cabbage genotypes (Fig. 4). The genotypes characterized by low responsiveness had no embryoids in this variant of the experiment (Table S2).

The use of the medium with different pH on four white cabbage genotypes showed a significant effect of medium acidity on embryoid yield. At the same time, all genotypes responded differently to the different pH (Table 3). In our experiments, we used nutrient media with the most common pH values for *in vitro* cultivation (Yuan, 2012) (5.8; 6.1; 6.4). Within each genotype, we observed a significant shift in embryoid yield relative to the pH of the medium. For genotypes No. 2403, 2405, 2407, the highest embryoid yield was achieved at pH 5.8. On the contrary, for genotype No. 2404, the medium with pH 6.1 and 6.4 increased the embryogenesis efficiency more than two times. Nutrient medium with pH 5.8 also promoted embryoid formation for this genotype but to a smaller extent.

For genotype No. 2403, an experiment on the combined effect of medium acidity and high-temperature treatment on embryogenesis was conducted. Two-factor analysis of variance showed that both the factor of nutrient medium acidity and the factor of high-temperature induction treatment, as well as the interaction of both factors have a significant effect on embryoids yield (Table S2). At the same time, the factor of temperature induction treatment is the main one (the influence share of 45 %). Apparently, this is the factor that triggers embryogenesis in the culture of isolated microspores. This is confirmed by the absence of embryoids in all three variants of the experiment with different nutrient medium acidity, if the microspores were not subjected to high-temperature stress and cultured at 25 °C. Temperature treatment at 32 °C promoted microspore reprogramming to the sporophytic developmental route and induced embryogenesis in all tested variants of nutrient medium acidity. At the same time, the highest embryoid yield was achieved after 48 h high-temperature treatment at pH 5.8. Increasing the duration of temperature treatment up to 2 days contributed to the increase of embryoid yield at pH 6.1, but in a smaller extent. However, at a pH value of 6.4, the duration of temperature treatment did not significantly affect embryoids yield (Table S2).

[1] Tables S1, S2 and Figures S1, S2 are available at: https://vavilov.elpub.ru/jour/manager/files/Suppl_Minej_Engl_29_4.pdf

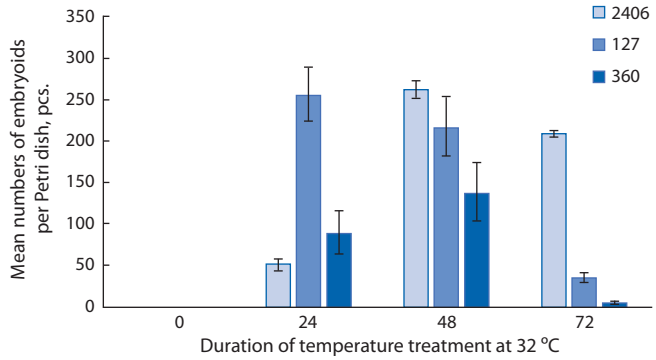


Fig. 4. The effect of the shock temperature treatment duration on embryoids yield in the highly responsive white cabbage genotypes No. 2406, 127, and 360 (from left to right: constant 25 °C, 24 h – 32 °C, 48 h – 32 °C, 72 h – 32 °C).

Table 3. The white cabbage embryoid yield in *in vitro* microspore culture depending on the nutrient medium pH

Genotype	Number of embryoids obtained at different pH, number/Petri dish (mean ± SE)		
	5.8	6.1	6.4
2403	13.0 ± 1.0a	3.5 ± 2.5b	1.0 ± 0.0c
2404	9.5 ± 2.5b	21.0 ± 8.0a	18.0 ± 1.0a
2405	4.5 ± 2.5a	2.3 ± 0.9a	0.7 ± 0.3b
2407	0.5 ± 0.3a	0.5 ± 0.3a	0b

The regeneration of DH plants from embryoids

Embryoids with cotyledons were first transplanted from liquid medium to solid MS (Fig. 5*a*), supplemented with 20 g/L sucrose, 7 g/L agar, BAP (1 mg/L), NAA (0.1 mg/L), and GA (0.1 mg/L) to induce shoot formation (Fig. 5*b*). One month later, the developing adventitious shoots were transplanted for rooting onto the solid hormone-free MS medium containing 20 g/L sucrose and 7 g/L agar or 3.5 g/L phytogel. When planted on the medium with phytogel, callus outgrowth at the base of the shoot with no root formation was observed. On nutrient media with agar, shoots quickly grew a well-developed root system (within 7–10 days) (Fig. 5*c*).

Identification of regenerant plant ploidy

The regenerant plant ploidy was determined at the seedling stage (5–8 leaves) by flow cytometry of isolated nuclei (Fig. S2). The analysis of 163 regenerant plants of the six

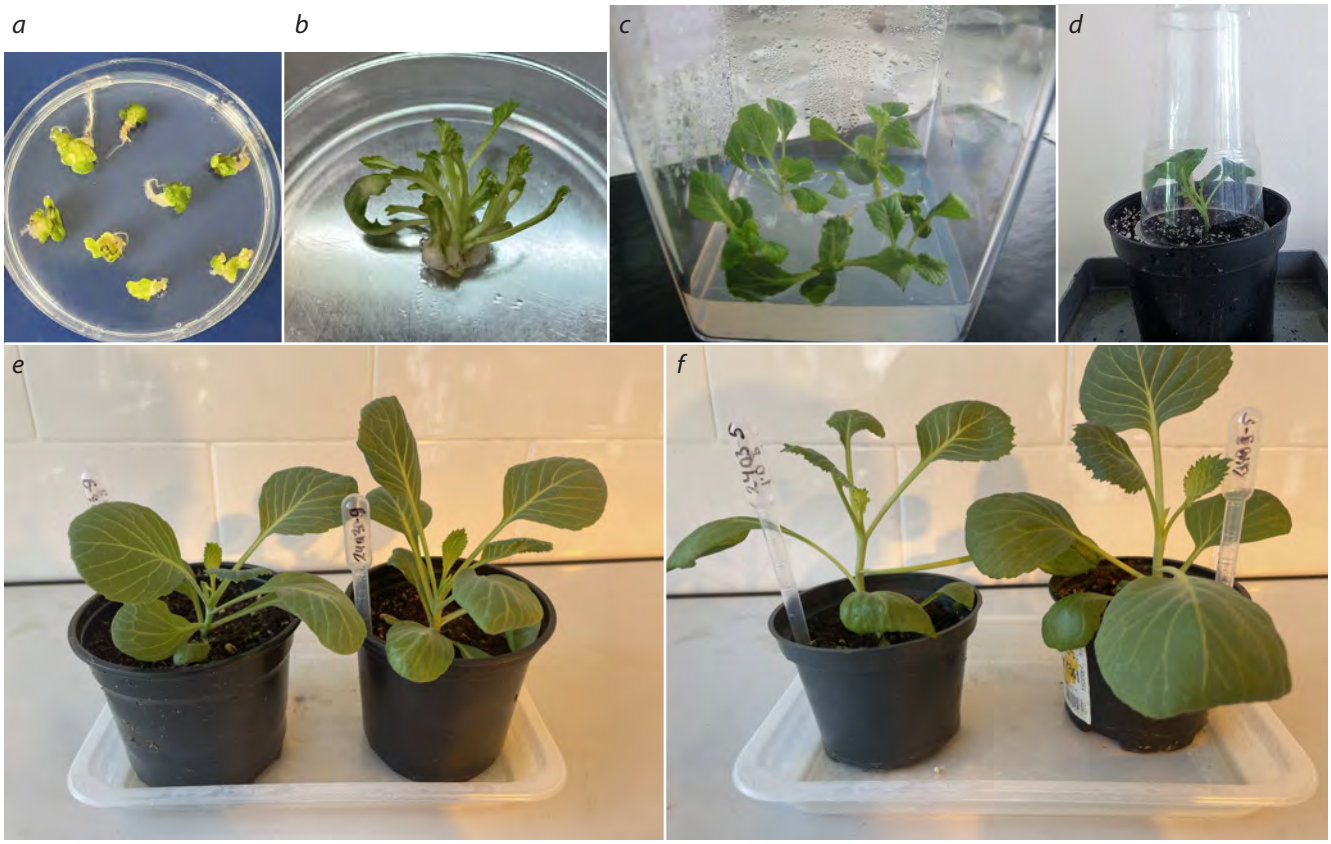


Fig. 5. The regeneration of white cabbage No. 2403 plants from embryoids and their adaptation to *ex vitro* conditions.

a – pembryoid outgrowth on MS medium; *b* – secondary embryogenesis; *c* – rhizogenesis; *d* – adaptation to *ex vitro* conditions; *e* – the diploid plants (2*n*) obtained from genotype No. 2403; *f* – the tetraploid plants (4*n*) obtained from genotype No. 2403.

Table 4. White cabbage regenerant plant ploidy

Genotype	Number of analyzed plants, pcs.	Ploidy level, (%)			
		n	2*n*	3*n*	4*n*
2403	50	0.0	88.0	0.0	12.0
2406	25	0.0	64.0	0.0	36.0
2404	13	7.8	46.1	0.0	46.1
127	20	5.0	70.0	5.0	20.0
303	30	30.0	66.7	0.0	3.3
360	25	8.0	84.0	4.0	4.0
Total	163	8.4	69.8	1.5	20.3

white cabbage genotypes that successfully passed the adaptation stage showed that the average percentage of haploids was 8.4 %, doubled haploids, 69.8 %, triploids, 1.5 %, tetraploids, 20.3 %. The genotypes noticeably differed by the occurrence of different ploidy (Table 4).

Confirmation of haploid origin of white cabbage regenerant plants

To confirm the haploid origin of diploid regenerated plants from four samples (No. 2403, 2406, 303, 360) obtained through the culture of isolated microspores, nine microsatellite loci (AJ427337, AF180355, AF241115(1), AF241115(2), AF458409, AF113918, BZ5223957, CC969431, U67451) that had been previously successfully used for genotyping of white cabbage were evaluated (Tonguç, Griffiths, 2004; Louarn et al., 2007; Domblides et al., 2020). Genetic differences between the regenerated plants and the donor genotypes were not observed in loci AJ427337, AF180355, and AF241115(1), where only one allele was obtained for all donor plant genotypes included in the study. More than four alleles were amplified

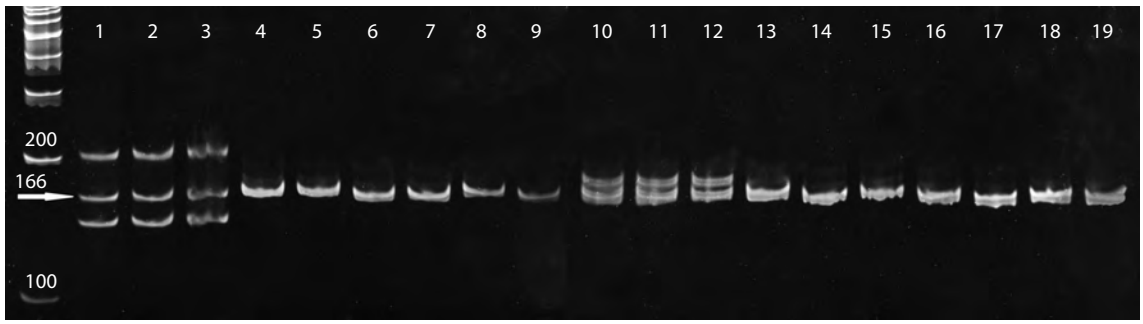


Fig. 6. The electrophoregram of DNA amplification results of white cabbage plants with microsatellite locus AF458409.

Here and in the Figure 7: the numbers: 1–3 – donor plants of genotype No. 2406; 4–9 – diploid regenerant plants obtained in the isolated microspore culture of No. 2406; 10–12 – donor plants of genotype No. 2403; 13–19 – diploid plants – regenerants obtained in the isolated microspore culture of No. 2403.

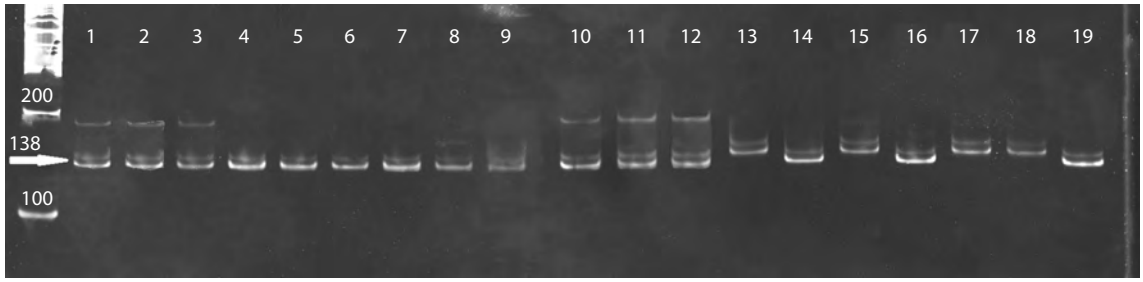


Fig. 7. The electrophoregram of DNA amplification results of white cabbage plants by microsatellite locus U67451.

with primers for the CC969431 locus, which was more than expected. Therefore, this marker was not taken for further difference assessment of the studied samples. Amplification with loci BZ5223957 and AF241115(2) revealed a polymorphism between donor plants and obtained regenerant plants only in genotypes No. 2406 and 303.

A polymorphism in all four genotypes was identified in loci AF458409, AF113918 and U67451. For example, amplification of the AF458409 locus revealed three alleles for No. 2403 and two alleles for No. 2406 in donor plants, whereas only one allele was observed in diploid regenerated plants (Fig. 6).

The U67451 locus showed two alleles in donor plants No. 2406, 303 and 360 and three alleles in donor plants No. 2403, while only one allele was present in regenerant plants (Fig. 7).

The absence of identity in SSR profiles of donor plants and regenerated plants was also observed for the microsatellite locus AF113918, where regenerated plants had a reduced number of alleles in SSR profiles compared to their donor plants.

Thus, genetic analysis with use of microsatellite markers enables to confirm the haploid origin of regenerated plants from microspores. When analyzing the profiles of three microsatellite loci AF458409, AF113918 and U67451, the absence of similarity in SSR markers between donor plants and regenerated plants was shown, and the reduced number of alleles in DH plants of white cabbage was also observed. In the case of implementation of *in vitro* culture of isolated anthers to obtain DH plants, this stage of the technology should be required to have 100 % accuracy, so that it is possible to separate true doubled haploids from diploid regenerated plants, which may have originated from anther somatic tissues.

Discussion

Over the last decades, significant progress has been made in the development of cell technologies for *Brassica* plants. They are based on the optimization of cultivation conditions and the application of manipulations that increase the embryoid yield.

The selection of buds at the optimal stage of microspore development determines the success of embryogenesis in any *in vitro* cultivation protocol. According to literature data, microspores at the late uninucleate and early bicellular stages have the highest probability of embryoid formation *in vitro* (Kott et al., 1988; Pechan, Keller, 1988). The stage of microspore development correlates with bud size, which allows us to select material based on this characteristic (Fan et al., 1988; Huang et al., 1990). Studies for different members of the Brassicaceae family have shown that the optimal bud size differs between species, different genotypes within one species and individual plants within one genotype. The optimal bud sizes of the studied genotypes were established experimentally – 4–5 mm for broccoli (Takahata, Keller, 1991), 4–6 mm for cauliflower (Gu et al., 2014), 4.1–5.0 mm for red cabbage (Mineykina et al., 2021), 3.0–3.1 mm for sarepta mustard (Ali et al., 2008), 2.5–3.5 mm for white cabbage (Yuan et al.,

2012; Tuncer et al., 2016), 4.5–4.6 mm for white cabbage of Indonesian origin (Winarto, da Silva, 2011).

In early studies, it was shown that a high percentage of non-embryogenic microspores in the culture resulted in increased levels of autotoxins in the medium affecting embryoids development. The negative effect of the toxins was correlated with the presence of bicellular microspores in the culture (Kott et al., 1988). In more recent studies (Duijs et al., 1992), good results were obtained in variants with 10 to 40 % of bicellular pollen in buds. We observed that the microspore developmental stage is a limiting factor in *in vitro* culture of white cabbage. Since cabbage exhibits pronounced asynchronous development of microspores within the bud, their selection with a certain limit of size variation maximizes the coverage of potentially embryogenic microspores.

An early study (Lighter, 1989) modified the microspore culturing protocol for the Brassicaceae family. The addition of activated charcoal to the cultured medium and the use of a shaker in microspore culture helped to minimize factors affecting cell destruction and suppression of cell division. Studies J.C. da Silva Dias, (1999) also indicate an increase in the embryogenesis of different cabbage genotypes when activated charcoal is used in the culture medium. A reduction in cultivation time by 1–4 days using a shaker platform was observed when Chinese cabbage microspores were cultured in liquid medium (Yuan et al., 2012).

In our studies, to increase the survival rate of white cabbage microspores and inhibit oxidation products that negatively affect cell division, we used activated charcoal in all variants as an integral element of the protocol. The use of a shaker platform allowed us to accelerate the embryoid development, thereby increasing the efficiency of the protocol by reducing the culture time in the liquid nutrient medium. As a result, more embryoids reached the cotyledonary stage 7–10 days faster, which is especially important for highly responsive genotypes that have a non-uniform embryoid maturation.

Stress is the most important condition during the transition of microspores from the gametophytic to the sporophytic pathway (Touraev et al., 1996). Stress can be applied both *in vivo* and *in vitro*. The most common type of stress for cabbage crops is exposure of buds and inflorescences to low positive temperatures (Gu et al., 2014) and short-term microspore heat shock in the *in vitro* culture (Custers et al., 1994).

Pretreatment with low positive temperatures does not promote reprogramming of microspore development but effectively maintains microspore viability (Żur et al., 2009). Heat treatment of isolated microspores in a range from 30 to 40 °C with time exposure from 1 to 3 days is most commonly used as an inducing factor of embryogenesis (Takahata, Keller, 1991; Duijs et al., 1992; Ferrie, Caswell, 2011). Some authors have noted that the duration of temperature treatment affects the number of developing embryoids (Telmer et al., 1992; Custers et al., 1994; Cordewener et al., 1995; Simmonds, Keller, 1999). For white cabbage of Indonesian origin (Winarto, da Silva, 2011), exposure to 30.5 °C for 48 h followed by continuous cultivation at 25 °C was found to be a successful approach. However, 30 % of embryoids in that study had abnormal cotyledon and no hypocotyls. In a study (Tuncer et al., 2016), the authors studied the effect of temperature shock on the induction of embryogenesis and embryoid development in *B. oleraceae* of Turkish origin. The authors observed a positive effect of 32 and 35 °C 2-day treatment on the induction of embryogenesis. However, the embryoid development was impeded in this study, and white cabbage plants did not regenerate.

To understand the molecular regulation of embryogenesis induced by high-temperature treatment, H. Su et al. (2019) performed a proteomic study. They found that the 32 °C high-temperature shock for 24 h induced changes in the expression of specific proteins in an *in vitro* culture of isolated cabbage microspores.

In addition to the fact that high-temperature stress is an effective trigger for switching microspore development to the sporophytic pathway, it also has a negative effect on cell division. Studies A. Zeng et al. (2015) showed massive white cabbage microspore death (80–90 %) after 3 days of *in vitro* cultivation. The authors observed microspore death after 24 h of shock temperature treatment at 32.5 °C, presumably caused by increased levels of reactive oxygen species (ROS) and the associated oxidative stress affecting cell viability and metabolism. According to the results of I. Żur et al. (2009), the lethal effect of high temperature during the induction of triticale microspore embryogenesis is associated with a sharp decrease in the enzymatic activity of all studied antioxidants. The authors suggest that high temperature stress induced oxidative stress, and cells in a nitrogen-carbohydrate starvation environment (Kyo, Harada, 1986) were unable to activate defense responses. The reduction of ROS levels was promoted by the use of ascorbic acid as an antioxidant in microspore culture (Zeng et al., 2015). However, the use of ascorbic acid in *in vitro* culture does not always have a positive effect and depends on its concentration (Rodriguez-Serrano et al., 2012; Hoseini et al., 2014).

In a study I. Barinova et al. (2004), the effects of carbohydrate stress induced by high pH values of the medium were observed in the cultivation of tobacco microspores. Analysis of sucrose metabolism at different pH showed that the activity of invertase (EC 3.2.1.26) in microspores was the highest at pH 5.0 and strongly decreased at higher pH, resulting in slower sucrose cleavage. These data suggested that isolated microspores cannot metabolize carbohydrates at higher pH and undergo starvation stress, which in turn triggers sporophytic development.

Our studies confirm the inducing effect of high-temperature stress at 32 °C on the white cabbage microspore development via the sporophytic pathway. The effect of the nutrient medium acidity on the embryoid yield depends on the duration of microspore high-temperature treatment.

Among the techniques used to obtain doubled haploids, only isolated microspore culture can ensure that the embryoids develop from haploid cells, whereas anther and unfertilized ovule cultures can include somatic tissues. The process of callus and embryoid formation as well as subsequent plant regeneration from somatic diploid tissues of anther walls is a well-known fact. Molecular analysis has been successfully used to distinguish true doubled haploid lines originating

from gametes from regenerant plants of somatic origin. The use of molecular markers to screen pepper plants obtained via *in vitro* culture is known from the literature (Gyulai et al., 2000). A. Cousin and M.N. Nelson (2009) used eight microsatellite markers to confirm homozygosity and the haploid cell origin of rape regenerant plants obtained in microspore culture.

SSR analysis has also been successfully used to assess heterozygosity and determine the origin (gametophytic or somatic) in lemon (Yahyaoui, Germanà, 2021), cucumber (Diao et al., 2009), melon (Malik et al., 2011) and Chinese cabbage (Adamus et al., 2021). Our results also allowed us to confirm the gametophytic origin of the regenerant plants using microsatellite markers. We showed that donor plant bands are different from the regenerant plant. Moreover, the number of alleles in DH plants of white cabbage decreased compared to the control.

Conclusion

In this study, we optimized the protocol for obtaining doubled haploids in the culture of isolated microspores *in vitro* for late-maturing white cabbage to achieve up to 273.6 ± 32.2 embryoids/Petri dish. A genotype-specific multifactorial approach should be used to achieve the best results. Before introduction into *in vitro* culture, determination of linear bud size is required. To obtain microspores at the optimal developmental stage for embryogenesis induction, the range of bud lengths in the sample should not exceed 0.5 mm. The final yield of white cabbage embryoids is significantly influenced by the duration of temperature induction treatment and the acidity of the nutrient medium. Cultivation of the induced microspores in the dark on a shaker platform at 40 revolutions/minute allows to significantly accelerate the development of embryoids to the cotyledonary stage and to shorten this step of the protocol by 7–10 days. Flow cytometry makes it possible to determine the ploidy of the regenerant plants at an early stage of development rather quickly. The following ploidy was observed: haploids (8.4 %), doubled haploids (69.8 %), triploids (1.5 %) and tetraploids (20.3 %).

The analysis with microsatellite markers confirmed the haploid origin of diploid regenerant plants. Three microsatellite loci AF458409, AF113918, and U67451 showed that the spectra of white cabbage donor plants and regenerant plants were not identical. Moreover, a smaller number of alleles was observed in DH plants.

References

Adamus A., Szklarczyk M., Kiełkowska A. Haploid and doubled haploid plant production in *Brassica rapa* L. subsp. *pekinensis* via microspore culture. *Methods Mol Biol*. 2021;2288:181-199. doi 10.1007/978-1-0716-1335-1_11

Alexander M.P. Differential staining of aborted and non-aborted pollen. *Stain Technol*. 1969;44(3):117-122. doi 10.3109/10520296909063335

Ali M.M., Khaleque M.A., Custers J.B.M., Khurram M.M.H. Microspore culture and the performance of microspore derived doubled haploid in *Brassica juncea* (L.). *Bangladesh J Agric. Res*. 2008; 33(4):571-578. doi 10.3329/bjar.v33i4.2290

Barinova I., Clément C., Martiny L., Baillieul F., Soukupova H., Heberle-Bors E., Touraev A. Regulation of developmental pathways in cultured microspores of tobacco and snapdragon by medium pH. *Planta*. 2004;219(1):141-146. doi 10.1007/s00425-003-1202-5

Barro F., Fernández-Escobar J., De La Vega M., Martín A. Modification of glucosinolate and erucic acid contents in doubled haploid lines of *Brassica carinata* by UV treatment of isolated microspores. *Euphytica*. 2003;129(1):1-6. doi 10.1023/A:1021578318098

Bhatia R., Dey S.S., Parkash C., Sharma K., Sood S., Kumar R. Modification of important factors for efficient microspore embryogenesis and doubled haploid production in field grown white cabbage (*Brassica oleracea* var. *capitata* L.) genotypes in India. *Sci Hortic*. 2018;233:178-187. doi 10.1016/j.scienta.2018.01.017

Cao M.Q., Charlot F., Dore C. Embryogenesis and plant regeneration in sauerkraut cabbage (*Brassica oleracea* L. subsp. *capitata*) by *in vitro* culture of isolated microspores. *Comptes Rendus de l'Academie des Sciences Serie III*. 1990;310:203-209. doi 10.1007/BF00231964

Cordewener J.H.G., Hause G., Görgen E., Busink R., Hause B., Dons H.J.M., Van Lammeren A.A.M., Van Lookeren Campagne M.M., Pechan P. Changes in synthesis and localization of members of the 70-kDa class of heat-shock proteins accompany the induction of embryogenesis in *Brassica napus* L. microspores. *Planta*. 1995;196:747-755. doi 10.1007/BF01106770

Cousin A., Nelson M.N. Twinned microspore-derived embryos of canola (*Brassica napus* L.) are genetically identical. *Plant Cell Rep*. 2009;28(5):831-835. doi 10.1007/s00299-009-0677-3

Cristea T.O., Iosob G.A., Brezeanu C., Brezeanu P.M. Effect of chemical composition of nutritive medium and explant size over androgenetic response in microspore culture of *Brassica oleracea* L. *Rev Chim*. 2020;71(10):131-136. doi 10.37358/RC.20.10.8357

Custers J.B.M., Cordewener J.H.G., Nöllen Y., Dons H.J.M., Van Lockeren Campagne M.M. Temperature controls both gametophytic and sporophytic development in microspore cultures of *Brassica napus*. *Plant Cell Rep*. 1994;13(5):267-271. doi 10.1007/BF00233317

da Silva Dias J.C. Effect of activated charcoal on *Brassica oleracea* microspore culture embryogenesis. *Euphytica*. 1999;108:65-69. doi 10.1023/A:1003634030835

da Silva Dias J.C. Protocol for broccoli microspore culture. In: Doubled Haploid Production in Crop Plants. Springer Netherlands, 2003; 195-204. doi 10.1007/978-94-017-1293-4_30

Diao W.P., Jia Y.Y., Song H., Zhang X.Q., Lou Q.F., Chen J.F. Efficient embryo induction in cucumber ovary culture and homozygous identification of the regenetants using SSR markers. *Sci Hortic*. 2009; 119(3):246-251. doi 10.1016/j.scienta.2008.08.016

Domblides A.S., Bondareva L.L., Pivovarov V.F. Assessment of genetic diversity among headed cabbage (*Brassica oleracea* L.) accessions by using SSR markers. *Sel'skokhozyajstvennaya Biologiya = Agricultural Biology*. 2020;55(5):890-900. doi 10.15389/agrobiology.2020.5.890rus (in Russian)

Duijs J.G., Voorrips R.E., Visser D.L., Custers J.B.M. Microspore culture is successful in most crop types of *Brassica oleracea* L. *Euphytica*. 1992;60:45-55. doi 10.1007/BF00022257

Fan Z., Armstrong K.C., Keller W.A. Development of microspores *in vivo* and *in vitro* in *Brassica napus* L. *Protoplasma*. 1988;147: 191-199. doi 10.1007/BF01403347

Ferrie A.M.R., Caswell K.L. Isolated microspore culture techniques and recent progress for haploid and doubled haploid plant production. *Plant Cell Tissue Organ Cult*. 2011;104:301-309. doi 10.1007/s11240-010-9800-y

Ferrie A.M.R., Epp D.J., Keller W.A. Evaluation of *Brassica rapa* L. genotypes for microspore culture response and identification of a

highly embryogenic line. *Plant Cell Rep*. 1995;14(9):580-584. doi 10.1007/BF00231942

Gu H.H., Zhou W.J., Hagberg P. High frequency spontaneous production of doubled haploid plants in microspore cultures of *Brassica rapa* ssp. *chinensis*. *Euphytica*. 2003;134:239-245. doi 10.1023/B:EUPH.0000004945.01455.6d

Gu H., Zhao Z., Sheng X., Yu H., Wang J. Efficient doubled haploid production in microspore culture of loose-curd cauliflower (*Brassica oleracea* var. *botrytis*). *Euphytica*. 2014;195:467-475. doi 10.1007/s10681-013-1008-x

Gyulai G., Gemesne J.A., Sagi Zs., Venczel G., Pinter P., Kristof Z., Torjek O., Heszky L., Bottka S., Kiss J., Zatyko L. Doubled haploid development and PCR-analysis of F_1 hybrid derived DH-R_2 paprika (*Capsicum annuum* L.) lines. *J Plant Physiol*. 2000;156(2):168-174. doi 10.1016/S0176-1617(00)80302-8

Hoseini M., Ghadimzadeh M., Ahmadi B., da Silva J.A. Effects of ascorbic acid, alpha-tocopherol, and glutathione on microspore embryogenesis in *Brassica napus* L. *In Vitro Cell Dev Biol Plant*. 2014;50:26-35. doi 10.1007/s11627-013-9579-8

Huang B., Bird S., Kemble R., Simmonds D., Keller W., Miki B. Effects of culture density, conditioned medium and feeder cultures on microspore embryogenesis in *Brassica napus* L. cv. Topas. *Plant Cell Rep*. 1990;8(10):594-597. doi 10.1007/BF00270061

Ji J., Su H., Cao W., Zhang X., Li H., Fang Z., Yang L., Zhang Ya., Zhuang M., Wang Yo., Taranov V., Lv H. Genetic analysis and mapping of QTLs for isolated microspore embryogenesis in cabbage. *Scientia Horticulturae*. 2023;313:111897. doi 10.1016/j.scienta.2023.111897

Kasha K.J. Chromosome doubling and recovery of doubled haploid plants. In: Don Palmer C., Keller W.A., Kasha K.J. (Eds) Haploids in Crop Improvement II. Biotechnology in Agriculture and Forestry. Vol. 56. Springer, 2005;123-152. doi 10.1007/3-540-26889-8_7

Kitashiba H., Taguchi K., Kaneko I., Inaba K., Yokoi S., Takahata Y., Nishio T. Identification of loci associated with embryo yield in microspore culture of *Brassica rapa* by segregation distortion analysis. *Plant Cell Rep*. 2016:35(10):2197-2204. doi 10.1007/s00299-016-2029-4. Epub 2016 Jul 20.

Kott L.S. Application of doubled haploid technology in breeding of oilseed *Brassica napus*. *AgBiotech News Inf*. 1998;10:69N-73N. doi 10.1533/9781908818478.183

Kott L.S., Polsoni L., Beversdorf W.D. Cytological aspects of isolated microspore culture of *Brassica napus*. *Can J Bot*. 1988;66(8):1658-1664. doi 10.1139/b88-226

Kozar E.V., Domblides E.A., Soldatenko A.V. Embryogenesis of European radish (*Raphanus sativus* L. subsp. *sativus* convar. radicula) in culture of isolated microspores in vitro. *Plants*. 2021;10(10):2117. doi 10.3390/plants10102117

Kozar E.V., Kozar E.G., Domblides E.A. Effect of the method of microspore isolation on the efficiency of isolated microspore culture in vitro for Brassicaceae family. *Horticulturae*. 2022;8(10):864. doi 10.3390/horticulturae8100864

Kyo M., Harada H. Control of the developmental pathway of tobacco pollen in vitro. *Planta*. 1986;168:427-432. doi 10.1007/BF00392260

Lichter R. Induction of haploid plants from isolated pollen of *Brassica napus*. *Zeitschrift für Pflanzenphysiologie*. 1982;105(5):427-434. doi 10.1016/S0044-328X(82)80040-8

Lighter R. Efficient yield of embryoids by culture of isolated microspores of different *Brassicaceae* species. *Plant Breeding*. 1989;103(2):119-123. doi 10.1111/j.1439-0523.1989.tb00359.x

Louarn S., Torp A.M., Holme I.B., Andersen S.B., Jensen B.D. Database derived microsatellite markers (SSRs) for cultivar differentiation in *Brassica oleracea*. *Genet Resour Crop Evol*. 2007;54(8):1717-1725. doi 10.1007/s10722-006-9181-6

Malik A.A., Cui L., Zhang S., Chen J. Efficiency of SSR markers for determining the origin of melon plantlets derived through unfertilized ovary culture. *Hort Sci*. 2011;38(1):27-34. doi 10.17221/47/2010-HORTSCI

Mineykina A., Bondareva L., Soldatenko A., Domblides E. Androgenesis of red cabbage inisolated microspore culture in vitro. *Plants*. 2021;10(9):1950. doi 10.3390/plants10091950

Murashige T., Skoog F. A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiol Plant*. 1962;15(3):473. doi 10.1111/j.1399-3054.1962.tb08052.x

Pechan P.M., Keller W.A. Identification of potentially embryogenic microspores in *Brassica napus*. *Physiol Plant*. 1988;74(2):377-384. doi 10.1111/j.1399-3054.1988.tb00646.x

Pechan P.M., Smykal P. Androgenesis: affecting the fate of the male gametophyte. *Physiol Plant*. 2001;111(1):1-8. doi 10.1034/j.1399-3054.2001.1110101.x

Prem D., Gupta K., Sarkar G., Agnihotri A. Activated charcoal induced high frequency microspore embryogenesis and efficient doubled haploid production in *Brassica juncea*. *Plant Cell Tissue Organ Cult*. 2008;93:269-282. doi 10.1007/s11240-008-9373-1

Rodriguez-Serrano M., Bárány I., Prem D., Coronado M.-J., Risueño M.C., Testillano P.S. NO, ROS, and cell death associated with caspase-like activity increase in stress-induced microspore embryogenesis of barley. *J Exp Bot*. 2012;63(5):2007-2024. doi 10.1093/jxb/err400

Rudolf K., Bohanec B., Hansen M. Microspore culture of white cabbage, *Brassica oleracea var. capitata* L.: genetic improvement of non-responsive cultivars and effect of genome doubling agents. *Plant Breed*. 1999;118(3):237-241. doi 10.1046/j.1439-0523.1999.118003237.x

Shumilina D., Kornyukhin D., Domblides E., Soldatenko A., Artemyeva A. Effects of genotype and culture conditions on microspore embryogenesis and plant regeneration in *Brassica rapa* ssp. *rapa* L. *Plants*. 2020;9(2):278. doi 10.3390/plants9020278

Simmonds D.H., Keller W.A. Significance of preprophase bands of microtubules in the induction of microspore embryogenesis of *Brassica napus*. *Planta*. 1999;208:383-391. doi 10.1007/s004250050573

Su H., Chen G., Yang L., Zhang Y., Wang Y., Fang Z., Lv H. Proteomic variations after short-term heat shock treatment reveal differentially expressed proteins involved in early microspore embryogenesis in cabbage (*Brassica oleracea*). *PeerJ*. 2020;8:e8897. doi 10.7717/peerj.8897

Takahata Y., Keller W.A. High frequency embryogenesis and plant regeneration in isolated microspore culture of *Brassica oleracea* L. *Plant Sci*. 1991;74(2):235-242. doi 10.1016/0168-9452(91)90051-9

Tang X., Liu Y., He Y., Ma L., Sun M. Exine dehiscing induces rape microspore polarity, which results in different daughter cell fate and fixes the apical-basal axis of the embryo. *J Exp Bot*. 2013;64(1):215-228. doi 10.1093/jxb/ers327

Telmer C.A., Simmonds D.H., Newcomb W. Determination of developmental stage to obtain high frequencies of embryogenic microspores in *Brassica napus*. *Physiol Plant*. 1992;84(3):417-424. doi 10.1111/j.1399-3054.1992.tb04685.x

Tonguç M., Griffiths P.D. Genetic relationships of *Brassica* vegetables determined using database derived simple sequence repeats. *Euphytica*. 2004;137(2):193-201. doi 10.1023/B:EUPH.0000041577.84388.43

Touraev A., Pfosser M., Vicente O., Heberle-Bors E. Stress as the major signal controlling the developmental fate of tobacco microspores: towards a unified model of induction of microspore/pollen embryogenesis. *Planta*. 1996;200:144-152. doi 10.1007/BF00196662

Tuncer B., Çğ A., Yanmaz R., Yaşar F. Effect of heat shock treatment on microspore embryogenesis in *Brassica oleracea* species. *J Agric Sci* (*Belihuloya*). 2016;22(4):548-554. doi 10.1501/Tarimbil_0000001413

Weber M., Davies J.J., Wittig D., Oakeley E.J., Haase M., Lam W.L., Schuebeler D. Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nat Genet*. 2005;37(8):853-862. doi 10.1038/ng1598

Winarto B., da Silva J.A.T. Microspore culture protocol for Indonesian *Brassica oleracea*. *Plant Cell Tissue Organ Cult*. 2011;107(2):305-315. doi 10.1007/S11240-011-9981-Z/TABLES/6

Yahyaoui E., Germanà M.A. Microspore embryogenesis in *Citrus*. In: Segui-Simarro J.M. (Ed.) Doubled Haploid Technology. Methods in Molecular Biology. Vol. 2289. New York, NY: Humana, 2021;149-166. doi 10.1007/978-1-0716-1331-3_10

Yang S., Liu X., Fu Y., Zhang X., Li Y., Liu Z., Feng H. The effect of culture shaking on microspore embryogenesis and embryonic development in Pakchoi (*Brassica rapa* L. ssp. *chinensis*). *Sci Hortic*. 2013;152:70-73. doi 10.1016/j.scienta.2013.01.019

Yuan S., Su Y., Liu Y., Fang Z., Yang L., Zhuang M., Zhang Y., Sun P. Effects of pH, MES, arabinogalactan-proteins on microspore cultures in white cabbage. *Plant Cell Tissue Organ Culture*. 2012;110(1):69-76. doi 10.1007/s11240-012-0131-z

Zeng A., Yan J., Song L., Gao B., Li J. Effects of ascorbic acid and embryogenic microspore selection on embryogenesis in white cabbage (*Brassica oleracea* L. var. *capitata*). *J Hortic Sci Biotechnol*. 2015;90(6):607-612. doi 10.1080/14620316.2015.11668722

Zhang F.L., Aoki S., Takahata Y. RAPD markers linked to microspore embryogenic ability in *Brassica* crops. *Euphytica*. 2003;131(2):207-213. doi 10.1023/A:1023955131523

Żur I., Dubas E., Golemiec E., Szechyńska-Hebda M., Gołębiowska G., Wędzony M. Stress-related variation in antioxidative enzymes activity and cell metabolism efficiency associated with embryogenesis induction in isolated microspore culture of triticale (x *Triticosecale* Wittm.). *Plant Cell Rep*. 2009;28(8):1279-1287. doi 10.1007/s00299-009-0730-2

Conflict of interest. The authors declare no conflict of interest.
Received November 13, 2024. Revised February 27, 2025. Accepted March 4, 2025.

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2025;29(4):530-538

doi 10.18699/vjgb-25-56

Drought tolerance of the photosynthetic apparatus of bread wheat (*Triticum aestivum* L.) lines with introgressions in chromosome 2D from *Aegilops tauschii* Coss.

S.V. Osipova [1, 2], A.V. Permyakov [1], A.V. Rudikovskii [1], E.G. Rudikovskaya [1], T.A. Pshenichnikova [3]

[1] Siberian Institute of Plant Physiology and Biochemistry of the Siberian Branch of the Russian Academy of Sciences, Irkutsk, Russia
[2] Irkutsk State University, Irkutsk, Russia
[3] Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
svetlanaosipova2@mail.ru

Abstract. One of the ways to increase yield stability of bread wheat under changing climatic conditions is through improving the photosynthesis efficiency. For this purpose, various genetic strategies are used. They include marker-assisted selection and the use of the genetic potential of wild wheat relatives. Previously, using introgression wheat lines carrying different segments of chromosome 2D from *Aegilops tauschii* in the genetic background of the wheat (*Triticum aestivum*) variety Chinese Spring (CS), we mapped QTLs associated with variability in shoot biomass and gas exchange under contrasting water supply conditions. In this work, by "splitting" the primary introgressions, we obtained secondary introgression CS lines with reduced segments of *Ae. tauschii* introgressions in the short and long arms of chromosomes 2D. The aim of this study was to investigate the tolerance of the photosynthetic apparatus to soil water deficit in these lines. We estimated the size of drought effect on shoot biomass, gas exchange parameters, photosynthetic pigment content, slow and fast chlorophyll fluorescence parameters, and fast light curve parameters. The results showed that line 1004 with an introgression in chromosome 2DS limited by microsatellite loci *Xgwm296* and *Xgwm261* was little affected by drought in respect of the chlorophyll (*a*+*b*)/carotenoid ratio and primary photosynthetic processes. In line 1005 with a single introgression in the region of the *Xgwm261* marker, the chlorophyll (*a*+*b*)/carotenoid ratio and indicators of the functional activity of photosystems significantly decreased under water deficiency. The chlorophyll (*a*+*b*)/carotenoid ratio, CO_2 assimilation rate, and chlorophyll fluorescence parameters remained stable in line 1034 with an introgression in chromosome 2DL near the *Xgwm1419* and *Xgwm157* loci. In line 1021 with an introgression in the region of the *Xgwm539* marker on the same chromosome, we observed a strong negative effect of drought on the rate of CO_2 assimilation and indicators of the functional activity of photosystems. The *Xgwm1419* and *Xgwm296* markers can be recommended for use in marker-assisted breeding for drought tolerance of bread wheat in the cases where *Ae. tauschii* acts as a donor of genetic material.

Key words: bread wheat; soil drought; shoot biomass; gas exchange; chlorophyll fluorescence; introgressions; molecular markers

For citation: Osipova S.V., Permyakov A.V., Rudikovskii A.V., Rudikovskaya E.G., Pshenichnikova T.A. Drought tolerance of the photosynthetic apparatus of bread wheat (*Triticum aestivum* L.) lines with introgressions in chromosome 2D from *Aegilops tauschii* Coss. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov J Genet Breed*. 2025;29(4):530-538. doi 10.18699/vjgb-25-56

Funding. The work was carried out within the framework of the basic project of the Siberian Institute of Plant Physiology and Biochemistry SB RAS No. 0277-2022-0006 and the basic project of the Institute of Cytology and Genetics SB RAS No. FWNR-2022-0017.

Acknowledgements. The studies were carried out on the equipment of the Bioanalytics Center of the SIPPB SB RAS, Irkutsk, Russia. We are grateful to M.D. Permyakova for technical assistance and E.V. Morozova for partial molecular analysis of plants.

Устойчивость к засухе фотосинтетического аппарата линий пшеницы *Triticum aestivum* L. с интрогрессиями от *Aegilops tauschii* Coss. в хромосоме 2D

С.В. Осипова [1, 2], А.В. Пермяков [1], А.В. Рудиковский [1], Е.Г. Рудиковская [1], Т.А. Пшеничникова [3]

[1] Сибирский институт физиологии и биохимии растений Сибирского отделения Российской академии наук, Иркутск, Россия
[2] Иркутский государственный университет, Иркутск, Россия
[3] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
svetlanaosipova2@mail.ru

© Osipova S.V., Permyakov A.V., Rudikovskii A.V., Rudikovskaya E.G., Pshenichnikova T.A., 2025
This work is licensed under a Creative Commons Attribution 4.0 License

Аннотация. Улучшение эффективности фотосинтеза в изменяющихся климатических условиях является одним из способов повышения стабильности урожая сельскохозяйственных растений. Для этого применяют различные генетические стратегии, в частности маркер-ориентированную селекцию, а также привлекают генетический потенциал диких сородичей пшеницы. Ранее, используя интрогрессивные линии пшеницы, содержащие различные сегменты хромосомы 2D от *Aegilops tauschii* в генетическом фоне пшеницы *Triticum aestivum* сорта Чайниз Спринг (ЧС), мы картировали QTL, ассоциированные с вариабельностью биомассы побега и газообмена в контрастных условиях водоснабжения. В данной работе путем «дробления» первичных интрогрессий мы получили вторичные интрогрессивные линии пшеницы ЧС с более короткими сегментами интрогрессий от *Ae. tauschii*. Целью исследования было изучить устойчивость фотосинтетического аппарата к дефициту воды в почве у вторичных интрогрессивных линий, содержащих редуцированные интрогрессии от *Ae. tauschii* в коротком и длинном плечах хромосомы 2D. Мы оценили размер эффекта засухи на биомассу побега, параметры газообмена, содержание фотосинтетических пигментов, параметры медленной и быстрой флуоресценции хлорофилла и параметры быстрых световых кривых. Результаты показали, что у линии 1004 с участком интрогрессии в хромосоме 2DS, ограниченном микросателлитными локусами *Xgwm296* и *Xgwm261*, засуха незначительно влияла на соотношение хлорофиллы *a*+*b*/каротиноиды и первичные процессы фотосинтеза. У линии 1005 с участком интрогрессии в районе маркера *Xgwm261* при дефиците воды значительно снижались соотношение хлорофиллы *a*+*b*/каротиноиды и показатели функциональной активности фотосистем. У линии 1034 с интрогрессией в хромосоме 2DL в районе локусов *Xgwm1419* и *Xgwm157* соотношение хлорофиллы *a*+*b*/каротиноиды, скорость ассимиляции CO_2 и параметры флуоресценции хлорофилла при засухе оставались стабильными. У линии 1021 с участком интрогрессии в районе маркера *Xgwm539* на этой же хромосоме мы наблюдали сильное негативное влияние засухи на скорость ассимиляции CO_2 и показатели функциональной активности фотосистем. Маркеры *Xgwm1419* и *Xgwm296* можно рекомендовать для использования в маркер-ориентированной селекции на засухоустойчивость мягкой пшеницы в случаях, когда донором генетического материала выступает *Ae. tauschii*.

Ключевые слова: мягкая пшеница; почвенная засуха; биомасса побега; газообмен; флуоресценция хлорофилла; интрогрессии; молекулярные маркеры

Introduction

Improving the efficiency of photosynthesis is considered one of the most important issues of breeding work aimed at increasing productivity of bread wheat (*Triticum aestivum* L.) through improving tolerance to unfavorable factors. Various genetic strategies are effective in achieving these goals, including the use of the genetic potential of wild relatives of wheat and marker-assisted selection (Reynolds et al., 2012).

Wild relatives represent a valuable gene pool for bread wheat improvement, since this crop has a limited genetic diversity for meeting the challenges of modern breeding. Various species of the genus *Aegilops* L., which is most closely related to the genus *Triticum* L., are considered a source of beneficial alleles for bread wheat fortification against abiotic stresses, pests, and diseases (Przewieslik-Allen et al., 2019; Pour-Aboughadareh et al., 2021). One such species is *Ae. tauschii*, known as the donor of the D genome of bread wheat and containing favorable allelic variations in genes associated with stress responses (Jia et al., 2013). Its homology with the D subgenome of bread wheat simplifies the introgression process during breeding and for genetic analysis. Therefore, *Ae. tauschii* is widely used in research aimed at improving the productivity and stability of wheat under various climatic conditions (Nyine et al., 2021; Ma et al., 2023).

An intermediate step in the transfer of genetic diversity from this genome is synthetic hexaploid wheats with the BBAADD genome, homologous to bread wheat. The first synthetic, called Synthetic 6x (Syn6x) (McFadden, Sears, 1946), was used to obtain single-chromosome substituted lines of Chinese Spring (CS)(Syn6x) (Nicholson et al., 1993). Subsequently, on the basis of substitution lines D-genome chromosomes, introgressive lines carrying single chromosome segments from *Ae. tauschii* of different sizes were obtained (Pestsova et al., 2001). Using this set of eighty introgressive lines CS(Syn6x), we mapped quantitative trait loci (major QTL) associated with variability in shoot biomass (SB) and gas exchange parameters under soil water deficit in two regions of chromosome 2D (Osipova et al., 2016). One of the regions was located on the short arm between microsatellite markers *Xgdm5* and *Xgwm296*, and the second was flanked by markers *Xgwm539* and *Xgwm1419* on the long arm. The size of the first region was 11.4 cM, and the second, 10.5 cM (Röder et al., 1998).

Further refinement of the position of loci associated with photosynthesis variability on chromosome 2D and search for putative candidate genes became possible by obtaining the lines with reduced segments of introgressions and studying the stability of the functioning of the photosynthetic apparatus in the new lines. Chlorophyll (Chl) fluorescence parameters are considered a reliable source of information about the physiological state of photosynthetic apparatus of plants (Goltsev et al., 2016). They have been successfully used in screening of adult bread wheat plants for drought tolerance in the field and in that of wheat seedlings in the laboratory (Botyanszka et al., 2020; Peršić et al., 2022).

The aim of this study was to investigate the tolerance of the photosynthetic apparatus to soil water deficiency in the introgressive lines containing short introgressions from *Ae. tauschii*.

Materials and methods

Genetic material, molecular analysis and experimental conditions. The two groups of secondary introgressive lines (SILs) obtained on the basis of two introgressive lines, Chinese Spring CS(Syn6x 2D-4) and CS(Syn6x 2D-6) (Pestsova et al., 2001), along with a recipient variety CS were used in the work. To narrow down the regions of introgressions, a "splitting" approach was used, consisting of hybridization

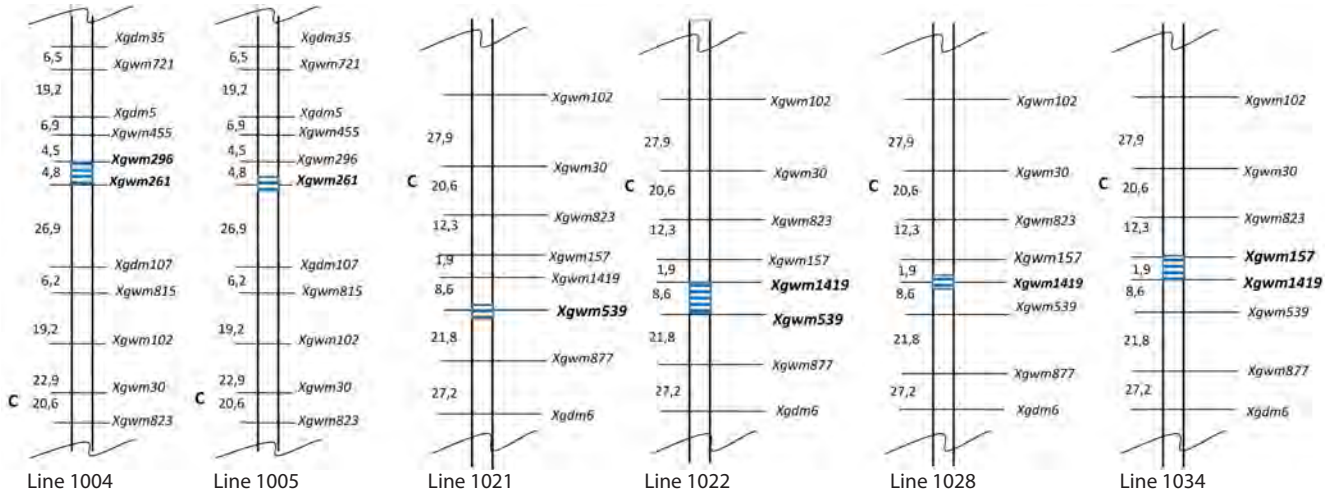


Fig. 1. Schematic arrangement of introgression regions of chromosome 2D in secondary introgressive lines CS(Syn6x 2D-4) (lines 1004 and 1005) and CS(Syn6x 2D-6) (lines 1021, 1022, 1028 and 1034).

Microsatellite markers of chromosome 2D represented by allelic variants of Syn6x and relevant to the previously identified positions of QTL clusters associated with drought response (Osipova et al., 2016) are shown in bold. The sequence of markers and the distances between them (not to scale) are presented according to the maps of M.S. Röder et al. (1998) and E.G. Pestsova et al. (2001).

of these two primary lines with the recipient CS. Secondary introgressive lines were obtained by a single backcrossing followed by subsequent self-pollination into F_2. The plants were then analyzed for their microsatellite marker composition. Those plants were selected that showed allelic differences in the target chromosome 2D regions where clusters of QTL loci associated with drought response had previously been detected (Osipova et al., 2016).

DNA extraction was performed according to the protocol of Plaschke et al. (1995). The obtained PCR products were separated in 3 % agarose gel and photographed under UV light using the Molecular Imager® Gel DocTM XR+ system (Bio-Rad Laboratories, Inc., California, USA). Two lines, numbered 1004 and 1005, were selected among F_2 plants from the cross between CS and IL CS(Syn6x 2D-4) (Fig. 1). These lines differed only in the allelic state of marker *Xgwm296*. Allelic variants of this marker obtained using the PCR reaction are presented in Figure S1 in Supplementary Material[1].

Four lines numbered 1021, 1022, 1028 and 1034 were selected among F_2 plants from the cross between CS and IL CS(Syn6x 2D-6). They differ in the allelic state of markers *Xgwm1419*, *Xgwm157* and *Xgwm539* (Fig. S2). The plants were grown under controlled conditions in a CLF PlantMaster climate chamber (CLF Plant Climatic GMBH, Germany) installed in the phytotron of SIPPB SB RAS, with a 16-hour photoperiod, a temperature of 23 °C during the day and 16 °C at night, air humidity of 60 % and a light intensity of 300 μmol/(m^2·s). Ten grains of each genotype were sown in two Mitscherlich pots filled with a mixture of humus, sand and peat (1:1:1). The moisture content of the soil in one pot was maintained at an optimal level (60 % of the total soil moisture capacity). In the second pot, watering was reduced by half, to 30 % of the total soil moisture capacity starting from the third leaf stage. The water regime was maintained gravimetrically. At the flowering stage, the gas exchange parameters and chlorophyll (Chl) fluorescence were measured in plants. At this stage, the main shoot mass was measured and samples were collected to determine the content of photosynthetic pigments.

Gas exchange, chlorophyll fluorescence and photosynthetic pigment content. Net photosynthesis rate (A), stomatal conductance (Gs) and transpiration rate (E) were measured using a portable leaf gas exchange system GFS-3000 (Heinz Walz, Germany). The following values of light intensity, CO_2 concentration, relative humidity, temperature and airflow rate were set: 800 μmol/(m^2·s), 400 μmol/mol, 60 %, 25 °C and 750 μmol/s, respectively. Water use efficiency (WUE) was calculated as A/E. The mean values and standard deviations for gas exchange parameters are given in Table S1.

Using a PAM 2500 fluorimeter (Heinz Walz, Germany) integrated with PamWin 3.05 software, the following parameters were measured: the kinetics of slow Chl fluorescence induction; the parameters of fast light curve; the kinetics of fast Chl fluorescence induction (OJIP test). To record the minimum Chl fluorescence yield in the dark-adapted state (F_0), the leaves were darkened for 30 min and then illuminated with modulated measuring light of low-frequency (5 Hz) and low-intensity (630 nm). The chlorophyll fluorescence intensity under conditions of closed reaction centers (Fm) was measured after exposure to a high-intensity light pulse of 25,000 μmol (photon)/(m^2·s), wavelength 630 nm, 0.1 s. Red actinic light (677 μmol photons/(m^2·s)) was used to maintain photosynthesis and achieve a steady state (F). Based on the measured values of chlorophyll fluorescence parameters, the PamWin 3.50 program calculated other parameters. We assessed the response to rapid irradiance increases (every 30 s) by exposing leaves to light intensities ranging from 0 to 1,935 μmol/(m^2·s) PAR photons and recorded the initial slope of the fast light response curve (α), maximum electron transfer rate (ETR_{max}) and minimum saturating irradiance (I_k). Chl fluorescence induced by strong light pulses was sampled in the range from 0.1 to 300 ms in the View instrument mode under the Fast Kinetics tab (Chen K. et al., 2013, Srivastava

[1] Figures S1–S3 and Tables S1, S2 are available at:
https://vavilov.elpub.ru/jour/manager/files/Suppl_Osipova_Engl.xlsx

et al., 2021). All Chl fluorescence parameters measured and calculated during the study, as well as the size of drought effect (SDE) on each parameter, are listed and described in Table S2.

Determination of photosynthetic pigments content. The preparation and measurement of the optical density of extracts containing photosynthetic pigments were carried out according to the previously described method (Osipova et al., 2024). To calculate the content of chlorophyll *a* (Chl *a*), chlorophyll *b* (Chl *b*) and carotenoids (Car) in the leaves, the formulas given in the work of D. Wettstein (1957) were used.

Search for coordinates of molecular markers and candidate genes that may participate in the formation of drought tolerance on chromosome 2D. To search for coordinates of markers *Xgwm261* and *Xgwm157*, the primer sequences presented in the GrainGenes information resource (https://wheat.pw.usda.gov/GG3) were used. For marker *Xgwm1419*, the primer sequences were provided by Martin Ganal (TraitGenetics GmbH). Using the BLASTN program, coordinates were determined for them in the wheat genome assembly Chinese Spring IWGSC RefSeq v2. The coordinates of markers *Xgwm296* and *Xgwm539* are specified in this assembly (https://wheat.pw.usda.gov, last accessed February 05, 2025). In the regions limited by these markers, a search for the most probable candidate genes was carried out. Candidate genes annotated by the International Wheat Genome Sequencing Consortium (IWGS, 2018) with a high degree of reliability were considered.

Statistics. A single plant was taken as a biological replicate. At the flowering stage, gas exchange parameters were measured in six plants of each genotype under each water regime. Chl fluorescence parameters were measured in three plants of each genotype. Then, the aboveground part of the main shoot of nine plants of each genotype was cut off and weighed. Three samples taken from the flag leaves of three plants were frozen with liquid nitrogen and stored at –70 °C for subsequent determination of pigment content. The pigment content is given in mg/g fresh leaf weight. The tables present the average values ± standard deviations. The effect of soil water deficit on chlorophyll fluorescence, pigment content, and shoot biomass was assessed using the size of drought effect (SDE) index (Hedges, Olkin, 1985). The formulas for calculating SDE are given in the work of S.V. Osipova et al. (2024). The higher the effect size, the greater the increase in the parameter under drought conditions compared to the control. Negative values indicate a decrease in the parameter compared to the control.

All calculations, including average values, pooled standard deviation, adjusted SDE value, and diagram plotting were performed in Microsoft Excel, version 14.0.7268.5000 (Microsoft Corporation, 2010). The significance of differences was assessed using Student's *t*-test. The fluorescence data pool was processed using nonmetric multidimensional scaling in Past, version 3.01 (Hammer et al., 2001).

Results

Shoot biomass and photosynthetic pigments content in leaves

The biomass of the main shoot in the lines varied from 2.6 to 3.5 g in the control and from 1.5 to 2 g under soil drought conditions (Table 1). Under normal watering, only one line, 1005, exceeded the recipient for this trait, and the lowest value was found in line 1021. Under water deficit conditions, the shoot biomass in two lines, 1028 and 1034, exceeded the values of this trait in CS, while in lines 1021 and 1022, on the contrary, it did not reach CS values. SDE was negative in all the studied genotypes, but the value of this indicator varied from –1.69 in line 1021 to –7.06 in line 1005. Line 1021 was distinguished by reduced values of shoot biomass, both under optimal irrigation and under drought conditions. The high shoot biomass of line 1005 in the control was significantly reduced under drought. Line 1004 was more stable than line 1005 in this trait.

Under drought conditions, the content of Chl *a*, Chl *b* and Car in CS decreased, and SDE on Chl *a*+*b*/Car ratio was close to zero (Table 2). Pigment content in leaves of the lines changed differently. In lines 1004 and 1034, Chl *a* content significantly increased under drought. Car content also increased in all lines, except for line 1022. The greatest positive effect of drought on Car content was observed in lines 1004, 1005 and 1034, in ascending order. Chl *a*+*b*/Car ratio changed insignificantly in all genotypes, except for line 1005. The negative effect of drought on this trait in line 1005 was due to the fact that Chl *a* content remained stable under different conditions, and Car content increased.

Table 1. Shoot biomass (g) and its size of drought effect (SDE) of the recipient Chinese Spring and the studied lines at the flowering stage in the control and under water deficit

Genotype	Line number	Control	Drought	SDE
Chinese Spring (CS)		3.0 ± 0.3	1.7 ± 0.1***	–4.74
		Secondary introgressive lines		
CS(Syn6x 2D-4)	1004	2.8 ± 0.7	1.6 ± 0.4***	–1.96
	1005	3.5 ± 0.3^{aa}	1.6 ± 0.3***	–7.06
CS(Syn6x 2D-6)	1021	2.6 ± 0.6	1.5 ± 0.1***	–1.69
	1022	2.9 ± 0.5	1.5 ± 0.1***	–3.82
	1028	2.7 ± 0.3	2.0 ± 0.2^{a}**	–2.85
	1034	2.9 ± 0.5	1.9 ± 0.1^{a}***	–3.04

** $p < 0.01$, *** $p < 0.001$ – significant differences between each genotype in control and drought conditions; ^{a} $p < 0.05$; ^{aa} $p < 0.01$ – significant differences between CS and lines.

Table 2. Content of photosynthetic pigments (mg/g of raw weight) and their size of drought effect in the leaves of the recipient Chinese Spring and the studied lines ($n = 9$) in the control and under water deficiency

Trait	Chinese Spring	1004	1005	1021	1022	1028	1034
Control							
Chlorophyll *a*	2.4 ± 0.2	2.9 ± 0.1	2.9 ± 0.1	2.7 ± 0.2	2.6 ± 0.1	–	2.3 ± 0.1
Chlorophyll *b*	1.1 ± 0.1	1.3 ± 0.1	1.2 ± 0.0	1.2 ± 0.1	1.2 ± 0.1	–	1.1 ± 0.1
Carotenoids	0.7 ± 0.1	0.6 ± 0.1	0.5 ± 0.0	0.6 ± 0.0	0.7 ± 0.0	–	0.4 ± 0.1
Chlorophyll *a*+*b*/Carotenoids	5.6 ± 0.3	11.6 ± 4.4	7.8 ± 0.3[a]	6.9 ± 0.2[a]	5.8 ± 0.3	–	10.6 ± 3.7
Drought							
Chlorophyll *a*	2.2 ± 0.1	3.4 ± 0.2*[a]	2.9 ± 0.1	3.1 ± 0.1[a]	2.9 ± 0.2	3.1 ± 0.2[a]	2.8 ± 0.2*
Chlorophyll *b*	1.0 ± 0.0	1.4 ± 0.1	1.3 ± 0.0[a]	1.4 ± 0.1[a]	1.0 ± 0.1	1.4 ± 0.1[a]	1.2 ± 0.1
Carotenoids	0.6 ± 0.0*	0.7 ± 0.0[a]	0.6 ± 0.0*	0.7 ± 0.1	0.7 ± 0.0	0.7 ± 0.0	0.6 ± 0.0**
Chlorophyll *a*+*b*/Carotenoids	5.7 ± 0.1	6.9 ± 0.1[a]	6.6 ± 0.2**	7.1 ± 0.7[a]	5.8 ± 0.4	6.3 ± 0.5	7.2 ± 0.1[a]
Size of drought effect							
Chlorophyll *a*	–1.2	3.2	0	2.2	2.2	–	3.0
Chlorophyll *b*	–1.3	1.2	2.1	1.8	–2.0	–	1.2
Carotenoids	–3.3	1.7	3.8	1.4	–0.5	–	5.3
Chlorophyll *a*+*b*/Carotenoids	0.3	–0.1	–4.4	0.4	0.0	–	–0.1

* $p < 0.05$, ** $p < 0.01$ – significant differences between each genotype in control and drought conditions; [a] $p < 0.05$; [aa] $p < 0.01$ – significant differences between CS and lines.

Gas exchange and chlorophyll fluorescence

Figure 2 shows SDE values for gas exchange parameters of CS and six secondary introgressive lines. Of all the studied genotypes, the recipient had the most stable gas exchange parameters, although its net photosynthesis rate significantly decreased under drought. The lines with introgression in the short arm of chromosome 2D showed similar changes for these traits. E and Gs significantly decreased, while net photosynthesis rate, to a lesser extent. As a result, WUE increased under drought. The lines with introgressions in the long arm of 2D chromosome showed various changes in gas exchange parameters. In line 1021, all gas exchange parameters, as well as WUE were significantly reduced. Lines 1022 and 1028 demonstrated atypical stomatal effects, increased E, Gs, and net photosynthesis rates under drought conditions. Line 1034 showed a classic adaptive response to water deficit, with decreased E and Gs, and stable net photosynthesis rate, resulting in a significant increase in WUE under drought.

To reveal the influence of introgressions on variability of Chl fluorescence parameters under water deficit conditions, we applied multidimensional nonmetric scaling of SDE indices for 39 Chl fluorescence parameters (Fig. S3). Three lines (1004, 1022, and 1028) formed a tight cluster with CS, indicating minor differences in SDE among these genotypes. This suggests that the existing introgression has a small effect on the structural and functional characteristics of photosynthetic apparatus under drought. Three other lines (1005, 1021, and 1034) were located at a significant distance from CS. This indicated significant differences in the responses of photosynthetic apparatus of these lines from others included in the cluster.

Next, we compared the size of drought effect on Chl fluorescence parameters in the recipient CS and lines 1004 and 1005 with introgression in the short arm of chromosome 2D (Fig. 3*a*) and in CS and lines 1021 and 1034 with introgressions in the long arm of the same chromosome (Fig. 3*b*). Figure 3 demonstrates that CS had relatively stable Chl fluorescence parameters under different irrigation conditions. The same is true for line 1004. Line 1005, on the contrary, demonstrated large differences for Chl fluorescence in the control and under drought indicating a stressed state of the photosynthetic apparatus under water deficiency. This conclusion follows from a statistically significant increase in F_0 and NPQ under drought, a decrease in F_v/F_m, Φ_{PSII}, F_v/F_0 and ETR, as well as the productivity indices PI_{abs} and PI_{tot}. Lines 1021 and 1034 with introgressions in the long arm of chromosome 2D differed significantly in their responses to drought (Fig. 3*b*). The photosynthetic apparatus of line 1034 adapted well to water deficit as indicated by an insignificant difference between the average values in the control and under drought and a zero value of SDE for PI_{tot}. In line 1021, Φ_{PSII}, qP, ETR and both productivity indices (PI_{abs} and PI_{tot}) significantly decreased under drought. These changes indicated a stressed state of the photosynthetic apparatus under water deficit.

Discussion

Effect of introgression from *Ae. tauschii* in the short arm of chromosome 2D on the stability of photosynthetic processes and shoot biomass. Line 1004 carried the introgression in chromosome 2D region flanked by markers *Xgwm296* and *Xgwm261*, which is limited by coordinates 2D:18085000–19623173 bp. Line 1005 carried the introgression in the region adjacent to marker *Xgwm261*. The lines differed substantially in the magnitude of SDE on chlorophyll and carotenoid content. This variability of photosynthetic pigments content and Chl *a*+*b*/Car ratio was more favorable for drought adaptation in line 1004 than in line 1005. In both lines, the size of PSII light-harvesting antenna increased

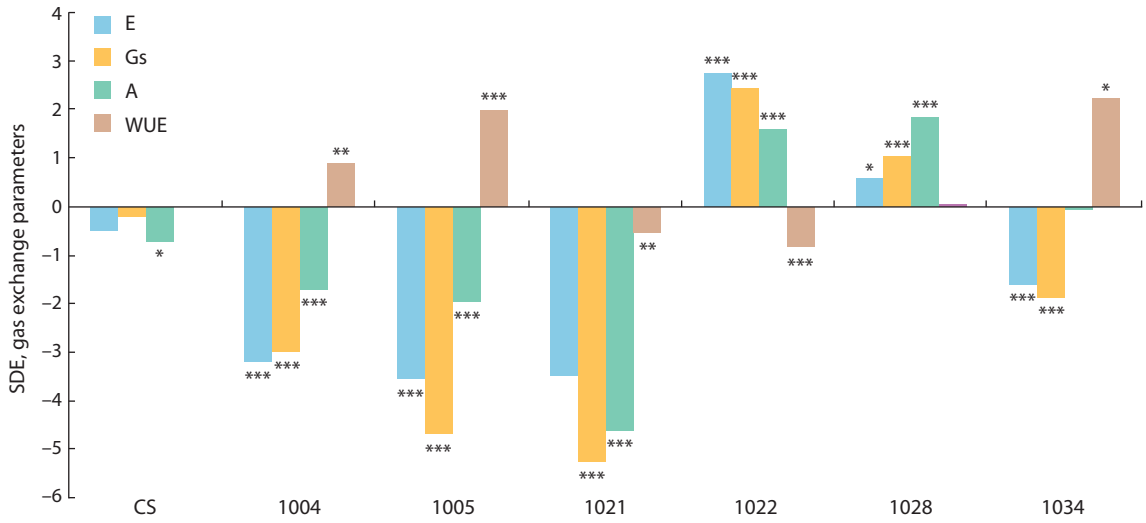


Fig. 2. Size of drought effect (SDE) on transpiration rate (E), stomatal conductance (Gs), photosynthetic rate (A) and water use efficiency (WUE) in CS and secondary recombinant introgressive lines.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ – significant differences between average values of the traits in control and drought conditions.

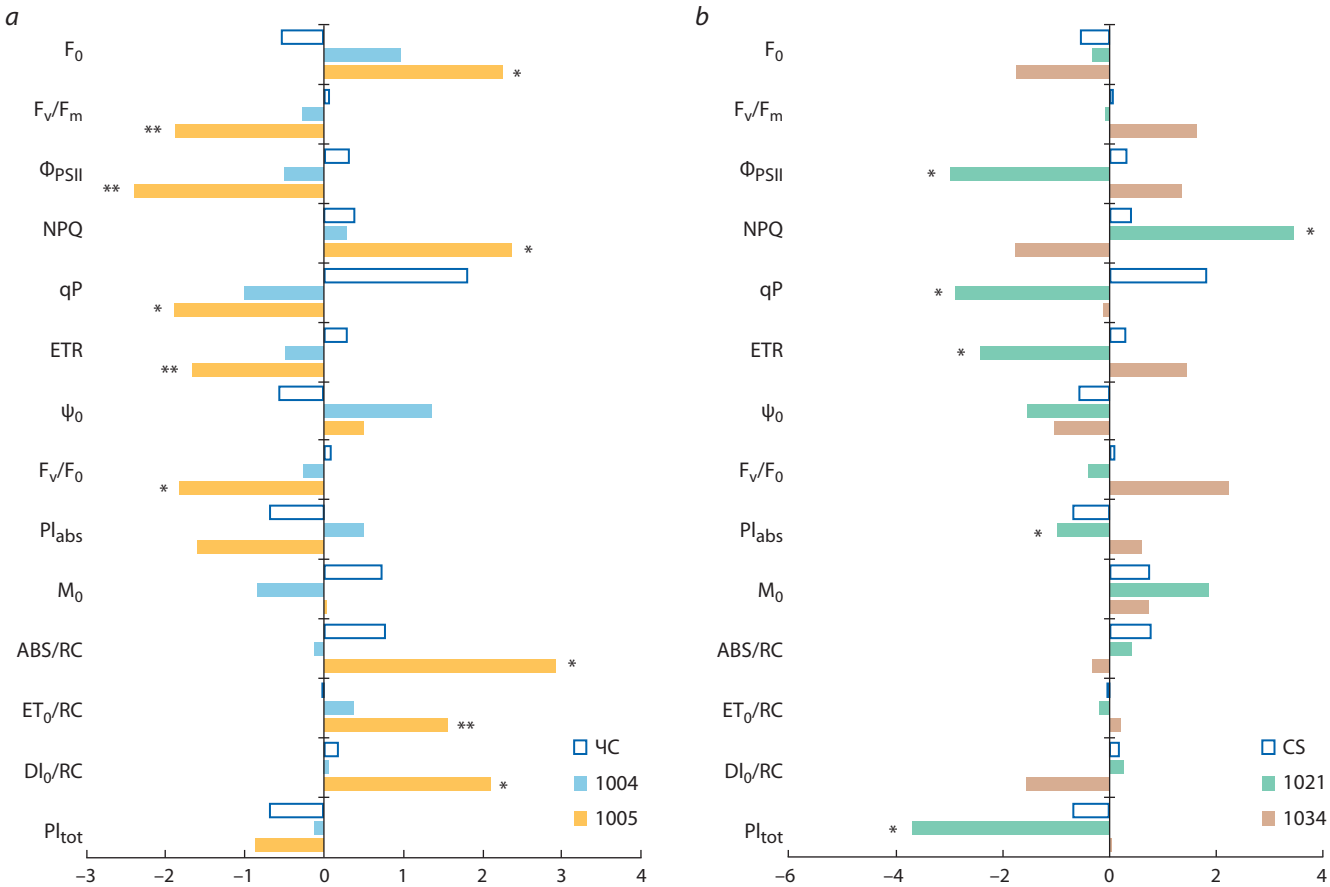


Fig. 3. Size of drought effect for Chl fluorescence parameters in CS and secondary introgressive lines 1004 and 1005 (*a*); 1021 and 1034 (*b*).

The Chl fluorescence parameters (after Goltsev et al., 2016): F_0 – the minimum fluorescence of dark-adapted leaves; F_v/F_m – the maximum photochemical activity of photosystem II (PSII); Φ_{PSII} – the effective quantum yield of PSII; NPQ – the non-photochemical fluorescence quenching; qP – the photochemical fluorescence quenching; ETR – the rate of linear electron transport through the photosystems; Ψ_0 – the efficiency with which an exciton captured by a reaction center moves an electron along the chain after QA; F_v/F_0 – the ratio of the rate constants of the primary photochemical reaction to the total rate of non-photochemical losses; PII_{abs} – an indicator of the functional activity of PSII related to the absorbed energy; M_0 – a parameter that reflects the rate of closure of the reaction centers of PSII; ABS/RC – the energy flow absorbed by one reaction center; DI_0/RC – the total amount of energy dissipated by one reaction center; PI_{tot} – an indicator of the functional activity of PSII, PSI and the electron transport chain between them.

* $p < 0.05$; ** $p < 0.01$, significant differences between average values of the traits in control and drought conditions.

during drought, but in line 1005, this adaptation did not result in efficient energy use. The energy flux absorbed by one PSII reaction center during drought (ABS/RC) increased in this line compared to the recipient CS, while energy dissipation (DI_0/RC parameter) increased. In line 1005, F_0 significantly increased under drought indicating disturbances in the excitation energy transfer in antenna and from the antenna to PSII reaction center (Goltsev et al., 2016). We observed a similar effect earlier on other wheat genetic material, in Saratovskaya 29 lines with modifications in the distal region of the short arm of chromosome 2A (Osipova et al., 2023), indicating the involvement of these chromosomal regions of the second homoeologous group in the control of primary photosynthesis processes.

The only difference between lines 1004 and 1005 was the introgression from *Ae. tauschii* in the region of marker *Xgwm296*, which was detected in line 1004 (Fig. 1). The genes associated with this marker probably determined the observed differences in drought tolerance of the two lines. The most likely candidate gene to explain them may be the gene *TraesCS2D03G0092600* encoding the plant-specific transcription factor (TF) TCP. The coordinates of this gene are 2D:18667052–18667318 bp (https://wheat.pw.usda.gov/cgi-bin/GG3). The TCP family of TFs regulate cell division, affect meristem growth (Cubas et al., 1999), and are involved in regulating responses to external signals (Danisman, 2016). The gene *TraesCS2D03G0092600* was likely involved in the formation of high shoot biomass, characteristic of line 1005 under optimal watering, an advantage that the line lost under drought. The same gene is presumably associated with large differences between lines 1004 and 1005 in Chl *a* accumulation under water deficit, since TCP TFs have been shown to regulate chlorophyll biosynthesis in *Arabidopsis* (Zhen et al., 2022). In general, the physiological differences between the two lines were that line 1005 exhibited an imbalance between growth and adaptation to water stress, while line 1004 remained relatively stable.

Effect of introgression from *Ae. tauschii* in the long arm of chromosome 2D on photosynthesis stability. We studied four lines with introgressions of different size in the region of the long arm of chromosome 2D, limited by markers *Xgwm157*, *Xgwm1419* and *Xgwm539*. Line 1034 with introgression in the region of markers *Xgwm1419* and *Xgwm157* and line 1021 with introgression in the region of marker *Xgwm539* were the most contrasting in stability. Line 1034 was distinguished by the stability of net photosynthesis and chlorophyll fluorescence indices, as well as by the relatively stable shoot biomass. In line 1021, on the contrary, the stability of photosynthetic parameters and shoot biomass were reduced compared to CS. The marker *Xgwm539* coordinates are 2D:515210161–515210309 bp. This region has a very high gene density. We believe that the most likely candidate gene for explaining the negative effect of introgression in line 1021 is the gene under the number *TraesCS2D03G008700* with coordinates 2D:515214093–515217180 bp (https://wheat.pw.usda.gov/cgi-bin/GG3). One of the two transcripts of this gene is annotated as corresponding to the homeodomain-like, Myb-containing protein. Its sequence is similar to that of the plant-specific GARP family of transcription factors (Hosoda et al., 2002). These proteins, including the Golden2-like proteins, play an important role throughout the plant's life cycle (Ohama, Yanagisawa, 2024). In particular, among other processes, they control the development of chloroplasts and determine the quantitative aspects of photosynthesis (Chen M. et al., 2016). Our results suggest that this gene plays an active role in adaptation of CS wheat to soil drought. The genetic modification of the chromosome segment in the region of its localization led to a significant decrease in the stability of photosynthesis and shoot biomass in line 1021. Golden2-like (GLK) transcription factors have recently been considered as potential candidates for improving photosynthesis in agricultural crops (Hernández-Verdeja, Lundgren, 2024). Our data support the idea that *GLK* genes may be a promising biotechnological tool for improving drought tolerance in bread wheat, if the donor genotype is properly selected.

Three lines (1022, 1028 and 1034) had segments from *Ae. tauschii* in the region of marker *Xgwm1419* in chromosome 2D. Additionally, line 1034 had introgression in the region of marker *Xgwm157* (Fig. 1). According to GrainGenes data, genes functionally significant for drought tolerance are not localized in the region of this marker. This is probably why lines 1028 and 1034 were similar in terms of stability of the shoot biomass. At the same time, lines 1022 and 1028 differed from lines 1034 and 1021 in the response of stomatal apparatus to water deficit. We suggest that this phenomenon is associated with the gene *TraesCS2D03G0081400* (2D:494675291–494678461 bp), localized relatively close to marker *Xgwm1419*. This gene encodes a protein, a member of the GTL1 family of transcription factors. GTL1 is known to be involved in the regulation of stomatal density, transpiration, stomatal conductance and, as a consequence, affects water use efficiency (Yoo et al., 2011). In addition, using RT-PCR, its significant expression was shown in many organs of wheat plants at the flowering stage, as well as an immediate (within 3 hours) response to osmotic stress (Zheng et al., 2016). The increase in transpiration and stomatal conductance in lines 1022 and 1028 and the decrease in WUE, especially in line 1022, could be associated with a change in the functionality of the *GTL1* gene.

Lines 1021 and 1034 differed contrastingly for the stability of chlorophyll fluorescence indices (Fig. 3). In line 1021, unlike 1034, the real efficiency of PSII, the rate of electron transport, the index of functional activity of PSII (PI_{abs}), and the integral index of functional activity of PSI and PSII (PI_{tot}) decreased under water deficiency. These differences are presumably due to introgression from *Ae. tauschii* into chromosome 2D. Line 1034 had the introgression in the region of marker *Xgwm1419*, which is located in coordinates 2D:472226450–472226470 bp, close to the *TraesCS2D03G0058100* gene (coordinates 2D:480941598–481111682 bp), encoding the PsbQ protein. The functions of this protein are associated with the coordination of the activities of the donor and acceptor functions of PSII and the stabilization of the active form of the light-harvesting complex of PSII (Ifuku et al., 2011). Introgression from *Ae. tauschii* in the region of marker *Xgwm1419* could have a positive effect on the functioning of photosystem II, which makes the main contribution to chlorophyll fluorescence.

Conclusion

A comparative study of the stability of photosynthesis and shoot biomass in wheat variety CS and secondary introgressive lines CS(Syin6x) for chromosome 2D showed significant diversity in these traits. Considering that the genotypes were grown under controlled conditions, the found differences in soil drought tolerance are presumably associated with introgressions from *Ae. tauschii*. Based on the results of the study, it can be concluded that the single introgression into the short arm of chromosome 2D limited by molecular markers *Xgwm296* and *Xgwm261* was favorable for drought tolerance. Introgression into the long arm of the same chromosome in the region of marker *Xgwm1419* also supported drought tolerance. Introgression in this chromosome arm restricted by marker *Xgwm539* was unfavorable for photosynthetic stability and shoot biomass. Markers *Xgwm296* and *Xgwm1419* can be recommended for the use in marker-assisted breeding of wheat for drought tolerance in cases where *Ae. tauschii* is used as a donor of genetic material

References

Botyanszka L., Zivcak M., Chovancek E., Sytar O., Barek V., Hauptvogel P., Halabuk A., Brestic M. Chlorophyll fluorescence kinetics may be useful to identify early drought and irrigation effects on photosynthetic apparatus in field-grown wheat. *Agronomy.* 2020;10(9):1275. doi 10.3390/agronomy10091275

Chen K., Chen L., Fan J.B., Fu J. Alleviation of heat damage to photosystem II by nitric oxide in tall fescue. *Photosynth Res.* 2013;116(1):21-31. doi 10.1007/s11120-013-9883-5

Chen M., Ji M., Wen B., Liu L., Li S., Chen X., Gao D., Li L. GOLDEN 2-LIKE transcription factors of plants. *Front Plant Sci.* 2016;7:1509. doi 10.3389/fpls.2016.01509

Cubas P., Lauter N., Doebley J., Coen E. The TCP domain: a motif found in proteins regulating plant growth and development. *Plant J.* 1999;18(2):215-222. doi 10.1046/j.1365-313x.1999.00444.x

Danisman S. TCP transcription factors at the interface between environmental challenges and the plant's growth responses. *Front Plant Sci.* 2016;7:1930. doi 10.3389/fpls.2016.01930

Goltsev V.N., Kalaji H.M., Paunov M., Bąba W., Horaczek T., Mojski J., Kocie H., Allakhverdiev S.I. Variable chlorophyll fluorescence and its use for assessing physiological condition of plant photosynthetic apparatus. *Russ J Plant Physiol*. 2016;63(6):869-893. doi 10.1134/S1021443716050058

Hammer Ø., Harper D.A.T., Ryan P.D. PAST: PAleontological STatistics software package for education and data analysis. *Palaeontol Electronica*. 2001;4(1):1-9. https://palaeo-electronica.org/2001_1/past/past.pdf

Hedges L.V., Olkin I. Estimation of a single effect size: parametric and nonparametric methods. In: Statistical Methods for Meta-Analysis. Amsterdam: Elsevier Academic Press, 1985;75-106. doi 10.1016/B978-0-08-057065-5.50010-5

Hernández-Verdeja T., Lundgren M.R. GOLDEN2-LIKE transcription factors: a golden ticket to improve crops? *Plants People Planet.* 2024;6(1):79-93. doi 10.1002/ppp3.10412

Hosoda K., Imamura A., Katoh E., Hatta T., Tachiki M., Yamada H., Mizuno T., Yamazaki T. Molecular structure of the GARP family of plant Myb-related DNA binding motifs of the Arabidopsis response regulators. *Plant Cell.* 2002;14(9):2015-2029. doi 10.1105/tpc.002733

Ifuku K., Ido K., Sato F. Molecular functions of PsbP and PsbQ proteins in the photosystem II supercomplex. *J Photochem Photobiol B.* 2011;104(1-2):158-164. doi 10.1016/j.jphotobiol.2011.02.006

IWGSC. Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science*. 2018;361(6403):eaar7191. doi 10.1126/science.aar7191

Jia J., Zhao S., Kong X., Li Y., Zhao G., He W., Appels R., … Yang H., Liu X., He Z., Mao L., Wang J. *Aegilops tauschii* draft genome sequence reveals a gene repertoire for wheat adaptation. *Nature.* 2013;496(7443):91-95. doi 10.1038/nature12028

Ma F., Li R., Guo G., Nie F., Zhu L., Liu W., Lyu L., Bai S., Zhao X., Li Z., Zhang D., Li H., Li S., Zhou Y., Song C.-P. Introgression of QTL from *Aegilops tauschii* enhances yield-related traits in common wheat. *Crop J.* 2023;11(5):1521-1532. doi 10.1016/j.cj.2023.05.001

McFadden E.S., Sears E.R. The origin of *Triticum spelta* and its free threshing hexaploid relative. *J Hered.* 1946;37(3):81-89. doi 10.1093/oxfordjournals.jhered.a105590

Nicholson P., Rezannor H.N., Worland A.J. Chromosomal location of resistance to *Septoria nodorum* in a synthetic hexaploid wheat determined by the study of chromosomal substitution lines in 'Chinese Spring' wheat. *Plant Breed.* 1993;110(3):177-184. doi 10.1111/j.1439-0523.1993.tb00575.x

Nyine M., Adhikari E., Clinesmith M., Aiken R., Betzen B., Wang W., Davidson D., Yu Z., Guo Y., He F., Akhunova A., Jordan K.W., Fritz A.K., Akhunov E. The haplotype-based analysis of *Aegilops tauschii* introgression into hard red winter wheat and its impact on productivity traits. *Front Plant Sci.* 2021;12:716955. doi 10.3389/fpls.2021.716955

Ohama N., Yanagisawa S. Role of GARP family transcription factors in the regulatory network for nitrogen and phosphorus acquisition. *J Plant Res.* 2024;137(3):331-341. doi 10.1007/s10265-023-01513-0

Osipova S., Permyakov A., Permyakova M., Pshenichnikova T., Verkhoturov V., Rudikovsky A., Rudikovskaya E., Shishparenok A., Doroshkov A., Börner A. Regions of the bread wheat D genome associated with variation in key photosynthesis traits and shoot biomass under both well-watered and water deficient conditions. *J Appl Genet.* 2016;57:151-163. doi 10.1007/s13353-015-0315-4

Osipova S., Permyakov A., Konstantinov D., Shchukina L., Rudikovskaya E., Permyakova M., Pshenichnikova T. Variability of photosynthesis parameters and yield in recombinant lines of bread wheat with introgressions from *Triticum timopheevii* into 2A chromosome under different water supply conditions. *Cereal Res Commun.* 2023;52:101-113. doi 10.1007/s42976-023-00372-8

Osipova S.V., Rudikovskii A.V., Permyakov A.V., Rudikovskaya E.G., Pomortsev A.V., Musalevskaya O.V., Pshenichnikova T.A. Using chlorophyll fluorescence parameters and antioxidant enzyme activity to assess drought tolerance of spring wheat. *Photosynthetica.* 2024;62(2):147-157. doi 10.32615/ps.2024.014

Peršić V., Ament A., Antunović Dunić J., Drezner G., Cesar V. PEG-induced physiological drought for screening winter wheat genotypes sensitivity – integrated biochemical and chlorophyll *a* fluorescence analysis. *Front Plant Sci.* 2022;13:987702. doi 10.3389/fpls.2022.987702

Pestsova E.G., Börner A., Röder M.S. Development of a set of *Triticum aestivum-Aegilops tauschii* introgression lines. *Hereditas.* 2001;135(2-3):139-143. doi 10.1111/j.1601-5223.2001.00139.x

Plaschke J., Ganal M.W., Röder M.S. Detection of genetic diversity in closely related bread wheat using microsatellite markers. *Theor Appl Genet.* 1995;91:1001-1007. doi 10.1007/BF00223912

Pour-Aboughadareh A., Kianersi F., Poczai P., Moradkhani H. Potential of wild relatives of wheat: ideal genetic resources for future breeding programs. *Agronomy.* 2021;11(8):1656. doi 10.3390/agronomy11081656

Przewieslik-Allen A.M., Burridge A.J., Wilkinson P.A., Winfield M.O., Shaw D.S., McAusland L., King J., King I.P., Edwards K.J., Barker G.L.A. Developing a high-throughput SNP-based marker system to facilitate the introgression of traits from *Aegilops* species into bread wheat (*Triticum aestivum*). *Front Plant Sci.* 2019;9:1993. doi 10.3389/fpls.2018.01993

Reynolds M., Foulkes J., Furbank R., Griffiths S., King J., Murchie E., Parry M., Slaffer G. Achieving yield gains in wheat. *Plant Cell Environ.* 2012;35(10):1799-1823. doi 10.1111/j.1365-3040.2012.02588.x

Röder M.S., Korzun V., Wendehake K., Plaschke J., Tixier M.H., Leroy P., Ganal M.W. A microsatellite map of wheat. *Genetics*. 1998;149(4):2007-2023. doi 10.1093/genetics/149.4.2007

Srivastava A., Biswas S., Yadav S., Kumar A., Rajaram H., Srivastava V., Mishra Y. Physiological and thylakoid proteome analyses of *Anabaena* sp. PCC 7120 for monitoring the photosynthetic responses under cadmium stress. *Algal Res*. 2021;54:102225. doi 10.1016/j.algal.2021.102225

Wettstein D. Chlorophyll-letale und der submikroskopische form wechsel der plastiden. *Exp Cell Res.* 1957;12(3):427-506. doi 10.1016/0014-4827(57)90165-9

Yoo C.Y., Hasegawa P.M., Mickelbart M.V. Regulation of stomatal density by the GTL1 transcription factor for improving water use efficiency. *Plant Signal Behav*. 2011;6(7):1069-1071. doi 10.4161/psb.6.7.15254

Zheng X., Liu H., Ji H., Wang Y., Dong B., Qiao Y., Liu M., Li X. The wheat GT factor TaGT2L1D negatively regulates drought tolerance and plant development. *Sci Rep.* 2016;6:27042. doi 10.1038/srep27042

Zheng X., Lan J., Yu H., Zhang J., Zhang Y., Qin Y., Su X.-D., Qin G. *Arabidopsis* transcription factor TCP4 represses chlorophyll biosynthesis to prevent petal greening. *Plant Commun*. 2022;3(4):100309. doi 10.1016/j.xplc.2022.100309

Conflict of interest. The authors declare no conflict of interest.
Received July 25, 2024. Revised February 21, 2025. Accepted March 2, 2025.

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2025;29(4):539-548

doi 10.18699/vjgb-25-57

Cytophysiological manifestations of wheat's defense reactions against stem rust induced by the biofungicide Novochizol

A.B. Shcherban [1, 2], L.Ya. Plotnikova [3] ✉, V.V. Knaub [3], E.S. Skolotneva [1, 2], V.V. Fomenko [4]

[1] Kurchatov Genomic Center of ICG SB RAS, Novosibirsk, Russia
[2] Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
[3] Omsk State Agrarian University named after P.A. Stolypin, Omsk, Russia
[4] N.N. Vorozhtsov Novosibirsk Institute of Organic Chemistry of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
✉ lya.plotnikova@omgau.org

Abstract. Biologization is a priority direction of agricultural production. One of the promising approaches to solve the biologization problem is the use of chitosan-based biopreparations to stimulate plant growth and protect plants from a wide range of pathogens. Currently, active work is underway to create and test new chitosan preparations. Novochizol was obtained as a result of intramolecular crosslinking of linear chitosan molecules and has a globular shape. Previously, a Novochizol-stimulating effect on the growth and development of common wheat was demonstrated. However, the induced resistance mechanisms against rust diseases have not been studied before. The reported studies have revealed the dose effect of the preparation on the development of wheat stem rust. The best results of visual estimation of plant reactions were obtained with 0.125 and 0.75 % Novochizol pretreatment four days before rust infection. After pretreatment of susceptible cv. Novosibirsk 29 seedlings, a resistant reaction appeared and the urediniopustule density was decreased. Cytophysiological studies have shown that 0.75 % Novochizol stimulated an intensive accumulation of hydrogen peroxide H_2O_2 in the leaves of the infected and healthy plants within 48 hours post inoculation (h p/in). During the period of 48–144 h p/in, H_2O_2 gradually disappeared from tissues, but its content increased significantly at the sporulation stage around pustules. However, Novochizol did not induce the hypersensitivity reaction in infected plants. The preparation induced an earlier and more intensive (compared with untreated plants) accumulation of phenolic substances with different autofluorescence in the zones around pathogen colonies. Novochizol induced a change in the ratio of phenols with different spectral characteristics towards compounds with an increased content of syringin derivatives. This work is the first stage in the study of Novochizol effects on wheat defense mechanisms against stem rust. The research will be continued using molecular genetics, biochemical and cytophysiological methods.
Key words: biopesticides; Novochizol; common wheat; stem rust; resistance mechanisms; ROS; phenols

For citation: Shcherban A.B., Plotnikova L.Ya., Knaub V.V., Skolotneva E.S., Fomenko V.V. Cytophysiological manifestations of wheat's defense reactions against stem rust induced by the biofungicide Novochizol. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov J Genet Breed*. 2025;29(4):539-548. doi 10.18699/vjgb-25-57

Funding. The work was supported by Russian Science Foundation (project No. 23-16-00119, https://rscf.ru/project/23-16-00119/).

Transparency of financial activities. The authors have no financial interest in the submitted materials or methods.

Цитофизиологические проявления защитных реакций пшеницы от стеблевой ржавчины, индуцируемые биофунгицидом Новохизолем

А.Б. Щербань [1, 2], Л.Я. Плотникова [3] ✉, В.В. Кнауб [3], Е.С. Сколотнева [1, 2], В.В. Фоменко [4]

[1] Курчатовский геномный центр ИЦиГ СО РАН, Новосибирск, Россия
[2] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
[3] Омский государственный аграрный университет им. П.А. Столыпина, Омск, Россия
[4] Новосибирский институт органической химии им. Н.Н. Ворожцова СО РАН, Новосибирск, Россия
✉ lya.plotnikova@omgau.org

Аннотация. Биологизация земледелия считается приоритетным направлением сельскохозяйственного производства. Одним из перспективных подходов к решению задачи биологизации является применение препаратов на основе хитозана для стимуляции роста и защиты растений от широкого круга патогенов. В настоящее время проводятся активные работы по созданию и испытанию новых форм хитозановых препаратов. Препарат «Новохизоль» получен в результате внутримолекулярных сшивок линейных молекул хитозана и имеет глобулярную

© Shcherban A.B., Plotnikova L.Ya., Knaub V.V., Skolotneva E.S., Fomenko V.V., 2025
This work is licensed under a Creative Commons Attribution 4.0 License

форму. Ранее установлено стимулирующее влияние Новохизоля на рост и развитие мягкой пшеницы, однако индуцируемые защитные механизмы против ржавчинных болезней не изучались. Проведенные исследования показали дозовый эффект препарата на развитие стеблевой ржавчины пшеницы. При обработке за четверо суток до заражения лучшие результаты по развитию устойчивой реакции растений, сокращению числа и размеров пустул были получены с Новохизолем в концентрации 0.125 и 0.75 %. После предобработки на проростках восприимчивого сорта Новосибирская 29 проявилась устойчивая реакция и снизилось число пустул. Цитофизиологические исследования показали, что обработка 0.75 % Новохизолем стимулировала интенсивное накопление пероксида водорода H_2O_2 в листьях инфицированных и здоровых растений в течение 48 ч после инокуляции. В период 48–144 ч после инокуляции H_2O_2 постепенно исчезал из тканей, но на стадии спороношения его содержание значительно возрастало в зоне колоний и пустул. Новохизоль не индуцировал развитие реакции сверхчувствительности в зараженных растениях. Применение препарата способствовало более раннему и интенсивному (по сравнению с необработанными растениями) накоплению фенольных веществ с разным спектром автофлуоресценции в зоне колоний патогена. Препарат повлиял на изменение соотношения фенолов с разными спектральными характеристиками в сторону соединений с повышенным содержанием остатков сирингина. Данная работа является первым этапом изучения действия Новохизоля на защитные механизмы пшеницы против стеблевой ржавчины. Исследования будут продолжены с применением молекулярно-генетических и биохимических методов.

Ключевые слова: биопестициды; Новохизоль; мягкая пшеница; стеблевая ржавчина; механизмы устойчивости; АФК; фенолы

Introduction

Due to the proposed rise in the world's population to 9.5 billion people by 2050, it is necessary to increase grain production by 1.7 times (USDA, 2016). An increase in wheat grain harvests can be achieved by breeding more productive and stress-resistant varieties, as well as reducing losses caused by abiotic and biotic factors. Synthetic pesticides are traditionally widely used to protect crops from diseases and pests. These protective agents are highly effective; however, they can be accumulated in plants and soils, having a negative effect on the ecological situation in agrocenoses and product quality (Sternshis et al., 2016). The use of biological pest management agents (BPMA) increases stress resistance mechanisms (Chandler et al., 2011).

BPMA based on natural compounds and beneficial microorganisms attract the attention of researchers and practitioners. These substances are often close to chemical pesticides in effectiveness, but do not have their disadvantages (Chakraborty et al., 2020). The range of biopesticides and their application schemes are very diverse, which is determined by the pathogens and pests' biology, as well as their interaction with plants. BPMA may inhibit the pathogens and pests directly or induce a complex of plant resistance reactions (Orzali et al., 2017; Yarullina et al., 2023).

Chitin and chitosan derivatives are widely used as BPMA (Tyuterev, 2015; Malerba, Cerana, 2016). Polymer carbohydrate chitin is widespread in nature, as components of integuments of arthropods (including crustaceans and insects) and fungi. Chitosan is produced by chitin hydrolysis and deacetylation. Chitosan-based preparations have a stimulating effect on plant growth and development, as well as enhance resistance to abiotic stresses (Haggag et al., 2014; Orzali et al., 2017). Chitosan derivatives are also of particular interest as inducers of resistance to fungal, bacterial, and viral diseases (Chakraborty et al., 2020; Shcherban, 2023).

Chitosan preparations may differ in their main characteristics: molecular weight, deacetylation degree, and polydispersity index (Richter et al., 2012). The effectiveness of chitosan derivatives can be significantly enhanced by their modification, such as the introduction of functional groups of Schiff bases, halogen atoms (Cl or F), metal nanoparticles, urea groups, etc. (Varlamov et al., 2020; Yarullina et al., 2023). Preparations based on conjugates of chitosan with phenolic hydroxycinnamic acids (ferulic and caffeic) have proven promising for protecting plants from fungal and viral diseases (Rkhaila et al., 2021; Yarullina et al., 2024a). A positive effect of combining chitosan preparations with other biologically active substances and beneficial microorganisms (Plant Growth Promoting Bacteria, PGPB) has been established (Rkhaila et al., 2021; Yarullina et al., 2024b). The protective effects strengthening is due to the synergistic action of different drug components (Tyuterev, 2015). Currently, a wide range of chitosan-based BPMA have been created in the world and their tests have been carried out on various cultures. A comparison of the results showed that their stimulating and protective effects depended on preparation compositions, as well as plant and pathogen species (Rabea et al., 2005; Orzali et al., 2017).

A number of complex chitosan preparations with the addition of biologically active substances have been developed in Russia, including "Narcissus" with succinic and glutamic acids; "Chitosar M" with salicylic acid (SA); "Chitosar F" with arachidonic acid; an agent with SA and vanillin, etc. (Tyuterev, 2015; Popova et al., 2018). The combined agents were effective against different pathogenic fungi, viruses and pests. Their application enhanced crop resistance to diseases, such as that of wheat to leaf rust, spot blotch and root rot; rice, to *Pyricularia*; tomatoes, to late blight and Fusarium fruit rot; potatoes, to late blight and Y virus; cucumbers, to downy mildew, etc. (Tyuterev, 2015; Badanova et al., 2016; Popova et al., 2018).

A promising new chitosan derivative is "Novochizol", obtained by intramolecular crosslinking of linear chitosan molecules. Novochizol has a globular shape, which gives it a number of advantages over chitosan, namely increased solubility in aqueous solutions, chemical stability, resistance

to biodegradation, high adhesion and ability to penetrate tissues. This form is able to absorb various substances and slowly release them into plants after application (Novochizol SA, www.novochizol.ch). These properties are important for creating promising combined agents with other biologically active substances. Novochizol has a growth-stimulating effect when processing seeds and leaves. It was shown that this substance enhanced common wheat seed germination, contributed to an increase in root and total plant weight (Teplyakova et al., 2022). The effectiveness of complex Novochizol preparations with usnic acid or Siberian pine bark extract for protecting wheat from root rot and *Septoria blotch* was proved in the field (Burlakova et al., 2025).

It is known that after plant recognition of non-specialized or avirulent pathogen effectors (elicitors), a set of defence reactions is activated. The earliest responses include the reactive oxygen species (ROS) and nitric oxide NO generation (Manjunatha et al., 2009; Singh et al., 2021; Plotnikova, Knaub, 2024). ROS ($O_2^{\bullet-}$, H_2O_2, $^{\bullet}OH$, $^1O_2$) accumulation leads to a splash of oxidative reactions, called an oxidative burst. The enzyme superoxide dismutase (SOD) converts the superoxide anion $O_2^{\bullet-}$ into the hydrogen peroxide H_2O_2 (Maksimov, Cherepanova, 2006). H_2O_2 has a toxic effect on pathogens, and is a messenger in NADP·H-oxidase signaling system implemented through a SA-dependent signaling cascade (Tarchevsky, 2000; Yarullina et al., 2023). As a result of SA-dependent cascade action, a complex of resistance mechanisms against biotrophic pathogens is implemented in the infection zone, including ROS generation, hypersensitive reaction (HR), defence PR proteins (Pathogenesis-Related Proteins) and phenolic substances synthesis. Defence reactions against necrotrophic pathogens are realized using a signaling cascade dependent on jasmonic acid (JA), abscisic acid and ethylene. The resistance to hemibiotrophs is ensured by the combined action of the SA- and JA-dependent cascades (Singh et al., 2021; Yarullina et al., 2023). The study of the chitosans' effects on defence reactions showed activation of the ROS and phenolic metabolism enzymes, PR proteins accumulation and cell wall strengthening with the lignin and callose (Orzali et al., 2017; Shcherban, 2023).

To develop BPMA technology, it is necessary to learn their effect on defence mechanisms and the development of the most devastating diseases. Novochizol action on wheat resistance mechanisms against rust diseases has not been studied before. The aim of the work was to study the Novochizol effect on the defence mechanisms of a susceptible common wheat variety infected with the stem rust fungus *Puccinia graminis* f. sp. *tritici* Erikss. et Henn.

Materials and methods

Plant material. The objects of the research were 10-day-old seedlings of the spring common wheat cv. Novosibirskaya 29 susceptible to stem rust. Plants were grown in pots with soil as recommended for experiments with rust fungi by international protocols (Woldeab et al., 2017). The seedlings were treated with Novochizol solutions at concentrations of 0.125, 0.75, 1.5, and 2.5 %. Solutions were applied to plants (15 ml per 100 plants) using a sprayer four days before infection with stem rust. Such a pretreatment period is sufficient to induce defensive effects by BPMA, including chitosan derivatives, against oomycetes and rust fungi (Faoro et al., 2008; Bellameche et al., 2021; Elsharkawy et al., 2022). Plants treated with bidistilled water served as a control.

The seedlings were inoculated with urediniospores of a mixed sample of the West Siberian population of *P. graminis* f. sp. *tritici* (*Pgt*), included isolates with avirulence/virulence genes to wheat genes *Sr11Sr24Sr30Sr31*/*Sr5Sr9aSr9b Sr9dSr9gSr10Sr17Sr38SrMcN*. The urediniospores were stored at –70 °C before the experiment and revitalized using susceptible common wheat cv. Khakasskaya (Rsaliyev A.S., Rsaliyev Sh.S., 2018). Urediniospore suspension at the concentration of 0.8 mg/ml Novec 7100 (Sørensen et al., 2016) was applied to seedlings using a sprayer. Inoculated plants were incubated for 24 h in a humid chamber in the dark at a temperature of 15–20 °C for maximal spore germination. After that, the plants were transferred to growth chambers and incubated under 16 h illumination with an intensity of 10,000 lux at a temperature of 26–28 °C. Such temperature is critical for full appressoria structure formation and pathogen penetration into the stomata, and infection hyphae development in the plant tissue (Roelfs et al., 1992).

Phytopathological assessment of plant reaction to infection. The effect of Novochizol was assessed by quantitative and qualitative characteristics used to describe the resistance of wheat seedlings to stem rust, such as pustule density (number per leaf, 10 plants per variant) and reaction type. Plant reaction (infection type, IT) was determined 12–14 days post inoculation (p/in) using a modified Stackman scale. The ITs "0", ";", "1", and "2" were interpreted as resistant (R), and "3", "3+" and "4", as susceptible (S) (Roelfs et al., 1992).

Cytological and cytochemical methods. The studies were carried with plants treated with 0.75 % Novochizol. The material was fixed at 0, 24, 96, 144 and 240 h p/in in lactophenol fixative (phenol, lactic acid, glycerin, distilled water, 96 % ethanol, in the ratio of 1 : 1 : 1 : 1 : 8) (Plotnikova, Meshkova, 2009). Infection structures on the surface and in plant tissues were detected using the fluorescent dye Uvitex 2B (Sigma-Aldrich, USA) by a modified method (Moldenhauer et al., 2006). For this, the material fixed in lactophenol was washed with distilled H_2O, and afterwards it was kept for 3 h in acetic alcohol (96 % ethanol and glacial acetic acid, in the ratio of 3 : 1). After washing with distilled H_2O, the leaf pieces were kept in a series of liquids, such as 50 % ethanol (20 min), 0.5N NaOH (30 min), distilled H_2O (5 min), 0.1M Tris-HCl buffer pH 5.8 (30 min), and distilled H_2O (5 min). Staining was carried out for 15 minutes in 0.1 % Uvitex 2B in 0.1M Tris-HCl (pH 5.8), preheated at 60 °C. To differentiate the colour, the material was kept in distilled water for 90 min. The observations were carried out in reflected light with an excitation wave of λ_{max} = 355 nm and an emission wave of λ_{max} = 420 nm. Undamaged fungal structures showed a blue fluorescence, the damaged plant cells and pathogen hyphae were light blue or white.

For hydrogen peroxide H_2O_2 localization in tissues, a vital staining of the material with 0.02 % 3,3′-diaminobenzidine tetrachloride (DAB, Sigma-Aldrich, USA) was implemented

before fixation (Plotnikova, Meshkova, 2009). The DAB solution was infused into the leaves by vacuum infiltration and incubated for 30 min. Insoluble cherry formazane was formed in the presence of H_2O_2.

Phenolic substances distribution in the leaves was studied using special reaction with aniline sulfate to common phenols (low molecular weight phenols and polymer lignin). The material was stained with 1 % aniline sulfate (aniline sulfate, glacial acetic acid, and 50 % ethyl alcohol, in the ratio of 1 : 2 : 97) for 1 h, followed by washing in distilled water (Japaridze, 1953). The lignins in the veins and plant cell walls in the infection zones were coloured yellow-brown. Additionally, the phenols autofluorescence in reflected light with an excitation wave of λ_{max} = 355 nm and emission of λ_{max} = 530 nm (green fluorescence) or λ_{max} = 605 nm (red fluorescence) was studied (Plotnikova, Meshkova, 2009). Cytological studies were carried out using an ARSTEK E62 light microscope (ARSTEK, China) with a Sony Alpha A6400 APS-C digital camera (the resolution of 24.2 MP/inch, Sony, Japan).

The results of 30–50 vbg*Pgt* urediniospores development in each of the five plants per variant were studied at each experiment stage, and were counted as repetitions. The areas of mycelium and urediniopustule (35–50 pcs. per variant) were measured after 240 h p/in, using the camera software. The mean values and standard errors were determined (in tables and graphs), and the least significant difference at $p \leq 0.05$ ($LSD_{0.05}$) was calculated.

Results

Visual assessment of the Novochizol effect on stem rust development

At the first stage of the work, the effect of different Novochizol concentrations on the disease development in the susceptible cv. Novosibirskaya 29 seedlings was studied. A wide range of preparation concentrations was used in the experiments, from 0.125 to 2.5 %. The pustules with IT "4" were formed on plants treated with water (control). Any Novochizol concentration influenced the disease development. It could be seen in the decrease of pustule density, pustule size reduction, and chlorosis appearance around the pustules (Table 1). IT decreased to the least extent when 2.5 % Novochizol solution was used (IT "3", "3+"). The treatments with 0.125, 0.75 and 1.5 % concentrations induced resistant reactions. The pustules sizes decreased to the greatest extent when the plants were treated with 0.75 % Novochizol (IT "2", "2–"). This experimental variant was used for studying plant defence reactions.

Results of cytophysiological studies of the Novochizol effect on pathogenesis

The Novochizol effects were assessed by *Pgt* development on the leaf surface and in the tissues, and plant reactions in the infection zone. After contact with the moistened plant surface, the urediniospores swelled and formed growing tubes (Fig. 1*a*). The appressoria were formed at the ends of most growing tubes, which were necessary for penetration into the stomata (Fig. 1*b*). A big part of the appressoria (73–78 %) were located on the stomata, and more than 93 % of them ensured pathogen penetration into the tissues. No significant differences in the development of *Pgt* on the surfaces of untreated and Novochizol-treated plants have been established (Table 2). The main appressoria proportion was formed 18–24 h p/in. After penetration into the stoma, the fungus formed infection hyphae with haustorial mother cells (Fig. 1*c*), and the first haustoria in mesophyll cells were formed 24–48 h p/in. *Pgt* formed large pustules with the next urediniospore generation 240 h p/in (Fig. 1*d*).

The localization of hydrogen peroxide and phenolic compounds in the leaves was studied to determine active reactions. A high H_2O_2 concentration was revealed in leaf cuts by DAB staining at the beginning of the experiment in each variant (control untreated uninfected, *Pgt*-infected, Novochizol-treated, and Novochizol-treated infected plants) (Fig. 1*e*, *n*). In areas far from the cuts, DAB staining was weak (Fig. 1*e*). At the end of the experiment, DAB staining decreased significantly at the ends of all leaves (Fig. 1*f*, *k*). Probably, H_2O_2 generation at leaf ends was associated with plant stress reaction to mechanical damage.

The distribution of total phenols in the leaves was firstly studied using a special aniline sulphate staining. The phenols were detected in the cytoplasm and plant cell walls in the zone of *Pgt* development, as well as in the vein cell walls, which corresponds to the presence of polymer lignin. The phenols were low in other leaf parts (Fig. 1*g*). Phenol autofluorescence coincided with their localization, determined by aniline sulphate staining. Under different observation modes, a bright green or red fluorescence appeared (with emission at λ_{max} = 530 nm or λ_{max} = 605 nm, respectively). Different fluorescence colour is associated with the presence of different phenol compounds. In the control plants, red fluorescence was brighter in the midveins, and green fluorescence was more active in small veins, in particular, in the walls of stomatal guard and mesophyll cells (Fig. 1*h*, *i*).

In the untreated plants, significant changes in the cells in the infection zones were not found during pathogenesis, up to the

Table 1. Results of a visual assessment of the Novochizol concentration effect on the development of *P. graminis* f. sp. *tritici* in wheat seedlings

Indicator	Control	Novochizol concentration, %			
		0.125	0.75	1.5	2.5
Reaction, IT	4	2	2, 2–	2, 2+	3, 3+
Average pustule number, pcs./leaf	20.4 ± 0.54	8.3 ± 0.35*	18.1 ± 0.32*	15.1 ± 0.28*	14.3 ± 0.28*

* Significant differences with the control at $p \leq 0.05$.

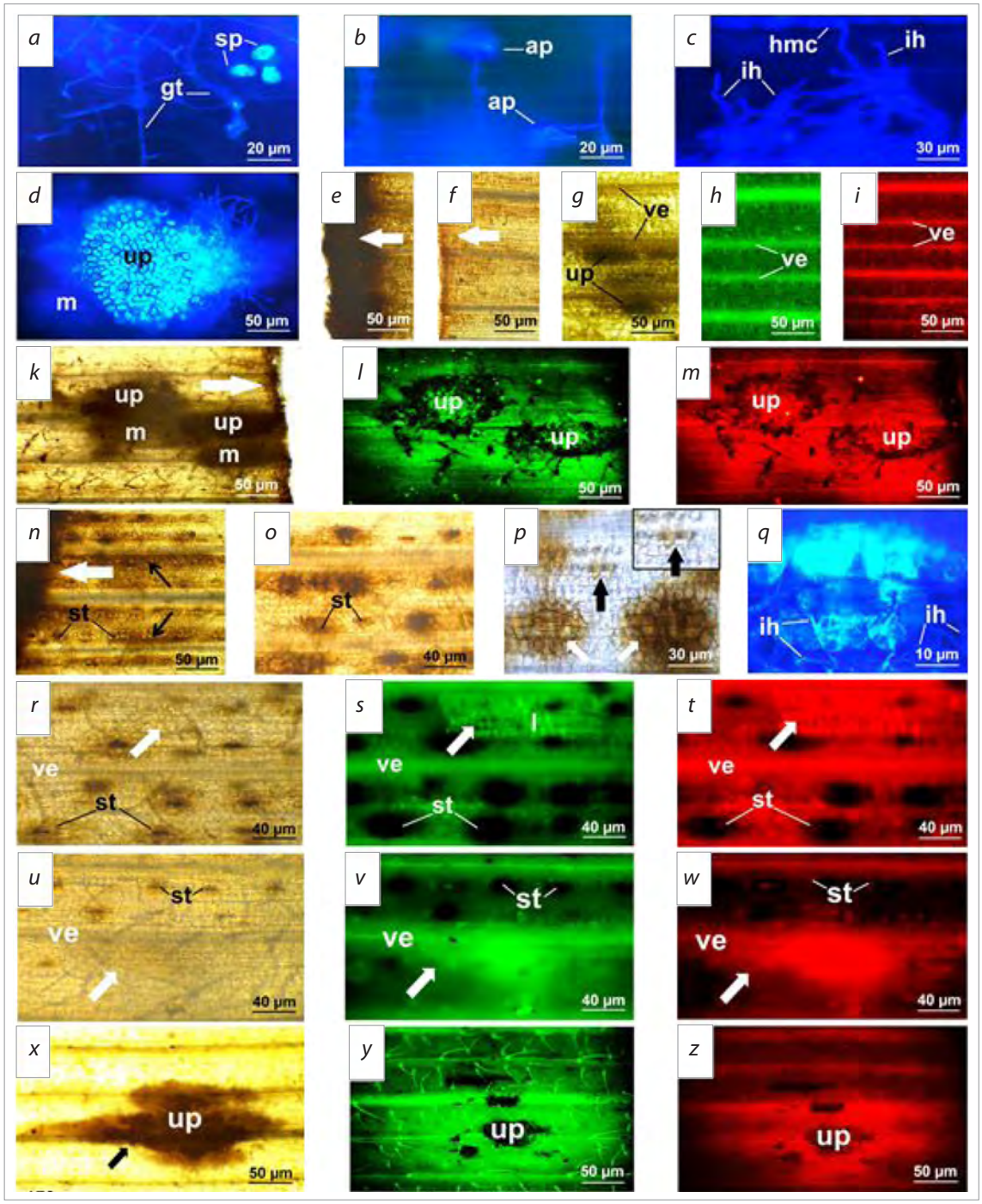


Fig. 1. Development of *P. graminis* f. sp. *tritici* and distribution of hydrogen peroxide and phenols in tissues. *a–d*, *g*, *k–m* – infected untreated plants; *e*, *f*, *h*, *i* – control uninfected plants; *n–z* – Novochizol-treated infected plants.

a – growing tubes development on the leaf surface; *b* – appressorium on stoma; *c* – infection hyphae and haustorial mother cell in the tissue; *d* – colony with urediniopustule; *e* – intensive H_2O_2 accumulation on the leaf cut of the control plant, 24 h p/in; *f* – weak H_2O_2 accumulation on the leaf section cut of the control plant, 240 h p/in; *g* – phenols in plant cytoplasm in the urediniopustule area and lignins in the parallel veins; *h*, *i* – phenols autofluorescence in the leaf of the control plant; *k* – H_2O_2 accumulation in the colony area with the urediniopustule and on the leaf cut (arrow), 240 h p/in; *l*, *m* – phenols autofluorescence in the tissues surrounding urediniopustules; *n* – intensive H_2O_2 accumulation on the leaf cut and in the plant tissue, 24 h p/in; *o* – H_2O_2 localization in the stomata area, 96 h p/in; *p* – empty appressorium shell on the plant stoma (black arrow, selected fragment) and intensive accumulation of H_2O_2 under other stomata (white arrows), 48 h p/in; *q* – autofluorescence of dead plant cells and fungal infection hyphae, damaged cells (light blue) and normal hyphae (blue), 48 h p/in; *r* – abortive colony (arrow); *s*, *t* – phenols and lignin autofluorescence in the same abortive colony (arrow) zone, 96 h p/in; *u* – actively developed colony (arrow); *v*, *w* – phenols autofluorescence in actively developed colony zone (arrows), 144 h p/in; *x* – intensive H_2O_2 accumulation (arrow) in the colony zone with urediniopustule, 240 h p/in; *y*, *z* – intensive phenols accumulation with different colour illumination in the colony with urediniopustule zone, 240 h p/in. Designations: ap – appressorium; ih – infection hypha; l – lignin; hmc – haustorial mother cell; m – mycelium; ve – vein; gt – growing tube; sp – spore; st – stoma; up – urediniopustule. Staining: *a–d*, *q* – Uvitex 2B; *e*, *f*, *k*, *n–p*, *r*, *u*, *x* – DAB; *g* – aniline sulfate; *h*, *l*, *s*, *v*, *y* – phenols autofluorescence at emission λ_{max} = 530 nm; *i*, *m*, *t*, *w*, *z* – phenols autofluorescence at emission λ_{max} = 605 nm.

Table 2. Development of *P. graminis* f. sp. *tritici* on the surface of wheat plants treated with Novochizol

Experimental variant	Germinated spore proportion, %	Proportion of appressoria, %		
		from the number of germinated spores	on stomata from their total number	penetrated into stomata
Control	77.2 ± 1.6	61.3 ± 5.1	72.7 ± 3.6	93.8 ± 1.3
Novochizol	80.2 ± 1.9	63.8 ± 3.9	78.0 ± 2.7	93.2 ± 1.5
$LSD_{0.05}$	3.2	3.4	6.2	2.1

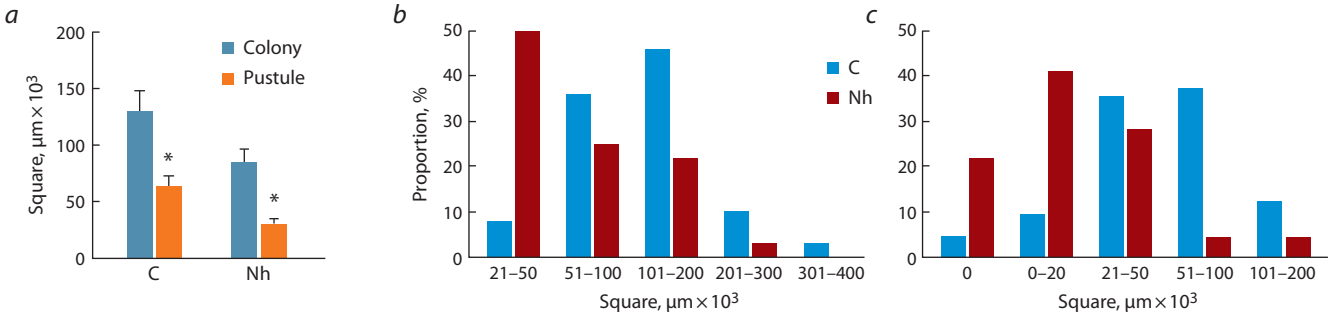


Fig. 2. The effect of Novochizol treatment on *P. graminis* f. sp. *tritici* colonies and pustules development.

a – caverage area; *b* – distribution of colonies by area; *c* – distribution of pustules by area. C – control; Nh – Novochizol. * Significant difference at $p \le 0.05$.

sporulation stage. High H_2O_2 accumulation was determined in the tissues under the pustules, and less in the surrounding mycelium area at 240 h p/in (Fig. 1*k*). A moderate accumulation of phenols with green fluorescence and that of phenols with brighter red autofluorescence were detected around the pustules in the mycelium zones (Fig. 1*l*, *m*).

In the Novochizol-treated uninfected plants, H_2O_2 accumulation of varying intensity was noted in the leaves in the form of spots for 48 h p/in. The H_2O_2 distribution was irregular, which may be due to uneven Novochizol distribution by spraying. The stomatal guard cells, as well as mesophyll cells under the stomata and between the veins, were strongly stained (Fig. 1*n*). The H_2O_2 gradually disappeared from the tissues after 96–144 h p/in, but remained in the guard cells and in small zones below them in small amounts (Fig. 1*o*, *r*, *u*).

In Novochizol-treated infected plants, the H_2O_2 content in tissues for 96 hours was similar to that described above. The fungus penetrated into the stomata between H_2O_2 accumulation zones without deviations, and empty appressoria shells remained on the surface of guard cells (Fig. 1*p*). In areas with a high ROS content, the cytoplasm of dead plant cells showed a white glow, while damaged ones showed light blue fluorescence. The dead fungal hyphae had white autofluorescence, and the intact ones had a blue colour (Fig. 1*q*). The colonies died (aborted) at early developmental stages in the ROS accumulating loci. The phenols in the cytoplasm and lignin on the cell walls accumulated in the zones of the dead colonies. These substances had a brighter green and a less pronounced red autofluorescence after 96 h p/in. The accumulation of green and red lignins was also enhanced in the adjacent vein regions. At the same time, phenols did not accumulate near the stomata area with a high H_2O_2 content (Fig. 1*s*, *t*).

Significant H_2O_2 generation was not detected near the actively developing colonies 144 h p/in, and simultaneously its content decreased in the stomatal zones (Fig. 1*u*). Phenols with brighter green and less vivid red fluorescence covered the mycelium area (Fig. 1*v*, *w*). Intensive H_2O_2 accumulation was determined in the large colony and pustule areas 240 h p/in (Fig. 1*x*), and strong phenol accumulation was noted around such colonies in a wider than H_2O_2 zone. Phenols and lignin with green autofluorescence were synthesized more intensively and spread over a larger area than ones with red colour (Fig. 1*y*, *z*).

A study of *Pgt* development showed that in Novochizol-treated plants, the average colony and pustule areas decreased (by 1.5 and 2.2 times to untreated, respectively) (Fig. 2*a*). The preparation's effect resulted in a significant change in distribution of the colonies and pustules by sizes, compared with untreated plants. The proportion of small colonies and pustules increased sharply, and some colonies (22 %) died before sporulation (Fig. 2*b*, *c*).

Discussion

The biological properties of the BPMA based on chitin and chitosan have been investigated since the 1980s. During this time, numerous tests have been carried out on the chitosan derivatives effects on pathogens (fungi, bacteria, and viruses) development and disease manifestations. The fungicidal effects of chitosans have been mainly studied on pathogens with

a necrotrophic or hemibiotrophic feeding type. These groups include the most harmful species of the *Botrytis*, *Fusarium*, *Alternaria*, *Colletotrichum*, *Phytophthora*, *Rhizoctonia* genera (Chakraborty et al., 2020; Zheng et al., 2021; Shcherban, 2023). The cultivation of these fungi is available on artificial media, which makes it possible to evaluate the drug's effects *in vitro*. Using different pathogen species, it was showed that the chitosan preparations manifested fungicidal effects by the suppression of spore germination, inhibition of growing tubes development, disruption of cell walls and membranes, and an impenetrable film formation around fungal cells (Ghaouth et al., 1994; Abd El-Kareem, Haggag, 2014). Growth suppression could also be associated with calcium and copper ions chelation and deposition of chelate complexes on cell surface, which reduced the metabolic activity of fungi (Chakraborty et al., 2020). The effects of preparations *in planta* are realized after chitosan recognition by plant receptors and signalling systems activation (Yarullina et al., 2023).

The mechanisms of Novochizol action on wheat stem rust development have been studied for the first time. The Novochizol concentration effect on the disease development has been revealed. This confirms the results of previous studies on the effect of drug doses on plant resistance reactions (Orzali et al., 2017; Varlamov et al., 2020). In the 0.75 % Novochizol variant, the greatest inhibiting effect on *Pgt* development was noted. In contrast to previous results, obtained with chitosan (Ghaouth et al., 1994; Abd El-Karee, Haggag, 2014), Novochizol had no negative effect on *Pgt* development on the leaf surface, as well as penetration into stomata.

In the 0.125 % Novochizol variant, a more intense pustule development suppression was found. The effect of the 0.125 % Novochizol on stem rust will be studied at the next research steps.

In the 2000s, a hypothesis of a two-level organization of plant immunity was formulated, called PTI-ETI (Gill et al., 2015). It was assumed that plants have PRRs (Pattern Recognition Receptors) that recognize molecules of non-pathogenic (MAMPs, Microbe-Associated Molecular Patterns) and non-specialized pathogenic microorganisms (PAMPs, Pathogen-Associated Molecular Patterns), as well as plant cell destruction products (DAMPs, Damage-Associated Molecular Patterns). As a result of the recognition of these molecules, the first level of PTI (PAMP-triggered immunity) defence is triggered. After PTI is overcome, the second resistance level is activated, associated with the recognition of specific effectors – ETI (Effector-Triggered Immunity). PTI corresponds to the response of non-host species, while ETI is similar to varietal resistance and is usually accompanied by hypersensitive reaction (Gill et al., 2015). Later, an improved model of plant immunity was proposed, according to which ETI is a PTI-dependent module for the reactions amplification, but not an isolated system (Yuan M. et al., 2021; Zhao et al., 2022).

Two peaks of ROS generation have been identified in resistant plants previously. The first peak occurs a few minutes after elicitor recognition and is associated with the activation of the NADP·H-oxidase enzyme, which is constitutively present in the membrane. NADP·H-oxidase produces the superoxide anion $O_2^{\bullet-}$, which is rapidly converted by the SOD enzyme to H_2O_2 (Boller, Keen, 2000). The second ROS peak appears 3–5 days later, and is associated with *de novo* synthesis of the pro/antioxidant system enzymes (peroxidases, oxalate oxidases). The pro/antioxidant system maintains optimal ROS levels in the tissues. Catalase cleaves H_2O_2 to water, and at the same time the peroxidases, polyphenol oxidase, and ascorbate oxidase utilize ROS in oxidative reactions (Maksimov, Cherepanova, 2006).

Previously, when studying the interactions between the rust fungi *P. triticina* and *P. coronata* with non-host species (oats and wheat, respectively), the $O_2^{\bullet-}$ generation by stomatal guard cells contacting with appressoria, which led to pathogen death, was revealed (Plotnikova, 2008). *Pgt* dies on the plant surface before penetration into the stomata of non-host species *Secale cereale* and *Thinopyrum ponticum*. When *Pgt* interacts with cultivars carrying resistance genes of non-hosts (*Sr31*, *Sr24*, *Sr25*, *Sr26*), the appressoria dies on the stomata after the peak of superoxide anion generation (Plotnikova et al., 2022, 2023). On the example of chitosan-treated rice, a similar NADP·H-dependent $O_2^{\bullet-}$ synthesis was shown (Lopez-Moya et al., 2021). An enhanced synthesis of enzymes involved in ROS accumulation was also found in millet plants treated with chitosan and infected with *Alternaria kikuchiana* (Meng et al., 2010). The chitosan application on barley induces an oxidative burst and synthesis of phenolic compounds, which increases the resistance to fungal diseases complex (Faoro et al., 2008).

Novochizol is similar to MAMPs in its origin. A histochemical study of Novochizol-treated plants revealed intensive H_2O_2 accumulation in tissues four days after its application. Obviously, this is due to the second peak of oxidative burst manifestation and confirms the inducing resistance activity of Novochizol. Zones with a high H_2O_2 content were found both in the stomata areas and between the veins. Such results may be explained by increased Novochizol ability to penetrate through the leaf epidermis and induce ROS production. The H_2O_2 content in the tissues decreased after 96–144 h p/in of *Pgt* inoculation, so did the traumatic ROS on leaf cuts to the end of the experiment in all variants. Such dynamics may be related to the synthesis of the antioxidant system components (both the enzymes and non-enzymatic substances) that utilize ROS. The activation of antioxidant enzymes following ROS accumulation was shown in potatoes treated with the chitin-ferulic acid conjugate and beneficial bacteria *Bacillus subtilis* (Yarullina et al., 2024a). It is also possible that the antioxidant activity increased with the age of the plants.

Defense reactions did not appeared before sporogenesis in infected untreated plants. In Novochizol-treated plants, a small number of host cells and mycelium fragments died in the areas with increased H_2O_2 content. At the same time, the dead plant cells did not exhibit the yellow fluorescence characteristic for HR (Vander et al., 1998). This indicates that Novochizol induces reactions that partially differ from those occurring during HR in resistant varieties.

Some colonies died at the early pathogenesis stages. H_2O_2 was not detected in the zones of abortive colonies 96 h p/in. ROS accumulation has also not been established in the areas of medium and large developing colonies before sporogenesis, and even a decrease in H_2O_2 near the colonies has been noted. The decrease in H_2O_2 content can be explained both by the accumulation of antioxidant plant enzymes and by the pathogen activity. Currently, it is known that biotrophic rust fungi secrete hundreds of effectors into plant cytoplasm and apoplast. The pathogens are able to suppress protective reactions, as well as to alter or reprogram host metabolism by the effectors. It was shown that a virulent isolate of wheat yellow rust pathogen *P. striiformis* f. sp. *tritici* secreted an effector catalase cleaving H_2O_2, which led to plant resistance suppression (Yuan P. et al., 2021). At the same time, Novochizol treatment stimulated increasing H_2O_2 accumulation in the colony zones at the stage of sporogenesis.

It has previously been shown that the phenols synthesis and the strengthening of cell walls with lignins after treatment with chitosans were the most typical protective reactions against necrotrophic and hemibiotrophic fungi (Orzali et al., 2017; Shcherban, 2023). In our experiments, it was found that Novochizol treatment stimulated an earlier and more intensive phenols accumulation than in untreated plants. For the first time, it was shown that Novochizol promotes the changing in the phenols ratio towards compounds with a green fluorescence, while phenols with a red light prevailed in untreated plants. It was previously determined that lignin with green autofluorescence includes syringin derivatives and accumulates in wheat tissues after treatment with SAR inducer Bion (Plotnikova, 2009). Previously, it was shown that in plants treated with the chitosans and infected with necrotrophic fungi, the PR proteins' (chitinases, glucanases, peroxidases, polyphenol oxidases, PR-1, PR-5, etc.) genes expression increased (Manjunatha et al., 2008, 2009; Nandeeshkumar et al., 2008; Orzali et al., 2014). Similar accumulation of PR proteins with different functions was also revealed in potatoes treated by chitosan conjugates with ferulic or caffeic acids (Yarullina et al., 2024a, b). Accumulation of PR-proteins in Novochizol-treated plants, which are not detectable by cytological methods, is also likely. The complex action of Novochizol-induced defence mechanisms led to the death of a significant part of the colonies at the early development stages, as well as to a significant reduction in the pustule density and suppression of the pathogen's reproduction.

The reported studies were the first stage of investigation of the Novochizol effect on the wheat resistance mechanisms against stem rust. At the next stage, detailed studies of the preparation's action on the pathogenesis will be carried out using molecular genetics, biochemical and cytophysiological methods.

Conclusions

Studies have shown that Novohizol can be used as a resistance inducer to wheat stem rust. The dose effect of the treatment was revealed, with the best results at 0.125 and 0.75 % concentrations.

Novochizol treatment of leaves at the 0.75 % concentration did not affect the urediniospore germination and fungal structures development on the plant surface, but led to a significant reduction in the number of colonies, as well as the mycelium and pustule sizes.

Intensive hydrogen peroxide accumulation in infected and uninfected plant tissues 4–8 days after Novochizol treatment was found (corresponds to 0–4 days after inoculation), which decreased by the end of the experiment.

Partial death of plant cells and pathogen mycelium was noted in the zones of intensive H_2O_2 accumulation. The dead plant cells did not show the autofluorescence characteristic for HR.

Novochizol stimulated earlier and more intensive phenols accumulation in infection zones, such as a change in the ratio of phenolic compounds towards substances with syringin derivatives.

References

Abd El-Kareem F., Haggag W. Chitosan and citral alone or in combination for controlling early blight disease of potato plants under field conditions. *Res J Pharm Biol Chem Sci.* 2014;5(6):941-949. https://rjpbcs.com/pdf/2014_5(6)/%5B141%5D.pdf

Badanova E.G., Davletbaev I.M., Sirotkin A.S. Preparations based on chitosan for agriculture. *Vestnik Tekhnologicheskogo Universiteta = Herald of Technological University.* 2016;19(16):89-95 (in Russian)

Bellameche F., Jasim M., Mauch-Mani B., Mascher F. Histopathological aspects of resistance in wheat to *Puccinia triticina*, induced by *Pseudomonas protegens* CHA0 and β-aminobutyric acid. *Phytopathol Mediterr*. 2021;60(3):441-453. doi 10.36253/phyto-13123

Boller T., Keen N.T. Perception and transduction of elisitor signals in host-pathogen interactions. In: Slusarenko A.J., Fraser R.S.S., van Loon L.C. (Eds) Mechanisms of Resistance to Plant Diseases. Dordrecht: Springer, 2000;189-230. doi 10.1007/978-94-011-3937-3_7

Burlakova S.V., Egorycheva M.T., Fomenko V.V., Salakhutdinov N.F., Shcherban A.B. Biological justification of the use of Novohisol with natural fungicides in the cultivation of bread wheat. *Chemistry for Sustainable Development*. 2025;3:303-314 (in Russian)

Chakraborty M., Hasanuzzaman M., Rahman M., Khan Md.A.R., Bhowmik P., Mahmud N.U., Tanveer M., Islam T. Mechanism of plant growth promotion and disease suppression by chitosan biopolymer. *Agriculture*. 2020;10:624. doi 10.3390/agriculture10120624

Chandler D., Bailey A.S., Tatchell G.M., Davidson G., Greaves J., Grant W.P. The development, regulation and use of biopesticides for integrated pest management. *Philos Trans R Soc Lond B Biol Sci*. 2011;366:1987-1998. doi 10.1098/rstb.2010.0390

Elsharkawy M.M., Omara R.I., Mostafa Y.S., Alamri S.A., Hashem M., Alrumman S.A., Ahmad A.A. Mechanism of wheat leaf rust control using chitosan nanoparticles and salicylic acid. *J Fungi*. 2022; 8(3):304. doi 10.3390/jof8030304

Faoro F., Maffi D., Cantu D., Iriti M. Chemical-induced resistance against powdery mildew in barley: the effects of chitosan and benzothiadiazole. *BioControl*. 2008;53(2):387-401. doi 10.1007/s10526-007-9091-3

Ghauoth A., Arul J., Grenier J., Benhamou N., Asselin A., Belanger G. Effect of chitosan on cucumber plants: suppression of *Pythium aphanidermatum* and induction of defense reaction. *Phytopathology*. 1994;84(3):313-320. doi 10.1094/PHYTO-84-31

Gill U.S., Lee S., Mysore K.S. Host versus nonhost resistance: distinct wars with similar arsenals. *Phytopathology*. 2015;105(5):580-587. doi 10.1094/PHYTO-11-14-0298-RVW

Haggag W.M.W., Hussein M.M., Medhat M.T., El Habbasha S.F. Enhancement of wheat resistant to diseases by elicitors. *Int J Sci Res*. 2014;3(11):1526-1530

Japaridze L.I. Practicum on Microscopic Chemistry of Plants. Moscow: Sovetskaya Nauka Publ., 1953 (in Russian)

Lopez-Moya F., Martin-Urdiroz M., Oses-Ruiz M., Were V.M., Fricker M.D., Littlejohn G.R., Lopez-Llorca L.V., Talbot N.J. Chitosan inhibits septin-mediated plant infection by the rice blast fungus *Magnaporthe oryzae* in a protein kinase C and Nox1 NADPH oxidase-dependent manner. *New Phytol*. 2021;230(4):1578-1593. doi 10.1111/nph.17268

Maksimov I.V., Cherepanova E.A. Pro-/antioxidant system and resistance of plants to pathogens. *Uspehi Sovremennoy Biologii = Advances in Current Biology*. 2006;126(3):250-261 (in Russian)

Malerba M., Cerana R. Chitosan effects on plant systems. *Int J Mol Sci*. 2016;17(7):996. doi 10.3390/ijms17070996

Manjunatha G., Roopa K.S., Prashanth G.N., Shetty H.S. Chitosan enhances disease resistance in pearl millet against downy mildew caused by *Sclerospora graminicola* and defence-related enzyme activation. *Pest Manag Sci*. 2008;64:1250-1257. doi 10.1002/ps.1626

Manjunatha G., Niranjan-Raj S., Prashanth G.N., Deepak S., Amruthesh K.N., Shetty H.S. Nitric oxide is involved in chitosan-induced systemic resistance in pearl millet against downy mildew disease. *Pest Manag Sci*. 2009;65(7):737-743. doi 10.1002/ps.1710

Meng X., Yang L., Kennedy J.F., Tian S. Effects of chitosan and oligochitosan on growth of two fungal pathogens and physiological properties in pear fruit. *Carbohydr Polym*. 2010;81(1):70-75. doi 10.1016/j.carbpol.2010.01.057

Moldenhauer J., Moerschbacher B.M., van der Westhuizen A.J. Histological investigation of stripe rust (*Puccinia striiformis* f. sp. *tritici*) development in resistant and susceptible wheat cultivars. *Plant Pathology*. 2006;55:469-474. doi 10.1111/j.1365-3059.2006.01385.x

Nandeeshkumar P., Sudisha J., Ramachandra K.K., Prakash H., Niranjana S., Shekar S.H. Chitosan induced resistance to downy mildew in sunflower caused by *Plasmopara halstedii*. *Physiol Mol Plant Pathol*. 2008;72(4-6):188-194. doi 10.1016/j.pmpp.2008.09.001

Orzali L., Forni C., Riccioni L. Effect of chitosan seed treatment as elicitor of resistance to *Fusarium graminearum* in wheat. *Seed Sci Technol*. 2014;42(2):132-149. doi 10.15258/sst.2014.42.2.03

Orzali L., Corsi B., Forni C., Riccioni L. Chitosan in agriculture: a new challenge for managing plant disease. In: Shalaby E.A. (Ed.) Biological Activities and Application of Marine Polysaccharides. *InTech*. 2017;87-96. doi 10.5772/66840

Plotnikova L.Ya. Influence of the surface features and physiological reactions of non-host species on the development of cellular structures of rust fungi. *Tsitologiia*. 2008;50(5):439-446 (in Russian)

Plotnikova L.Ya. Effect of benzothiadiazole, an inducer of systemic acquired resistance, on the pathogenesis of wheat brown rust. *Russian Journal of Plant Physiology*. 2009;56(4):517-526. doi 10.1134/S1021443709040116

Plotnikova L., Knaub V. Exploitation of the genetic potential of *Thinopyrum* and *Agropyron* genera to protect wheat from diseases and environmental stresses. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov J Genet Breed*. 2024;28(5):536-553. doi 10.18699/vjgb-24-60

Plotnikova L.Y., Meshkova L.V. Evolution of cytophysiological relationships between leaf rust causal agent and common wheat in the process of overcoming of resistance determined by the gene *Lr19*. *Mikologiya i Fitopatologiya = Mycology and Phytopathology*. 2009;43(4):343-357 (in Russian)

Plotnikova L., Pozherukova V., Knaub V., Kashuba Y. What was the reason for the durable effect of *Sr31* against wheat stem rust? *Agriculture*. 2022;12:2116. doi 10.3390/ agriculture12122116

Plotnikova L., Knaub V., Pozherukova V. Nonhost resistance of *Thinopyrum ponticum* to *Puccinia graminis* f. sp. *tritici* and the effects of the *Sr24*, *Sr25*, and *Sr26* genes introgressed to wheat. *Int J Plant Biol*. 2023;14:435-457. doi 10.3390/ijpb14020034

Popova E.V., Domnina N.S., Kovalenko N.M., Sokornova S.V., Tyuterev S.L. Influence of chitosan hybrid derivatives on induced wheat resistance to pathogens with different nutrition strategies. *Applied Biochemistry and Microbiology*. 2018;54(5):535-539. doi 10.1134/S0003683818050150

Rabea E.I., Badawy M.E., Rogge T.M., Stevens C.V., Höfte M., Steurbaut W., Smagghe G. Insecticidal and fungicidal activity of new synthesized chitosan derivatives. *Pest Manag Sci*. 2005;61(10):951-960. doi 10.1002/ps.1085

Richter T., Gulich M., Richter K. Quality control and good manufacturing practice (GMP) for chitosan-based biopharmaceutical products. In: Sarmento B., das Neves J. (Eds) Chitosan-Based Systems for Biopharmaceuticals: Delivery, Targeting and Polymer Therapeutics. John Wiley & Sons, 2012;503-542. doi 10.1002/9781119962977.ch26

Rkhaila A., Chtouki T., Erguig H., El Haloui N., Ounine K. Chemical proprieties of biopolymers (chitin/chitosan) and their synergic effects with endophytic bacillus species: unlimited applications in agriculture. *Molecules*. 2021;26(4):1117. doi 10.3390/molecules26041117

Roelfs A.P., Singh R.P., Saari E.E. Rust diseases of wheat: concepts and methods of disease management. Cimmyt, Mexico DF, Mexico, 1992.

Rsaliyev A.S., Rsaliyev Sh.S. Principal approaches and achievements in studying race composition of wheat stem rust. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov J Genet Breed*. 2018;22(8): 967-977. doi 10.18699/VJ18.439 (in Russian)

Shcherban A.B. Chitosan and its derivatives as promising plant protection tools. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov J Genet Breed*. 2023;27(8):1010-1021. doi 10.18699/VJGB-23-116

Singh Y., Nair A.M., Verma P.K. Surviving the odds: from perception to survival of fungal phytopathogens under host-generated oxidative burst. *Plant Commun*. 2021;2:100142. doi 10.1016/j.xplc.2021.100142

Sørensen C.K., Thach T., Hovmøller M.S. Evaluation of spray and point inoculation methods for the phenotyping of *Puccinia striiformis* on wheat. *Plant Disease*. 2016;100(6):1064-1070. doi 10.1094/PDIS-12-15-1477-RE

Sternshis M.V., Belyaev A.A., Tsvetkova V.P., Shpatova T.V., Lelyak A.A., Bakhvalov S.A. Biopreparations Based on Bacteria of the Genus *Bacillus* for Plant Health Management. Novosibirsk: Publishing House of SB RAS, 2016 (in Russian)

Tarchevsky I.A. Elicitor-induced signaling pathways and their interaction. *Russian Journal of Plant Physiology*. 2000;47(2):285-294

Teplyakova O.I., Fomenko V.V., Salakhutdinov N.F., Vlasenko N.G. Novochizol™ seed treatment: effects on germination, growth and development in soft spring wheat. *Nat Prod Chem Res*. 2022;10(5):1-4. doi 10.35248/naturalproducts.10.5.1-04

Tyuterev S.L. Ecologically safe inducers of plant resistance to diseases and physiological stresses. *Vestnik Zashchity Rasteniy = Plant Protection News*. 2015;1(83):3-13 (in Russian)

USDA. World Agricultural Production; USDA Foreign Agricultural Service. Washington, DC, USA, 2016.

Vander P., Vårum K.M., Domard A., El Gueddari N.E., Moerschbacher B.M. Comparison of the ability of partially N-acetylated chitosans and chitooligosaccharides to elicit resistance reactions in wheat leaves. *Plant Physiol*. 1998;118(4):1353-1359. doi 10.1104/pp.118.4.1353

Varlamov V.P., Ilyina A.V., Shagdarova B.T., Lunkov A.P., Mysyakina I.S. Chitin/chitosan and its derivatives: fundamental problems and practical approaches. *Biochemistry*. 2020;85:154-176. doi 10.1134/S0006297920140084

Woldeab G., Hailu E., Bacha N. Protocols for race analysis of wheat stem rust (*Puccinia graminis* f. sp. *tritici*). Ambo, Ethiopia: EIAR, 2017.

Yarullina L.G., Kalatskaja J.N., Cherepanova E.A., Yalouskaya N.A., Tsvetkov V.O., Ovchinnikov I.A., Burkhanova G.F., Rybinskaya K.I.,

Sorokan A.V., Herasimovich K.M., Zaikina E.A., Nikolaichuk V.V., Hileuskaya K.S., Mardanshin I.S. Approaches to improving biological activity of agricultural formulations based on bacteria of the genus *Bacillus* and chitosan nanocomposites (review). *Applied Biochemistry and Microbiology.* 2023;59(5)549-560. doi 10.1134/s0003683823050186

Yarullina L.G., Burkhanova G.F., Tsvetkov V.O., Cherepanova E.A., Sorokan A.V., Zaikina E.A., Mardanshin I.S., Fatkullin I.Y., Maksimov I.V., Kalatskaja J.N., Yalouskaya N.A., Rybinskay E.I. The effect of chitosan conjugates with hydroxycinnamic acids and *Bacillus subtilis* bacteria on the activity of protective proteins and resistance of potato plants to *Phytophthora infestans*. *Appl Biochem Microbiol*. 2024a;60(2):231-240. doi 10.1134/S0003683824020194

Yarullina L., Kalatskaja J., Tsvetkov V., Burkhanova G., Yalouskaya N., Rybinskaya K., Zaikina E., Cherepanova E., Hileuskaya K., Nikalaichuk V. The influence of chitosan derivatives in combination with *Bacillus subtilis* bacteria on the development of systemic resistance in potato plants with viral infection and drought. *Plants*. 2024b;13:2210. doi 10.3390/plants13162210

Yuan M., Pok B., Ngou M., Ding P., Xin X.-F. PTI-ETI crosstalk: an integrative view of plant immunity. *Curr Opin Plant Biol*. 2021;62:102030. doi 10.1016/j.pbi.2021.102030

Yuan P., Qian W., Jiang L., Jia C., Ma X., Kang Z., Liu J. A secreted catalase contributes to *Puccinia striiformis* resistance to host-derived oxidative stress. *Stress Biol*. 2021;1(1):22. doi 10.1007/s44154-021-00021-2

Zhao Y., Zhu X., Chen X., Zhou J. From plant immunity to crop disease resistance. *J Genet Genom*. 2022;49(8):693-703. doi 10.1016/j.jgg.2022.06.003

Zheng K., Lu J., Li J., Yu Y., Zhang J., He Z., Ismail O.M., Wu J., Xie X., Li X., Xu G., Dou D., Wang X. Efficiency of chitosan application against *Phytophthora infestans* and the activation of defence mechanisms in potato. *Int J Biol Macromol*. 2021;182:1670-1680. doi 10.1016/j.ijbiomac.2021.05.097

Conflict of interest. The authors declare no conflict of interest.
Received October 23, 2024. Revised December 9, 2024. Accepted December 9, 2024.

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2025;29(4):549-558

doi 10.18699/vjgb-25-58

Receptor-like leucine-rich repeat kinases of subfamily III are involved in the recognition of *Pectobacterium* spp. by Solanaceae plants

E.V. Shrub, N.V. Kalubaka, P.V. Vychyk, O.A. Badalyan, Y.A. Nikolaichik

Belarusian State University, Minsk, Belarus
nikolaichik@bsu.by

Abstract. The genomes of Solanaceae plants contain over 600 receptor-like protein kinase genes with leucine-rich repeats (LRR-RLK), many likely associated with pathogen detection, but very few functionally characterized. *Pectobacterium* spp. are the major bacterial pathogens of agricultural crops, particularly potatoes and other Solanaceae plants. For relevant potato pathogens from the genus *Pectobacterium*, specific immune receptors have not been described in Solanaceae. However, in *Malus* × *domestica*, four LRR-RLK from the LRRIII subfamily (DIPM1-4) have been characterized as receptors for the related pathogen *Erwinia amylovora*. DIPMs specifically interact with the effector protein DspE and are involved in *E. amylovora* recognition. Since the DspE ortholog is also the main effector in *Pectobacterium* spp., we performed a phylogenetic analysis of LRRIII subfamily receptors in the most relevant Solanaceae representatives together with a much better characterized LRR-RLKIII of *Arabidopsis thaliana* and identified nine clusters of related RLKs. Clustering followed by analysis of published data allowed us to functionally characterize this RLK family and suggest the most likely candidates for checking interactions with the main effector of pectobacteria, DspE. Testing the kinase domains of representative cluster members in a yeast two-hybrid system revealed four Solanaceae RLKs interacting with the DspE effector from *Pectobacterium versatile*. Virus-induced silencing of these RLK genes demonstrated their involvement in *P. versatile* recognition. The *RLK6* gene from *Solanum bulbocastanum*, which is not an ortholog of the DIPM proteins in apple, seems to be the most promising potential resistance gene. This work expands our understanding of LRR-RLKIII subfamily RLKs and their role in plant immunity, providing a foundation for future development of disease-resistant Solanaceae varieties.

Key words: receptor-like protein kinase; Solanaceae; *Pectobacterium*; effector; plant immunity

For citation: Shrub E.V., Kalubaka N.V., Vychyk P.V., Badalyan O.A., Nikolaichik Y.A. Receptor-like leucine-rich repeat kinases of subfamily III are involved in the recognition of *Pectobacterium* spp. by Solanaceae plants. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov J Genet Breed*. 2025;29(4):549-558. doi 10.18699/vjgb-25-58

Funding. This work was supported by Belarusian Republican Foundation for Basic Research (project No. B24M-035) and by the Ministry of Education of Belarus (project No. 20241129).

Рецепторподобные киназы с лейцин-богатыми повторами подсемейства III участвуют в распознавании *Pectobacterium* spp. растениями семейства Solanaceae

Е.В. Шруб, А.В. Колубако, П.В. Вычик, О.А. Бадалян, Е.А. Николайчик

Белорусский государственный университет, Минск, Беларусь
nikolaichik@bsu.by

Аннотация. Геномы растений семейства Пасленовые содержат более 600 генов рецепторных протеинкиназ с лейцин-богатыми повторами (LRR-RLK), многие из которых, вероятно, связаны с детекцией патогенов, но лишь некоторые были функционально охарактеризованы. Энтеробактерии рода *Pectobacterium* – основные бактериальные патогены многих сельскохозяйственных культур, в том числе картофеля и других растений семейства Пасленовые. Для актуальных патогенов из рода *Pectobacterium* специфические иммунные рецепторы растений не описаны. Однако у *Malus* × *domestica* охарактеризовано четыре LRR-RLK из подсемейства LRRIII (DIPM1-4), специфически взаимодействующих с эффекторным белком DspE и участвующих в распознавании родственного энтеробактериального фитопатогена *Erwinia amylovora*. Поскольку ортолог DspE является основным эффектором и у *Pectobacterium* spp., мы выполнили филогенетический анализ RLK-LRRIII растений семейства Пасленовые совместно с более полно охарактеризованными LRR-RLKIII у *Arabidopsis thaliana* и выделили девять кластеров родственных RLK. Кластеризация и анализ опубликованных данных позволили функционально охарактеризовать это семейство RLK и предложить наиболее вероятных кандидатов для проверки

© Shrub E.V., Kalubaka N.V., Vychyk P.V., Badalyan O.A., Nikolaichik Y.A., 2025
This work is licensed under a Creative Commons Attribution 4.0 License

взаимодействия с основным эффектором пектобактерий DspE. Тестирование киназных доменов репрезентативных представителей разных кластеров в дрожжевой двухгибридной системе выявило четыре RLK растений семейства Пасленовые, которые взаимодействуют с эффектором DspE из *Pectobacterium versatile* (*Pve*). Уровень экспрессии генов этих RLK и их ортологов у разных растений семейства варьировал, но в целом был очень низким. При этом обнаружена сильная DspE-зависимая супрессия генов *RLK2* и *RLK5* у инфицированных *Pve* растений картофеля, а инактивация их ортологов предотвращала развитие сверхчувствительной реакции в листьях растений, инфильтрованных суспензиями *Pve*. Данная работа расширяет понимание разнообразия RLK подсемейства LRR-RLKIII и их роли в иммунитете растений и может способствовать селекции устойчивых к бактериозам сортов растений семейства Пасленовые.

Ключевые слова: рецепторподобные протеинкиназы; Solanaceae; *Pectobacterium*; эффектор; растительный иммунитет

Introduction

Pectobacterium spp. are major bacterial pathogens of a range of important crops, particularly potatoes. *P. atrosepticum*, *P. carotovorum*, *P. parmentieri*, *P. brasiliense*, and *P. versatile* are the most relevant pathogenic species for potato. Depending on conditions and strain characteristics, pectobacteria can infect underground or above-ground parts of the plant, causing soft rot of tubers, blackleg, or aerial stem rot. The hallmark of *Pectobacterium* infections is the massive production of around 30 exoenzymes (pectolytic, cellulolytic, and proteolytic), leading to the characteristic softening of infected tissues (Chatterjee et al., 1995; Pérombelon, 2002).

Most potato varieties are susceptible to *Pectobacterium* infections. Although relatively tolerant varieties are known (Kwenda et al., 2016), potato plants resistant to *Pectobacterium* have not yet been developed. This situation can be largely explained by an insufficient understanding of the mechanisms by which plants recognize these pathogens, especially at the earliest stage of the infection.

Pectobacterium spp. were long considered typical necrotrophs minimally interacting with their hosts. However, since the advent of the genomic era, a substantial amount of data has accumulated, indicating the complex nature of molecular communication between *Pectobacterium* spp. and their hosts. A notable feature of this communication is the ability of *Pectobacterium* to coexist with its hosts in a "stealth" mode – without developing a systemic infection and without causing significant damage to plant tissues (Toth, Birch, 2005; Gorshkov et al., 2018).

The switch between these two fundamentally different phases of infection, latent and symptomatic, depends on regulatory events that are still poorly understood, particularly the early stages of plant-pathogen interaction. For instance, the pathogen's quorum sensing system activates the synthesis of numerous virulence factors only at high population densities (Liu H. et al., 2008). The specialized repressor of pectin degradation, KdgR, is inactivated by the products of polygalacturonate hydrolysis (Liu Y. et al., 1999; Skoblyakov et al., 2004), which occurs only after significant cell wall degradation. Additionally, the PhoPQ-dependent switch in the pathogen's metabolism and transmembrane transport is triggered by the release of divalent cations during cell wall degradation (Kravchenko et al., 2021). These three regulatory mechanisms are well studied but operate relatively late in the infection process; by this time some damage to the plant is already inevitable, making them less effective as targets for pathogen control. Therefore, the initial stage of pathogen recognition by the plant, mediated by membrane receptors, appears more promising. However, specific plant receptors for *Pectobacterium* spp. have not yet been described.

The classical model of plant immunity (Jones, Dangl, 2006) is based on the specific recognition of pathogen-associated molecular patterns (PAMP/MAMP, Pathogen/Microbe-Associated Molecular Pattern) and effectors, which induce interconnected pathways of PAMP-triggered immunity (PTI) and effector-triggered immunity (ETI). PAMPs are often conserved proteins (flagellin, translation elongation factor Tu) (Gómez-Gómez, Boller, 2000; Zipfel et al., 2006), lipids (e. g., 3-hydroxydecanoic acid) (Kutschera et al., 2019), peptidoglycan (Willmann et al., 2011), and polysaccharides (Kawaharada et al., 2015).

Effectors are proteins translocated by the pathogen directly into plant cells. Some effectors, such as Avr proteins of the fungal pathogen *Cladosporium fulvum* (*Fulvia fulva*), act outside the cell (Rooney et al., 2005). The primary function of effector proteins is to disrupt plant signaling pathways responsible for pathogen recognition and immune response activation, with effector mechanisms being quite diverse (Giraldo, Valent, 2013; Macho, Zipfel, 2015; Zhang S. et al., 2022). Despite the key role of effectors in pathogen adaptation to its host, the presence of a plant immune receptor specific to a particular effector (encoded by an *R*-gene) activates ETI and provides resistance to infection.

PAMP/MAMP detection is carried out by membrane receptor complexes, while effector receptors triggering ETI can be either cytoplasmic or membrane-bound (Böhm et al., 2014; Couto, Zipfel, 2016; Bentham et al., 2020; Sun, Zhang J., 2020). Membrane receptors for MAMP/PAMP and cytoplasmic receptors for effectors can physically interact with each other (Qi et al., 2011). Despite different initial components, the subsequent signaling pathways and activated immune responses largely overlap, so the difference between PTI/MTI and ETI is quantitative rather than qualitative (Navarro et al., 2004; Thomma et al., 2011; Yuan et al., 2021).

Receptor complexes typically include co-receptors, which can be part of many receptor complexes. Receptors and co-receptors belong to several protein families, but most are part of the leucine-rich repeat (LRR) receptor domain family (Shiu, Bleecker, 2003; Chakraborty et al., 2019; Dievart et al., 2020). Receptors usually have more than 10 LRRs, while co-receptors have fewer than 9. Specific receptors may or may not have a cytoplasmic kinase domain and are called receptor-like kinases (RLK) or receptor-like proteins (RLP). Co-receptors typically have a kinase domain and can phosphorylate other

components of the receptor complex, including the receptor itself. A receptor complex may include several co-receptors (mandatory in receptor complexes involving RLP) (Huang, Joosten, 2025).

Applying the classical zigzag model of immunity (Jones, Dangl, 2006) to *Pectobacterium* is challenging due to the limited data on PAMP-triggered responses (Kröner et al., 2011; Kuzmich et al., 2014), and the fact that, to date, only one effector protein, DspE (DspA), has been identified for *Pectobacterium* spp. (Nikolaichik et al., 2005; Kim J.-G. et al., 2011). DspE belongs to the AvrE superfamily of type III secretion system (T3SS) effectors (Nikolaichik et al., 2005; Degrave et al., 2015). Effectors in this family are named in different species after "avirulence", "disease-specific protein" and "water-soaking" according to phenotype induced in plants (e. g. AvrE, DspE, and WtsE) and are considered critical for pathogens with a small number of effectors (*Erwinia* spp., *Pantoea* spp., and *Pectobacterium* spp.) (Gaudriault et al., 1997; Frederick et al., 2001; Mor et al., 2001; Kim H.-S. et al., 2011).

DspE of *P. versatile* (*Pve*) is required for successful infection of the host plant. It is also the main inducer of the hypersensitive response (a typical sign of ETI) in non-hosts. DspE is delivered into plant cells via T3SS and can be detected within host cells as early as 3 hours after infection with a non-induced culture (Nikolaichik et al., 2005). DspE triggers local and systemic defense responses, indicating the plant's ability to detect this effector protein (Nikolaichik, 2009).

The first direct evidence of a plant's ability to specifically recognize DspE was obtained by studying the ortholog of this effector from *Erwinia amylovora*, the causative agent of apple fire blight. In yeast two-hybrid screen, four receptor-like kinases (DIPM1-4) were identified in *Malus* × *domestica* plants, the intracellular domains of which specifically interacted with DspE (Meng et al., 2006). DIPMs (DspE-Interacting Proteins from *Malus*) are receptor-like kinases with leucine-rich repeats in the sensory domain (LRR-RLK), belonging to the third subfamily (LRR-RLKIII).

Inactivation of certain DIPMs via silencing or genome editing increased plant resistance to fire blight (Borejsza-Wysocka, 2006; Pompili et al., 2020). A similar approach allowed us to identify three receptor-like kinases in tomato and tobacco that interact with DspE from *P. versatile*. Silencing the *RLK2* and *RLK5* genes in *Nicotiana benthamiana* reduced the plant's ability to recognize *P. versatile*, leading to a weakened hypersensitive response (Nikolaichik et al., 2012; Badalyan, Nikolaichik, 2014). Three receptor-like kinases (WIP3-5) that specifically interact with WtsE, an ortholog of DspE from *Pantoea stewartii* subsp. *stewartii*, were also identified in *Zea mays*, but their functions have not yet been studied *in planta* (Jin et al., 2016).

The phenotype of apple and *N. benthamiana* plants with inactivated DIPM1-4 and RLK2/5 indicates that these receptor-like protein kinases are responsible for plant sensitivity to DspE-producing pathogens, meaning they can be considered *S*-genes. Inactivation of *S*-genes can be used to create resistant plants. However, the characterized DspE-interacting kinases partially duplicate each other's functions, and their full spectrum is unknown, so complete elimination of pathogen sensitivity through the inactivation of a single or even a couple of these receptor genes seems unlikely. On the other hand, to ensure resistance to the pathogen, a single "suitable" *R*-gene might be sufficient, and such a gene encoding an LRR-RLKIII has been described for another pathosystem (Zhao et al., 2019).

The discussion above shows that LRR-RLKIII can be candidates for *S*- and *R*-genes against various pathogens and can also perform functions related to plant growth and development. This work summarizes and classifies available data on LRR-RLKIII with a focus on plants of the Solanaceae family, and may simplify the search for promising resistance genes for use in breeding programs of Solanaceae plants. We also provide an example of using this LRR-RLKIII classification to identify a potential resistance gene to pectobacterial infection.

Materials and methods

Plant material and microorganism strains. *Solanum tuberosum* cv. Ragneda and *N. benthamiana* plants were grown in non-sterile nutrient soil at 20 °C with a 16-hour photoperiod. The following microbial strains were used: *P. versatile* JN42 (*mcrB*::ISPcc2, Δ*fliTEFG*, Cm^R (Tn9), Rif^R), VKE (JN42 *dspE*) (Nikolaichik et al., 2005); *A. tumefaciens* GV3101 (Rif^R, Gm^R, vir^+) (Arabidopsis Biological Resource Center); *Saccharomyces cerevisiae* SKY48 (*MATα*, *trp1*, *his3*, *ura3*, *lexAop-LEU2*, *cIop-LYS2*), *SKY473* (*MATa*, *his3*, *leu2*, *trp1*, *ura3*, *lexAop-LEU2*, *cIop-LYS2*) (Serebriiskii et al., 2005). The JN42 strain is derived from the natural *P. versatile* isolate 3-2, which lacks flagella due to a deletion within the *fli* cluster (GenBank CP024842). Cultures of *P. versatile* and *A. tumefaciens* were grown on LB medium, while *S. cerevisiae* was cultured on YPD medium at 28 °C.

Nucleic acids. The following plasmids were used: pTRV2, p1039, p1044, p1046 (Liu Y. et al., 2002), obtained from the Arabidopsis Biological Resource Center; pTRV2::*RLK2*, pTRV2::*RLK5*; pJG4-5; pJG4-5::*dspF*, pJK202::*dspE* (Nikolaichik et al., 2012); pJG4-5::*′slRLK2*, pJG4-5::*′slRLK5*, pJG4-5::*′ntRLK5* (Badalyan, Nikolaichik, 2014). The oligonucleotide sequences for RT-qPCR are listed in Table S1 (Supplementary Materials)[1].

Molecular cloning. A fragment of the *C00T013379* (*sbRLK6*) gene from *S. bulbocastanum* was amplified using the primers 5′-ccgaattcggtttatttcctggtaagat-3′ and 5′-cgcctcga gggccaactcattgagaatcag-3′ and cloned into the pJG4-5 and pTRV2 vectors using EcoRI and XhoI restriction sites.

Protein-protein interaction analysis. Protein-protein interactions were analyzed using the LexA-based yeast two-hybrid system (Serebriiskii et al., 2007). The bait plasmid had a fragment of the *dspE* gene cloned into the pJK202 vector. For positive control of interaction with DspE, the secretory chaperone DspF, specific to DspE, was used (Valentovich et al., 2008). Cells of the *S. cerevisiae* SKY473 strain were transformed with pJG4-5 derivatives. Cells of the opposite mating type SKY48 were transformed with pJK202::*dspE*. To detect protein-protein interactions, diploid *S. cerevisiae* cells obtained by crossing SKY473 and SKY48 strains were

[1] Supplementary Tables S1 and S2 are available at: https://vavilov.elpub.ru/jour/manager/files/Suppl_Shrub_Engl_29_4.pdf

cultured on a selective medium for 2–4 days, followed by a "blue-white" test for β-galactosidase activity.

Protein sequence analysis. Genomic assemblies and annotations of the following versions were used: *S. tuberosum* DM v.6.1 (Pham et al., 2020), *S. bulbocastanum* (Tang et al., 2022), *S. lycopersicum* cv. Micro-Tom v.1.2.1 (Kudo et al., 2017), *Arabidopsis thaliana* Araport11 as of 2022-09-14 (Cheng et al., 2017).

Identification of receptor-like kinases in proteomes and their classification into families was performed using iTAK v. 1.2 (Zheng et al., 2016). Redundancy in sequences was reduced using the easy-cluster algorithm of the MMseqs2 program (Steinegger, Söding, 2017). For the alignment of kinase domain amino acid sequences, the MAFFT web service was used (Rozewicki et al., 2019). Alignment correction was performed using Jalview 2.11.2.7 (Waterhouse et al., 2009). A maximum likelihood dendrogram was constructed using ITREE v. 2.1.3 based on the Edge-linked partition model (von Haeseler et al., 2014; Chernomor et al., 2016), and graphical visualization was performed using the iTOL 6.8 service (Letunic, Bork, 2021).

Virus-induced gene silencing (VIGS). VIGS in *N. benthamiana* plants was performed using the TRV2 vector as described by Y. Liu et al. (2002). The hypersensitivity test was conducted 40 days after VIGS induction by infiltrating plant leaves with suspensions of *P. versatile* strains in 0.85 % NaCl at a density of 1.5×10^8 cells/ml using a needle-less syringe. At least 10 plants of each type were subjected to infiltration with all types of suspensions and NaCl solution as a control. For each type of suspension or NaCl solution, two or more leaves per plant were used. Results were recorded and samples were collected 24 hours after infection.

RT-qPCR. Leaf samples from *N. benthamiana* were collected 24 hours after infiltration (as described above) with *P. versatile* cell suspensions. Potato tubers were inoculated with an automatic pipette using suspensions of the same density (1.5×10^8 cells/ml) in a volume of 10 µl. Potato tissue samples were collected at the maceration zone boundary 48 hours after inoculation. RNA extraction and RT-qPCR were performed in six biological replicates as described (Nikolaichik et al., 2009). Gene expression levels were determined by RT-qPCR relative to the reference genes *CAC*, *EF1A*, *TBP* for *N. benthamiana*, and *SAND*, *CAC*, *EF1a* for *S. tuberosum* and calculated with the REST2009 2.0.13 software (Qiagen, USA).

Results

Phylogenetic analysis identifies distinct structural and functional subgroups within LRR-RLKIII

Since all known RLKs (Receptor-Like Kinases) that interact with AvrE-like effectors belong to the LRR-RLKIII family, we analyzed the spectrum of RLKs in this family across various members of the Solanaceae family and compared them with the characterized LRR-RLKIII (mostly from *A. thaliana*). The iTAK classifier assigns between 45 and 77 RLKs to LRR-RLKIII in different Solanaceae species and 47 RLKs in *A. thaliana* (see the Table), leaving many options for identifying potential receptors involved in immunity regulation and pathogen detection, including *P. versatile*.

Phylogenetic analysis allowed us to identify nine clusters of related LRR-RLKIII (Fig. 1). For RLKs in clusters I–V and VIII, published information does not show a connection to immunity. In the remaining three clusters (VI, VII, and IX), kinases capable of binding AvrE-like effector proteins are present but immunity-related functions have only been demonstrated for members of clusters VII and IX. More detailed information on experimentally characterized LRR-RLKIII is provided in Table S2.

Identification of a NEw DspE-interacting LRR-RLKIII in *S. bulbocastanum*

So far, no *R*-genes have been identified among those encoding DspE-interacting RLKs. Most known immunity-related genes from this family, such as *mdRLK4*, *slTARK1*, *nbEIR1*, and *ntRLK2*, can be classified as *S*-genes. However, at least one clear *R*-gene in this family (*RLK902* of *A. thaliana*, required for resistance to *Hyaloperonospora arabidopsidis*) has been described (ten Hove et al., 2011), suggesting the possibility of identifying resistance genes to other pathogens in this family. RLKs with a demonstrated role in immunity (includ-

Quantitative analysis of RLKs in plants of the Solanaceae family

Species	Total RLKs	Number of RLK classes (iTAK)	Rank of LRR-RLKIII among other classes	Number of LRR-RLKIII members
Capsicum annuum var. *glabirisculum*	1,151	121	III	59
S. melongena	1,189	123	VI	45
S. bulbocastanum	1,892	123	VI	69
S. tuberosum	1,328	124	VII	46
S. lycopersicum cv. Heinz	1,020	123	III	45
S. lycopersicum cv. Micro-Tom	2,073	123	VIII	61
N. benthamiana	1,329	123	IV	77
A. thaliana	1,028	123	III	47

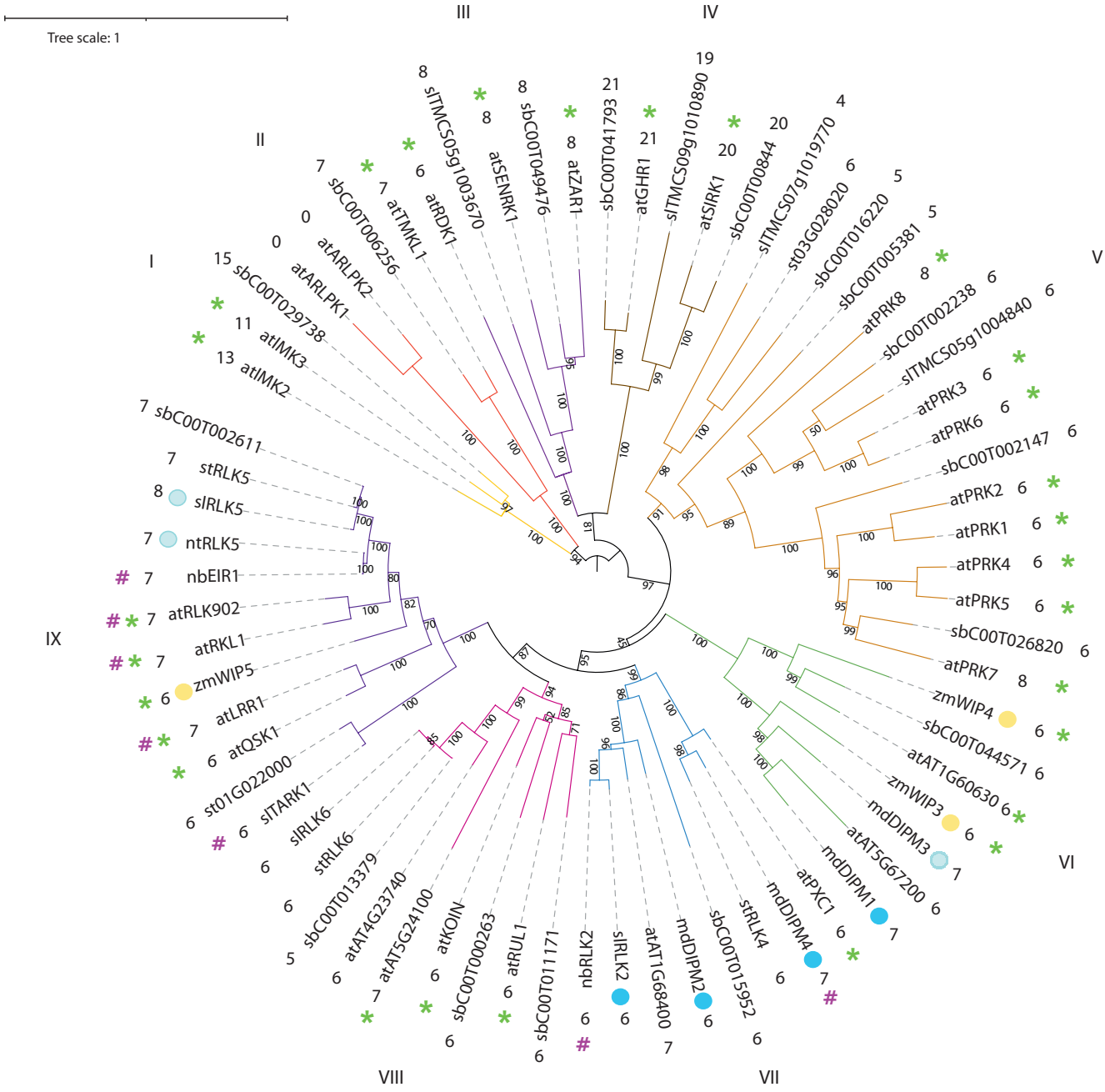


Fig. 1. Dendrogram of the LRR-RLKIII family members.

Only kinase domain sequences were used. Species are indicated by two-letter codes: at – *A. thaliana*, md – *Malus* × *domestica*, nb – *N. benthamiana*, nt – *N. tabacum*, sb – *S. bulbocastanum*, sl – *S. lycopersicum*, st – *S. tuberosum*, zm – *Z. mays*. Clusters of similar RLKs are highlighted in color and labeled with Roman numerals. LRR-RLKs interacting with AvrE family effectors (DspA/E, WtsE) are marked with circles, where the shading intensity is proportional to the interaction strength, and the color (blue or yellow) indicates whether the blue-white test was performed. Numbers next to the identifiers show the number of leucine-rich repeats (LRR) within the sensory domain. Experimentally studied LRR-RLKs are marked with "*" if they are involved in growth and development regulation and/or "#" if they are involved in immunity control.

ing those interacting with AvrE-type effectors) are present only in clusters VII and IX. The connection to immunity of DspE/WtsE-interacting RLKs from cluster VI has not yet been shown but seems plausible. Cluster VIII is the closest to clusters VII, IX, and VI, but none of its members was shown to interact with DspE. Therefore, we selected a representative from cluster VIII for experimental analysis: the *C00T013379* gene from *S. bulbocastanum*, a species often used as a source of *R*-genes suitable for potato resistance breeding.

The cloned fragment of *C00T013379*, encoding the cytoplasmic part of the RLK, gave a positive result in the interaction test with DspE in the yeast two-hybrid system (Fig. 2). Based on the intensity of growth and coloration of yeast macro-colonies on selective medium, the interaction of DspE

Fig. 2. Interaction of the effector protein DspE with kinase domains of RLKs. Growth of diploids on leucine-deficient medium with X-gal.
All cells contain plasmids pJK202-DspE, pDR8 with *lacZ*, and derivatives of plasmid pJG4-5 with an insertion of the reading frame of the indicated gene.

Fig. 3. Hypersensitive response of *N. benthamiana* plants subjected to *RLK* gene silencing.
Control plants were infected with TRV containing a neutral insert (GFP).

with kinase domain amplified from *S. bulbocastanum* was at the level of slRLK2 and the positive control (DspF) and significantly exceeded the interaction intensity with slRLK5 and ntRLK5 (orthologs of nbEIR1). By analogy with the previously characterized LRR-RLKIII, we designated this gene as *sbRLK6*.

Silencing the ortholog of *sbRLK6* in *N. benthamiana*, unlike silencing *nbRLK2* and *nbRLK5*, did not affect the intensity of the hypersensitive response upon infiltration of leaves with *Pve* JN42 cell suspension (Fig. 3).

P. versatile suppresses plant LRR-RLK genes

To understand the consequences of DspE recognition by LRR-RLKIII, we assessed the expression levels of key immunity marker genes in both host plants (*S. tuberosum*) and non-host plants (*N. benthamiana*). Radical changes were observed for genes involved in salicylic acid and jasmonic acid signaling, as well as for the RLK genes (Fig. 4).

In *Pve*-infected *N. benthamiana* plants, expression of salicylic acid-dependent genes *PR1A* and *SIPK* was significantly reduced (Fig. 4*a*). In contrast, the jasmonic acid-dependent transcriptional activator gene *JAZ3* was induced, while *COI1* encoding the inhibitor of this pathway was repressed, and the marker genes *WIPK* and *PR3* were strongly induced. Importantly, both suppression and induction of these genes were dependent on DspE, as the response of plants to inoculation with *dspE* mutant bacteria was much weaker. Similar DspE-dependent suppression of the salicylic acid pathway marker *PR1A* was observed in *Pve*-infected potato plants (Fig. 4*b*). *COI1* was slightly repressed, while *JAZ3* – strongly induced, but the effect was independent of DspE.

Expression of *RLK2* and *RLK5* orthologs in both plants, as well as that of *RLK4* in *N. benthamiana*, was also reduced in a DspE-dependent manner (Fig. 4*a*, *b*). The very low expression level of *stRLK6* did not allow for an assessment of its change.

To see how typical the observed expression pattern is during *Pve* infection, we checked the expression of two well-studied RLK genes from other families in *N. benthamiana*: *FLS2*, encoding the flagellin receptor, and *WAK1*, encoding a cell wall-associated kinase involved in oligogalacturonate perception. *FLS2* responded to contact with wild-type *Pve* and the *dspE* mutant similarly to *RLK2* and *RLK5* (Fig. 4*a*). *WAK1* in *Pve*-infected plants showed a barely noticeable (approximately twofold) but reproducible suppression, independent of DspE.

Discussion

The phylogenetic analysis allowed us to divide LRR-RLKIII into two distinct functional groups. According to the available information, we conclude that members of clusters I–V are

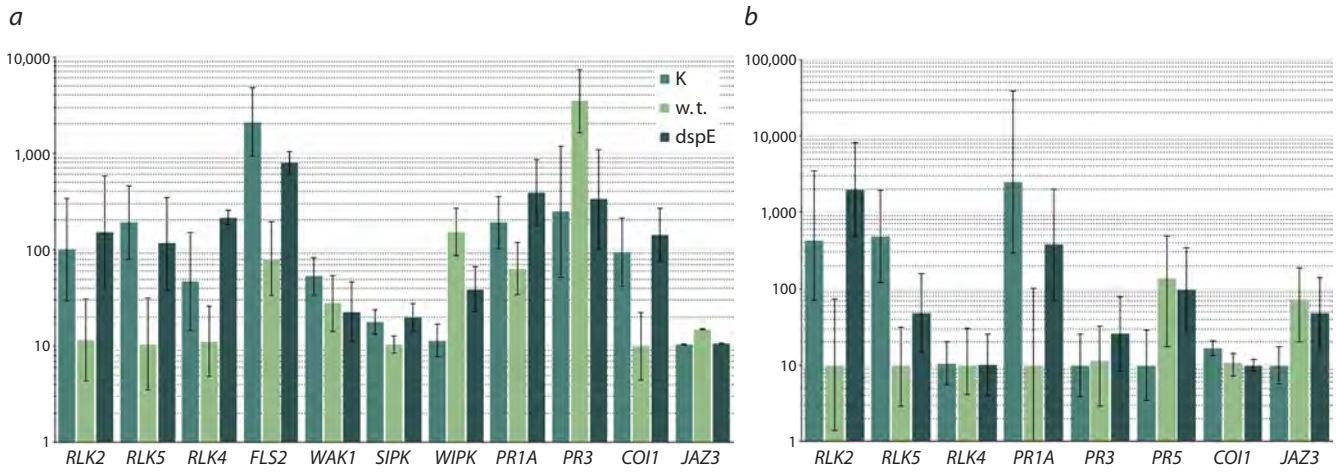


Fig. 4. Changes in gene expression levels in *N. benthamiana* (*a*) and *S. tuberosum* (*b*) plants inoculated with suspensions of wild-type *P. versatile* (w.t.) and *dspE* mutant (dspE) or 0.85 % NaCl solution (K).

Average values (in arbitrary units) of six measurements with 95 % confidence intervals are shown.

primarily involved in controlling plant growth and development, while clusters VI–IX are enriched with RLKs associated with the regulation of immune responses. Most LRR-RLKIII, including all those interacting with DspE, lack a key aspartate residue in the conserved catalytic loop motif (HRDXXXXN) and are therefore classified as pseudokinases. However, many pseudokinases retain some kinase activity and can function as part of receptor complexes containing an active kinase (Rodriguez-Furlan et al., 2022). Additionally, LRR-RLKIII generally have a small number (5–7) of leucine-rich repeats in their sensor domain, meaning they can only perform their signaling function as part of complex receptor systems that must include two additional critical components: a kinase and a receptor (or receptor-like protein) with a full set (more than 10, typically around 20) of leucine-rich repeats.

Known DspE-interacting proteins belong to clusters VI, VII, and IX. Among these, the role in pathogen recognition has been established for four proteins: mdDIPM4 and nbRLK2 from cluster VII and ntRLK5 and nbEIR1 from cluster IX. In all these cases, suppression of the RLK gene increased plant resistance (Borejsza-Wysocka et al., 2006; Nikolaichik et al., 2012; Badalyan, Nikolaichik, 2014; Pompili et al., 2020). Based on these results, *mdDIPM4*, *nbRLK2*, and *nbEIR1* can be considered *S*-genes. However, we note that due to the cross-regulation of RLK genes reported in both apple (Borejsza-Wysocka et al., 2006) and tobacco (Badalyan, Nikolaichik, 2014), the observed phenotype cannot be unequivocally linked to the suppressed gene. Another member of cluster IX, TARK1, is also a product of an *S*-gene, as its over-expression enhances, while its inactivation weakens disease symptoms (Kim J.-G. et al., 2009; Campos, 2020; Guzman et al., 2020).

The *sbRLK6* gene, described for the first time in this study, encodes an LRR-RLKIII belonging to cluster VIII, where no DspE-interacting RLKs had been previously described. Silencing the ortholog of *sbRLK6* in *N. benthamiana* plants did not (unlike silencing *RLK2* and *RLK5*) weaken the hypersensitive response of plants upon contact with *Pve* bacteria. For necrotrophs (including *Pve*), necrosis accompanying the hypersensitive response is favorable for expanding the infection zone and further colonization of the plant, so we classify *RLK2* and *RLK5* as *S*-genes. However, *sbRLK6* cannot currently be considered an *S*-gene. The question of whether *RLK6* can function as an *R*-gene requires its stable inactivation or over-expression followed by testing plant resistance when infecting organs (tubers and stems) that are typical targets of pectobacteria but are poorly suited for the virus-induced gene silencing technology used here.

Since the genes of co-receptor-like RLK2, RLK5, and the newly described RLK6 exhibit an expression pattern during *Pve* infection similar to that of the pattern recognition receptor gene *FLS2*, and the expression of all four genes shows signs of common control through the salicylic acid signaling pathway, we can hypothesize that DspE-interacting LRR-RLKIII are components of complex receptor systems involved in detecting *Pve*. Other components of such complexes and the cytoplasmic signaling proteins interacting with them may be considered promising candidates for identifying resistance genes to *Pve* and other bacteria.

Conclusion

So far, specific resistance genes to *Pectobacterium spp.* have not been described, and targeted breeding of potatoes for resistance to pectobacteriosis is not being conducted, not least due to the lack of information on *R*-genes that could provide such resistance. We propose to use receptor protein kinases of the RLK-LRRIII family, which specifically recognize the main effector protein of pectobacteria, DspE, for this purpose. As a first step in this direction, this study has compiled information on experimentally studied members of this family, identified promising subfamilies (clusters) for further research, and identified a new receptor-like kinase that specifically recognizes DspE and has properties distinct from those previously described.

References

Badalyan O.A., Nikolaichik Y.A. Receptor-like kinases RLK2 and RLK5 of *Nicotiana benthamiana* are involved in regulation of gene expression of key plant immune system components during the contact with *Pectobacterium carotovorum*. *Izvestiya NAN Belarusi. Seriya Biologicheskikh Nauk = Proceedings of the National Academy of Sciences of Belarus. Biological Series*. 2014;4:75-80 (in Russian)

Bentham A.R., De la Concepcion J.C., Mukhi N., Zdrzałek R., Draeger M., Gorenkin D., Hughes R.K., Banfield M.J. A molecular roadmap to the plant immune system. *J Biol Chem*. 2020;295(44): 14916-14935. doi 10.1074/jbc.REV120.010852

Böhm H., Albert I., Fan L., Reinhard A., Nürnberger T. Immune receptor complexes at the plant cell surface. *Curr Opin Plant Biol*. 2014;20:47-54. doi 10.1016/j.pbi.2014.04.007

Borejsza-Wysocka E.E., Malnoy M., Aldwinckle H.S., Meng X., Bonasera J.M., Nissinen R.M., Kim J.F., Beer S.V. The fire blight resistance of apple clones in which DspE-interacting proteins are silenced. *Acta Hortic*. 2006;704:509-514. doi 10.17660/ActaHortic.2006.704.80

Campos M.L. A novel regulator of stomatal immunity in tomato. *Plant Physiol*. 2020;183(3):820-821. doi 10.1104/pp.20.00655

Chakraborty S., Nguyen B., Wasti S.D., Xu G. Plant leucine-rich repeat receptor kinase (LRR-RK): structure, ligand perception, and activation mechanism. *Molecules*. 2019;24(17):3081. doi 10.3390/molecules24173081

Chatterjee A., Cui Y., Liu Y., Dumenyo C.K., Chatterjee A.K. Inactivation of *rsmA* leads to overproduction of extracellular pectinases, cellulases, and proteases in *Erwinia carotovora* subsp. *carotovora* in the absence of the starvation/cell density-sensing signal, N-(3-oxohexanoyl)-L-homoserine lactone. *Appl Environ Microbiol*. 1995; 61(5):1959-1967. doi 10.1128/aem.61.5.1959-1967.1995

Cheng C.-Y., Krishnakumar V., Chan A.P., Thibaud-Nissen F., Schobel S., Town C.D. Araport11: a complete reannotation of the *Arabidopsis thaliana* reference genome. *Plant J*. 2017;89(4):789-804. doi 10.1111/tpj.13415

Chernomor O., von Haeseler A., Minh B.Q. Terrace aware data structure for phylogenomic inference from supermatrices. *Syst Biol*. 2016;65(6):997-1008. doi 10.1093/sysbio/syw037

Couto D., Zipfel C. Regulation of pattern recognition receptor signalling in plants. *Nat Rev Immunol*. 2016;16(9):537-552. doi 10.1038/nri.2016.77

Degrave A., Siamer S., Boureau T., Barny M.-A. The AvrE superfamily: ancestral type III effectors involved in suppression of pathogen-associated molecular pattern-triggered immunity. *Mol Plant Pathol*. 2015;16(8):899-905. doi 10.1111/mpp.12237

Dievart A., Gottin C., Périn C., Ranwez V., Chantret N. Origin and diversity of plant receptor-like kinases. *Annu Rev Plant Biol*. 2020;71: 131-156. doi 10.1146/annurev-arplant-073019-025927

Frederick R.D., Ahmad M., Majerczak D.R., Arroyo-Rodríguez A.S., Manulis S., Coplin D.L. Genetic organization of the *Pantoea stewartii* subsp. *stewartii hrp* gene cluster and sequence analysis of the *hrpA, hrpC, hrpN,* and *wtsE* operons. *Mol Plant Microbe Interact*. 2001;14(10):1213-1222. doi 10.1094/MPMI.2001.14.10.1213

Gaudriault S., Malandrin L., Paulin J.-P., Barny M.-A. DspA, an essential pathogenicity factor of *Erwinia amylovora* showing homology with AvrE of *Pseudomonas syringae*, is secreted via the Hrp secretion pathway in a DspB-dependent way. *Mol Microbiol*. 1997; 26(5):1057-1069. doi 10.1046/j.1365-2958.1997.6442015.x

Giraldo M.C., Valent B. Filamentous plant pathogen effectors in action. *Nat Rev Microbiol*. 2013;11(11):800-814. doi 10.1038/nrmicro3119

Gómez-Gómez L., Boller T. FLS2: An LRR receptor-like kinase involved in the perception of the bacterial elicitor flagellin in *Arabidopsis*. *Mol Cell*. 2000;5(6):1003-1011. doi 10.1016/S1097-2765(00)80265-8

Gorshkov V., Gubaev R., Petrova O., Daminova A., Gogoleva N., Ageeva M., Parfirova O., Prokchorchik M., Nikolaichik Y., Gogolev Y. Transcriptome profiling helps to identify potential and true molecular switches of stealth to brute force behavior in *Pectobacterium atrosepticum* during systemic colonization of tobacco plants. *Eur J Plant Pathol*. 2018;152(4):957-976. doi 10.1007/s10658-018-1496-6

Guzman A.R., Kim J.-G., Taylor K.W., Lanver D., Mudgett M.B. Tomato atypical receptor kinase1 is involved in the regulation of preinvasion defense. *Plant Physiol*. 2020;183(3):1306-1318. doi 10.1104/pp.19.01400

Huang W.R.H., Joosten M.H.A.J. Immune signaling: receptor-like proteins make the difference. *Trends Plant Sci*. 2025;30(1):54-68. doi 10.1016/j.tplants.2024.03.012

Jin L., Ham J.H., Hage R., Zhao W., Soto-Hernández J., Lee S.Y., Paek S.-M., Kim M.G., Boone C., Coplin D.L., Mackey D. Direct and indirect targeting of PP2A by conserved bacterial type-III effector proteins. *PLoS Pathog*. 2016;12(5):e1005609. doi 10.1371/journal.ppat.1005609

Jones J.D.G., Dangl J.L. The plant immune system. *Nature*. 2006; 444(7117):323-329. doi 10.1038/nature05286

Kawaharada Y., Kelly S., Nielsen M.W., Hjuler C.T., Gysel K., Muszyński A., Carlson R.W., … Jensen K.J., Ronson C.W., Blaise M., Radutoiu S., Stougaard J. Receptor-mediated exopolysaccharide perception controls bacterial infection. *Nature*. 2015;523(7560): 308-312. doi 10.1038/nature14611

Kim J.-G., Li X., Roden J.A., Taylor K.W., Aakre C.D., Su B., Lalonde S., Kirik A., Chen Y., Baranage G., McLane H., Martin G.B., Mudgett M.B. *Xanthomonas* T3S effector XopN suppresses PAMP-triggered immunity and interacts with a tomato atypical receptor-like kinase and TFT1. *Plant Cell*. 2009;21(4):1305-1323. doi 10.1105/tpc.108.063123

Kim H.-S., Thammarat P., Lommel S.A., Hogan C.S., Charkowski A.O. *Pectobacterium carotovorum* elicits plant cell death with DspE/F but the *P. carotovorum* DspE does not suppress callose or induce expression of plant genes early in plant-microbe interactions. *Mol Plant Microbe Interact*. 2011;24(7):773-786. doi 10.1094/MPMI-06-10-0143

Kravchenko U., Gogoleva N., Kalubaka N., Kruk A., Diubo Y., Gogolev Y., Nikolaichik Y. The PhoPQ two-component system is the major regulator of cell surface properties, stress responses and plant-derived substrate utilisation during development of *Pectobacterium versatile*-host plant pathosystems. *Front Microbiol*. 2021;11: 621391. doi 10.3389/fmicb.2020.621391

Kröner A., Hamelin G., Andrivon D., Val F. Quantitative resistance of potato to *Pectobacterium atrosepticum* and *Phytophthora infestans*: Integrating PAMP-triggered response and pathogen growth. *PLoS One*. 2011;6(8):e23331. doi 10.1371/journal.pone.0023331

Kudo T., Kobayashi M., Terashima S., Katayama M., Ozaki S., Kanno M., Saito M., Yokoyama K., Ohyanagi H., Aoki K., Kubo Y., Yano K. TOMATOMICS: a web database for integrated omics information in tomato. *Plant Cell Physiol*. 2017;58(1):e8. doi 10.1093/pcp/pcw207

Kutschera A., Dawid C., Gisch N., Schmid C., Raasch L., Gerster T., Schäffer M., … Ernst R.K., Dorey S., Hückelhoven R., Hofmann T., Ranf S. Bacterial medium-chain 3-hydroxy fatty acid metabolites trigger immunity in *Arabidopsis* plants. *Science*. 2019;364(6436): 178-181. doi 10.1126/science.aau1279

Kuzmich S.V., Badalyan O.A., Nikolaychik E.A. Analysis of induction and suppression of MAMP-induced immunity of *Nicotiana benthamiana* plants upon contact with *Pectobacterium atrosepticum*. *Vestnik Belorusskogo Gosudarstvennogo Universiteta. Seriya 2: Khimiya. Biologiya. Geografiya = Bulletin of the Belarusian State University. Series 2: Chemistry, Biology, Geography.* 2014;(2):36-40.

Kwenda S., Motlolometsi T.V., Birch P.R.J., Moleleki L.N. RNA-seq profiling reveals defense responses in a tolerant potato cultivar to

stem infection by *Pectobacterium carotovorum* ssp. *brasiliense*. *Front Plant Sci*. 2016;7:1905. doi 10.3389/fpls.2016.01905

Letunic I., Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res*. 2021;49(W1):W293-W296. doi 10.1093/nar/gkab301

Liu H., Coulthurst S.J., Pritchard L., Hedley P.E., Ravensdale M., Humphris S., Burr T., Takle G., Brurberg M.-B., Birch P.R.J., Salmond G.P.C., Toth I.K. Quorum sensing coordinates brute force and stealth modes of infection in the plant pathogen *Pectobacterium atrosepticum*. *PLoS Pathog*. 2008;4(6):e1000093. doi 10.1371/journal.ppat.1000093

Liu Y., Jiang G., Cui Y., Mukherjee A., Ma W.L., Chatterjee A.K. $kdgR_{Ecc}$ negatively regulates genes for pectinases, cellulase, protease, $Harpin_{Ecc}$, and a global RNA regulator in *Erwinia carotovora* subsp. *carotovora*. *J Bacteriol*. 1999;181(8):2411-2421. doi 10.1128/jb.181.8.2411-2421.1999

Liu Y., Schiff M., Dinesh-Kumar S.P. Virus-induced gene silencing in tomato. *Plant J*. 2002;31(6):777-786. doi 10.1046/j.1365-313X.2002.01394.x

Macho A.P., Zipfel C. Targeting of plant pattern recognition receptor-triggered immunity by bacterial type-III secretion system effectors. *Curr Opin Microbiol*. 2015;23:14-22. doi 10.1016/j.mib.2014.10.009

Meng X., Bonasera J.M., Kim J.F., Nissinen R.M., Beer S.V. Apple proteins that interact with DspA/E, a pathogenicity effector of *Erwinia amylovora*, the fire blight pathogen. *Mol Plant Microbe Interact*. 2006;19(1):53-61. doi 10.1094/MPMI-19-0053

Mor H., Manulis S., Zuck M., Nizan R., Coplin D.L., Barash I. Genetic organization of the *hrp* gene cluster and *dspAE/BF* operon in *Erwinia herbicola* pv. *gypsophilae*. *Mol Plant Microbe Interact*. 2001;14(3):431-436. doi 10.1094/MPMI.2001.14.3.431

Navarro L., Zipfel C., Rowland O., Keller I., Robatzek S., Boller T., Jones J.D.G. The transcriptional innate immune response to flg22. Interplay and overlap with Avr gene-dependent defense responses and bacterial pathogenesis. *Plant Physiol*. 2004;135(2):1113-1128. doi 10.1104/pp.103.036749

Nikolaichik Y.A. Systemic induction of PR genes in *Solanum lycopersicum* plants upon contact with *Pectobacterium carotovorum* bacteria: the role of the *DspE* gene. *Trudy Belorusskogo Gosudarstvennogo Universiteta = Proceedings of the Belarusian State University*. 2009;4(2):215-220 (in Russian)

Nikolaichik Y.A., Ovchinnikova T.V., Valentovich L.N., Gubich O.I., Sholukh M.V., Evtushenkov A.N. DspE protein is translocated by phytopathogenic bacteria *Erwinia carotovora* subsp. *atroseptica* into the cells of *Nicotiana tabacum* and is required for the induction of the hypersensitive reaction. *Doklady Nacional'noj Akademii Nauk Belarusi = Doklady of the National Academy of Sciences of Belarus*. 2005;49(5):81-85 (in Russian)

Nikolaichik Y.A., Homskaya L.L., Ignatenko Y.I. The plant pathogen *Pectobacterium carotovorum* employs its Type III secretion system for blocking the systemic defense response in the host plant. *Trudy Belorusskogo Gosudarstvennogo Universiteta – Proceedings of the Belarusian State University*. 2009;4(1):193-200 (in Russian)

Nikolaichik Y.A., Kulik E.V., Badalyan O.A., Valentovich L.N., Kuzmich S.V., Evtushenkov A.N. Receptor-like transmembrane kinase of Solanaceae plants controls interaction with plant pathogen *Pectobacterium carotovorum*. *Doklady Nacional'noj Akademii Nauk Belarusi = Doklady of the National Academy of Sciences of Belarus*. 2012;56(1):106-112 (in Russian)

Pérombelon M.C.M. Potato diseases caused by soft rot erwinias: an overview of pathogenesis. *Plant Pathol*. 2002;51(1):1-12. doi 10.1046/j.0032-0862.2001.Shorttitle.doc.x

Pham G.M., Hamilton J.P., Wood J.C., Burke J.T., Zhao H., Vaillancourt B., Ou S., Jiang J., Buell C.R. Construction of a chromosome-scale long-read reference genome assembly for potato. *GigaScience*. 2020;9(9):giaa100. doi 10.1093/gigascience/giaa100

Pompili V., Dalla Costa L., Piazza S., Pindo M., Malnoy M. Reduced fire blight susceptibility in apple cultivars using a high-efficiency CRISPR/Cas9-FLP/FRT-based gene editing system. *Plant Biotechnol J*. 2020;18(3):845-858. doi 10.1111/pbi.13253

Qi Y., Tsuda K., Nguyen L.V., Wang X., Lin J., Murphy A.S., Glazebrook J., Thordal-Christensen H., Katagiri F. Physical association of *Arabidopsis* hypersensitive induced reaction proteins (HIRs) with the immune receptor RPS2. *J Biol Chem*. 2011;286(36):31297-31307. doi 10.1074/jbc.M110.211615

Rodriguez-Furlan C., Campos R., Toth J.N., Van Norman J.M. Distinct mechanisms orchestrate the contra-polarity of IRK and KOIN, two LRR-receptor-kinases controlling root cell division. *Nat Commun*. 2022;13(1):235. doi 10.1038/s41467-021-27913-1

Rooney H.C., Van't Klooster J.W., van der Hoorn R.A., Joosten M.H., Jones J.D., de Wit P.J. Cladosporium Avr2 inhibits tomato Rcr3 protease required for *Cf-2*-dependent disease resistance. *Science*. 2005;308(5729):1783-1786. doi 10.1126/science.1111404

Rozewicki J., Li S., Amada K.M., Standley D.M., Katoh K. MAFFT-DASH: integrated protein sequence and structural alignment. *Nucleic Acids Res*. 2019;47(W1):W5-W10. doi 10.1093/nar/gkz342

Serebriiskii I.G., Golemis E.A., Uetz P. The yeast two-hybrid system for detecting interacting proteins. In: Walker J.M. (Ed.) The Proteomics Protocols Handbook. Springer Protocols Handbooks. Humana Press, 2005;653-682. doi 10.1385/1-59259-890-0:653

Shiu S.-H., Bleecker A.B. Expansion of the receptor-like kinase/Pelle gene family and receptor-like proteins in Arabidopsis. *Plant Physiol*. 2003;132(2):530-543. doi 10.1104/pp.103.021964

Skoblyakov S.A., Miamin V.E., Lagonenko A.L., Nikolaichik Y.A., Pesnyakevich A.G. The effect of mutations in the *peIW* and *kdgR* genes on the production of pectate lyases in *Erwinia carotovora* subsp. *atroseptica*. *Vestnik Belorusskogo Gosudarstvennogo Universiteta. Seriya 2: Khimiya. Biologiya. Geografiya = Bulletin of the Belarusian State University. Series 2: Chemistry, Biology, Geography*. 2004;(2):40-44 (in Russian)

Steinegger M., Söding J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat Biotechnol*. 2017;35(11):1026-1028. doi 10.1038/nbt.3988

Sun L., Zhang J. Regulatory role of receptor-like cytoplasmic kinases in early immune signaling events in plants. *FEMS Microbiol Rev*. 2020;44(6):845-856. doi 10.1093/femsre/fuaa035

Tang D., Jia Y., Zhang J., Li H., Cheng L., Wang P., Bao Z., Liu Z., Feng S., Zhu X., Li D., Zhu G., Wang H., Zhou Ya., Zhou Yo., Bryan G.J., Buell C.R., Zhang C., Huang S. Genome evolution and diversity of wild and cultivated potatoes. *Nature*. 2022;606(7914):535-541. doi 10.1038/s41586-022-04822-x

ten Hove C.A., de Jong M., Lapin D., Andel A., Sanchez-Perez G.F., Tarutani Y., Suzuki Y., Heidstra R., van den Ackerveken G. Transrepression of gene activity upstream of T-DNA tagged *RLK902* links *Arabidopsis* root growth inhibition and downy mildew resistance. *PLoS One*. 2011;6(4):e19028. doi 10.1371/journal.pone.0019028

Thomma B.P., Nürnberger T., Joosten M.H. Of PAMPs and effectors: The blurred PTI-ETI dichotomy. *Plant Cell*. 2011;23(1):4-15. doi 10.1105/tpc.110.082602

Toth I.K., Birch P.R. Rotting softly and stealthily. *Curr Opin Plant Biol*. 2005;8(4):424-429. doi 10.1016/j.pbi.2005.04.001

Valentovich L.N., Gubich O.I., Nikolaichik Y.A. The role of the DspF protein of *Erwinia carotovora* supsp. *atroseptica* in the functioning of the type III secretion system. *Doklady Nacional'noj Akademii Nauk Belarusi = Doklady of the National Academy of Sciences of Belarus*. 2008;52(5):79-85 (in Russian)

von Haeseler A., Schmidt H.A., Bui M.Q., Nguyen L.T. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol*. 2015;32(1):268-274. doi 10.1093/molbev/msu300

Waterhouse A.M., Procter J.B., Martin D.M.A., Clamp M., Barton G.J. Jalview Version 2 – a multiple sequence alignment editor and analy-

sis workbench. *Bioinformatics*. 2009;25(9):1189-1191. doi 10.1093/bioinformatics/btp033

Willmann R., Lajunen H.M., Erbs G., Newman M.-A., Kolb D., Tsuda K., Katagiri F., … Kulik A., Molinaro A., Lipka V., Gust A.A., Nürnberger T. *Arabidopsis* lysin-motif proteins LYM1 LYM3 CERK1 mediate bacterial peptidoglycan sensing and immunity to bacterial infection. *Proc Natl Acad Sci USA*. 2011;108(49):19824-19829. doi 10.1073/pnas.1112862108

Yuan M., Ngou B.P.M., Ding P., Xin X.-F. PTI-ETI crosstalk: an integrative view of plant immunity. *Curr Opin Plant Biol*. 2021;62:102030. doi 10.1016/j.pbi.2021.102030

Zhang S., Li C., Si J., Han Z., Chen D. Action mechanisms of effectors in plant-pathogen interaction. *Int J Mol Sci*. 2022;23(12):6758. doi 10.3390/ijms23126758

Zhao Y., Wu G., Shi H., Tang D. RECEPTOR-LIKE KINASE 902 associates with and Phosphorylates BRASSINOSTEROID-SIGNALING KINASE1 to regulate plant immunity. *Mol Plant*. 2019;12(1):59-70. doi 10.1016/j.molp.2018.10.008

Zheng Y., Jiao C., Sun H., Rosli H.G., Pombo M.A., Zhang P., Banf M., Dai X., Martin G.B., Giovannoni J.J., Zhao P.X., Rhee S.Y., Fei Z. iTAK: a program for genome-wide prediction and classification of plant transcription factors, transcriptional regulators, and protein kinases. *Mol Plant*. 2016;9(12):1667-1670. doi 10.1016/j.molp.2016.09.014

Zipfel C., Kunze G., Chinchilla D., Caniard A., Jones J.D.G., Boller T., Felix G. Perception of the bacterial PAMP EF-Tu by the receptor EFR restricts *Agrobacterium*-mediated transformation. *Cell*. 2006;125(4):749-760. doi 10.1016/j.cell.2006.03.037

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

Received November 15, 2024. Revised February 6, 2025. Accepted February 10, 2025.

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2025;29(4):559-567

doi 10.18699/vjgb-25-59

Influence of selected rootstock on growth parameters, accumulation of IAA and vitamins in scions of *Cucumis sativus* L. and *Cucumis melo* L.

A.Zh. Shoibekova [1], S.K. Jantassov [1], A.S. Jantassova [1], A.T. Samatov [1], T.S. Sagindykov [1], A.N. Karimova [1], G.A. Serikbayeva [1], M.R. Toishimanov [1, 2], G.T. Bari [1]

[1] Kazakh National Agrarian Research University, Almaty, Kazakhstan
[2] Institute of Plant Biology and Biotechnology, Almaty, Kazakhstan
baracuda.co@mail.ru

Abstract. Grafting with resistant rootstocks is one of the most effective methods to prevent soil-borne diseases, and it can influence vegetative growth, flowering, maturation periods, and fruit quality, thereby ensuring high yields. In this study, four species from the family Cucurbitaceae were tested as potential candidates for grafting cucumber and melon: *Cucurbita ficifolia* Bouché, *Cucurbita moschata* L., *Cucurbita pepo* L. and *Cucurbita maxima* Duch. The study focused on the grafting methods that optimize growth parameters and the accumulation of hormones and vitamins in rootstock. The results indicated that *Cucurbita maxima* Duch. is the most suitable rootstock material for grafting to *Cucumis sativus* L. and *Cucumis melo* L., as it exhibited superior plant and root mass. Among the two grafting methods tested, the tongue approach ('X') demonstrated the best results in terms of growth parameters and the accumulation of indole-3-acetic acid (IAA) and vitamins in the scion leaves. IAA and vitamin concentrations were measured using HPLC in grafted samples at 2, 4 and 6 weeks of age. In the 'X' method, IAA accumulation from the end of the second week was twice as high compared to control plants. This method also showed higher vitamin content, with increased levels of B vitamins and vitamin C at the end of the 4th week (25.2–135.1 and 52.3–67.0 %, respectively), and vitamins A, E, D_3, K starting from the 2nd week (1.5–2 times higher). Conversely, the insertion or slant cut grafting method ('Y') did not show any significant increase in the analyzed parameters and was comparable to the control. The 'X' method for grafting both *Cucumis sativus* L. and *Cucumis melo* L. onto *Cucurbita maxima* Duch. plants demonstrated the best results and is recommended for production.

Key words: morphometric analysis; grafting; *C. maxima*; HPLC; IAA; vitamins

For citation: Shoibekova A.Zh., Jantassov S.K., Jantassova A.S., Samatov A.T., Sagindykov T.S., Karimova A.N., Serikbayeva G.A., Toishimanov M.R., Bari G.T. Influence of selected rootstock on growth parameters, accumulation of IAA and vitamins in scions of *Cucumis sativus* L. and *Cucumis melo* L. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov J Genet Breed*. 2025;29(4):559-567. doi 10.18699/vjgb-25-59

Funding. This research was funded by the Committee of Science of the Ministry of Science and Higher Education of the Republic of Kazakhstan, grant number AP19679681.

Влияние отобранного подвоя на параметры роста, накопление ИУК и витаминов в привоях *Cucumis sativus* L. и *Cucumis melo* L.

А.Ж. Шойбекова [1], С.К. Джантасов [1], А.С. Джантасова [1], А.Т. Саматов [1], Т.С. Сагиндыков [1], А.Н. Каримова [1], Г.А. Серикбаева [1], М.Р. Тойшиманов [1, 2], Г.Т. Бари [1]

[1] Казахский национальный аграрный исследовательский университет, Алматы, Казахстан
[2] Институт биологии и биотехнологии растений, Алматы, Казахстан
baracuda.co@mail.ru

Аннотация. Прививка с помощью устойчивых подвоев – один из наиболее эффективных методов предотвращения болезней, передающихся через почву, который может влиять на вегетативный рост, цветение, периоды созревания и качество плодов, тем самым обеспечивая высокую урожайность. В настоящем исследовании четыре вида из семейства тыквенных были протестированы в качестве потенциальных кандидатов для прививки огурца и дыни: *Cucurbita ficifolia* Bouché, *Cucurbita moschata* L., *Cucurbita pepo* L. и *Cucurbita maxima* Duch. Исследование было сосредоточено на методах прививки, которые оптимизируют параметры роста и накопление гормонов и витаминов в привое. Согласно полученным результатам, отобранный вид *Cucurbita maxima* Duch. является наиболее подходящим материалом подвоя для прививки к *Cucumis sativus* L. и *Cucumis*

© Shoibekova A.Zh., Jantassov S.K., Jantassova A.S., Samatov A.T., Sagindykov T.S., Karimova A.N., Serikbayeva G.A., Toishimanov M.R., Bari G.T., 2025
This work is licensed under a Creative Commons Attribution 4.0 License

melo L., поскольку он показал наилучшую массу растения и корней. Среди двух протестированных методов прививки язычковый подход ('X') продемонстрировал лучшие результаты с точки зрения параметров роста, накопления индолил-3-уксусной кислоты (ИУК) и витаминов в листьях привоя. Концентрации ИУК и витаминов измеряли с помощью ВЭЖХ в образцах привоя в возрасте 2, 4 и 6 недель. В методе 'X' накопление ИУК с конца второй недели было в два раза выше по сравнению с контрольными растениями. Этот метод также показал более высокое содержание витаминов, с повышенным уровнем витаминов группы B и витамина C в конце четвертой недели (25.2–135.1 и 52.3–67.0 % соответственно), а витаминов A, E, D_3, K – начиная со второй недели (в 1.5–2 раза выше). Напротив, метод прививки вставкой или косым срезом ('Y') не показал значительного увеличения анализируемых параметров и был сопоставим с контролем. Метод прививки 'X' как для *Cucumis sativus* L., так и для *Cucumis melo* L. на растения *Cucurbita maxima* Duch. продемонстрировал наилучшие результаты и рекомендуется для производства.

Ключевые слова: морфометрический анализ; прививка; *C. maxima*; ВЭЖХ; ИУК; витамины

Introduction

Grafting cucumber and melon (as scions) onto pumpkin rootstocks is one of the most effective methods to prevent soilborne diseases, influence vegetative growth, flowering, maturation periods, and fruit quality, thereby ensuring high yields of these crops (Mauro et al., 2022). Grafting vegetable crops is an important cultivation method in many countries where intensive and continuous cultivation is practiced (Farhadi et al., 2016), particularly in greenhouses with controlled conditions, where rootstocks can extend the fruiting period.

In global practice, rootstocks such as fig leaf gourd *Cucurbita ficifolia* Bouché (*C. ficifolia*) (El-Eslamboly, Deabes, 2014), winter squash landrace *Cucurbita moschata* L. (*C. moschata*) (Traka-Mavrona et al., 2000; Noor et al., 2019; Li X. et al., 2023), pumpkin *Cucurbita maxima* Duch. (*C. maxima*) (Farhadi et al., 2016), summer squash landrace *Cucurbita pepo* L. (*C. pepo*) (Noor et al., 2019), and combinations of *C. maxima* × *C. moschata* (Bekhradi et al., 2009; Toporek, Keinath, 2020) are used.

Scion-rootstock combinations affect pH, taste, sugar content, color, carotenoid content, fruit texture, resistance to low soil temperatures and salinity, and nutrient and water uptake. Studies have shown that RNA, proteins, and small molecules, some of which are involved in signal transduction, can move from the rootstock to the scion, directly affecting scion physiology (Mauro et al., 2022). This practice is also applied to other vegetable crops (Tsaballa et al., 2021). Such functional interdependence includes a complex relationship between the two plants, involving the exchange of water, nutrients, hormones, and other metabolites (Albacete et al., 2015).

Auxins play a central role in root formation. They induce the initiation of root primordia and influence the growth of newly formed roots. Plants produce indole-3-acetic acid (IAA) in shoot tips and young leaves, but exogenous auxin is important for successful rooting. There is no direct evidence that synthetic auxins can replace natural ones in cells, but they help in the overall accumulation of IAA in the plant, thereby promoting the formation of adventitious roots (Stefancic et al., 2007). The percentage of rooted cuttings positively correlates with the concentration of exogenous auxin, but only up to a certain point – at high concentrations, rooting stops or even decreases. Therefore, the presence of endogenous auxin in the plant is important (Stefancic et al., 2006).

The quality of the initial material plays a very important role in the formation of adventitious roots. The optimal physiological state of the initial plants can significantly improve the rooting of cuttings.

It is especially important to consider the accumulation of one of the main growth hormones, IAA (Balliu, Sallaku, 2017; Bunsangiam et al., 2021; Tang et al., 2023), in this case in scion-rootstock combinations (Noda et al., 2000; Li W. et al., 2017; Lam et al., 2020; Bantis et al., 2021) of cucumber and melon with pumpkins. In addition to resistance to biotic and abiotic factors, grafted plants need a good scion-rootstock union, rapid growth, and high productivity in a shorter time.

Moreover, the accumulation of vitamins in plants as a biochemical indicator plays a significant role (Asensi-Fabado, Munné-Bosch, 2010; Abbas et al., 2023), particularly if these are water-soluble and fat-soluble vitamins with antioxidant properties (Asensi-Fabado, Munné-Bosch, 2010). Previously, the importance of their role in plants, various organs, and subcellular locations, as well as their main biosynthetic pathways, were described by the authors. In this context, it is necessary to study the influence of rootstock on scion in vitamin accumulation over post-grafting periods, as such studies have not been conducted previously according to the literature data. For optimal quantitative determination of IAA and vitamins, the high-performance liquid chromatography (HPLC) method is used (Battal, Tileklioğlu, 2001; Aslam at al., 2008; Keskin et al., 2022).

The aim of this study was to select the most suitable candidate from *C. ficifolia*, *C. moschata*, *C. pepo*, and *C. maxima* as rootstock for grafting of *C. sativus* and *C. melo* as scion, and to select the grafting method that ensures optimal growth parameters, measurement of IAA and vitamins in the scion.

Materials and methods

Plant material. The plant material used for the study consisted of the following Cucurbitaceae species: cucumber (*C. sativus*) cultivar Asylim, melon (*C. melo*) cultivar Valet, fig leaf gourd (*C. ficifolia*) cultivar Arbuzny, winter squash landrace (*C. moschata*) cultivar Aphrodite, summer squash landrace (*C. pepo*) cultivar Danaya, and large-fruited pumpkin (*C. maxima*) cultivar Karina, from both Kazakhstan and global selections. The cultivation of cucurbits was conducted on neutralized peat with a pH of 6.0 (Kekkila™) in 1-liter containers with expanded clay drainage. Seeds of the cucurbits were planted in the peat-filled containers and watered daily with a nutrient solution of mineral salts at a rate of 100 mL per plant. After seed germination, the plants were illuminated with LED lamps at 5,000 lux for one week, followed by 10,000 lux for the subsequent six weeks. Morphometric analysis measures included plant mass and root mass separately, number and area of leaves, and stem thickness. Table 1

presents the composition of mineral salts and trace elements (According to the nutrient system of General Hydroponics, https://generalhydroponics.com).

Every two weeks, the ppm values of the solution were increased by 500. To achieve concentrations of 500, 1,000, and 1,500 ppm in the nutrient solutions, 1.3 mL, 2.7 mL and 4 mL, respectively, were taken from each stock solution per liter (Table 1). The pH was adjusted to 6.0 using a 1M solution of NaOH or KOH. The total concentration of the nutrient solution was measured using a TDS meter.

HPLC analysis of IAA, fat- and water-soluble vitamins determination. The chromatographic separation was performed using a Shimadzu Prominence LC-20 system (Shimadzu, Japan) equipped with a UV detector (SPD-20A) and a fluorescent detector (RF-10AXL). The HPLC system was equipped with a binary pump (LC-20AD), an autosampler (SIL-20AC), a degasser (DGU-20A5) and a column oven (CTO-20A) controlled by LCSolution. For IAA, fat- and water-soluble vitamins determination were used the fresh plants.

Fat-soluble vitamins analysis. Stock solutions of vitamins A, D_3, E, K 10 mg (Sigma Aldrich, USA) were dissolved in 10 mL of methanol in each falcon tube. Next, the working calibration standard was prepared with seven concentration ranges of 0.48–250 µg/mL for D_3, 1.95–1,000 µg/mL for vitamin E, 0.195–100 µg/mL for vitamin K and 0.39–200 µg/mL for vitamin A. All standards were stored at –20 °C and protected from light. The concentration of each vitamin was selected based on the sensitivity of the detector.

All working calibration standard solutions and samples were analyzed using the column Shimpack ODS XR (75 mm × 3.0 mm × 2.2 µm) (Shimadzu, Japan) and the following HPLC conditions: column oven temperature 35 °C; eluent Acetonitrile/methanol/dichloromethane/H_2O (70:15:10:5 %), and flow rate programmed using the following conditions: 0–10 min of 0.5 mL/min, then for 11–19 min the flow was increased to 1.0 mL/min, at 20 min it was decreased to 0.5 mL/min. Working calibration standards and samples were determined at a UV detector at 280 nm for 0–14 min, then the UV wavelength was switched to 295 nm. All these chromatographic parameters allow to better separate mixed standard calibrations.

In 1 g of the mixed sample, 100 µg ascorbic acid, 10 mL ethanol, 3 mL KOH (50 %) were added, stirred, and refluxed for 50 min using water bath at 80 °C. Extracts were neutralized with distilled water twice and then dehydrated using anhydrous sodium sulfate. Extracts were concentrated using rotary evaporator at 50 °C (IKA HB-8 basic, IKA, Germany), diluted by 5 mL acetonitrile, filtered (Aslam at al., 2008) using 0.45 µm membrane (Chromafil AO-45/25, Macherey-Nagel, Germany) and finally analyzed using HPLC.

Water-soluble vitamins analysis. The mobile phase consisted of 100 % acetonitrile and 99 % deionized water with 0.1 % orthophosphoric acid and 25 mM sodium dihydrogen phosphate. Flow rate was isocratic – 0.5 mL/min. The separation of vitamins was carried out in a Supelco Ascentis C18 column (250 mm – 4.6 mm – 5 µm) at 35 °C. Preparation of stock standard samples of B vitamins (B_1, B_2, B_3, B_5, B_6, B_7, B_9, B_{12}) was dissolved in deionized water at a concentration of 1 mg/mL. All water-soluble vitamins were purchased from Titan Biotech Ltd. All vitamins were pure and pharma-grade (purity at least ≥ 99 %). The solubility of vitamins B_2 and B_9 in water is limited, so a separate aqueous solution was prepared in 5 mM KOH and 20 mM $KHCO_3$, respectively. Working standard samples will be prepared for B_1, B_5, B_7 at a concentration range of 50–200 µg/mL, B_3, B_6 at 25–100 µg/mL, B_{12} at 12.5–50 µg/mL, B_2 and B_9 at 2.5–10 µg/mL, and then all will be combined into one single standard for further calibration (Aslam et al., 2008).

Table 1. Minimum allowable composition of nutrient elements in stock solution, %

Grow (vegation)	Bloom (flowering)	Micro (microelements)
Total N – 3	Total P_2O_5 – 5	Total CaO – 5
Total P_2O_5 – 1	Total K_2O – 4	Total N – 5
Total K_2O – 6	Total MgO – 3	Total K_2O – 1
Total MgO – 0.8	Total SO_4 – 5	Boron (B) – 0.01
		Molybdenum (Mo) – 0.0008
		Cobalt (Co) – 0.0005
		Cu chelate EDTA – 0.01
		Zn chelate EDTA – 0.015
		Mn chelate EDTA – 0.05
		Fe chelate EDTA – 0.10

For preparing an extraction solution, 50 mL of acetonitrile was mixed with 10 mL of acetic acid, and the final volume was made up to 1,000 mL with deionized water. 1 g samples were weighed and homogenized. After that, the samples were transferred into a conical flask where 10 mL of extraction solution was added. A water bath was set at 70 °C for 30 min. Afterwards, the sample was cooled down and finally filtered with filter trips (0.45 µm) and 20 µL aliquots solution was injected into the HPLC (Mozumder et al., 2019).

IAA analysis. Samples and standards were separated on a Restek Ultra C18 HPLC column 150 mm × 4 mm, 5 µm (Bellefonte, PA, USA) at 40 °C. UV detection was performed at 269 nm. The flow rate of the mobile phase was 0.8 mL/min. Mobile phase A consisted of 100 % HPLC grade acetonitrile, mobile phase B consisted of 99.9 % HPLC grade water and 0.1 % formic acid using the gradient elution as follows: 95 % B, 0 min; 70 % B, 13 min; 95 % B, 15 min. The flow rate of the mobile phase was 0.8 mL/min (Battal, Tilekliоğlu, 2001; Keskin et al., 2022). 10 mg of IAA standard solution was dissolved in 1 mL 1N NaOH, then filled with 9 ml deionized water in a 10 mL tube.

1 g samples were weighed and homogenized. After that, the samples were transferred into 10-mL centrifuge tubes and 10 mL of acetonitrile was added: deionized water (9:1, v/v) under dim light conditions. Then, 100 µL formic acid was added and shaken. Homogenates were incubated for 2 h, and centrifuged at 12,000g at 4 °C for 20 min. The upper supernatants were then collected and dried in a vacuum evaporator (Biobase, China). The residues obtained by drying were dissolved in 2 mL of acetonitrile and purified by a 200 mg/3 mL C18 solid phase extraction cartridge (Strata, Phenomenex, Torrance, USA). The cartridge was prepared by successively passing 2 mL of water and 2 mL of ethanol using SPE tube

vacuum manifold (Biobase, China), where the vacuum valve was set at negative pressure – 0.01 MPa. The liquid that passed through the cartridge was discarded, and the IAA was washed off with 2 mL of a mixture of ethanol–water–formic acid (80:20:0.5 %; v/v) into a 10 mL centrifugal tube. The collected residuals were transferred to a 2 mL Eppendorf tube, evaporated using a sample concentrator (NDK200-2N, Miulab, China), then dissolved with 2 mL acetonitrile. Finally, 20 µL aliquots solution was injected to HPLC.

Grafting methods. Two of the most common grafting methods were used for plant grafting. The first method, termed 'Y', involved insertion grafting or slant-cut grafting. One cotyledon leaf was left on the rootstock to enhance grafting success (approximately to 90 %). All leaves, including cotyledon leaves, were left on the scion. The graft junction was secured with a clip.

The second method, termed 'X', involved tongue approach grafting. Longitudinal cuts at an angle of 20–30° were made on the stems of both the rootstock and the scion, with the cut on the rootstock directed downward and the cut on the scion directed upward. The plants were then joined by their tongues and secured with clips. Initially, both root systems were used. After 10 days, the vegetative part of the rootstock and the root system of the scion were pruned.

The grafted plants were then placed in a climate chamber with controlled temperature, humidity, and light conditions for further grafting. The grafted plants were grown in the growth chamber at 20 °C with 92 % humidity in complete darkness for four days. Subsequently, the grafted plants were grown under 12,000 lux lighting with an 8/16-hours light/dark cycle for 10 days (Lee et al., 2010). *C. sativus* and *C. melo* were used as control and self-grafted. All data were statistically assessed using Duncan's test.

Results

Morphometric analysis of the Cucurbitaceae family

For setting up the experiment on growth parameters (plant and root mass), seeds of rootstocks *C. maxima*, *C. pepo*, *C. ficifolia*, and *C. moschata*, and scions *C. sativus* and *C. melo* were sown for further morphometric analysis and extraction (for quantitative determination of auxins and vitamins) over a pe-

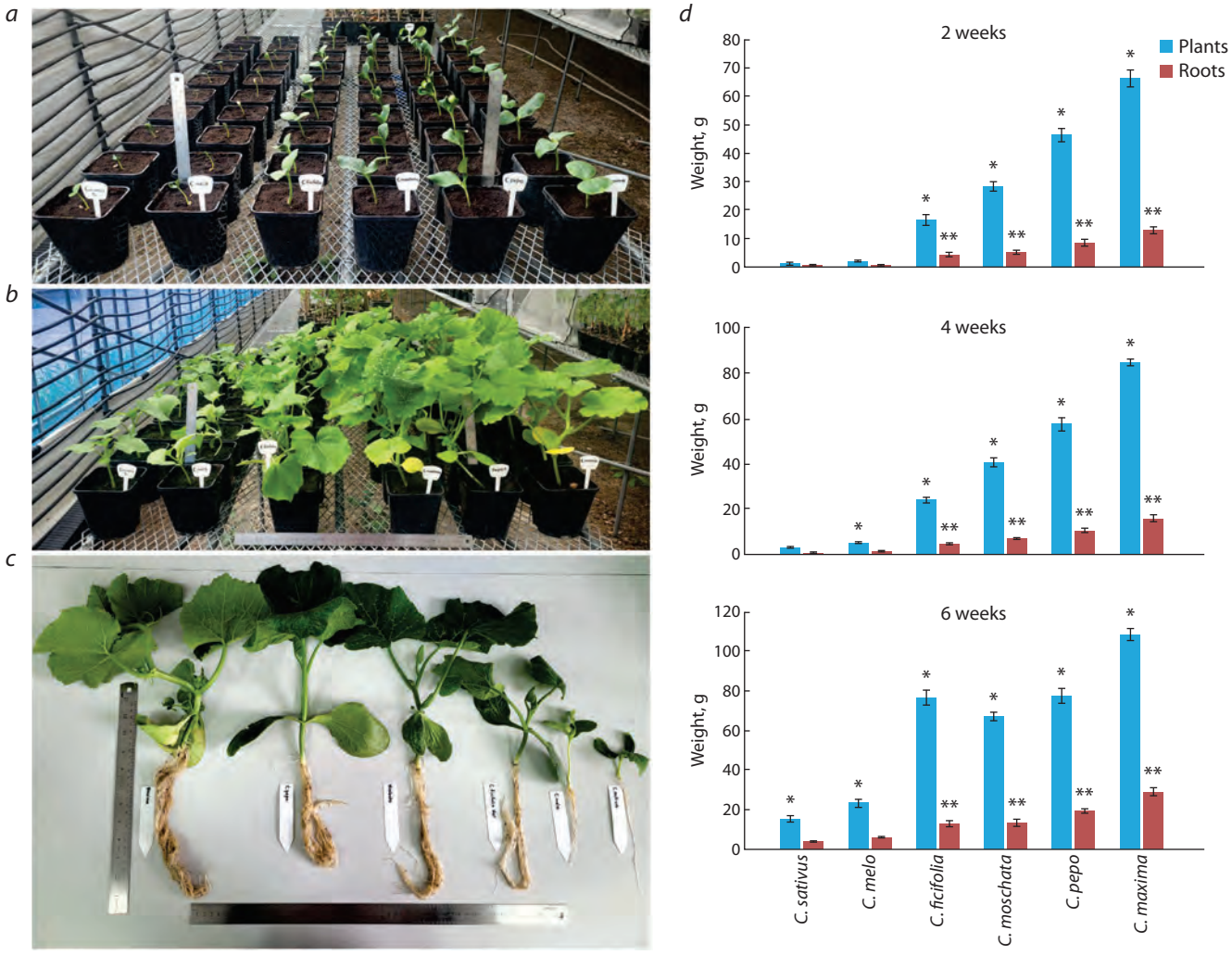


Fig. 1. Growth of Cucurbitaceae plants: *a*, one-week seedlings after germination: *b*, four-week plants; *c*, morphometric analysis of the Cucurbitaceae plants; *d*, 2-, 4- and 6-week plants and roots' weight in gram of fresh plants.

* Statistically significant results between the Cucurbitaceae plants at $p \leq 0.05$. ** Statistically significant results between the Cucurbitaceae roots at $p \leq 0.05$.

riod of 2 to 6 weeks (Fig. 1*a*). For reliable results, experiments were conducted in three replicates under identical conditions: seedlings were planted in 1,000 mL containers and irrigated with a nutrient solution at a concentration of 500 ppm. During the third and fourth weeks, the concentration of the nutrient solution was increased to 1,000 ppm, and from the fifth week to the end of the sixth week, it was increased to 1,500 ppm. The results showed a significant difference across the four types of rootstocks and two samples of scions.

As an example, Figure 1*b* shows the plants at four weeks post-germination and the morphometric analysis (Fig. 1*c*). The morphometric analysis focused on the mass of the entire plant and the root. The results of the morphometric analysis are presented in Figure 1*d* as bar charts.

At 2, 4, and 6 weeks post-germination, plant and root mass parameters were measured using digital scales adapted from the methods. Based on the plant and root mass measurements at all intervals (2, 4, and 6 weeks), *C. maxima* exhibited the highest values. *C. pepo* ranked second in these metrics. However, at the 6-week mark, *C. ficifolia* matched *C. pepo* in plant mass, though all its values remained significantly lower than those of *C. maxima*. The lowest values were observed in *C. sativus* and *C. melo*. Morphometric analysis identified *C. maxima* as the most suitable candidate for grafting of *C. sativus* and *C. melo*. *C. sativus* plant did not show statistical significance, except for the 6th week. *C. melo*, *C. ficifolia*, *C. moschata*, *C. pepo*, and *C. maxima* plants showed significant differences, with *C. maxima* having the highest plant biomass over time. *C. sativus* and *C. melo* roots weight had no statistical significance at all. *C. ficifolia*, *C. moschata*, *C. pepo*, and *C. maxima* root weights showed significant differences.

IAA, water- and fat-soluble vitamins in grafted plants

In the second stage, grafting of one-week-old plants of *C. sativus* and *C. melo* onto the pumpkin rootstock *C. maxima* was performed. The grafting method is illustrated in Figure 2*a*. The tongue approach grafting (Fig. 2*a*, first from the left) is designated as 'X' based on the shape of stem fusion. The second grafting method (second from the left) involved hole insertion grafting (scion stem with an oblique cut) to the growth point of the rootstock and is conditionally designated as 'Y'. For the 'Y' grafting method, the true leaf buds of the rootstock were removed, and a cut was made along the stem where the scion with an oblique stem cut was attached. The grafting methods 'X' and 'Y' are marked with red arrows in Figure 2*a*. The third grafting method was the control, where a tongue cut was made on the *C. sativus* and *C. melo* plants to ensure the experiment was conducted under the same conditions. The self-grafted scion plants of *C. sativus* and *C. melo* are designated as Control.

Two weeks after grafting, active plant growth commenced. Differences in growth rate and development between the grafted plant variants were observed at the end of the third week. Figure 2*b* illustrates the results for plants at four weeks

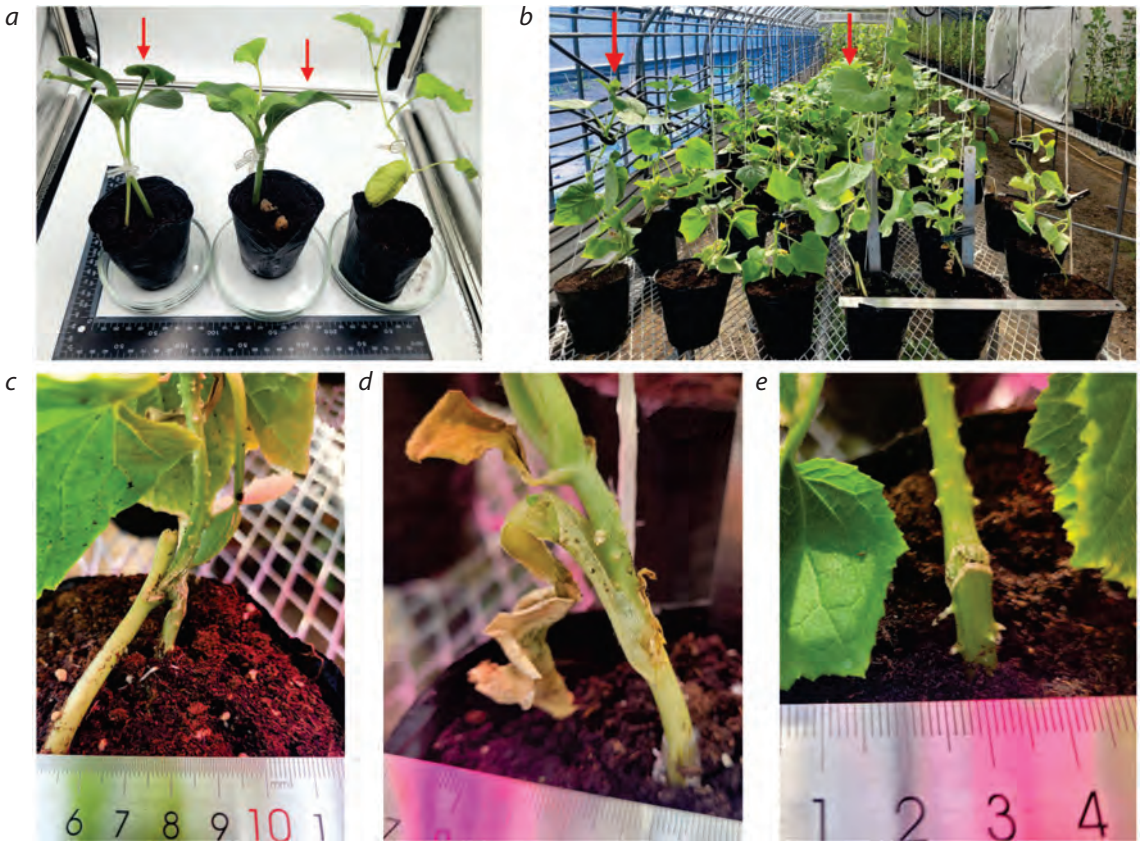


Fig. 2. Grafting of *C. sativus* and *C. melo* on *C. maxima*.

a, From left to right X-type grafting, Y-type grafting and self-grafting accordingly; *b*, grafted and self-grafted Control plants; *c*, X-type grafted connection between scion and rootstock; *d*, Y-type grafted connection between scion and rootstock; *e*, self-grafted Control.

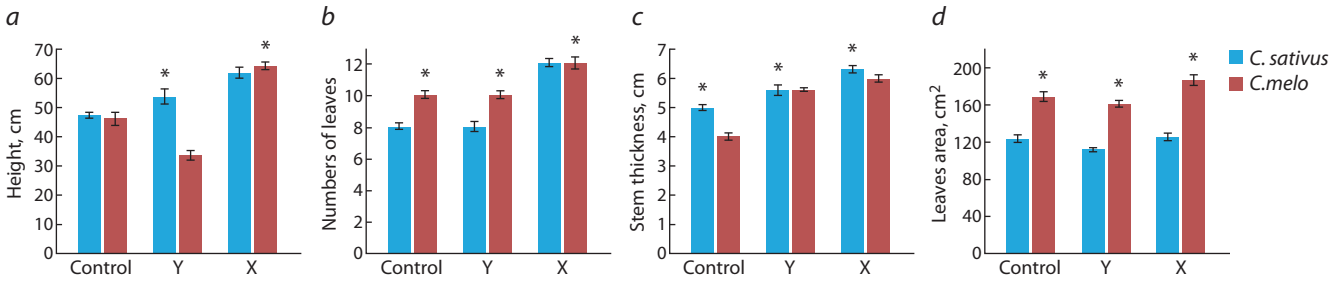


Fig. 3. Morphometric analysis of 6-week-old grafted plants (fresh plants): *a*, plant height; *b*, number of every grafted plant leaf; *c*, stem thickness; *d*, leaves area of grafted plants.

* Statistically significant results between the control and experimental groups at $p \leq 0.05$.

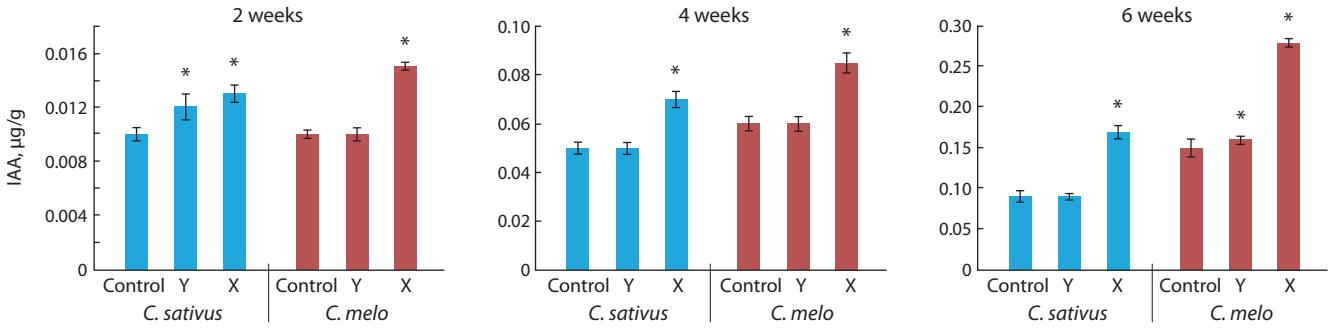


Fig. 4. IAA content in *C. sativus* and *C. melo* depending on grafting methods (from fresh plants).

* Statistically significant results between the control and experimental groups at $p \leq 0.05$.

of age marked with red arrows. Additionally, the healed graft unions are shown in Figure 2*c–e* for graft variants 'X', 'Y', and Control, respectively. From the start of grafting until the end of the sixth week post-grafting, it was visually observed that the graft variant 'X' did not lag behind the control plants and the 'Y' variant. In the graft variant 'Y', a slowdown in growth and development was noted. Upon reaching six weeks of age, a structural analysis was conducted, which identified the graft variant 'X' as the best among the others (Fig. 3). Six weeks post-grafting, the 'X' variant exhibited a 96 % survival rate, while the 'Y' variant showed only a 43 % survival rate of grafted plants. The control variant also demonstrated a high survival rate of 97 %.

Upon reaching six weeks of age, a structural analysis of the grafted plants was conducted, revealing that the grafting variant 'X' was superior to the others (Fig. 3). This variant excelled in plant height, leaf number, stem thickness, and leaf area for both *C. sativus* and *C. melo*. Grafting variant 'Y' ranked second in stem thickness and matched the Control in leaf number and leaf area. When comparing *C. sativus* and *C. melo*, no significant differences were found within the 'X' variant of grafting, for plant length, stem thickness and number of leaves, except for leaves area according to Figure 3. The IAA content in the grafted plant variants is presented in Figure 4. The extracts of leaves and stems were filtered through solid phase extraction (SPE) (Supplementary Fig. S1)[1]. Upon completion of the extraction, IAA detection was performed, which was expressed as a chromatographic peak (Fig. S1). Also, chromatographic peaks of water and fat-soluble vitamins quantification of the Cucurbitaceae family are shown in Supplementary Figure S2. And peaks of pure standard substances of IAA, water/fat-soluble vitamins are provided in Supplementary Figure S3.

Based on the results of the HPLC analysis, it was found that the content of IAA and vitamins is higher in the leaves than in the stems. Consequently, we conducted further targeted biochemical analyses using the leaves. As a result, the highest accumulation of IAA was observed in the graft variant 'X'. As shown in Figure 4, the accumulation of IAA in grafted plants increased in the 'X' variant from the end of the second week, and by the end of the sixth week, the difference with the control plants was almost doubled.

Vitamins identification using chromatographic detection was performed (Tables 2, 3). The quantitative values of vitamins are given in µg/g, recalculated to dry weights of stem and leaf samples. The differences in the content of water- and fat-soluble vitamins in the grafted plant variants are presented in Tables 2, 3. A significant increase in vitamin content was observed from the end of the fourth week.

An increased content of all B vitamins and vitamin C was noted in the grafting variant 'X' at the end of the 4th week – 25.2–135.1 and 52.3–67.0 %, respectively, with all indicators being higher in the grafted *C. sativus* than in the *C. melo*. In *C. melo*, vitamin B_2 levels exceeded the Control by 122.7 %, and those of vitamin B_{12}, by 135.1 %. In the grafting variant 'Y', all indicators were at the control level.

After 6 weeks, the indicators in the grafting variant 'X' became more balanced and amounted to 17.5–61.8 % for B vi-

[1] Supplementary Figures S1–S3 are available at:
https://vavilovj-icg.ru/download/pict-2025-29/appx20.pdf

Table 2. Water-soluble vitamins accumulation in grafted plants (from fresh plants), µg/g

Vitamins	*C. sativus*			*C. melo*			*C. sativus*			*C. melo*			*C. sativus*			*C. melo*			*p*-value		
	2 weeks						4 weeks						6 weeks						by weeks	by grafting	by plants
	C	Y	X	C	Y	X	C	Y	X	C	Y	X	C	Y	X	C	Y	X			
B_1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0	0	0
B_2	2.2	1.9	2.3	2.3	2.2	2.4	13.2	13.8	29.4	30.6	28.9	54.0	33.4	34.5	50.7	81.2	80.0	110.2	0.0055	NS	NS
B_3	16.9	15.8	17.2	21.9	22.0	23.0	65.1	66.4	90.2	63.2	60.7	79.1	96.2	98.1	140.3	100.1	101.2	133.4	0.0001	NS	NS
B_5	43.1	44.0	44.5	37.0	37.2	37.7	102.0	105.5	190.6	102.2	92.5	145.8	179.0	171.5	210.3	204.6	206.7	291.8	0.0004	0.0450	NS
B_6	27.0	27.5	27.9	3.7	3.6	3.9	70.3	71.4	128.5	10.1	9.9	18.2	88.7	90.6	134.4	59.0	58.6	81.4	0.0021	0.0384	0.0149
B_7	71.0	72.1	72.9	109.7	109.8	111.4	266.9	270.0	438.1	185.8	180.4	260.6	525.6	530.0	790.3	239.4	240.5	320.5	0.0409	NS	NS
B_9	8.1	8.1	8.3	15.3	15.4	16.2	20.5	20.3	39.9	65.2	66.5	110.3	44.0	42.5	71.2	101.1	100.9	120.6	0.0185	NS	0.0288
B_{12}	7.6	7.5	7.9	3.7	3.8	3.9	11.1	12.3	26.1	5.0	4.9	8.2	42.0	40.5	63.7	13.9	14.1	19.4	0.0104	NS	0.0394
C	17.2	17.4	17.9	11.3	11.1	12.0	25.6	25.3	39.0	18.8	18.0	31.4	60.1	58.9	91.1	34.1	35.2	49.4	0.0019	0.0336	NS

Note. Here and in the Table 3: C – control, Y – 'Y' type of grafting, X – 'X' type of grafting.

Table 3. Fat-soluble vitamins accumulation in grafted plants (from fresh plants), µg/g

Vitamins	*C. sativus*			*C. melo*			*C. sativus*			*C. melo*			*C. sativus*			*C. melo*			*p*-value		
	2 weeks						4 weeks						6 weeks						by week	by grafting	by plants
	C	Y	X	C	Y	X	C	Y	X	C	Y	X	C	Y	X	C	Y	X			
A	2.112	1.785	3.055	1.984	1.102	2.164	5.136	6.492	8.175	3.512	4.250	5.019	6.103	6.171	10.295	6.053	6.112	9.069	0.0014	NS	NS
E	0.609	0.585	0.918	0.175	0.116	0.253	1.281	1.597	2.141	0.285	0.352	1.234	3.014	2.997	5.413	0.091	0.102	1.714	0.0272	NS	0.0116
D_3	0.001	0.001	0.004	0.001	0.001	0.003	0.004	0.019	0.037	0.005	0.006	0.015	0.017	0.016	0.031	0.014	0.014	0.026	0.0212	0.0434	NS
K	0.007	0.006	0.008	0.012	0.011	0.017	0.041	0.054	0.069	0.011	0.015	0.028	0.045	0.044	0.061	0.051	0.052	0.095	0.0192	NS	NS

tamins and 51.6 % for vitamin C in *C. sativus*; 19.3–42.6 % for B vitamins and 44.9 % for vitamin C in *C. melo*. In the grafting variant 'Y', the indicators were also at the standard Control level.

For vitamins A, E, D_3, and K, starting from the 2nd week, a stable difference in their accumulation in leaves was observed in the grafting variant 'X', being 1.5–2 times higher compared to the Control variant.

Additionally, in the 'X' variant, the accumulation of IAA compared to the control plants in Control also increased from the end of the second week and was almost twice as high. This pattern with the grafting variant 'X' was observed both in the *C. sativus* grafting variant and in the melon grafting variant.

Discussion

The influence of the root system of the rootstock on the scion, and vice versa, i. e., the impact of the more developed biomass of the rootstock on the further development of the scion and improved growth of melon plants grafted onto Cucurbitaceae species, was identified in a previous study (Martínez-Ballesta et al., 2010). This study confirms that the union of the rootstock and scion and the differentiation of new vascular tissue from callus cells, as well as the resumption of scion biomass growth, begin within 2 weeks post-grafting, as clearly demonstrated in our research.

Based on the obtained results, we observe that the grafting method 'X' is the most acceptable among other methods. As expected, this method allows plants to resume growth processes more quickly and better accumulate IAA and vitamins in the leaves. According to (Noor et al., 2019), the use of the tongue grafting method showed high compatibility with hybrid cucumber scions compared to other grafting methods and non-grafted plants.

The novel idea of the study was to identify the optimal type of grafting that will result in the fastest recovery of the grafted plant, as well as a stimulating effect of the rootstock on the scion, where an increase in IAA and vitamin content occurs. Other researchers have noted that the influence of cucurbit rootstock on cucumber scion provides salt tolerance and increases fruit yield by improving morpho-physio-biochemical and ionic properties, specifically increasing the content of the following substances in grafted plants: superoxide dismutase, catalase and peroxidase enzymes, antioxidant scavenging activity, ionic ↑K and Ca, ↓Na (Abbas et al., 2023). We obtained similar results. It can be assumed that the mechanism of vitamin accumulation in scions occurs similarly to the mechanism of IAA accumulation in scions post-grafting onto the rootstock. In other words, the growth factors of the rootstock in the form of IAA may stimulate the accumulation of hormones and vitamins in the scion. The conducted studies, including morphometric analysis of grafted plants, show that parameters such as plant height, number of leaves on grafted plants, stem thickness, leaf area of grafted plants, as well as chromatographic data on IAA and vitamin accumulation, are superior in grafting method 'X' compared to the control and grafting method 'Y'.

The next stage should include studies throughout the entire physiological development of the plant, up to full maturation and harvest.

Conclusion

Analyzing the data, we concluded that among four species of Cucurbitaceae: *C. ficifolia*, *C. moschata*, *C. pepo* and *C. maxima*, the plants of the species *C. maxima* are the best candidate for use as rootstock material for grafting of *C. sativus* and *C. melo*. This is due to their superior performance in terms of plant mass increase, root mass, and stem thickness at the root base for both *C. sativus* and *C. melo*. The second-best candidates are plants of the species *C. pepo*.

Among the two different grafting methods tested, the grafting method 'X' showed the best results in terms of growth factors and the accumulation of IAA and vitamins in the leaves of the rootstock. In method 'X', the IAA accumulation from the end of the second week was twice as high compared to the Control plants. Regarding vitamins, this method also exceeded the control, with increased levels of all B vitamins and vitamin C at the end of the fourth week by 25.2–135.1 and 52.3–67.0 %, respectively, and vitamins A, E, D_3, and K starting from the second week by 1.5–2 times. In contrast, the grafting method 'Y' did not show any significant increase in any of the analyzed parameters and was at the Control level.

Therefore, it is recommended to graft both *C. sativus* and *C. melo* onto *C. maxima* plants using the tongue approach grafting method 'X'.

References

Abbas F., Faried H.N., Akhtar G., Ullah S., Javed T., Shehzad M.A., Ziaf K., Razzaq K., Amin M., Wattoo F.M., Hafeez A., Rahimi M., Abeed A.H.A. Cucumber grafting on indigenous cucurbit landraces confers salt tolerance and improves fruit yield by enhancing morpho-physio-biochemical and ionic attributes. *Sci Rep.* 2023;13(1): 21697. doi 10.1038/s41598-023-48947-z

Albacete A., Martínez-Andújar C., Martínez-Pérez A., Thompson A.J., Dodd I.C., Pérez-Alfocea F. Unravelling rootstock×scion interactions to improve food security. *J Exp Bot.* 2015;66(8):2211-2226. doi 10.1093/jxb/erv027

Asensi-Fabado M.A., Munné-Bosch S. Vitamins in plants: occurrence, biosynthesis and antioxidant function. *Trends Plant Sci.* 2010; 15(10):582-592. doi 10.1016/j.tplants.2010.07.003

Aslam J., Mohajir M.S., Khan S.A., Khan A.Q. HPLC analysis of water-soluble vitamins (B_1, B_2, B_3, B_5, B_6) in *in vitro* and *ex vitro* germinated chickpea (*Cicer arietinum* L.). *Afr J Biotechnol.* 2008; 7(14):2310-2314. doi 10.5897/AJB2008.000-5058

Balliu A., Sallaku G. Exogenous auxin improves root morphology and restores growth of grafted cucumber seedlings. *Hortic Sci.* 2017; 44(2):82-90. doi 10.17221/53/2016-HORTSCI

Bantis F., Panteris E., Dangitsis C., Carrera E., Koukounaras A. Blue light promotes vascular reconnection, while red light boosts the physiological response and quality of grafted watermelon seedlings. *Sci Rep.* 2021;11(1):21754. doi 10.1038/s41598-021-01158-w

Battal P., Tileklioğlu B. The effects of different mineral nutrients on the levels of cytokinins in maize (*Zea mays* L.). *Turk J Bot.* 2001; 25(3):123-130

Bekhradi F., Kashi A.K., Delshad M. Effect of different cucurbits rootstocks on vegetative and yield of watermelon. *Acta Hortic.* 2009; 807:649-654. doi 10.17660/ActaHortic.2009.807.97

Bunsangiam S., Thongpae N., Limtong S., Nantana S. Large scale production of indole-3-acetic acid and evaluation of the inhibitory effect of indole-3-acetic acid on weed growth. *Sci Rep.* 2021;11(1):13094. doi 10.1038/s41598-021-92305-w

El-Eslamboly A.A.S.A., Deabes A.A.A. Grafting cucumber onto some rootstocks for controlling root-knot nematodes. *Minufiya J Agric Res.* 2014;39(3):1109-1129

Farhadi A., Aroeii H., Nemati H., Salehi R., Giuffrida F. The effectiveness of different rootstocks for improving yield and growth of cucumber cultivated hydroponically in a greenhouse. *Horticulturae.* 2016;2(1):1. doi 10.3390/horticulturae2010001

Keskin N., Kaya O., Ates F., Turan M., Gutiérrez-Gamboa G. Drying grapes after the application of different dipping solutions: effects on hormones, minerals, vitamins, and antioxidant enzymes in Gök Üzüm (*Vitis vinifera* L.) raisins. *Plants.* 2022;11(4):529. doi 10.3390/plants11040529

Lam V.P., Lee M.H., Park J.S. Optimization of indole-3-acetic acid concentration in a nutrient solution for increasing bioactive compound accumulation and production of *Agastache rugosa* in a plant factory. *Agriculture.* 2020;10(8):343. doi 10.3390/agriculture10080343

Lee J.M., Kubotab C., Tsaoc S.J., Bied Z., Echevarriae P.H., Morraf L., Oda M. Current status of vegetable grafting: diffusion, grafting techniques, automation. *Sci Hortic.* 2010;127(2):93-105. doi 10.1016/j.scienta.2010.08.003

Li W., Fang C., Krishnan S., Chen J., Yu H., Murphy A.S., Merewitz E., Katin-Grazzini L., McAvoy R.J., Deng Z., Zale J., Li Y. Elevated auxin and reduced cytokinin contents in rootstocks improve their performance and grafting success. *Plant Biotechnol J.* 2017;15(12):1556-1565. doi 10.1111/pbi.12738

Li X., Sun Y., Yuan X., Ma Z., Hong Y., Chen S. Impact of *Cucurbita moschata* resistant rootstocks on *Cucumis sativus* fruit and *Meloidogyne incognita* development. *Plant Dis.* 2023;107(12):3851-3857. doi 10.1094/PDIS-02-22-0319-RE

Martínez-Ballesta M.C., Alcaraz-Lypez C., Muries B., Mota-Cadenas C., Carvajal M. Physiological aspects of rootstock-scion interactions. *Sci Hortic.* 2010;127(2):112-118. doi 10.1016/j.scienta.2010.08.002

Mauro R.P., Pérez-Alfocea F., Cookson S.J., Ollat N., Vitale A. Physiological and molecular aspects of plant rootstock-scion interactions. *Front Plant Sci.* 2022;13:852518. doi 10.3389/fpls.2022.852518

Mozumder N.H.M.R., Akhter M.J., Anwara A.K., Rokibuzzaman M., Akhtaruzzaman M. Estimation of water-soluble vitamin B-complex in selected leafy and non-leafy vegetables by HPLC method. *Orient J Chem.* 2019;35(4):1344-1351. doi 10.13005/ojc/350414

Noda K., Okuda H., Iwagaki I. Indole acetic acid and abscisic acid levels in new shoots and fibrous roots of citrus scion-rootstock combinations. *Sci Hortic.* 2000;84(3):245-254. doi 10.1016/S0304-4238(99)00080-1

Noor R.S., Wang Z., Umair M., Yaseen M., Ameen M., Rehman S.U., Khan M.U., Imran M., Ahmed W., Sun Y. Interactive effects of grafting techniques and scion-rootstocks combinations on vegetative growth, yield and quality of cucumber (*Cucumis sativus* L.). *Agronomy.* 2019;9(6):288. doi 10.3390/agronomy9060288

Stefancic M., Stampar F., Osterc G. Influence of endogenous IAA levels and exogenous IBA on rooting and quality of leafy cuttings of *Prunus* ‘GiSelA 5’. *J Hortic Sci Biotechnol.* 2006;81(3):508-512. doi 10.1080/14620316.2006.11512095

Stefancic M., Stampar F., Veberic R., Osterc G. The levels of IAA, IAAsp and some phenolics in cherry rootstock ‘GiSelA 5’ leafy cuttings pretreated with IAA and IBA. *Sci Hortic.* 2007;112(4):399-405. doi 10.1016/j.scienta.2007.01.004

Tang J., Li Y., Zhang L., Mu J., Jiang Y., Fu H., Zhang Y., Cui H., Yu X., Ye Z. Biosynthetic pathways and functions of indole-3-acetic acid in microorganisms. *Microorganisms.* 2023;11(8):2077. doi 10.3390/microorganisms11082077

Toporek S.M., Keinath A.P. Evaluating cucurbit rootstocks to prevent disease caused by *Pythium aphanidermatum* and *P. myriotylum* on watermelon. *Plant Dis.* 2020;104(11):3019-3025. doi 10.1094/PDIS-03-20-0474-RE

Traka-Mavrona E., Koutsika-Sotiriou M., Pritsa T. Response of squash (*Cucurbita* spp.) as rootstock for melon (*Cucumis melo* L.). *Sci Hortic.* 2000;83(3-4):353-362. doi 10.1016/S0304-4238(99)00088-6

Tsaballa A., Xanthopoulou A., Madesis P., Tsaftaris A., Nianiou-Obeidat I. Vegetable grafting from a molecular point of view: the involvement of epigenetics in rootstock-scion interactions. *Front Plant Sci.* 2021;11:621999. doi 10.3389/fpls.2020.621999

Conflict of interest. The authors declare no conflict of interest.
Received November 4, 2024. Revised January 17, 2025. Accepted March 25, 2025.

POPULATION GENETICS
Original article

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2025;29(4):568-577

doi 10.18699/vjgb-25-60

Mitochondrial DNA data allow distinguishing the subpopulations in the widespread Demoiselle crane (*Anthropoides virgo*)

E.A. Mudrik [1] ✉, E.I. Ilyashenko [1, 2], P.A. Kazimirov [1], K.D. Kondrakova [1, 2], T.P. Archimaeva [3], L.D. Bazarov [4], O.A. Goroshko [5, 6], Ts.Z. Dorzhiev [7, 8], A.N. Kuksin [3], K.A. Postelnykh [9], V.V. Shurkina [10], V.Yu. Ilyashenko [2], A.V. Shatokhina [1], D.V. Politov [1]

[1] Vavilov Institute of General Genetics of the Russian Academy of Sciences, Moscow, Russia
[2] Severtsov Institute of Ecology and Evolution of the Russian Academy of Sciences, Moscow, Russia
[3] Tuvinian Institute for Exploration of Natural Resources of the Siberian Branch of the Russian Academy of Sciences, Kyzyl, Russia
[4] Tunkinsky National Park, Kyren, Russia
[5] Institute of Nature Resources, Ecology and Cryology of the Siberian Branch of the Russian Academy of Sciences, Chita, Russia
[6] Daursky State Nature Biosphere Reserve, Nizhny Tsasuchey, Russia
[7] Buryat State University named after Dorji Banzarov, Ulan-Ude, Russia
[8] Institute of General and Experimental Biology of the Siberian Branch of the Russian Academy of Sciences, Ulan-Ude, Russia
[9] Oka State Nature Biosphere Reserve, Brykin Bor, Russia
[10] Khakassky State Nature Reserve, Abakan, Russia
✉ mudrik@vigg.ru

Abstract. The polymorphism of the mtDNA cytochrome *b* (cyt *b*) gene's partial sequences has been studied in the Demoiselle crane (*Anthropoides virgo* Linnaeus, 1778) for the first time. Based on cyt *b* variability, the population genetic structure of the species was characterized within most of its range in Russia. Among 157 individuals we identified 18 haplotypes, nine of which were unique. In the European samples, we observed greater haplotype and nucleotide diversity and stronger genetic differentiation than in the Asian ones. Gene flow between different parts of the Demoiselle crane range is probably mediated by birds breeding in the Trans-Urals. The overall genetic subdivision of the species as estimated by F_{ST} was 0.265 ($p < 0.001$). The structure of the gene pool is formed by three main haplotypes, one of which predominates in the Azov-Black Sea region, the second in the Caspian and Volga-Ural regions, and the third is most common in the Asian samples. Based on the correspondence of intraspecific genetic differentiation of the Demoiselle cranes from different parts of the range to their flyways, we propose to distinguish the following subpopulations: (1) Azov-Black Sea/Chadian; (2) Caspian/Sudanese; (3) Trans-Ural/Indian; (4) South Siberian/Indian; (5) Baikal/Indian and (6) Trans-Baikal/Indian. The obtained data create the basis for monitoring the genetic diversity of the Demoiselle crane and developing a scientific background for measures to protect the gene pool of the species as a whole and its subpopulations.

Key words: Gruidae; gene pool; cytochrome *b* (cyt *b*); haplotype; genetic diversity; genetic differentiation; population-genetic structure; flyway

For citation: Mudrik E.A., Ilyashenko E.I., Kazimirov P.A., Kondrakova K.D., Archimaeva T.P., Bazarov L.D., Goroshko O.A., Dorzhiev Ts.Z., Kuksin A.N., Postelnykh K.A., Shurkina V.V., Ilyashenko V.Yu., Shatokhina A.V., Politov D.V. Mitochondrial DNA data allow distinguishing the subpopulations in the widespread Demoiselle crane (*Anthropoides virgo*). *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov J Genet Breed*. 2025;29(4):568-577. doi 10.18699/vjgb-25-60

Funding. This study was supported by a grant from the Russian Science Foundation, No. 23-24-00613, agreement from January 13, 2023, https://rscf.ru/project/23-24-00613/

Acknowledgements. We thank Abushin A.A., Babichev Yu.V., Badmaeva E.N., Balzhimaeva S.B., Belik V.P., Gugueva E.V., Davygora A.V., Dzhamirzoev G.S., Kartashov N.D., Kyzyl-ool V.A., Mikhailovsky V.M., Pyzhyanov S.V., Skripnichenko A.Yu., and V.N. Fedosov for assistance in the collection of samples in the field.

Данные митохондриальной ДНК позволяют выделить субпопуляции широкоареального вида журавлей красавки (*Anthropoides virgo*)

Е.А. Мудрик [1] ✉, Е.И. Ильяшенко [1, 2], П.А. Казимиров [1], К.Д. Кондракова [1, 2], Т.П. Арчимаева [3], Л.Д. Базаров [4], О.А. Горошко [5, 6], Ц.З. Доржиев [7, 8], А.Н. Куксин [3], К.А. Постельных [9], В.В. Шуркина [10], В.Ю. Ильяшенко [2], А.В. Шатохина [1], Д.В. Политов [1]

© Mudrik E.A., Ilyashenko E.I., Kazimirov P.A., Kondrakova K.D., Archimaeva T.P., Bazarov L.D., Goroshko O.A., Dorzhiev Ts.Z., Kuksin A.N., Postelnykh K.A., Shurkina V.V., Ilyashenko V.Yu., Shatokhina A.V., Politov D.V., 2025

This work is licensed under a Creative Commons Attribution 4.0 License

[1] Институт общей генетики им. Н.И. Вавилова Российской академии наук, Москва, Россия
[2] Институт проблем экологии и эволюции им. А.Н. Северцова Российской академии наук, Москва, Россия
[3] Тувинский институт комплексного освоения природных ресурсов Сибирского отделения Российской академии наук, Кызыл, Россия
[4] Национальный парк «Тункинский», Кырен, Россия
[5] Институт природных ресурсов, экологии и криологии Сибирского отделения Российской академии наук, Чита, Россия
[6] Государственный природный биосферный заповедник «Даурский», Нижний Цасучей, Россия
[7] Бурятский государственный университет им. Доржи Банзарова, Улан-Удэ, Россия
[8] Институт общей и экспериментальной биологии Сибирского отделения Российской академии наук, Улан-Удэ, Россия
[9] Окский государственный природный биосферный заповедник, Брыкин Бор, Россия
[10] Государственный природный заповедник «Хакасский», Абакан, Россия
mudrik@vigg.ru

Аннотация. Впервые изучен полиморфизм последовательностей фрагмента гена мтДНК цитохрома *b* и на его основе охарактеризована популяционно-генетическая структура журавля красавки (*Anthropoides virgo* Linnaeus, 1778) на большей части ареала в России. Для 157 особей идентифицировано 18 гаплотипов, девять из которых оказались уникальными. Европейские выборки характеризовались большей гаплотипической и нуклеотидной изменчивостью и более сильной генетической дифференциацией, чем азиатские. Поток генов между разными частями ареала красавки, вероятно, осуществляется через птиц, гнездящихся в Зауралье. Общая генетическая дифференциация вида составила >20 % (F_{ST} = 0.265, $p < 0.001$). Структура генофонда сформирована тремя основными гаплотипами, один из которых преобладает в Азово-Черноморском регионе, второй – в Прикаспийском и Волго-Уральском, а третий наиболее распространен в азиатских выборках. Исходя из соответствия внутривидовой генетической дифференциации красавки пролетным путям птиц из разных частей ареала, мы предлагаем в структуре вида выделить следующие субпопуляции: 1) азово-черноморско-чадскую; 2) прикаспийско-суданскую; 3) зауральско-индийскую; 4) южносибирско-индийскую; 5) байкальско-индийскую; 6) забайкальско-индийскую. Полученные данные создают основу для мониторинга генетического разнообразия красавки и разработки научного обоснования мер охраны генофонда как вида в целом, так и его отдельных субпопуляций.
Ключевые слова: Gruidae; генофонд; цитохром *b*; гаплотип; генетическое разнообразие; генетическая дифференциация; популяционно-генетическая структура; пролетный путь

Introduction

The Demoiselle crane (*Anthropoides virgo*, Linneaus, 1778) is a Eurasian crane species, the gene pool of which has been studied only fragmentarily. The breeding part of the Demoiselle crane range extends across the steppe and semi-desert zones from the Azov-Black Sea region of Russia eastward to North-Eastern China. In Europe, the species is categorized as endangered (BirdLife International, 2021) due to habitat degradation, periods of prolonged drought, and a steady population number decline caused, among other things, by hunting at migration routes and wintering grounds, while on a global scale its status is evaluated as least concern (BirdLife International, 2018). The Demoiselle crane is listed in the Red Book of the Russian Federation (Ilyashenko, 2021). The remaining European breeding groups are almost completely localized in the territory of the south of the European part of Russia. As for the Asian part of the range, the core of which is centered in Kazakhstan and Mongolia, its northern border runs along the southern regions of the Trans-Urals and Siberia (Ilyashenko, 2019).

The first and only data on the population genetic structure of the Demoiselle crane to date were obtained by us using microsatellite loci and sequences of the Control Region (CR) of mtDNA. A high level of genetic diversity was revealed for both types of markers in all parts of the range. European groupings have been shown to be more subdivided than Asian ones, and in general, the genetic differentiation of the species is low (Mudrik et al., 2018, 2022). However, these studies were carried out on a small number of individuals from the wild, especially as far as the Asian part of the range is concerned. In addition, some biomaterial from zoo birds was also included. At the same time, the analysis of the CR, justified by the high variability of this non-coding region of mtDNA, may not reflect the structure of the gene pool formed by protein-coding genes, among which cytochrome *b* is recognized as one of the most reliable markers (Zardoya, Meyer, 1996). The overwhelming majority of population genetic studies of cranes were performed on the CR, while information on the polymorphism of the sequences of the more conservative cytochrome *b* in this group of birds is quite scarce (we found only a few identical Demoiselle crane cyt *b* sequences in the NCBI Genbank). In this regard, we set the goal of assessing for the first time at the population level and on a large geographic scale the polymorphism of the mitochondrial cytochrome *b* gene in the Demoiselle crane and characterizing its gene pool in different parts of the range using more representative biomaterial from nature than in previous studies, primarily from previously unstudied Asian groups.

Material and methods

Biological sample collection. In this research, we used biological specimens from 157 individuals of the Demoiselle crane. The study was approved by the Local Bioethics Committee at the Vavilov Institute of General Genetics of the Russian Academy of Sciences (protocols No. 1 dated May 15, 2017 and No. 1 dated May 18, 2023). The source of DNA was blood (or, less often, epidermis) from plucked feathers from the chest or neck area of chicks aged 15–35 days. Feather samples were collected in 2016–2024 during our own expeditions to Demoiselle crane breeding locations run during the seasons when the chicks were yet unable to fly (June–July). The chicks were caught by hand in accordance with permits from the Federal Service for Supervision of Natural Resources of the Russian Federation No. 43 (2016); No. 104, 105, 106

(2017); No. 52, 56 (2018); No. 9, 60 (2019); No. 21 (2023); and No. 78 (2024). After collecting the biomaterial and tagging, which usually takes 5–10 min, the chicks were released and monitored until they rejoined their parents. Plucked feathers were placed in Longmire's preservative solution in screw tubes, transported to the laboratory at room temperature, and then stored in a freezer at –20 ºC. Normally, a Demoiselle crane brood consists of two chicks, and if biomaterial from both sibs was available, a specimen from only one sibling of each pair was included in the analysis.

To designate samples in the European part of the range, we followed the established division into breeding groups (Belik et al., 2011), while for the Asian part, we assigned topographic names to the samples. So, we analyzed 156 unrelated individuals from 10 samples covering most of the species range in Russia: Azov-Black Sea, Caspian, Volga-Ural (which includes several individuals from Western Kazakhstan), Cis-Ural (European part); Trans-Ural, Khakass, Altai, Tyvan, Baikal and Trans-Baikal (Asian part). Some samples (Cis-Ural, Trans-Ural, Khakass, Altai) were represented by a small but the maximum available number of birds due to the low density of Demoiselle cranes in the corresponding study areas and/or low success of their reproduction during the years of field work. In order to increase the Altai sample, we additionally sequenced the biomaterial of an individual kept in the Barnaul Zoo, which, according to documents, originated from the nature of the Altai Krai.

Molecular genetic analysis. Genomic DNA was extracted from plucked feathers using the K-sorb kit (Syntol, Russia) according to the manufacturer's protocol. Amplification of the cytochrome *b* gene fragment was carried out using forward (F: CTACTACTAGCYGCACACTA) and reverse (R: AGG TTGGCGGTTAGGGTTC) oligonucleotide primers (Sun et al., 2020) and the GenPak PCR Core Reagent kit (Isogen Laboratory LTD, Russia) on a GeneExplorer amplifier, model GE-96G (Bioer Technology Co LTD, China). The amplification program consisted of pre-denaturation (94 °C for 5 min), 30 cycles (94 °C for 30 s, 55 °C for 30 s, 72 °C for 1 min) and final elongation (72 °C for 10 min) (Sun et al., 2020). The size and quality of the amplification products were checked by electrophoresis in 1.5 % agarose gel, then they were purified using Cleanup St PCR kits (Evrogen, Russia) and sequenced in the forward direction on an ABI 3130 Genetic Analyzer (Applied Biosystems, USA) at Evrogen Joint Stock Company (Russia).

Analysis of molecular genetic data. Alignment of approximately 900 bp cytochrome *b* sequences obtained from Sanger sequencing was performed against each other and the only complete sequence of this gene of the Demoiselle crane in NCBI Genbank (NC_020573) using the MAFFT algorithm (Katoh et al., 2002) in Geneious v. 9.1.8 (Kearse et al., 2012). Nucleotide diversity, selective neutrality tests, pairwise and total estimates of genetic subdivision of G_{ST} and F_{ST}, female gene flow (number of female migrants per generation) *N*m were calculated using DnaSP v. 6.11.01 (Librado, Rozas, 2009). AMOVA analysis of molecular variability and construction of the median haplotype network using the TCS algorithm (Clement et al., 2002) were performed in PopART (Leigh, Bryant, 2015). Maximum Likelihood haplotype trees were constructed using the IQTree service (Trifinopoulos et al., 2016; Kalyaanamoorthy et al., 2017; Minh et al., 2020) based on the HKY+F (Hasegawa–Kishino–Yano) nucleotide substitution model (Hasegawa et al., 1985), selected as optimal according to the Bayesian criterion (BIC). Branch node support was calculated using the UltraFast Bootstrap method for 1,000 replications (Hoang et al., 2017). The cytochrome *b* sequence of the Demoiselle crane's closest relative, the Blue crane (*Anthropoides paradiseus*) (Genbank accession number U27557), was used as an outgroup.

Graphical visualization of trees was performed in the R environment (R Core Team, 2022) using the ggtree (Yu et al., 2017, 2018; Yu, 2020, 2022), ggtreeExtra (Xu et al., 2021; Yu, 2022), tidytree (Yu, 2022), ggplot2 (Wickham, 2016), pals (Wright, 2024), and ggnewscale (Campitelli, 2024) packages. A heat map of haplotype similarity based on nucleotide substitutions was constructed in the R environment using the algorithm described in the article (Toparslan et al., 2020). To create maps with the geographic localization of haplotypes, the following packages were used: ggmap (Kahle, Wickham, 2013), ggrepel (Slowikowski, 2024), smoothr (Strimas-Mackey, 2023), sp (Pebesma, Bivand, 2005; Bivand et al., 2013) and pals, as well as basic methods of the R environment.

Results

Haplotype distribution patterns

After alignment, the size of the analyzed sequences was 771 bp. In the total sample of 157 birds, 18 haplotypes were identified (Table 1, Fig. 1*a*) (Genbank accession numbers PQ663762–PQ663779). Nine of them (h1, h2, h3, h5, h7, h12, h14, h15, h18) were found in at least two breeding groups. The most frequent (in 50.9 % of individuals) was haplotype h18; it was present in all samples except Cis-Ural. Also, haplotypes h7 (except Cis-Ural and Khakass, 23.6 % of individuals) and h5 (except Azov-Black Sea and Altai, 11.5 % of individuals) were spread almost throughout the entire range. Unique haplotypes were found in the Caspian (h6, h10, h13), Trans-Ural (h4, h16), Tyvan (h8, h17), Baikal (h9) and Trans-Baikal (h11) samples. No unique haplotypes were found in the Azov-Black Sea and Volga-Ural samples; however, the most frequent ones were h7 and h5, respectively, but not h18, as in the others. The only individual from the Cis-Ural sample had a non-unique haplotype: the same was present in the Caspian sample (h2). The haplotype of the Demoiselle crane from the Barnaul Zoo turned out to be the same as in the Azov-Black Sea region (h15) (Table 1), and since the reliability of the origin of this bird was not obvious, we excluded it from the subsequent population genetic analysis, as well as the only available Cis-Ural individual.

Genetic diversity and differentiation

In general, the European and Asian samples were comparable in the number of individuals analyzed, and the same number of haplotypes (11) and segregating sites (10) of cytochrome *b* were found in them (Table 2). The samples from the western (Azov-Black Sea) and eastern (Trans-Baikal) boundaries of the range showed the lowest haplotype (*Hd*) and nucleotide (π) diversity compared to other samples and the average values for

Table 1. Distribution of cytochrome *b* haplotypes in the Demoiselle crane samples

Sample	Haplotype																		Sample size
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
Azov-Black Sea (AB)																			21
Republic of Crimea			1				14								1			3	
Krasnodar Krai							2												
Caspian (CP)																			32
Republic of Kalmykia	1	1			5		1			1			1					11	
Republic of Dagestan	1		1		1	1	1											4	
Stavropol Krai																		2	
Volga-Ural (VU)																			22
Volgograd Region	3				4		3							1				3	
West Kazakhstan Region					3		3							1				1	
Cis-Ural (CU)																			1
Orenburg Region, Sol-Iletsk District		1																	
Trans-Ural (TU)																			7
Orenburg Region, Svetlinsky District	1			1			1									1		2	
Kostanay Region																		1	
Khakass (KH)																			4
Republic of Khakassia												1						3	
Altai (AL)																			6
Altai Republic	1						1											3	
Altai Krai*															1				
Tyvan (TV)																			24
Republic of Tyva					2		9	1									1	11	
Baikal (BK)																			20
Irkutsk Region																		2	
Republic of Buryatia					2		2		1			3						10	
Trans-Baikal (TB)																			20
Zabaikalsky Krai					1						1	4						14	
Sample size	7	2	2	1	18	1	37	1	1	1	1	8	1	2	2	1	1	70	157

* Bird from the Barnaul Zoo, presumably from the Altai Krai.

Europe and Asia and the species as a whole. Reduced values of these indices compared to the average were also found in the Khakass sample, which was the northernmost of those studied.

The maximal number of haplotypes (9) was found in the Caspian sample, while the Volga-Ural and Trans-Ural samples had the highest haplotype diversity. In general, the values of haplotype and nucleotide diversity indices as well as the average number of nucleotide differences were higher in the European part of the range ($Hd = 0.768 \pm 0.027$; $\pi = 0.00178 \pm 00018$; $k = 1.371$) compared to the Asian part ($Hd = 0.635 \pm 0.054$; $\pi = 0.00130 \pm 0.00017$; $k = 1.001$).

Analysis of the similarity and spatial distribution of haplotypes

The heat map of haplotype similarity showed two clusters (h1–h5 and h6–h18), within which haplotypes h5, h7, h6, and h18 demonstrated the greatest similarity to haplotypes from outside of their own group (Fig. 2). This is probably due to

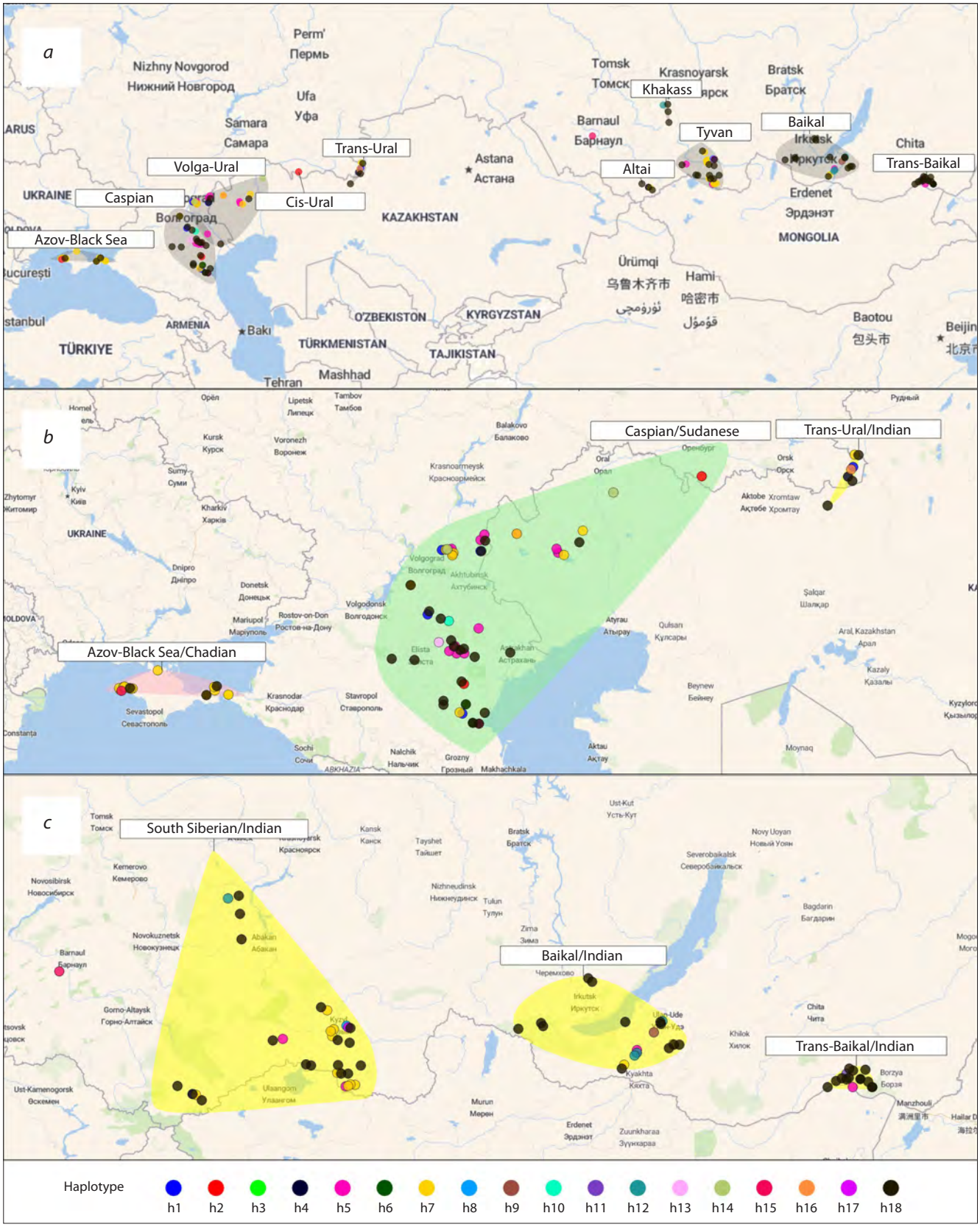


Fig. 1. Location of the Demoiselle crane cytochrome *b* haplotypes in the studied samples (*a*) and subpopulations identified in the European (*b*) and Asian (*b*, *c*) parts of the range.

The contours of the subpopulations are conditional, since they only outline the points of material collection.

Table 2. Summary statistics of polymorphism and genetic differentiation of the Demoiselle crane samples according to cytochrome *b* data

Sample	*N*	*Nh*	S	*Hd*	π	*k*	F_{ST}	G_{ST}	*Nm*
AB	21	4	5	0.414 ± 0.124	0.00091 ± 0.00037	0.705			
CP	32	9	9	0.692 ± 0.079	0.00208 ± 0.00035	1.601			
VU	22	5	4	0.801 ± 0.043	0.00176 ± 0.00024	1.355			
Europe	75	11	10	0.768 ± 0.027	0.00178 ± 0.0018	1.371	0.105	0.158	2.13
TU	7	5	5	0.857 ± 0.137	0.00247 ± 0.00062	1.905			
KH	4	2	1	0.500 ± 0.265	0.00065 ± 0.00034	0.500			
AL	5	3	3	0.700 ± 0.218	0.00182 ± 0.00074	1.400			
TV	24	5	4	0.606 ± 0.062	0.00089 ± 0.00015	0.844			
BK	20	5	5	0.626 ± 0.110	0.00129 ± 0.00033	0.995			
TB	20	4	4	0.489 ± 0.117	0.00090 ± 0.00029	0.695			
Asia	80	11	10	0.635 ± 0.054	0.00130 ± 0.00017	1.001	0.032	0.027	7.66
Average	Total 155	18	16	0.732 ± 0.027	0.00170 ± 0.00014	1.263	0.116	0.128	1.91

Note. *N* – sample size; *Nh* – haplotype number; *S* – segregating site number; *Hd* – haplotype diversity; π – nucleotide diversity; *k* – average number of nucleotide differences; F_{ST} and G_{ST} – genetic subdivision estimates; *Nm* – gene flow. Samples: AB – Azov-Black Sea, CP – Caspian, VU – Volga-Ural, CU – Cis-Ural, TU – Trans-Ural, KH – Khakass, AL – Altai, TV – Tyvan, BK – Baikal, TB – Trans-Baikal.

the fact that h5, h7, and h18 were the most widespread haplotypes, occurring in almost all samples from the studied part of the range of the Demoiselle crane. On the median network, haplotype h7 (and its derivative h6) was located between h5 and h18 (Fig. 3).

The cluster formed by h5 included haplotypes of European samples and the geographically close Trans-Ural sample, and the cluster in which the central haplotype was h18 was distributed throughout the entire studied part of the species range. This is also confirmed by the clustering of individuals on the ML-tree, which demonstrates the intermediate position of individuals with the h7 haplotype relative to h5 and h18 with a high degree of bootstrap support (Fig. 4*a*).

The tree constructed using the outgroup indicated that the cytochrome *b* haplotypes did not form a single monophyletic group, and the h7 haplotype was putatively ancestral to the other two most frequent haplotypes, h5 and h18, and their derivatives (Fig. 4*b*).

Genetic differentiation and gene flow

Genetic differences between the studied samples generally reflected their relative geographic location. The highest genetic differences were found between the most distant Azov-Black Sea and Trans-Baikal (F_{ST} = 0.4675), as well as between the northernmost Khakass and two European samples – Azov-Black Sea and Volga-Ural (Table 3). There was no genetic differentiation detected between the most geographically close Volga-Ural and Trans-Ural; Altai and Tyvan; Baikal and Trans-Baikal groups, as well as between some other samples within the European and Asian parts of the range. The Trans-Ural sample, geographically close to the European ones, but belonging to the Asian group, was genetically indiscernible from the Caspian one in Europe and the Altai and Tyvan ones in Asia,

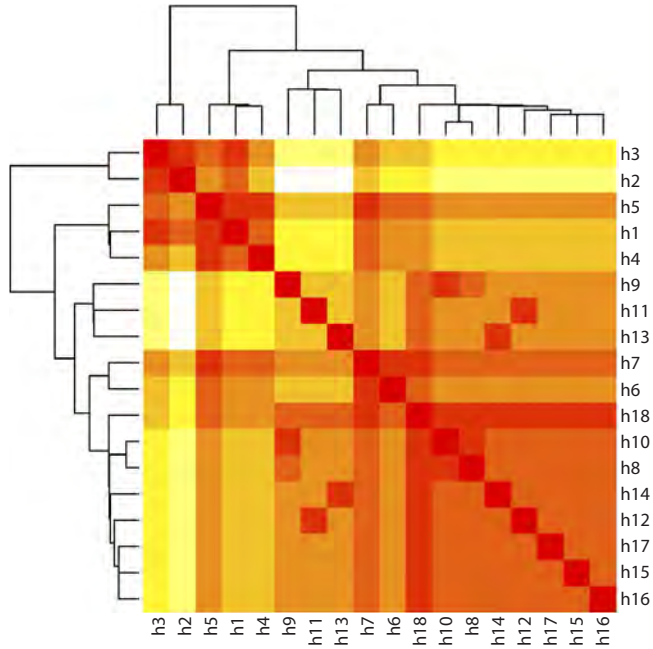


Fig. 2. Heat map of nucleotide differences between Demoiselle crane cytochrome *b* haplotypes.

Color intensity indicates degree of similarity (decreasing from darkest to lightest tint).

and differed only slightly from all other studied samples, which was probably due to its westernmost position in the Asian part of the range. Genetic subdivision within the European samples (F_{ST} = 0.105, G_{ST} = 0.158) was more pronounced than that among the Asian ones (F_{ST} = 0.032, G_{ST} = 0.027), which

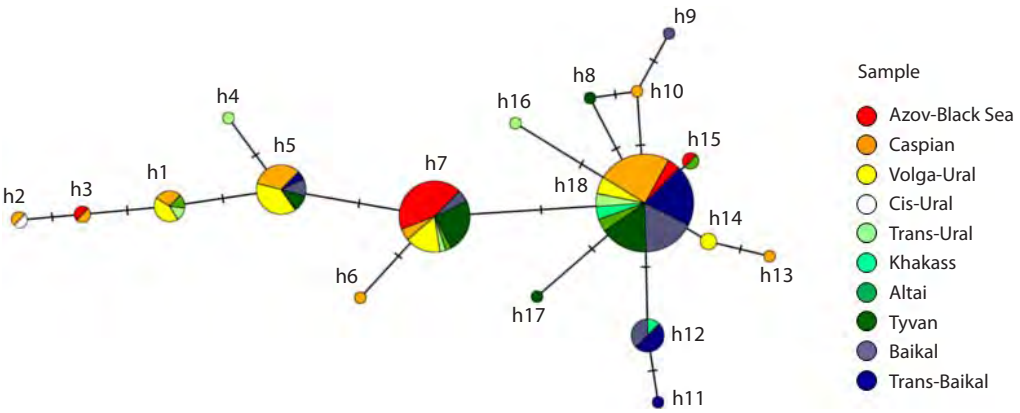


Fig. 3. Median network of Demoiselle crane cytochrome b haplotypes constructed using the TCS algorithm.

The circle size is proportional to the number of individuals, the length of the branches corresponds to the genetic distances, the notches indicate the number of mutation events, and the pie charts display the frequencies of haplotypes in the samples.

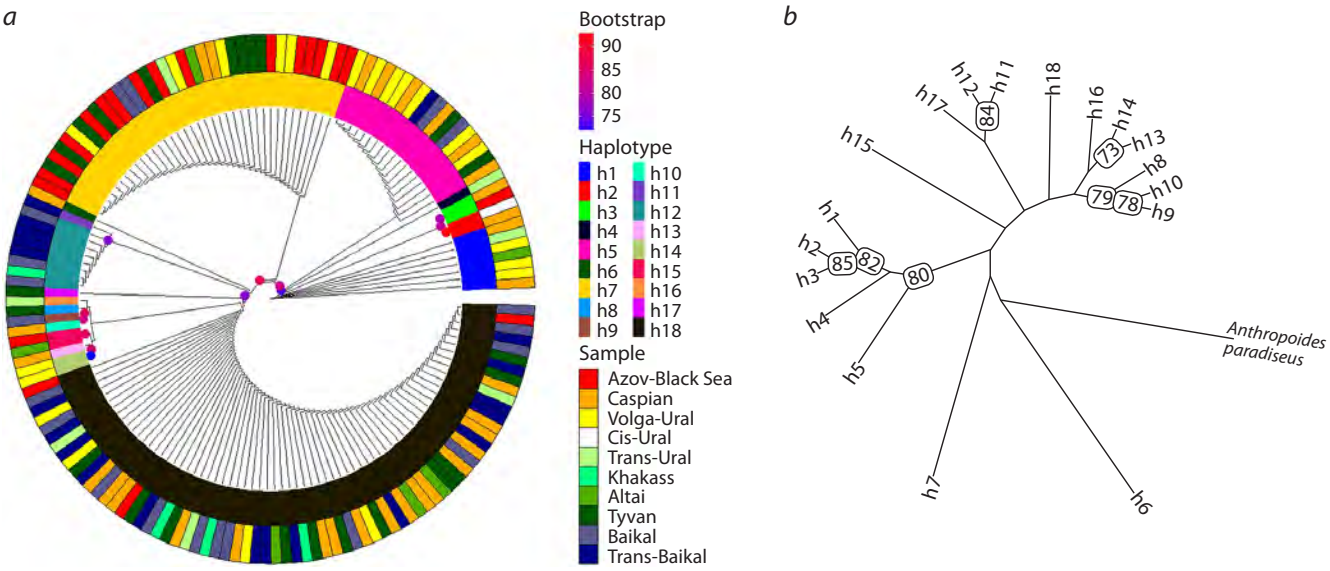


Fig. 4. Maximum likelihood clustering (ML-trees) of individuals (*a*) and haplotypes (*b*) of the Demoiselle crane by nucleotide sequences of cytochrome *b*.

In the left figure, the outer circle illustrates the belonging of individuals to samples, the inner one illustrates the belonging of individuals to haplotypes.

Table 3. Pairwise values of genetic subdivision statistic F_{ST} between the Demoiselle crane samples based on cytochrome *b* sequence data

Sample	AB	CP	VU	TU	KH	AL	TV	BK	TB
AB	–								
CP	0.1505	–							
VU	0.1237	0.0507	–						
TU	0.0734	–0.0669	0	–					
KH	0.5181	0.1996	0.4409	0.1367	–				
AL	0.0555	–0.1012	0.0096	–0.1566	0.0952	–			
TV	0.1148	0.0310	0.1487	–0.0354	0.2318	–0.0947	–		
BK	0.3015	0.0625	0.2551	0.0099	–0.0308	–0.0411	0.0549	–	
TB	0.4675	0.1746	0.3939	0.1166	–0.1378	0.0812	0.2074	0.0003	–

Note. AB – Azov-Black Sea, CP – Caspian, VU – Volga-Ural, CU – Cis-Ural, TU – Trans-Ural, KH – Khakass, AL – Altai, TV – Tyvan, BK – Baikal, TB – Trans-Baikal samples.

Table 4. Results of AMOVA in the total sample of the Demoiselle crane and when divided into European and Asian groups according to cytochrome *b* data

Source	df	Variance, %	F_{ST}
I. Total			0.18570***
Among individuals	9	18.56972	
Within individuals	147	81.43028	
II. Europe and Asia			0.26524***
Between groups	1	16.82040	
Among individuals	8	9.70327	
Within individuals	147	73.47633	

Note. df – number of degrees of freedom; *** $p < 0.001$.

was putatively associated with a more limited gene flow in Europe (*N*m = 2.13) compared to Asia (*N*m = 7.66) (Table 2). The average values of these parameters for the species were estimated as: $F_{ST} = 0.116$, $G_{ST} = 0.128$, *N*m = 1.91. Selective neutrality test values for cytochrome *b* were slightly negative and statistically insignificant ($D = –1.514$, $F = –1.618$), indicating the absence of the Demoiselle crane population expansion in the recent evolutionary past, just as we have shown previously for the CR sequence analysis (Mudrik et al., 2018, 2022).

Hierarchical analysis of molecular variance AMOVA showed that the genetic differentiation of the total studied sample of the Demoiselle crane was 18.57 % (Level I: $F_{ST} = 0.1875$, $p < 0.001$), and when divided into European and Asian groups, it was 26.52 % (Level II: $F_{ST} = 0.26524$, $p < 0.001$) (Table 4).

Discussion

Analysis of the nucleotide sequences of cytochrome *b* in the Demoiselle crane on a large geographical scale and a representative sample of birds from nature revealed polymorphism of this gene and a more pronounced population genetic structuring of the species compared to the data obtained previously for the Control Region of mtDNA (Mudrik et al., 2018, 2022). The lowest haplotype and nucleotide diversity were found in the westernmost (Azov-Black Sea), easternmost (Trans-Baikal) and northernmost (Khakass) samples, which was probably due to their attribution to marginal populations surviving at the edges of the species range. The highest values of these parameters were found eastwards of the Volga River and on both sides of the Urals, i. e. in the Volga-Ural and Trans-Ural samples. The largest number of haplotypes, including unique ones, was found in the Caspian region.

The cytochrome *b* gene pool structure of the Demoiselle crane was formed by the three most frequent haplotypes h5, h7 and h18, of which the most frequent haplotype from the Azov-Black Sea region, h7, was presumably ancestral. It is interesting that the Azov-Black Sea birds differ from other European and especially Asian ones in their migration routes over the Black and Mediterranean Seas and their wintering ground in the Republic of Chad at the junction of North and Central Africa, which was recently discovered using GPS-GSM telemetry (Ilyashenko et al., 2021). Probably, such isolation and greater similarity to the outgroup (the related African species Blue crane) compared to other haplotypes has an evolutionary basis, which needs to be studied further using genomic methods.

The most frequent haplotype in the total studied sample, h18, found in more than half of the individuals and equally common with higher frequencies in the most remote Asian samples (Tyva, Buryatia and Transbaikalia) and prevailing in Altai and Khakassia (Fig. 1*a*), formed a “star” from which most other cytochrome *b* haplotypes, including unique ones, originated (Fig. 3). Birds from all these samples use a common wintering ground in the states of Rajasthan and Gujarat in India, and most of them (except for the Trans-Ural ones) make loop migrations, crossing the Himalayas in autumn and skirting the Tien Shan from the west in spring, sharing a significant part of the flyway (Ilyashenko et al., 2021). All this putatively contributes to gene flow among local breeding groups and a decrease in genetic subdivision of Demoiselle cranes in this part of the range. Genetic differences between Trans-Baikal and Baikal; Altai and Tyvan; Baikal, Altai and Khakass samples were practically absent (Table 3).

Finally, the third of the above-mentioned structure-forming haplotypes h5, lying on the median network on the other side than h18 from the central haplotype h7, was the most frequent in the Trans-Volga region (Volga-Ural sample) and formed the branch of “European” haplotypes, which also included haplotypes from the Cis-Ural and Trans-Ural samples. It should be noted that the previously identified Volga-Ural and Caspian breeding groups (Belik et al., 2011) are essentially a single genetically homogeneous (Table 3, Fig. 1*b*) subpopulation using common migration routes over the Arabian Peninsula and the Red Sea to wintering grounds in Sudan and partly Ethiopia in Africa (Ilyashenko et al., 2021). The only individual from the Cis-Ural sample had the same haplotype as the bird from Kalmykia (Caspian sample) (Table 1) and used the same flyway and pre-migratory gathering site in the Manych Valley as the Caspian and Volga-Ural cranes (Ilyashenko et al., 2021, 2024), which allows it to be classified as part of this subpopulation. It is noteworthy that in the “European” group of haplotypes, half of the haplotypes from the Trans-Ural sample are present. Although the Trans-Ural

sample is geographically close to the European ones in the breeding part of the range (Fig. 1*a*), it uses wintering site in India, like all other Asian Demoiselle cranes. However, birds from the Trans-Ural sample fly in both autumn and spring through Kazakhstan, Uzbekistan, Tajikistan and Pakistan, without making a loop migration (Ilyashenko et al., 2021). According to the F_{ST} values, the Trans-Ural cranes have no genetic differences from the Volga-Ural and Caspian samples in the west, and with the Altai, Tyvan and Baikal samples in the east of the range, and with the geographically marginal ones (Azov-Black Sea, Trans-Baikal and northern Khakass), the proportion of differences was within 7–13 % (Table 3). Thus, we assume that the Trans-Ural integrates the Demoiselle crane gene pool to a certain extent, possibly due to the gene flow between the European and Asian parts of the range through Central and Eastern Kazakhstan, which requires further study in the future using a set of various DNA markers and verification by independent methods. For a more complete understanding of the Demoiselle crane gene pool structure, population genetic studies need to be undertaken in Kazakhstan and Mongolia, the countries with the largest Demoiselle crane population number.

Conclusion

So, we have demonstrated the effectiveness of using sequences of the mitochondrial cytochrome *b* gene, which is less variable than the Control Region, but exhibits a higher degree of interpopulation differentiation, to identify the population genetic structure of the Demoiselle crane. Based on the definition of the term "subpopulation" (interbreeding individuals with highly limited gene flow with adjacent subpopulations) and the correspondence of the intraspecific genetic differentiation data of the Demoiselle crane flyways characterized using remote tracking (Ilyashenko et al., 2021), we propose to distinguish subpopulations in the species structure, reflecting their breeding and wintering grounds in their names: 1) Azov-Black Sea/Chadian (Azov-Black Sea region – Chad); 2) Caspian/Sudanese (Caspian region, Trans-Volga, Cis-Urals – Sudan); 3) Trans-Ural/Indian (East of the Orenburg Region, Northern Kazakhstan and presumably the Chelyabinsk Region – India); 4) South Siberian/Indian (Altai, Khakassia, Tyva – India); 5) Baikal/Indian (Buryatia, Irkutsk Region – India) and 6) Trans-Baikal/Indian (Zabaikalsky Krai – India) (Fig. 1*b*, *c*).

The obtained results create a basis for monitoring the genetic diversity of the Demoiselle crane and developing a scientific justification for its protection measures at the level of species, subpopulations and local breeding groups. Further comprehensive studies (remote tracking and molecular genetic analysis) in other parts of the range will contribute to a more complete understanding of the factors of isolation and integration of the gene pool of this crane species.

References

Belik V.P., Guguyeva E.V., Vetrov V.V., Milobog Y.V. The Demoiselle Crane in the northwestern Caspian lowland: distribution, number, and breeding success. In: Cranes of Eurasia (Biology, Distribution, Migrations, Management). Vol. 4. Moscow, 2011;157-174 (in Russian)

BirdLife International. *Anthropoides virgo*. The IUCN Red List of Threatened Species. 2018. Available at: https://dx.doi.org/10.2305/IUCN.UK.2018-2.RLTS.T22692081A131927771.en

BirdLife International. *Anthropoides virgo* (Europe assessment). The IUCN Red List of Threatened Species. 2021. Available at: https://dx.doi.org/10.2305/IUCN.UK.2021-3.RLTS.T22692081A166235355.en

Bivand R.S., Pebesma E., Gomez-Rubio V. Applied Spatial Data Analysis with R. NY: Springer, 2013. doi 10.1007/978-1-4614-7618-4

Campitelli E. ggnewscale: Multiple Fill and Colour Scales in 'ggplot2'. R package version 0.5.0.9000. 2024. doi 10.5281/zenodo.2543762

Clement M., Snell Q., Walke P., Posada D., Crandall K. TCS: estimating gene genealogies. In: Proceedings 16th International Parallel and Distributed Processing Symposium. IEEE, 2002;184. doi 10.1109/IPDPS.2002.1016585

Hasegawa M., Kishino H., Yano T.-A. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J Mol Evol*. 1985; 22(2):160-174. doi 10.1007/BF02101694

Hoang D.T., Chernomor O., von Haeseler A., Minh B.Q., Vinh L.S. UFBoot2: improving the ultrafast bootstrap approximation. *Mol Biol Evol*. 2018;35(2):518-522. doi 10.1093/molbev/msx281

Ilyashenko E.I. Demoiselle Crane (*Anthropoides virgo*). In: Mirande C.M., Harris J.T. (Eds) Crane Conservation Strategy. Baraboo, Wisconsin, USA: International Crane Foundation, 2019;383-396

Ilyashenko E.I. Demoiselle crane. In: The Red Book of the Russian Federation. Animals. Moscow: VNII Ecologiya Publ., 2021;689-691 (in Russian)

Ilyashenko E.I., Mudrik E.A., Andryushchenko Yu.A., Belik V.P., Belyalov O.V., Wikelski M., Gavrilov A.E., Goroshko O.A., Guguyeva E.V., Korepov M.V., Mnatsekanov R.A., Politov D.V., Postelnykh K.A., Lei C., Ilyashenko V.Yu. Migrations of the Demoiselle Crane (*Anthropoides virgo*, Gruiformes): remote tracking along flyways and at wintering grounds. *Biol Bull*. 2022;49(7):863-888. doi 10.1134/S1062359022070068

Ilyashenko E.I., Kondrakova K.D., Mudrik E.A., Wikelski M., Lei S., Ilyashenko V.Yu. The feature of the use by the Demoiselle Crane (*Anthropoides virgo*, Linneaus 1758) the European part of the range in the spring-summer and the pre-migratory periods. *Arid Ecosystems*. 2024;14(2):209-217. doi 10.1134/S2079096124700100

Kahle D., Wickham H. ggmap: spatial visualization with ggplot2. *R J*. 2013;5(1):144-161. doi 10.32614/RJ-2013-014

Kalyaanamoorthy S., Minh B.Q., Wong T.K.F., von Haeseler A., Jermiin L.S. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods*. 2017;14(6):587-589. doi 10.1038/nmeth.4285

Katoh K., Misawa K., Kuma K., Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res*. 2002;30(14):3059-3066. doi 10.1093/nar/gkf436

Kearse M., Moir R., Wilson A., Stones-Havas S., Cheung M., Sturrock S., Buxton S., Cooper A., Markowitz S., Duran C., Thierer T., Ashton B., Mentjies P., Drummond A. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. 2012;28(12):1647-1649. doi 10.1093/bioinformatics/bts199

Leigh J.W., Bryant D. PopART: full-feature software for haplotype network construction. *Methods Ecol Evol*. 2015;6(9):1110-1116. doi 10.1111/2041-210X.12410

Librado P., Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*. 2009;25(11):1451-1452. doi 10.1093/bioinformatics/btp187

Minh B.Q., Schmidt H.A., Chernomor O., Schrempf D., Woodhams M.D., von Haeseler A., Lanfear R. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol*. 2020;7(5):1530-1534. doi 10.1093/molbev/msaa015

Mudrik E.A., Ilyashenko E.I., Goroshko O.A., Kashentseva T.A., Korepov M.V., Sikorskiy I.A., Dzhamirzoev G.S., Ilyashenko V.Yu., Politov D.V. The Demoiselle crane (*Anthropoides virgo*) population genetic structure in Russia. *Vavilov J Genet Breed*. 2018;22(5):586-592. doi 10.18699/VJ18.398

Mudrik E.A., Ilyashenko E.I., Ilyashenko V.Yu., Postelnykh K.A., Kashentseva T.A., Korepov M.V., Goroshko O.A., Nechaeva A.V., Politov D.V. Genetic diversity and differentiation of the widespread migratory Demoiselle Crane, *Grus virgo*, on the northern edge of the species' distribution. *J Ornithol*. 2022;163(1):291-299. doi 10.1007/s10336-021-01919-4

Pebesma E.J., Bivand R. Classes and methods for spatial data in R. *R News*. 2005;5(2):9-13

R Core Team. R: A Language and Environment for Statistical Computing. Manual. Vienna, Austria: R Foundation for Statistical Computing, 2022

Slowikowski K. ggrepel: Automatically Position Non-Overlapping Text Labels with 'ggplot2'. 2024. Available at: https://ggrepel.slowkow.com/

Strimas-Mackey M. smoothr: Smooth and Tidy Spatial Features. 2023. Available at: https://github.com/mstrimas/smoothr

Sun C.-H., Liu H.-Y., Xu P., Lu C.-H. Genetic diversity of wild wintering red-crowned crane (*Grus japonensis*) by microsatellite markers and mitochondrial *Cyt B* gene sequence in the Yancheng reserve. *Anim Biotechnol.* 2020;32(5):531-536. doi 10.1080/10495398.2020.1725538

Toparslan E., Karabag K., Bilge U. A workflow with R: phylogenetic analyses and visualizations using mitochondrial cytochrome *b* gene sequences. *PLoS One*. 2020;15(12):e0243927. doi 10.1371/journal.pone.0243927

Trifinopoulos J., Nguyen L.-T., von Haeseler A., Minh B.Q. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res*. 2016;44(1):W232-W235. doi 10.1093/nar/gkw256

Wickham H. ggplot2: Elegant Graphics for Data Analysis. Springer, 2016. doi 10.1007/978-3-319-24277-4

Wright K. pals: Color Palettes, Colormaps, and Tools to Evaluate Them. R package version 1.9. 2024. Available at: https://kwstat.github.io/pals/

Xu S., Dai Z., Guo P., Fu X., Liu S., Zhou L., Tang W., Feng T., Chen M., Zhan L., Wu T., Hu E., Jiang Y., Bo X., Yu G. ggtreeExtra: compact visualization of richly annotated phylogenetic data. *Mol Biol Evol*. 2021;38(9):4039-4042. doi 10.1093/molbev/msab166

Yu G. Using ggtree to visualize data on tree-like structures. *Curr Protoc Bioinformatics*. 2020;69(1):e96. doi 10.1002/cpbi.96

Yu G. Data Integration, Manipulation and Visualization of Phylogenetic Treess. Chapman and Hall, 2022. doi 10.1201/9781003279242

Yu G., Smith D., Zhu H., Guan Y., Lam T.T.-Y. ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol Evol*. 2017;8(1):28-36. doi 10.1111/2041-210X.12628

Yu G., Lam T.T.-Y., Zhu H., Guan Y. Two methods for mapping and visualizing associated data on phylogeny using ggtree. *Mol Biol Evol*. 2018;35(2):3041-3043. doi 10.1093/molbev/msy194

Zardoya R., Meyer A. Phylogenetic performance of mitochondrial protein-coding genes in resolving relationships among vertebrates. *Mol Biol Evol*. 1996;13(7):933-942. doi 10.1093/oxfordjournals.molbev.a025661

Conflict of interest. The authors declare no conflict of interest.
Received October 3, 2024. Revised December 10, 2024. Accepted December 10, 2024.

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2025;29(4):578-593

doi 10.18699/vjgb-25-61

Genetic variation and phylogeography of the magpie's genus *Pica* in the Holarctic

A.P. Kryukov

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

Abstract. The theory of Pleistocene refugia is often used to explain the population genetic structure of species. However, it does not fully account for the diversity of species-specific characteristics and natural conditions. The genus *Pica*, which is widespread in the Holarctic, provides an ideal model for studying phylogeographic patterns in order to better understand processes of diversification and speciation. Markers of mitochondrial DNA remain widely used in phylogeographic studies, despite advances of whole genome techniques. We have summarized published research on the mitochondrial DNA Control Region (CR) variation, based on data from 279 samples which represent the majority of extant taxa across the entire distribution range of the genus. In the phylogenetic trees and networks, we found several cases of reciprocal monophyly among most allopatric species and subspecies, and in addition some examples of paraphyly and polyphyly. Bayesian skyline plots were calculated to explore population dynamics over time. They showed varying longevity of the lineages since their origin or after experiencing a bottleneck, e. g., in the case of the Kamchatka population, as well as unequal rates of expansion. In most cases, speciation followed a geographic model involving expansion and vicariance, sometimes with divergence in refugia. Somewhere, peripatric speciation may have happened due to separation of a marginal populations. By comparing haplotype composition among populations, we traced the origin of the recently established populations on Hokkaido and Kyushu islands from a limited number of colonizers from the mainland. Isolated cases of species *in statu nascendi* were identified through evidence of incomplete lineage sorting, leading to paraphyly, or signs of limited unidirectional interspecies introgression of nuclear genes in secondary contact zones. Several hypotheses regarding the formation of the magpie´s range are proposed. Various evolutionary scenarios found in the genus *Pica* were compared to those reported for the other bird species in a number of literature sources.
Key words: mitochondrial DNA; Control Region; speciation; refugia; range; Pleistocene

For citation: Kryukov A.P. Genetic variation and phylogeography of the magpie's genus Pica in the Holarctic. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov J Genet Breed*. 2025;29(4):578-593. doi 10.18699/vjgb-25-61

Funding. The research was carried out within the governmental work program of the Ministry of Science and Higher Education of the Russian Federation (theme No. 121031500274-4).

Acknowledgements. I am much grateful to the following colleagues for active participation in several projects that laid the foundation for this paper: L.N. Spiridonova, E. Haring, H. Suzuki, O.A. Goroshko, S. Edwards, K.A. Kryukov, S.-I. Lee, B. Dorda, S. Mori, E.G. Lobkov, A.Yu. Arkhipov, K. Collier, B. Fang, A.P. Tyunin. Many thanks to Ya.A. Red'kin for providing the map of the magpie range. I am deeply grateful to E. Haring and D. Bazykin for their in-depth editing of the English manuscript.

Генетическая изменчивость и филогеография сорок рода *Pica* Голарктики

А.П. Крюков

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

Аннотация. Для объяснения истории формирования популяционно-генетической структуры видов часто привлекают теорию плейстоценовых рефугиумов. Однако она не может охватить все многообразие видоспецифических особенностей и природных ситуаций. Широко распространенный в Голарктике род сорок *Pica* оказался удобным для построения картины филогеографии с целью познания процессов диверсификации и видообразования. Маркеры митохондриальной ДНК по-прежнему широко используются в филогеографических исследованиях, несмотря на прогресс методов полногеномного секвенирования. Представлен обзор результатов анализа изменчивости контрольного региона (CR) митохондриальной ДНК по опубликованным нами данным от 279 образцов, представляющих подавляющее большинство таксонов сорок. На филогенетических деревьях и сетях

© Kryukov A.P., 2025
This work is licensed under a Creative Commons Attribution 4.0 License

гаплотипов мы обнаружили, помимо реципрокной монофилии аллопатрических видов и подвидов, примеры парафилии и полифилии. Контурные диаграммы демографии популяций показали разную продолжительность жизни линий после их основания либо прохождения «бутылочного горлышка», как в камчатской популяции, и неодинаковую интенсивность экспансий. Видообразование сорок проходило, вероятно, по географической модели за счет расселения и викарирования, в том числе с изоляцией и дивергенцией в рефугиумах. В ряде случаев предполагается перипатрическое видообразование за счет отделения краевых изолятов. По гаплотипическому составу молодых популяций островов Хоккайдо и Кюсю прослежены материковые источники их происхождения. Случаи незавершенного видообразования выявлены по наличию неполной сортировки линий, приводящей к парафилии, либо современной межвидовой интрогрессии ядерных генов. Предложены гипотезы формирования ареалов некоторых таксонов сорок. Привлечение большого объема литературы позволило сопоставить отмеченные в роде *Pica* разнообразные эволюционные сценарии с описанными для других видов птиц.
Ключевые слова: митохондриальная ДНК; контрольный регион; видообразование; рефугиум; ареал; плейстоцен

Introduction

The current distribution and genetic structure of species are primarily shaped by processes which took place during the Quaternary (Avise, Walker, 1998; Hewitt, 2000). However, this perspective often underestimates overlapping processes, such as invasions, and shifts the species' range boundaries, ecological and anthropogenic changes, population size fluctuations and secondary contacts with or without hybridization. Distinguishing the genetic consequences of these factors is crucial for a better understanding of diverse processes driving diversification and speciation. Widespread polytypic species – or complexes of closely related species – are of particular interest for building hypotheses of range formation and for learning the divergence mechanisms. Modern phylogeographic approaches have been adopted to address a wide range of evolutionary and genetic problems (Avise, 2000; Bannikova, 2004; Abramson, 2007; Zink, Barrowclough, 2008; Kholodova, 2009; Edwards et al., 2015, 2016a, b, 2022).

Phylogeographic studies of birds and other animals have made significant advances (reviews: Zink, 1996; Joseph, Omland, 2009; Hickerson et al., 2010; Toews, Brelsford, 2012; McCormack et al., 2013; Ottenburghs et al., 2019; Pârâu, Wink, 2021; Fu, Wen, 2023). Multilocus and genomic databases are expanding, analytic approaches and hypotheses testing methods are becoming more sophisticated, species distribution and ecological niches are being modeled, comparative as well as statistical phylogeography develops. Phylogeographic structures of many bird species, first European and American, were investigated, primarily by using traditional mitochondrial DNA markers. These structures and speciation ways are usually associated with the refugial phenomenon, in which recurring glaciation cycles forced populations to retreat southward and form isolated populations (Taberlet et al., 1998; Hewitt, 2000, 2004). Within such refugia, populations diverged due to genetic drift or/and local selection. In the case of long enough isolation over several glacial cycles, speciation could occur. Populations that underwent a bottleneck in a refugium, suffered a loss of diversity. On the other hand, fusion of the diverged populations within a refugium could increase variation.

However, refugial theory cannot fully explain all of the diverse cases found in nature. Range expansions occurred during the brief interglacials and especially after the Last Glacial Maximum (LGM), when advancing populations experienced new environmental conditions which could drive divergence. Consequently, the postglacial expansion hypothesis provides an alternative pathway of speciation within a short time frame, rather than through refugial isolation over a set of glacial cycles (Hansson et al., 2008). The case of exceptionally fast speciation was reported for the genus *Junco*, where five genetically distinct morphotypes of the species level evolved within ~10,000 years, during a single postglacial expansion (Mila et al., 2007). This challenges the idea that speciation occurred throughout the entire Pleistocene (Avise, Walker, 1998) or at least over the last 250 thousand years (Johnson, Cicero, 2004). Other reports suggest that the principal diversification and speciation events occurred as early as Pliocene and concluded in Pleistocene (Klicka, Zink, 1997). In some cases, phylogeographic breaks may appear within a continuous range, even without any geographic barriers to gene flow, particularly when individual dispersion distances are limited or population size diminished as was demonstrated for the greenish warbler *Phylloscopus trochiloides* (Irwin, 2002). In addition, zones of secondary contact and hybridization appeared in other cases of postglacial expansion. These processes are diverse, usually species-specific and insufficiently studied.

Despite being one of the best-known birds, "the magpie" keeps still many mysteries. Species of the genus *Pica* are widely distributed across the Holarctic from Western Europe to North America and from the arctic tundra to the Arabian deserts (Fig. 1). The genus includes forms with varying degrees of relatedness. Aside from the "good" allopatric species, several subspecies-level forms intergradate in Eurasia, while others form isolates. This naturally causes disputes over their taxonomic rank: whether they should be classified as separate species or subspecies. Interestingly, no magpie species are sympatric. For a long time, all magpies were classified as one species *P. pica* (Linnaeus, 1758) with 9–15 subspecies; however, this species was divided into several species after genetic approaches were adopted. Current taxonomic classifications accept seven species of magpies (Song et al., 2018; Madge et al., 2020; Gill et al., 2021), although the taxonomy of the genus remains a subject of debate.

Isolated populations of North Africa, Arabian Peninsula and Central China are accepted as distinct species based on analyses of mitochondrial and, to some extent, nuclear markers: Maghreb magpie (*P. mauritanica* Malherbe, 1845), Asir magpie (*P. asirensis* Bates, 1936), and black-rumped or Tibetan magpie (*P. bottanensis* Delessert, 1840), respectively (Song et al., 2018). Likewise, two allopatric forms of North America – black-billed magpie (*P. hudsonia* (Sabine, 1823)) and yellow-billed magpie (*P. nuttalli* (Audubon, 1837)) – are

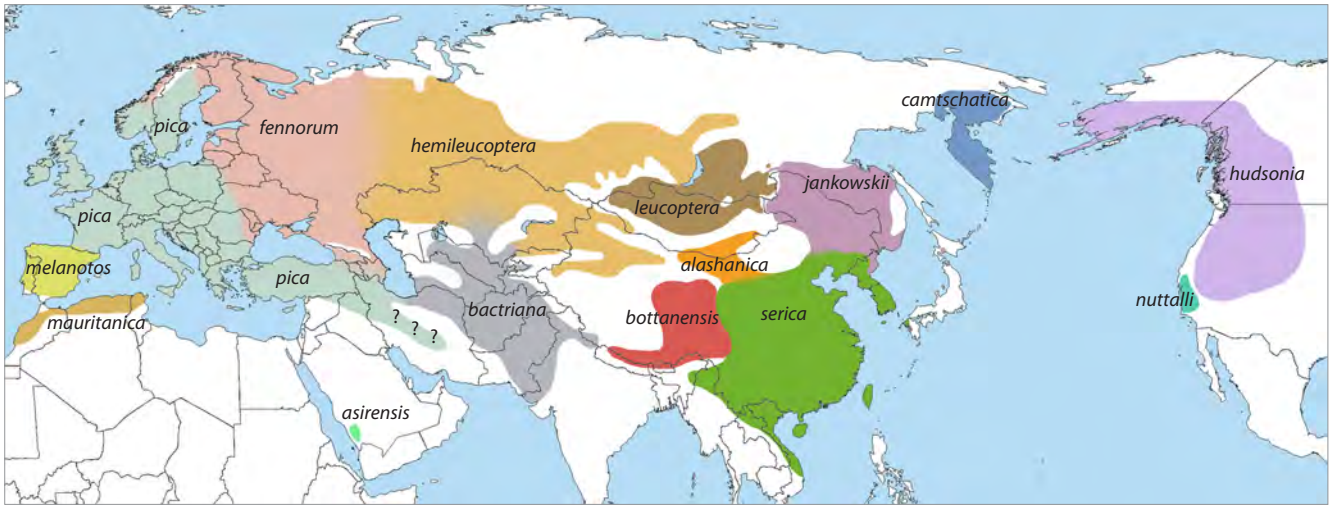


Fig. 1. Distribution of magpies *P. pica*. From (Kryukov et al. 2022), with changes.

clearly distinct by phenotype and genetically diverse. This classification is further supported by reciprocal monophyly in the phylogenetic trees based on single genes (Song et al., 2018) and by complete mitochondrial genome analysis (Kryukov et al., 2020, 2024). The isolated Kamchatka population has been identified as a distinct lineage among *Pica* subspecies (Lee S. et al., 2003) and traditionally treated as subspecies *P. pica camtschatica*, but its potential elevation to species rank is a matter of discussion.

The known gap between the western and eastern subspecies groups in South Siberia has not been previously investigated genetically. They differ by both phenotype and calls (Ebels, 2003; Kryukov et al., 2017). Our analysis of the mt *cyt b* gene and CR estimated *p*-distances between them as 4–5 % (Kryukov et al., 2004, 2017; Haring et al., 2007). These findings became the background for separating the eastern magpie *P. serica* Gould, 1845 from the former single species *P. pica* (Song et al., 2018; Madge et al., 2020). Thus, the modern taxonomic scheme of the genus includes five monotypic species *P. mauritanica*, *P. asirensis*, *P. bottanensis*, *P. hudsonia* and *P. nuttalli*, along with the polytypic species *P. pica* comprising the subspecies *P. p. pica* (Linnaeus, 1758), *P. p. fennorum* Lönnberg, 1927, *P. p. hemileucoptera* Stegmann, 1928, *P. p. bactriana* Bonaparte, 1850, *P. p. leucoptera* Gould, 1862, *P. p. melanotos* A.E. Brehm, 1857 and *P. p. camtschatica* Stejneger, 1884; and *P. serica* including the subspecies *P. s. serica* Gould, 1845, *P. s. jankowskii* Stegmann, 1928 and *P. s. alashanica* Stegmann, 1928 (Winkler et al., 2020; Gill et al., 2021, with small corrections in subspecies).

The above-mentioned range gap between *P. pica* and *P. serica* deserves special attention. The gap was reported by the ornithologists as early as the last century (Stegmann, 1932; Rustamov, 1954), but was ignored in most major studies. The range of *P. pica* therefore appeared to stretch continuously from the Iberian Peninsula to the Sea of Okhotsk (Goodwin, 1986; del Hoyo, Collar, 2016). We established that the gap in fact exists and coincides with a discontinuity in mtDNA. However, it is gradually filling up before our eyes due to the range expansion of the eastern subspecies *P. s. jankowskii* westward along the Amur River valley and the Siberian subspecies *P. p. leucoptera* moving in the opposite direction (Goroshko et al., 2018). It was discovered that a few decades ago these populations came into contact and hybridization started. This zone was the subject of our recent integrative study (Kryukov et al., 2022). It revealed asymmetric introgression using nuclear single nucleotide polymorphism (SNP) analyses. Furthermore, a statistically significant decrease in breeding success was found in a hybrid population in Eastern Mongolia. This implies selection against hybrids and, consequently, limitation of introgression (Kryukov, 2019; Kryukov, Goroshko, 2025).

Despite extensive research on distribution, ecology and variability of magpies throughout the genus' range, a comprehensive understanding of the relationships and origins of the taxa – and the genus in general – is still lacking. The aim of the current study is to summarize both our own and previously published data on genetic variation, phylogeography and population dynamics of (nearly) all *Pica* taxa and to propose a hypothesis on the formation of their ranges. As our main genetic marker, we used the mitochondrial Control Region (CR) which is well-known as a frequently used marker at lower taxonomic levels, widely applied in phylogeographic studies. This noncoding region is one of the most variable and phylogenetically informative regions of mtDNA (Baker, Marshall, 1997; Saunders, Edwards, 2000; Barker et al., 2012). In total, we obtained and analysed 279 sequences ranging in length from 1,298 to 1,310 nucleotide pairs from the samples representing almost all taxa of the genus *Pica* (Kryukov et al., 2004, 2017, 2022; Haring et al., 2007). The origin of the samples, museum numbers and GenBank accession numbers are presented in Table S1[1]. In addition, fragments of mt CR of *P. hudsonia* and *P. nuttalli* were extracted from their total mitogenomes published by Kryukov et al. (2024). We applied commonly accepted methods to analyse nucleotide and haplotype variation, perform neutrality tests, model population dynamics, construct haplotype networks and phylogenetic

[1] Table S1 and Figure S1 are available at:
https://vavilov.elpub.ru/jour/manager/files/Suppl_Kryukov_Engl_29_4.pdf

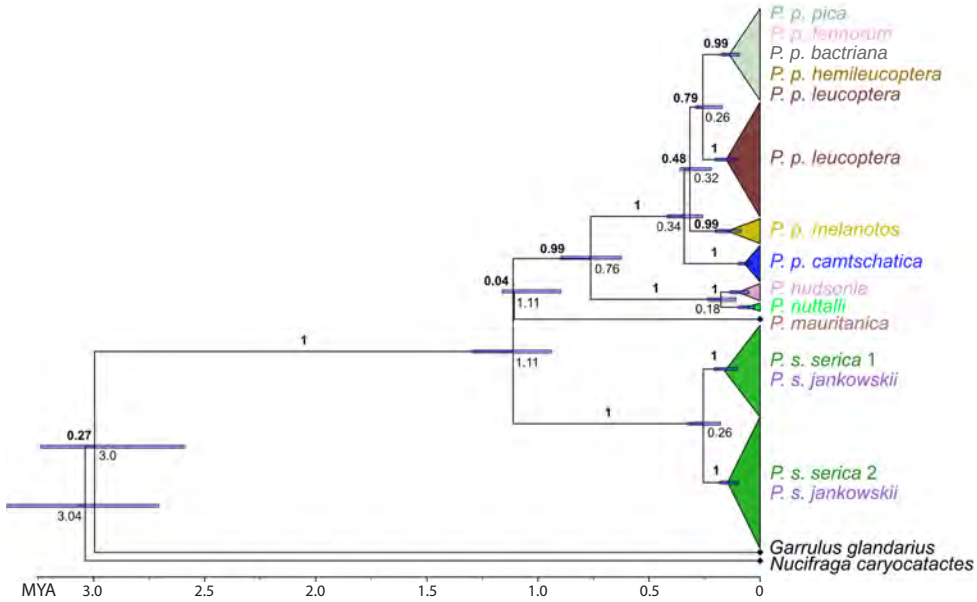


Fig. 2. Time-calibrated Bayesian inference tree based on complete mitochondrial Control Region sequences of *Pica* species and outgroup.

Bold figures indicate Bayesian posterior probabilities. Values below the nodes indicate divergence time estimates (in millions of years), aligned with the time scale below. Blue bars next to the nodes indicate 95 % credibility intervals for age estimates. Triangle widths reflect specimen numbers. Colors of triangles and taxon names correspond to those at the map and network.

trees, and estimate divergence times. These methodologies had been described in details in our studies mentioned above. Published data on both mitochondrial and nuclear genes were considered for the discussion of our findings.

Origin of the magpie genus *Pica*

Origin of the genus *Pica* and the classification of its close relatives remain uncertain. Molecular phylogenetic reconstructions suggest that the ancestral forms of Corvides (previously referred to as "core Corvoidea") diversified during a period of insular isolation as a result of creation of the proto-Papua archipelago after its separation from Australia in the late Eocene–Oligocene (Jønsson et al., 2011; Aggerbeck et al., 2014). Subsequently these birds spread across Asia and other continents. The family Corvidae is believed to have originated in Southeast Asia (Ericson et al., 2005). However, the phylogenetic position of genus *Pica* remains unresolved, and a range of studies indicate various closely related and sister genera. Proposed sister genera include *Ptilostomus* and *Podoces* based on sequences of one mitochondrial and two nuclear genes (Ericson et al., 2005); *Zavattariornis* based on the mitochondrial *cytochrome b* gene (Ekman, Ericson, 2006); *Nucifraga* and *Perisoreus* based on the mitochondrial Control Region (Haring et al., 2012); *Podoces* and *Carrulus* based on the complete mitogenome (Iqbal et al., 2020). However, no study has provided a comprehensive analysis of all possible related genera, leaving the evolutionary origins of the genus *Pica* speculative.

The basal split of the eastern magpie *P. serica* in the phylogenetic tree (Fig. 2) supports the hypothesis that the genus *Pica* originated in Southeast Asia. Its subsequent expansion across the continent was likely linked to the Holocene agricultural centers of south China and Mesopotamia (Nazarenko, 1982). However, our molecular dating suggests that the main divergence events within the genus occurred earlier, between 1 million years ago (hereafter Mya) and 200 thousand years ago (hereafter kya) (Fig. 2). There is a reason to believe that the evolution of magpies was clearly associated with grazing mammals. These animals provide a steady food source in the form of ectoparasites, as well as insects and other small animals disturbed while grazing. Some researchers even proposed a mutualistic relationship with ungulates (Londei, 2018). The long-stepped tail of magpies may have originally functioned as a balancer for perching on the backs of moving ungulates. The role in maneuverable flight among trees may have developed secondarily (Londei, 2018). Magpies primarily forage in short grass and likely spread mainly across grasslands and pastures. With the emergence and expansion of human populations, their high adaptability allowed them to occupy anthropogenic landscapes where they successfully reproduced. Magpies are mainly sedentary, but display a tendency to vagrancy including even hitchhiking on ships, as discussed below.

Phylogeny of the magpies

Despite analyzing only a rather short part of each mitogenome, a high-resolution phylogenetic tree was obtained for all the principal branches, representing nearly all taxa of the genus *Pica* (Fig. 2). Deep divergence revealed between all main lineages generally corresponds to current taxonomic scheme at the species level (Fig. 2 and 3). The three main branches of the tree form a polytomy with deep divergence: 1) eastern magpie *P. serica*, 2) *P. pica* including its subspecies and the related *P. hudsonia* and *P. nuttalli*, and 3) the North-African *P. mauritanica*. The species *P. mauritanica*, *P. hudsonia*, and

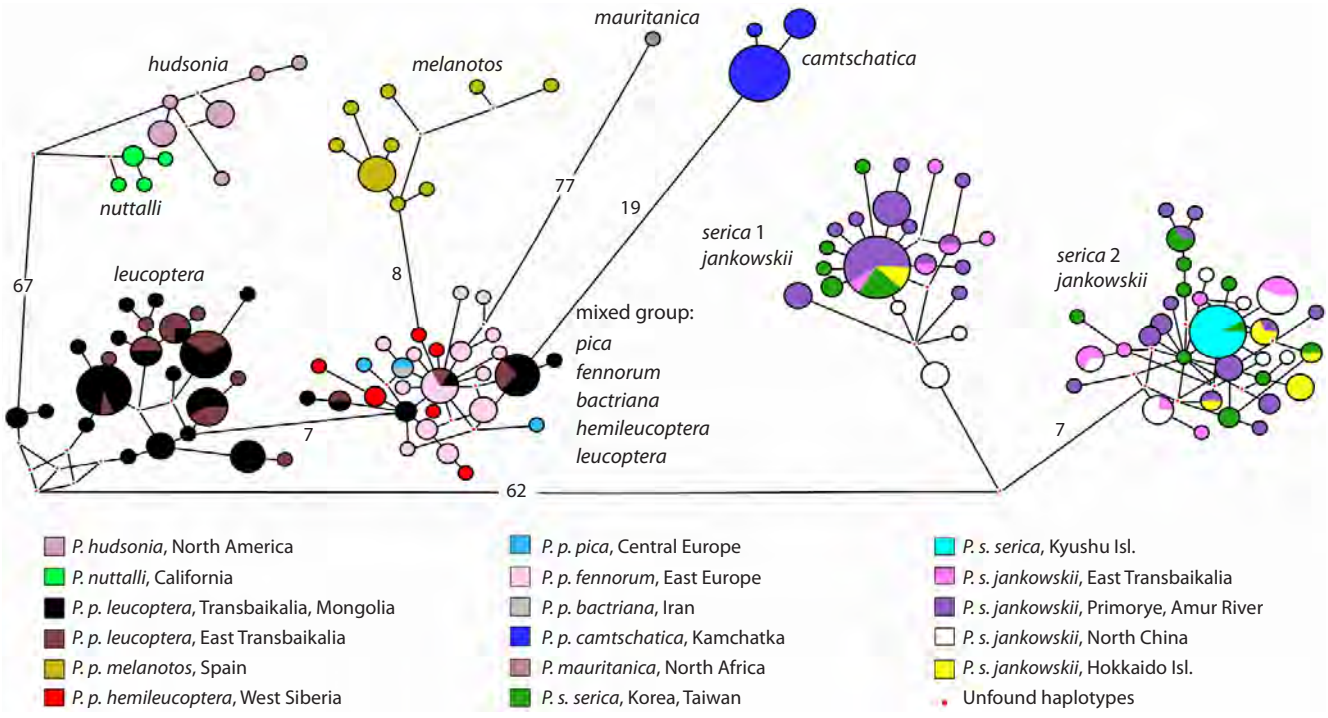


Fig. 3. Phylogenetic Median Joining network based on complete mitochondrial Control Region sequences.

Sizes of circles correspond to the number of birds sharing this haplotype; branch lengths are proportional to the number of substitutions and those over 6 are shown at the nodes.

Table 1. Parameters of variation, neutrality tests and times for the most recent common ancestor (tMRCA) estimations for eight haplogroups of the genus *Pica*

Haplogroup/ taxa	*N*	*S*	*k*	$\pi \pm$ SD, %	*h*	Hd ± SD	D	Fs	R_2	*r*	tMRCA, kyr	
											by the BSP curves	by BEAST +Tracer (95 % HPD)
melanotos	14	19	4.088	0.312 ± 0.081	9	0.835 ± 0.101	–1.314	–2.046	0.081	0.024	54	113.0 (54–179)
leucoptera	61	28	4.861	0.371 ± 0.019	23	0.920 ± 0.017	–0.782	–6.871	0.081	**0.047**	66	118.2 (66–178)
pica, *fennorum* *bactriana*, *hemileucoptera*, *leucoptera*	49	34	3.301	0.252 ± 0.019	28	0.955 ± 0.017	–1.912*	–22.798***	0.043***	**0.031**	36	70.0 (36–111)
serica 1	49	26	2.592	0.198 ± 0.027	21	0.864 ± 0.042	–1.829*	–13.445***	0.044**	0.023	32	76.4 (32–132)
serica 2	70	30	2.694	0.207 ± 0.016	30	0.942 ± 0.017	–1.803*	–25.189***	0.042*	**0.037**	37	73.9 (37–116)
camtschatica	20	5	1.032	0.079 ± 0.022	4	0.489 ± 0.117	–0.820	0.063	0.120	**0.284**	6	35.7 (6–72)
hudsonia	10	9	3.067	0.236 ± 0.000	6	0.867 ± 0.085	–0.158	–0.763	0.162	**0.085**	29	75.0 (29–131)
nuttalli	5	4	1.600	0.123 ± 0.000	4	0.900 ± 0.161	–1.094	–1.405	0.187	**0.150**	6	34.2 (6–72)

Note. *N* – sample size; *S* – number of polymorphic sites; *k* – average number of pairwise nucleotide differences; $\pi \pm$ SD – nucleotide diversity with standard deviation; *h* – number of haplotypes; Hd – haplotype diversity with standard deviation. Neutrality tests: D – Tajima's; Fs – Fu's; R_2 – Ramos-Onsins & Rozas's; their *p*-values: * $p \le 0.05$, ** $p \le 0.01$, *** $p \le 0.001$. *r* – Harpending's raggedness index and its *p*-value, with insignificant values ($p > 0.05$) given in bold. tMRCA – times for the most recent common ancestor estimated by BSP plots and from Tracer, with 95 % HPD range, all in thousands of years.

P. nuttalli along with the subspecies *P. p. camtschatica* and *P. p. melanotos* are reciprocally monophyletic. *P. serica* is monophyletic, but consists of two lineages: *serica* + *jankowskii* 1 and *serica* + *jankowskii* 2 (further briefly *serica* 1 and *serica* 2). Subspecies *P. p. leucoptera* is paraphyletic regarding the other subspecies of *P. pica*.

The highest nucleotide variation and number of pairwise differences were found in the *leucoptera* lineage, while the lowest occurred in the *camtschatica* lineage (Table 1). Haplotype variation is close in all lineages, with the exception of the lowered variation in *camtschatica*. Interspecies nucleotide substitution level was from 4 to 77 (1–6 % *p*-distance), while

Table 2. Average number of nucleotide substitutions per site (D_{xy}) and *p*-distances between *Pica* haplogroups (in %)

Haplogroup	1	2	3	4	5	6	7	8
1. *leucoptera*, *n* = 62		0.997	1.494	1.977	4.990	4.471	4.859	4.829
2. Mixed group, *n* = 49	0.998		0.986	1.618	4.708	4.631	5.060	5.040
3. *melanotos*, *n* = 14	1.499	0.987		1.814	4.486	4.492	4.902	5.096
4. *camtschatica*, *n* = 20	1.976	1.616	1.815		5.373	5.274	5.770	5.720
5. *hudsonia*, *n* = 10	4.491	4.499	4.290	5.123		0.921	5.894	6.309
6. *nuttalli*, *n* = 5	4.276	4.275	4.295	5.030	0.919		5.635	6.014
7. *serica* 1, *n* = 49	4.665	4.850	4.706	5.524	5.526	5.286		1.081
8. *serica* 2, *n* = 70	4.701	4.892	4.953	5.532	6.051	5.779	1.081	

Note. Average number of pairwise nucleotide substitutions per site between haplogroups (D_{xy}) above diagonal and uncorrected *p*-distances below diagonal.

subspecies differences ranged from 0 to 19 substitutions (up to 2 %) (Fig. 3, Table 2). *P. hudsonia* and *P. nuttalli* where the closest related taxa. Some lineages comprised several taxa, such as the subspecies *P. p. pica*, *P. p. fennorum*, *P. p. bactriana*, *P. p. hemileucoptera* and *P. p. leucoptera* in one "mixed" group. In contrast, *P. serica* is represented by two highly distinct clades, each of them mixed regarding subspecies composition. A mutation rate of 0.025 substitutions/site/million years was used to calculate divergence times, consistent with the commonly applied substitution rate in CR of bird mtDNA (Freeland, Boag, 1999; Fok et al., 2002; Omland et al., 2006). Based on these calibrations, divergence of the main lineages of magpies took place in the mid-Pleistocene, about 1.1 Mya (Fig. 2), which is more recent than the previously proposed estimate of 2.5–3.1 Mya (Song et al., 2018). The most recent common ancestor (MRCA) of each lineage appeared in the late Pleistocene, approximately 66 kya or later (Table 1).

Deep divergence of the southeast lineage *serica* from the others was identified earlier by mitochondrial genes coding for *16S rRNA*, *tRNA-Leu* and *ND1* (Lee S. et al., 2003), as well as by *cyt b* (Kryukov et al., 2004), and was later confirmed by CR (Haring et al., 2007; Kryukov et al., 2017). At the same time, the position of the lineage of *P. p. camtschatica* remote from the eastern magpie was demonstrated (Lee S. et al., 2003). The pronounced divergence between the *camtschatica* lineage and the common *hudsonia* and *nuttalli* lineage (Fig. 2 and 3) contradicts the supposition of origin of American magpies form the Kamchatka subspecies (Lee S. at al., 2003). Instead, both American species seem to share a common ancestry with the south Siberian populations *P. p. leucoptera*. In contrast, the African *P. mauritanica* and the Iberian *P. p. melanotos* are more closely linked to the European-Siberian haplogroup (Fig. 3). The relationships among the remaining subspecies of *P. pica* are weakly resolved and appear as a polytomy in the tree (Fig. 2). However, based on a haplotype network and subspecies distribution, there is reason to assume that *P. p. leucoptera* may represent the ancestral haplogroup of all other subspecies of common magpie.

The subspecies *P. p. leucoptera* is paraphyletic regarding the subspecies group *pica*, *fennorum*, *bactriana* and *hemileucoptera* (Fig. 2 and 3). Species-level paraphyly is common in phylogenies based on animal mitochondrial genes and causes inconsistencies in taxon delimitation and discrepancies between gene and species phylogenies. A survey of 2,319 bird species revealed 23 % paraphyletic or polyphyletic for mtDNA (Funk, Omland, 2003). Mitochondrial paraphyly is distributed in 44 % of Australian bird species (Joseph, Omland, 2009). Misinterpretation of paraphyly may lead to false evolutionary inferences. There are numerous examples of erroneous taxonomy, and elevating subspecies status to species status can sometimes eliminate paraphyly. For examples, raising the rank of *Corvus corax clarionensis* to species status solved the problem of paraphyly in American ravens (McKay, Zink, 2010). During divergence from a common ancestor, lineages typically progress through phases of polyphyly, paraphyly, and ultimately reciprocal monophyly, driven by stochastic gene sorting (Avise, 2000). Therefore, a common cause of paraphyly is the incomplete lineage sorting due to recent speciation (Funk, Omland, 2003). In addition, introgressive hybridization, ancient or recent, may contribute to paraphyly, but distinguishing it from incomplete lineage sorting requires nuclear gene analysis involving coalescence models (Peters et al., 2007). In magpies, incomplete lineage sorting is the most likely explanation for subspecies-level paraphyly. This is supported by the observation that in early-stage divergence, common haplotypes are mostly in the centre of a clade, while taxon-specific haplotypes occupy the periphery (Omland et al., 2006). This pattern is clearly visible in the "mixed" haplogroup of the network (Fig. 3).

The presence of well-differentiated haplogroups is well confirmed by the networks we constructed. Each group corresponds to one or several taxa. The network constructed by the NeighborNet method with the SplitsTree software clearly shows the close affinity of *P. hudsonia* and *P. nuttalli* and the sister group relationship between *serica* 1 and *serica* 2 (Fig. S1). *P. p. camtschatica* appears to be most closely related to the "mixed" group. The Median Joining network provides a more detailed picture. Intergroup distances reach 77 substitutions (Fig. 3). The subspecies *P. p. leucoptera* is present in two haplogroups: in the "mixed" group and in the one including *leucoptera* only. The "mixed" group has a star-like structure with the single central haplotype shared by three subspecies. Notably, *P. p. camtschatica* is related to the mixed group, while both American species, *P. hudsonia* and *P. nuttalli*, are

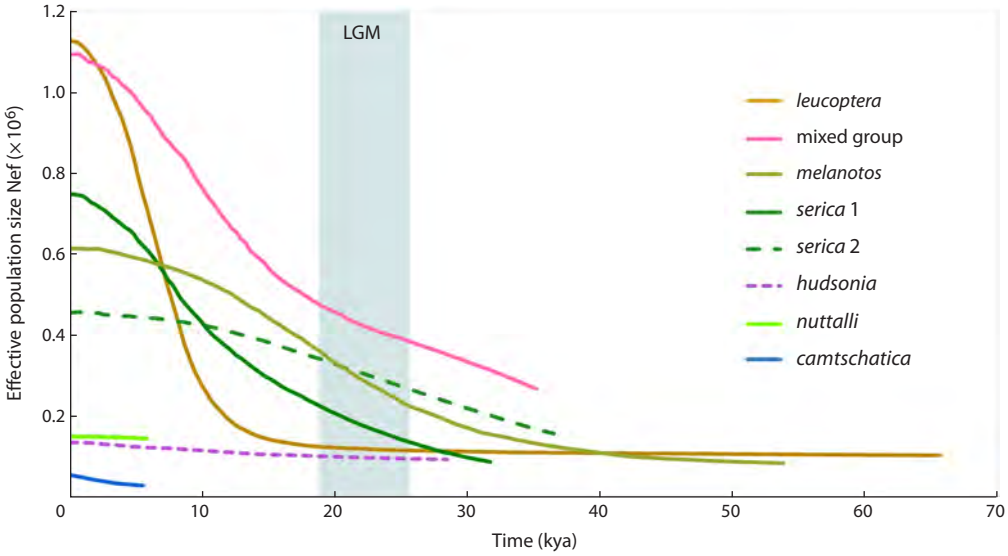


Fig. 4. Bayesian skyline plots for population dynamics over time for eight haplogroups, based on the mitochondrial Control Region.

Curves of the skyline plot represent median values of effective female population size in millions (Nef). Colored column depicts the LGM period.

closer to the Siberian subspecies *P. p. leucoptera*. The groups *serica* 1 and *serica* 2 differ by 10 or more nucleotide substitutions, corresponding to a p-distance of 1.1 %. Representatives of the same populations from both subspecies *P. s. serica* and *P. s. jankowskii* were observed in each group without any apparent geographic pattern. The central haplotype of group *serica* 1 is shared by both subspecies (altogether from four populations), while in group *serica* 2, the main haplotype represents a sample from Korea. Thus, the phylogenetic tree and the haplotype networks show mutually complementary patterns offering insights into lineage divergence and ongoing evolutionary processes.

Population size dynamics

The skyline plots based on Bayesian analysis of mitochondrial haplotypes reflect the dynamics of effective populations sizes for maternal populations and time since the most recent common ancestor (tMRCA). The earliest lineage emergence or bottleneck event was identified for the *leucoptera* population from Transbaikalia and Mongolia, while the latest was observed for *nuttalli* and *camtschatica* (Fig. 4). These two populations, together with *hudsonia*, exhibit relatively stable population sizes, while all the other populations show signs of population growth (Fig. 4). Among these, only three lineages ("mixed" and both *serica* lineages) showed significant support for population growth according to the three neutrality tests (Table 1). Yet the *r*-index suggests that population growth cannot be excluded for most lineages, except for *melanotos* and the *serica* 1 lineage. Pairwise nucleotide difference plots (not shown) displayed single leftwards peaks for all populations except *melanotos*, which does not contradict the hypothesis of population growth.

The diversity of contour diagram patterns depicting lineage or population dynamics suggests the following conclusions. The *leucoptera* lineage appears to have been formed earlier than the others, and the model predicts its most rapid growth after the LGM (Fig. 4). The "mixed" lineage, which has a star-like structure in the haplotype network, also underwent significant growth. Its growth started earlier than that of *leucoptera* and went in parallel with the expansion of the *melanotos* lineage. In eastern Eurasia, among the two *P. serica* lineages, the *serica* 1 lineage (represented by less samples) grew faster, corresponding to a star-like pattern with multiple representations of the common central haplotype (Fig. 3) and a shorter growth curve (Fig. 4). The recent growth of the "mixed" lineage as well as both of the *serica* lineages is supported by three neutrality tests (Table 1). The North American sister species *P. hudsonia* and *P. nuttalli* show population stability. *P. nuttalli*, which inhabits the extreme south of the American part of the genus range, diverged from common ancestor with *P. hudsonia* very recently. Short lifespan of the former species (Fig. 4) is supported by high haplotype and low nucleotide diversity (Table 1) which may indicate a founder effect. Short lifespan as well as low haplotype and nucleotide diversity is also observed in *P. p. camtschatica* (Fig. 4, Table 1), likely indicating a bottleneck rather than a founder effect. Generally, the pattern of population dynamics aligns with the estimated growth of the East Chinese clade after 100 kya (Zhang R. et al., 2012), as well as with the expansion of the east-Asian lineage around 60 kya, Eurasian lineage around 40 kya, and American lineage around 20 kya (Song et al., 2018).

Phylogeography of magpies compared to other birds

The phylogeographic structure of species is primarily manifested in the presence of genetic clades (haplotype groups) that are distributed allopatrically or parapatrically and have been mainly identified by mtDNA data. For example, 14 species have been recorded in the Western Palearctic, which display clear distinctions between geographic lineages within species (Pârâu, Wink, 2021). These are mostly sedentary species. Three allopatric haplogroups corresponding to subspecies

were revealed in the green woodpecker *Picus viridis* (Pons et al., 2011). In the middle spotted woodpecker *Dendrocoptes medius*, two groups were found, each associated with several separate refugia during the LGM (Kamp et al., 2019). Similar patterns have been observed in other species: three haplogroups in the Arctic warbler *Phylloscopus borealis* (Saitoh et al., 2010), three in the black-throated tit *Aegithalos concinnus* in east China (Dai et al., 2011), and three lineages with associated morphotypes in the Steller's jay *Cyanocitta stelleri* (Cicero et al., 2022). Presence of three well-supported haplogroups, originated from three South-European Pleistocene refugia, was shown for the tawny owl *Strix aluco* (Brito, 2005). Five monophyletic groups with deep divergence in the early-middle Pleistocene and expansion before the LGM were discovered in the great tit *Parus major* (Zhao et al., 2012). The dipper *Cinclus cinclus* exhibits a complex structure with five lineages derived from two main refugia, Italian and Balkano-Karpatian, which were isolated during interglacials (Hourlay et al., 2008). In a number of examples, the presence of such divergent clades, often supported by additional distinct features as well, led to proposals for recognizing them as species. This applies to the horned lark *Eremophila alpestris* (Drovetski et al., 2014), the winter wren *Troglodytes troglodytes* (Toews, Irwin, 2008), the long-tailed rosefinch *Carpodacus sibiricus* (Liu et al., 2020) and the Arctic warbler *Phylloscopus borealis* (Alström et al., 2011). On the other hand, there are also cases where taxa were merged rather than split, for example, the lumping of three species of rosy-finches of the genus *Leucosticte* into a single species (Drovetski et al., 2009).

More common, however, is the lack of a clear genetic structuring of species across their ranges. For example, 90 % of the 145 analysed bird species of the western Palearctic show either a high degree of panmixia (46 species) or are only weakly differentiated throughout their ranges (85 species) (Pârâu, Wink, 2021). This finding was attributed to admixing of populations during both their retreat into southern refugia during glaciation and their subsequent post-glacial expansion. Overlapping haplogroup ranges were reported in several species, e. g., for the bunting *Emberiza schoeniclus* (Zink et al., 2008), the common rosefinch *Carpodacus erythrinus* (Pavlova et al., 2005) and the bearded vulture *Gypaetus barbatus* (Godoy et al., 2004). Four groups were identified in the long-tailed tit (*Aegithalos caudatus* complex), two of them occur allopatrically in southern China, while the other two, widespread across the northern Palearctic, overlap (Song et al., 2016). Similarly, among the four distinct clades of the wagtail *Motacilla alba*, three (N, SE and SW) partially overlap (Li X. et al., 2016). The European turtle dove *Streptopelia turtur* does not exhibit panmixia in the European part of its range. However, the three most frequent haplotypes were found in samples of all populations, from Greece to Spain and Great Britain (Calderon et al., 2016). These haplotypes differ by 2–6 substitutions only, which is a much smaller difference than that between the overlapping groups *serica* 1 and 2 (16 substitutions between the centers of the haplogroups in the network (Fig. 3)). Non-strict phylogeographic structure was revealed in the stonechat *Saxicola torquata* complex, which consists of three highly diverged and partially overlapping clades in the Palearctic (Zink et al., 2009). In the Chinese hwamei *Leucodioptron canorum*, three clades partially overlap in east China, with intensive gene flow between offspring of different refugia maintaining a high effective population size (Li S.H. et al., 2009). Similarly, in the vinous-throated parrotbill *Paradoxornis webbianus*, two lineages partly overlap as result of recent gene flow (Qu et al., 2012). In the common raven *Corvus corax*, two groups with divergence level of 4 % overlap in the western United States; this secondary contact allows for frequent interbreeding (Webb et al., 2011).

The examples of species exhibiting wide gene flow includes the hoopoe *Upupa epops* in Europe (Wang et al., 2017), the willow tit *Parus montanus* (Kvist et al., 2001; Pavlova et al., 2006) and the common sandpiper *Actitis hypoleucos* (Zink et al., 2008). Panmixia has been observed in the great spotted woodpecker *Dendrocopos major* throughout the central Palearctic (Perktaş, Quintero, 2013), and in the marsh warbler *Acrocephalus palustris* throughout Europe (Arbabi et al., 2014). In most of the examples cited, bottlenecking or expansion from a single refugium occurred. The genetic affinity among western magpie subspecies, except for the Iberian subspecies, can be explained by recent gene flow between populations (Fig. 2 and 3, Fig. S1).

The significant phylogeographic break in the magpie's range was discovered in South Siberia (Kryukov et al., 2004), which led to the separation of the eastern magpie (*P. serica*) as a distinct species from the previously single species *P. pica sensu lato*. Such a division between closely related western and eastern taxa is observed rather often in species with wide trans-Palearctic ranges and is often reflected in clear mtDNA divergence. Examples of this East/West pattern include: the azure-winged magpie *Cyanopica cyanus* (Fok et al., 2002; Kryukov et al., 2004), the rook *Corvus frugilegus* (Haring et al., 2007; Salinas et al., 2021), the flycatcher *Ficedula parva* and the skylark *Alauda arvensis* (Zink et al., 2008). Apart from birds, similar breaks have been observed in other animals, e. g., in the narrow-headed vole *Microtus gregalis* (Abramson et al., 2006), the Siberian newt *Salamandrella keyserlingii* (Berman et al., 2005) and the wasp spider *Argiope bruennichi* (Krehenwinkel et al., 2016). Interestingly, localization of these phylogeographic breaks rarely coincides across species. Complex structure with division into a western clade (Europe and Caucasus) and an eastern clade (central, eastern Asia and Sino-Himalayas) was found in the Eurasian wren *Nannus troglodytes* (Albrecht et al., 2020). In the carrion crow *Corvus corone* and the hooded crow *C. cornix*, the divergence between mtDNA clades does not coincide with subdivision into subspecies (Kryukov, Suzuki, 2000; Haring et al., 2007). In other examples, subspecific division coincides with phylogeographic breaks: the two subspecies of the black kite *Milvus migrans* show clear mtDNA divergence, with a broad zone of intergradation in Siberia (Andreyenkova et al., 2021). At the same time, other widely distributed species have no such phylogeographic breaks, which may indicate their more recent evolutionary history or/and current gene flow.

History of the formation of the magpie range

The evolutionary history of the genus *Pica* throughout its vast trans-Holarctic range has undergone multiple stages, and it seems difficult to completely reconstruct it. However, several key points may be identified. Vicariance, as a result of fragmentation of previously extensive ranges, together

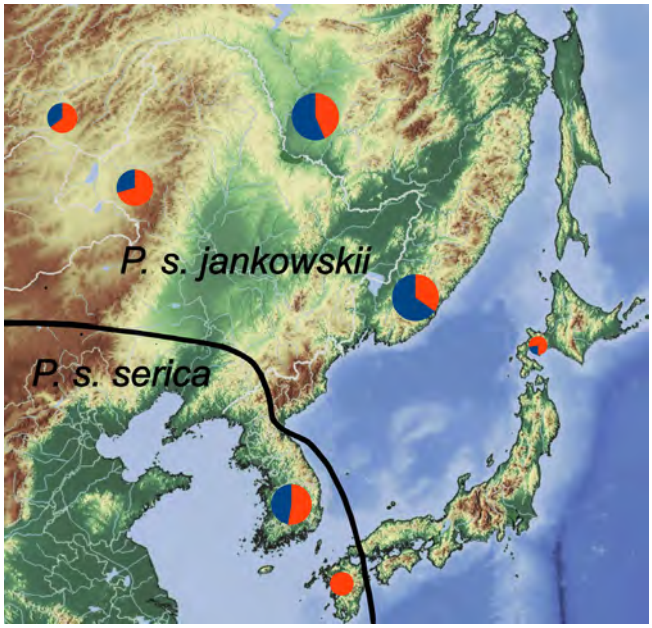


Fig. 5. Genotypic content of the population of the eastern magpie *P. serica.*

Proportion of *serica* 1 haplogroup representatives is shown in blue, and that of *serica* 2 in red.

with local adaptations appears to be the main mechanism of speciation. Marginal populations might have diversified from isolated border population, in accordance with the peripatric speciation model, a variation of the geographic speciation model (Mayr, 1963). This hypothesis was discussed in details for the endemic species *P. nuttalli* in a study analyzing complete mitogenomes of magpies (Kryukov et al., 2024). A similar process may have led to the diversification of the isolated, small-range taxa, e. g., *P. mauritanica*, *P. asirensis* and *P. p. camtschatica*. The Asir magpie *P. asirensis* represents a remote isolate and is still barely studied genetically (Song et al., 2018). However, a recent study of two mitochondrial genes revealed its sister group relation to the Tibetan magpie *P. bottanensis* (Song et al., 2018). This observation suggests a hypothetical spreading of magpies along the belt of steppes and semideserts from East Asia to Arabia and further into North Africa in multiple waves, driven by repeated cooling and warming periods. Range fragmentation may have led to the formation of isolates such as the Asir and Maghreb magpies, which are identified as the most basal lineages according to mitochondrial phylogenetic reconstructions (Song et al., 2018). It is unlikely that the magpie carried on from North Africa to Europe since the Strait of Gibraltar separated them at the early Pliocene, ~5 Mya (Krijgsman, 2002). According to our data, the Maghreb and Iberian magpies are deeply diverged in both phenotypes and mtDNA, and the former species may represent a dead end to that hypothetical route. This scenario, which involves a few remote extant isolates, implies the past elimination of intermediate forms.

The eastern magpie *P. serica* diverged from the other lineages in the middle Pleistocene (Fig. 2). The subspecies *P. s. serica* and *P. s. jankowskii* are geographically separated (Fig. 1). They are similar by plumage coloration, but clearly distinct by size and proportions (Red'kin et al., 2021). The finding of two haplogroups with genetic divergence at *p*-distance of 1.1 % within this species was unexpected (Fig. 2 and 3). Notably, both haplogroups *serica* 1 and *serica* 2 coexist in populations across the species' entire range, from Eastern Transbaikalia in the west to Korea and Hokkaido in the east (Fig. 5). The proportions of both haplogroups vary between population, but not significantly. The only exception is the homogenic population of Kyushu Island, discussed below.

A similar sympatric pattern was reported for the red-backed shrike *Lanius collurio*, where two clearly diverged haplogroups with a genetic distance of 2.8 % coexist throughout Europe (Pârâu et al., 2019). A further example is the common redstart *Phoenicurus phoenicurus*: two haplogroups diverged by 5.1 % are sympatric in the whole of Western Europe (Hogner et al., 2012). In both cases, the most likely explanation of this rare phenomenon is the impact of recurrent glacial cycles throughout the Pleistocene. Populations retreated to the south in cooling periods and intermixed in refugia, e. g. in Iberia and the Balkans, while during the warming cycles they spread northward with repeated mixing. A similar pattern became evident in the European bee-eater *Merops apiaster*, where two star-like haplogroups, differing by only one substitution, connect haplogroups of populations from South Africa to Western Europe and China, illustrating the panmixia resulting from lineage mixing both after refugial isolation as well as in present times (Moura et al., 2019). A different case was observed in the great reed warbler *Acrocephalus arundinaceus* in Europe: two mitochondrial clades diverged 65–87 kya and partly overlapped on a wide range (Hansson et al., 2008). Presumably, they originated independently in two refugia located in Southern Europe and Middle East, respectively. The first expansion wave may have resulted in the occupation of the total range, while the second one was limited to its southern part only (Hansson et al., 2008). These clades are not isolated reproductively, because the length of isolation within refugia proved not to be long enough for establishing complete isolation barriers. It supports the hypothesis of postglacial expansion as a reason for speciation in a rather short period of time. Some other examples were listed above at the section Phylogeography.

Origination and establishment of the haplogroups *serica* 1 and *serica* 2 in the studied parts of Asia likely occurred in two refugia, which is supported by the post-Pleistocene divergence times and tMRCA estimates (Fig. 2, Table 1). While the southern part of the *P. serica* range was not surveyed by us, previous research on two nuclear genes did not reveal any genetic structure in Eastern China (Zhang R. et al., 2012). This suggests a spread from a single refugium and subsequent gene flow between populations. The star-like haplotype structures and the results of neutrality tests, for both haplogroups *serica* 1 and *serica* 2, do not contradict the hypothesis of population expansion, which might have started even before the LGM (Fig. 3 and 4, Table 1). A similar pattern has been observed in the great tit *Parus major*, where a clade in Eastern Asia started a wide expansion ~50 kya showing no impact from the LGM (Song et al., 2020).

In general, the region of eastern temperate Asia might have included many large and small refugia, not necessary the same for different species, in contrast to Europe, where the main refugia were identified in Iberia, the Apennines and the Balkans (Hewitt, 1996; Fu, Wen, 2023). Several species in Far East Asia demonstrate a deep divergence between haplogroups, from the Korean peninsula, on one side, and North-Eastern China and the Primorsky region in the Russian Far East, on the other. Such a pattern was found in the Siberian chipmunk *Tamias sibiricus* with a divergence level of 11 % (Lee M.Y. et al., 2008), the Asian wood mouse *Apodemus peninsulae* (Serizawa et al., 2002; Kim H.R., Park, 2015; Chelomina et al., 2024), the Korean field mouse *Apodemus agrarius* (Sakka et al., 2010), the tree frogs of the group *Hyla japonica* (Dufresnes et al., 2016), the Asiatic toad *Bufo gargarizans* (Borzée et al., 2017) and partly in the raccoon dog *Nyctereutes procyonoides* (Kim S.-I. et al., 2013). Refugia in both Eastern China and Korea were recognized for the black-spotted frog *Pelophylax nigromaculata*, with two lineages having diverged by 7.7 % (Zhang H. et al., 2008). Two refugia were also proposed for the bamboo partridge *Bambusicola thoracica* in China (Huang et al., 2010). Climatic oscillations having occurred in that region did not lead to total glaciation, and average temperatures decreased in the Korean peninsula by only 5–6 °C (Yi, Kim, 2010). Many species survived the LGM in their refugia and subsequently expanded northward, but not exclusively (Fu, Wen, 2023).

The refugial hypothesis in magpies and specifically those of two refugia in the east easily explains the existence of two rather deeply genetically diverged haplogroups: *serica* 1 and *serica* 2. One refugium may have been located in the Korean Peninsula, the most known refugium for Eastern Asia, as the central haplotype of group *serica* 2 in the network originated from Korea (Fig. 3). The ancestral population of group *serica* 1 likely has formed later and expanded faster (Fig. 4). The central haplotype of group *serica* 1 is widely distributed from Hokkaido to Transbaikalia, suggesting that its corresponding refugium might have been located in Primorye or Manchuria, similar to those described above for other species. The divergence level of these groups of 1.1 % (Table 2) is lower than in the examples presented above. This low level of divergence is likely due to a rather short period of isolation in their respective refugia with insufficient time for establishing reproductive isolation in case of secondary contact. In Northeast China, the coldest period of the Pleistocene was not the LGM, as in Europe and America, but the Dali glaciation (corresponding to Würm in Europe), which started 54–44 kya (Li J.J. et al., 2004; Zhang H. et al., 2008), when the refugia could have formed. This dating is close to our estimate of the post-glacial expansion of group *serica* 2 from the Korean Peninsula (37 kya), and a more recent and rapid expansion of *serica* 1 from Primorye (32 kya) (Table 1, Fig. 4). An alternative hypothesis could be that the two haplogroups originated from the same population as a result of ecological speciation. Yet such an explanation appears unlikely. First, it would require the existence of different ecological preferences; however, magpies are eurybionts. Second, it would contradict the differentiation of the two subspecies in their geographic ranges.

After merging, both lineages continued to spread without undergoing lineage sorting (Fig. 5). The population introduced to Kyushu is the only homogenic population among group *serica* 2 (Fig. 3), as explained below. The later and currently ongoing expansion westward along the Amur River valley and eastward to Hokkaido originated from a common population carrying both haplogroups. Overall, the formation of the current genetic structure within the range of *P. serica* implies a process of vicariant divergence in refugia, followed by expansion and subsequent admixing of representatives of both groups.

Magpies are generally sedentary birds but exhibit nomadic tendencies. The initial diversification of the ancient lineage occurred before 1 Mya (Fig. 2), presumably during a period of expansion. This expansion from the original range of *P. pica* in Southeast Asia could have taken two routes: a southern path, south of the deserts and mountains of Central Asia to Arabia and then up to Northern Africa, and a northern path through South Siberia continuing westward. The southern route might have left relic populations, such as ancestors for the Asir and Maghreb magpies. Establishment of the modern populations completed much later, in the late Pleistocene. During the most recent interglacial since ~126 kya, South Siberian forests were replaced by steppe, albeit some forests and forest-steppe landscapes still persisted even during the maximal glaciations (Nazarenko, 1982; Granoszewski et al., 2005; Allen et al., 2010). The presence of sufficient herbaceous vegetation supported ungulates such as the saiga, which may have facilitated the spread of magpies. As magpies spread along the northern path, partial lineage sorting may have led to the formation of two haplogroups in South Siberia (diverged by 1 %) (Table 2). One of them (*leucoptera*) emerged around 66 kya, after the cold period of Marine Isotope Stage (MIS) 4 (71–57 kya). Over time it accumulated considerable nucleotide and haplotype diversity (Table 1). This lineage kept a stable population size until the LGM, after which fast numeric and range growth (Fig. 4) accompanied its wide distribution throughout Siberia and adjacent regions. Descendants of this lineage apparently migrated to Alaska across the Bering Strait and gave rise to the two American species *P. hudsonia* and *P. nuttalli*. The other ("mixed") group might have been formed later, about 36 kya during a relatively warm period MIS 3 (Fig. 4). This may have occurred in the Altai-Sayan refugium (Pavelková Řičánková et al., 2014), also named the "center of spread" (de Lattin, 1957), or in the Hentei subcenter (Nazarenko, 1982). The star-like pattern in the haplotype network (Fig. 3) suggests that it underwent a bottleneck stage, followed by a rapid population growth, surpassing that of its sister lineage (Fig. 4). While spreading to the West, this lineage gave rise to a series of subspecies, ranging from *P. p. hemileucoptera* to the nominate *P. p. pica* (Fig. 1). The lack of lineage sorting in this process resulted in paraphyly of *P. p. leucoptera* (Fig. 2 and 3). In this line of subspecies, clinal variation in size and coloration follows an "isolation by distance" pattern (Cramp, Perrins, 1994), yet the close genetic affinity among these subspecies is indisputable (Fig. 3, Table 2). Physical barriers, such as the Ural Mountains, evidently do not prevent the gene flow between them. The subspecies *P. p. melanotos* presumably originated from this same "mixed" group but

underwent deep divergence beyond the Pyrenees, as is discussed below. Additionally, the "mixed" group lineage gave rise to the Kamchatka subspecies, independently form the American lineage.

Forming the populations of the islands and peninsulas

The homogenic population of Kyushu Island originated from a small number of birds introduced by people from Korea about 400 years ago (Eguchi, Kubo, 1992). It is likely, that just by chance, among the few founders of the population there were no representatives of the other haplogroup, namely *serica* 1. The population has been long protected and had a very restricted range, and only in the last 40 years it started to spread to the north of the island (Eguchi, 2016). The extreme genetic homogeneity of the Kyushu population illustrates the founder effect. Its origin is confirmed by a haplotype from Korea, which is identical to that of all birds from Kyushu (Fig. 3), as well as by analysis of six microsatellite loci (Mori et al., 2014).

The same study demonstrated that the Hokkaido Island population originated from Primorye or Korea rather than from Kyushu. Hokkaido and Primorye populations are closest in allele composition. In addition, mitochondrial haplotypes found in Hokkaido are as diverse as those from the presumably parental population of Primorye, and both haplogroups are present (Fig. 3). This indicates a fairly large number of founders of the Hokkaido population. The first nesting pairs of magpies were met in the port cities of south-western Hokkaido since 1993 (Horimoto, 2004), and the population has since grown to over 200 pairs (O. Hasegawa, personal com.). Notably, magpies do not breed on the neighboring islands of Sakhalin and Honshu. Their wings are not adapted for long-distance flights across open water. The most likely way of their arrival at Hokkaido is the occasional invasion with logging and other ships in the 1980–1990s, when cargo traffic between Primorye and Hokkaido was common. Observations of ornithologists suggest that magpies are attracted to ships in harbors as overnight roosting sites (Kryukov et al., 2017). It is likely that magpies use the same way to make it to Australia (GWA, 2017), Mauritius (Reinegger, Bhanda, 2024), and the eastern USA (Ebels, 2003). The same holds for the widely introduced house crows *Corvus splendens* (Ryall, 2016).

The Kamchatka magpie population shows a close affinity to the western forms and, according to genetic data and morphology, by means of migration from Siberia, and not from the southern *P. serica*. Magpies may have inhabited Kamchatka as early as the Pleistocene. During much of the last stage of the ice age, at least 40 % of the peninsula was covered by ice (Kamchatka…, 1974). However, magpies could have survived the harshest period in refugia of tree and bush vegetation in the Central Kamchatka depression. Postglacial expansion to the north and beyond the peninsula may have been influenced by human activity, including the development of settlements and reindeer herding. So far, the magpie penetrated the anthropogenic landscape of Kamchatka poorly, and in cities it only inhabits parks. The population's extremely low nucleotide and haplotype diversity (Table 1) and the very short curve of the skyline plot (Fig. 4) indicate that it underwent a severe bottleneck in the recent past.

The Iberian Peninsula is recognized as one of the main European refugia alongside the Apennines and the Balkans (Hewitt, 1996). The magpie population in Iberia probably originated from the North, but became isolated behind the Pyrenees during glacial advances, leading to divergence within the refugium. Magpies have since occupied almost the entire peninsula and a recent population growth was noted, which fits our genetic analyses (Fig. 4). Haplotype diversity in *P. p. melanotos* is comparable to that of other widely distributed lineages (Table 1) suggesting the existence of not one but several refugia within the Iberian Peninsula. This aligns with the "refugium within refugium" concept (Gómez, Lunt, 2007; Abellán, Svenning, 2014). Genetic evidence for multiple refugia within Iberia has been reported for other species, e. g., the red-legged partridge *Alectoris rufa* (Ferrero et al., 2011), the ocellated lizard *Lacerta lepida* (Miraldo et al., 2011), and some fish and amphibian species (Gómez, Lunt, 2007). In no less than seven cases, the refugial ranges coincide across different species (Hewitt, 2011). We found no shared haplotypes between *P. p. melanotos* and the nominate subspecies *P. p. pica*, thus there is no clear evidence of gene flow beyond the peninsula. Nevertheless, this cannot be ruled out due to a lack of specimens from the Pyrenees, where intermediate *melanotos* × *pica* phenotypes have been reported (Martínez, 2016). On the other hand, in the Pyrenees there are contacts between the ranges of several animal and plant species (Hewitt, 2011; Poschel et al., 2018; Pons et al., 2019), which allows for the attribution of this ridge to "suture" hybrid zones (Remington, 1968). Unlike magpies, the Iberian rook population (*Corvus frugilegus*) has contributed to the northern populations, although it retains its genetic distinctness as shown by mtDNA and microsatellite analyses (Salinas et al., 2021). Generally, the Pyrenees appear to act as a barrier primarily for sedentary bird species (Neto et al., 2012), and for a few low-mobility amphibians and reptiles.

Conclusion

The magpie genus *Pica* is of great interest for phylogeographic research due to its wide Holarctic distribution and notable phenotypic diversity. Despite its broad range, the genetic variation of magpies remains insufficiently investigated, particularly concerning nuclear genes. At the same time, exploring the highly variable, non-coding Control Region of mtDNA has proven effective for population genetic studies, revealing significant phylogeographic breaks among major genetic lineages. Interestingly, these breaks are not always consistent with the current taxonomic classification of the genus. The degree of reproductive isolation and divergence among genetic lineages in magpies varies considerably, from strict isolation in allopatry (all but one pair of species) to secondary contact with limited gene flow and selection against hybridization, as is the case in *P. pica* × *P. serica*. In contrast, other lineages that presumably diverged in separate refugia, such as *serica* 1 and *serica* 2, have completely merged. Also, the degree of reproductive isolation is weakly correlated with the level of mtDNA divergence. Despite substantial genetic differences, *P. pica* and *P. serica* interbreed rather successfully, and conversely, the parapatric species *P. hudsonia* and *P. nuttalli*, which are closely related in mtDNA, are fully reproductively isolated.

The speciation of magpies appears to have predominantly followed an allopatric model driven by dispersal and vicariance, including separation of marginal isolates (peripatric speciation). The hypothesis combining divergence in two refugia with partial lineage sorting and present gene flow between western subspecies is proposed to explain the paraphyly in the white-winged magpie *P. p. leucoptera*. The presence of two genetically distinct sympatric haplogroups *serica* 1 and *serica* 2 in the eastern magpie *P. serica* might be similarly explained by the hypothesis of divergence in refugia followed by mutual introgression and current expansion. The ongoing process of speciation is also evident in Transbaikalia and Mongolia, where incomplete reproductive isolation of *P. pica* and *P. serica* leads to limited asymmetric introgression of nuclear genes. The gene pool of the young insular populations of Kyushu and Hokkaido reflects the genetic makeup of their parental populations, while the Kamchatka population presumably experienced a glaciation and underwent a bottleneck. Both historical processes and current dynamics of species ranges shape the phylogeographic structure of magpies. Notably, all these phenomena can be detected by rather conventional analysis of the mtDNA Control Region. In total, a widespread and trivial taxon – the common magpie – presents us with a uniquely variable set of microevolutionary processes and their outcomes.

References

Abellán P., Svenning J.-C. Refugia within refugia – patterns in endemism and genetic divergence are linked to Late Quaternary climate stability in the Iberian Peninsula. *Biol J Linn Soc.* 2014;113(1):13-28. doi 10.1111/bij.12309

Abramson N.I. Phylogeography: results, issues and perspectives. *Informatsionnyy Vestnik VOGiS*. 2007;11(2):307-331 (in Russian)

Abramson N.I., Kostygov A.Yu., Gambaryan N.G. Phylogeography of narrow-skulled vole (*Microtus gregalis*, Cricetidae, Rodentia) inferred from the variation of mitochondrial *cyt b* and a number of nuclear genes. *Hystrix It J Mamm* (*N.s.*). Supp. 10th Int. Conf. Rodens & Spatium. 2006;155-156

Aggerbeck M., Fjeldså J., Christidis L., Fabre P.-H., Jønsson K.A. Resolving deep lineage divergences in core corvoid passerine birds supports a proto-Papuan island origin. *Mol Phylogenet Evol.* 2014;70:272-285. doi 10.1016/j.ympev.2013.09.027

Albrecht F., Hering J., Fuchs E. Illera J.C., Ihlow F., Shannon T.J., Collinson J.M., Wink M., Martens J., Päckert M. Phylogeny of the Eurasian Wren *Nannus troglodytes* (Aves: Passeriformes: Troglodytidae) reveals deep and complex diversification patterns of Ibero-Maghrebian and Cyrenaican populations. *PLoS One.* 2020;15(3):e0230151. doi 10.1371/journal.pone.0230151

Allen J.R., Hickler T., Singarayer J.S., Sykes M.T., Valdes P.J., Huntley B. Last glacial vegetation of northern Eurasia. *Quaternary Sci Rev.* 2010;29(19-20):2604-2618. doi 10.1016/j.quascirev.2010.05.031

Alström P., Saitoh T., Williams D., Nishiumi I., Shigeta Y., Ueda K., Irestedt M., Björklund M., Olsson U. The Arctic Warbler *Phylloscopus borealis* – three anciently separated cryptic species revealed. *Ibis.* 2011;153(2):395-410. doi 10.1111/j.1474-919X.2011.01116.x

Andreyenkova N.G., Karyakin I.V., Starikov I.J., Sauer-Gürth H., Literák I., Andreyenkov O.V., Shnayder E.P., Bekmansurov R.H., Alexeyenko M.N., Wink M., Zhimulev I.F. Phylogeography and demographic history of the black kite *Milvus migrans*, a widespread raptor in Eurasia, Australia and Africa. *J Avian Biol.* 2021;52(10):e02822. doi 10.1111/jav.02822

Arbabi T., Gonzalez J., Wink M. Mitochondrial evidence for genetic diversity and low phylogeographic differentiation in the Marsh Warbler *Acrocephalus palustris* (Aves: Acrocephalidae). *Org Divers Evol.* 2014;14:409-417. doi 10.1007/s13127-014-0177-3

Avise J.C. Phylogeography: the history and formation of species. Boston, MA: Harvard Univ. Press, 2000

Avise J.C., Walker D. Pleistocene phylogeographic effects on avian populations and the speciation process. *Proc R Soc B Biol Sci*. 1998;265(1395):457-463. doi 10.1098/rspb.1998.0317

Baker A.J., Marshall H.D. Mitochondrial control region sequences as tools for understanding evolution. In: Mindell D.P. (Ed.) Avian Molecular Evolution and Systematics. San Diego, California: Acad. Pr., 1997;51-82

Bannikova A.A. Molecular markers and modern phylogenetics of mammals. *Zhurnal Obshchei Biologii = Journal of General Biology.* 2004;65(4):278-305 (in Russian)

Barker F.K., Benesh M.K., Vandergon A.J., Lanyon S.M. Contrasting evolutionary dynamics and information content of the avian mitochondrial control region and ND2 gene. *PLoS One.* 2012;7(10):e46403. doi 10.1371/journal.pone.0046403

Berman D.I., Derenko M.V., Malyarchuk B.A., Grzybowski T., Kryukov A.P., Miscicka-Sliwka D. Genetic polymorphism of Siberian newt (*Salamandrella keyserlingii*, Caudata, Amphibia) in its range and the cryptic species of the newt *S. schrenckii* from Primorie. *Doklady Biological Sciences.* 2005;403(1-6):275-278. doi 10.1007/s10630-005-0110-1

Borzée A., Santos J.L., Sánchez-RamÍrez S., Bae Y., Heo K., Jang Y., Jowers M.J. Phylogeographic and population insights of the Asian common toad (*Bufo gargarizans*) in Korea and China: population isolation and expansions as response to the ice ages. *PeerJ.* 2017;5:e4044. doi 10.7717/peerj.4044

Brito P.H. The influence of Pleistocene glacial refugia on tawny owl genetic diversity and phylogeography in Western Europe. *Mol Ecol.* 2005;14(10):3077-3094. doi 10.1111/j.1365-294X.2005.02663.x

Calderon L., Campagna L., Wilke T., Lormee H., Eraud C., Dunn J.C., Rocha G., Zehtindjiev P., Bakaloudis D.E., Metzger B., Cecere J.G. Genomic evidence of demographic fluctuations and lack of genetic structure across flyways in a long distance migrant, the European turtle dove. *BMC Evol Biol.* 2016;16(1):237. doi 10.1186/s12862-016-0817-7

Chelomina G.N., Meschersky I.G., Gajduchenko H., Borisov Y.M. Phylogeography of Korean field mouse *Apodemus peninsulae* (Rodentia: Muridae): an update. *Zool J Linn Soc.* 2024;zlae016. doi 10.1093/zoolinnean/zlae016

Cicero C., Mason N.A., Oong Z., Title P.O., Morales M.E., Feldheim K.A., Koo M.S., Bowie R.C. Deep ecomorphological and genetic divergence in Steller's Jays (*Cyanocitta stelleri*, Aves: Corvidae). *Ecol Evol.* 2022;12(12):e9517. doi 10.1002/ece3.9517

Cramp S., Perrins C.M. (Eds) Handbook of the Birds of Europe the Middle East and North Africa. The Birds of the Western Palearctic. Vol. VIII. Crows to Finches. Oxford: Oxford Univ. Press, 1994

Dai C., Zhao N., Wang W., Lin C., Gao B., Yang X., Zhang Z., Lei F. Profound climatic effects on two East Asian black-throated tits (Aves: Aegithalidae), revealed by ecological niche models and phylogeographic analysis. *PLoS One.* 2011;6(12):e29329. doi 10.1371/journal.pone.0029329

de Lattin G. Die Ausbreitungszentren der Holarktischen Landtierwelt. In: Pflugfelder O. (Ed.) Verhandlungen der Deutschen Zoologischen Gesellschaft, vom 21. bis 26. Mai 1956 in Hamburg. Zoologischer Anzeiger, 20. Supplementband. 1957;380-410

del Hoyo J., Collar N.J. HBW and BirdLife International Illustrated Checklist of the Birds of the World. 2. Passerines. Barcelona: Lynx Edicions, 2016

Drovetski S.V., Zink R.M., Mode N.A. Patchy distributions belie morphological and genetic homogeneity in rosy-finches. *Mol Phyloget Evol.* 2009;50(3):437-445. doi 10.1016/j.ympev.2008.12.002

Drovetski S.V., Raković M., Semenov G., Fadeev I.V., Red'kin Y.A. Limited phylogeographic signal in sex-linked and autosomal loci despite geographically, ecologically, and phenotypically concordant structure of mtDNA variation in the Holarctic avian genus *Eremophila*. *PLoS One*. 2014;9(1):e87570. doi 10.1371/journal.pone.0087570

Dufresnes C., Litvinchuk S.N., Borzée A., Jang Y., Li J.T., Miura I., Perrin N., Stöck M. Phylogeography reveals an ancient cryptic radiation in East-Asian tree frogs (*Hyla japonica* group) and complex relationships between continental and island lineages. *BMC Evol Biol*. 2016;16(1):253. doi 10.1186/s12862-016-0814-x

Ebels E.B. Speciation in *Pica* magpies. *Dutch Birding*. 2003;25(2): 103-116

Edwards S., Schultz A., Campbell-Staton S. Next-generation sequencing and the expanding domain of phylogeography. *Folia Zool*. 2015; 64(3):187-206. doi 10.25225/fozo.v64.i3.a2.2015

Edwards S.V., Potter S., Schmitt C.J., Bragg J.G., Moritz C. Reticulation, divergence, and the phylogeography-phylogenetics continuum. *Proc Natl Acad Sci USA*. 2016a;113(29):8025-8032. doi 10.1073/pnas.1601066113

Edwards S.V., Xi Z., Janke A., Faircloth B.C., McCormack J.E., Glenn T.C., Zhong B., Wu S., Lemmon E.M., Lemmon A.R., Leaché A.D. Implementing and testing the multispecies coalescent model: a valuable paradigm for phylogenomics. *Mol Phylogenet Evol*. 2016b;94:447-462. doi 10.1016/j.ympev.2015.10.027

Edwards S.V., Robin V.V., Ferrand N., Moritz C. The evolution of comparative phylogeography: putting the geography (and more) into comparative population genomics. *Genome Biol Evol*. 2022;14(1): evab176. doi 10.1093/gbe/evab176

Eguchi K. The Eurasian Magpie. *Jpn J Ornithol*. 2016;65(1):5-30

Eguchi K., Kubo H. The origin of the Magpie *Pica pica sericea* in Japan – an investigation of historical records. *J Yamashina Inst Ornithol*. 1992;24:32-39 (in Japanese with English abstract)

Ekman J., Ericson P.G.P. Out of Gondwanaland; the evolutionary history of cooperative breeding and social behaviour among crows, magpies, jays and allies. *Proc Biol Sci*. 2006;273(1590):1117-1125. doi 10.1098/rspb.2005.3431

Ericson P.G.P., Jansen A.-L., Johansson U.S., Ekman J. Inter-generic relationships of the crows, jays, magpies and allied groups (Aves: Corvidae) based on nucleotide sequence data. *J Avian Biol*. 2005; 36(3):222-234. doi 10.1111/j.0908-8857.2001.03409.x

Ferrero M.E., Blanco-Aguiar J.A., Lougheed S.C., Sánchez-Darbudo I., De Nova P.J., Villafuerte R., Dávila J.A. Phylogeography and genetic structure of the red-legged partridge (*Alectoris rufa*): more evidence for refugia within the Iberian glacial refugium. *Mol Ecol*. 2011;20(12):2628-2642. doi 10.1111/j.1365-294X.2011.05111.x

Fok K.W., Wade C.M., Parkin D.T. Inferring the phylogeny of disjunct populations of the azure-winged magpie *Cyanopica cyanus* from mitochondrial control region sequences. *Proc Biol Sci*. 2002; 269(1501):1671-1679. doi 10.1098/rspb.2002.2057

Freeland J.R., Boag P.T. Phylogenetics of Darwin's finches: paraphyly in the tree-finches, and two divergent lineages in the Warbler Finch. *The Auk*. 1999;116(3):577-588

Fu J., Wen L. Impacts of Quaternary glaciation, geological history and geography on animal species history in continental East Asia: a phylogeographic review. *Mol Ecol*. 2023;32(16):4497-4514. doi 10.1111/mec.17053

Funk D.J., Omland K.E. Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Ann Rev Ecol Evol Syst*. 2003;34(1):397-423. doi 10.1146/annurev.ecolsys.34.011802.132421

Gill F., Donsker D., Rasmussen P. (Eds) IOC World Bird List (v 11.2). 2021. doi 10.14344/IOC.ML.11.2

Godoy J.A., Negro J.J., Hiraldo F., Donázar J.A. Phylogeography, genetic structure and diversity in the endangered bearded vulture (*Gypaetus barbatus*, L.) as revealed by mitochondrial DNA. *Mol Ecol*. 2004;13(2):371-390. doi 10.1046/j.1365-294x.2003.02075.x

Gómez A., Lunt D.H. Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula. In: Weiss S., Ferrand N. (Eds) Phylogeography of Southern European Refugia. Dordrecht: Springer, 2007;155-188. doi 10.1007/1-4020-4904-8_5

Goodwin D. Crows of the World. Seattle, WA, 1986

Goroshko O.A., Kryukov A.P., Liu Songtao, Dou Huashan, Bazyrool B.K. On distribution, subspecies and taxonomic rank of the magpie (*Pica pica*) in the Hailar-Argun' river basin (North-East China and Transbaikalia, Russia). *Baykal'skiy Zoologicheskiy Zhurnal = Baikal Zoological Journal*. 2018;2(23):38-45 (in Russian)

Granoszewski W., Demske D., Nita M., Heumann G., Andreev A.A. Vegetation and climate variability during the Last Interglacial evidenced in the pollen record from Lake Baikal. *Global Planet Change*. 2005;46(1-4):187-198. doi 10.1016/j.gloplacha.2004.09.017

GWA (Government of Western Australia). Keep eyes peeled for unusual birds at ports. 2017. Available at https://www.agric.wa.gov.au/news/media-releases/keep-eyes-peeled-unusual-birds-ports. Accessed January 8, 2024

Hansson B., Hasselquist D., Tarka M., Zehtindjiev P., Bensch S. Postglacial colonisation patterns and the role of isolation and expansion in driving diversification in a passerine bird. *PLoS One*. 2008;3(7):e2794. doi 10.1371/journal.pone.0002794

Haring E., Gamauf A., Kryukov A. Phylogeographic patterns in widespread corvid birds. *Mol Phylogenet Evol*. 2007;45:840-862. doi 10.1016/j.ympev.2007.06.016

Haring E., Däubl B., Pinsker W., Kryukov A., Gamauf A. Genetic divergences and intraspecific variation in corvids of the genus *Corvus* (Aves: Passeriformes: Corvidae) – a first survey based on museum specimens. *J Zool Syst Evol Res*. 2012;50(3):230-246. doi 10.1111/j.1439-0469.2012.00664.x

Hewitt G.M. Some genetic consequences of ice ages, and their role in divergence and speciation. *Biol J Linn Soc*. 1996;58(3):247-276. doi 10.1006/bijl.1996.0035

Hewitt G. The genetic legacy of the Quaternary ice ages. *Nature*. 2000; 405(6789):907-913. doi 10.1038/35016000

Hewitt G.M. Genetic consequences of climatic oscillations in the Quaternary. *Philos Trans R Soc Lond B Biol Sci*. 2004;359(1442): 183-195. doi 10.1098/rstb.2003.1388

Hewitt G.M. Quaternary phylogeography: the roots of hybrid zones. *Genetica*. 2011;139(5):617-638. doi 10.1007/s10709-011-9547-3

Hickerson M.J., Carstens B.C., Cavender-Bares J., Crandall K.A., Graham C.H., Johnson J.B., Rissler L., Victoriano P.F., Yoder A.D. Phylogeography's past, present, and future: 10 years after Avise, 2000. *Mol Phylogenet Evol*. 2010;54(1):291-301. doi 10.1016/j.ympev.2009.09.016

Hogner S., Laskemoen T., Lifjeld J.T., Porkert J., Kleven O., Albayrak T., Kabasakal B., Johnsen A. Deep sympatric mitochondrial divergence without reproductive isolation in the common redstart *Phoenicurus phoenicurus*. *Ecol Evol*. 2012;2(12):2974-2988. doi 10.1002/ece3.398

Horimoto T. Records of magpie *Pica pica* in Iburi district, Southwestern Hokkaido. *J Yamashina Inst Ornithol*. 2004;36:87-90 (in Japanese with English abstract)

Hourlay F., Libois R., D'Amico F., Sarà M., O'Halloran J., Michaux J.R. Evidence of a highly complex phylogeographic structure on a specialist river bird species, the dipper (*Cinclus cinclus*). *Mol Phylogenet Evol*. 2008;49(2):435-444. doi 10.1016/j.ympev.2008.07.025

Huang Z., Liu N., Liang W., Zhang Y., Liao X., Ruan L., Yang Z. Phylogeography of Chinese bamboo partridge, *Bambusicola thoracica thoracica* (Aves: Galliformes) in south China: inference from mitochondrial DNA control-region sequences. *Mol Phylogenet Evol*. 2010;56(1):273-280. doi 10.1016/j.ympev.2010.01.028

Hung C.M., Drovetski S.V., Zink R.M. Recent allopatric divergence and niche evolution in a widespread Palearctic bird, the common rosefinch (*Carpodacus erythrinus*). *Mol Phylogenet Evol*. 2013; 66(1):103-111. doi 10.1016/j.ympev.2012.09.012

Iqbal F., Ayub O., Song B.K., Wilson R., Fahim M., Rahman S. Sequence and phylogeny of the complete mitochondrial genome of the Himalayan jungle crow (Corvidae: *Corvus macrorhynchos intermedius*) from Pakistan. *Mitochondrial DNA B Resour.* 2020;5(1): 348-350. doi 10.1080/23802359.2019.1704637

Irwin D.E. Phylogeographic breaks without geographic barriers to gene flow. *Evolution*. 2002;56(12):2383-2394. doi 10.1111/j.0014-3820.2002.tb00164.x

Johnson N.K., Cicero C. New mitochondrial DNA data affirm the importance of Pleistocene speciation in North American birds. *Evolution*. 2004;58(5):1122-1130. doi 10.1111/j.0014-3820.2004.tb00445.x

Jønsson K.A., Fabre P.-H., Ricklefs R.E., Fjeldså J. Major global radiation of corvoid birds originated in the proto-Papuan archipelago. *Proc Natl Acad Sci USA.* 2011;108(6):2328-2333. doi 10.1073/pnas.1018956108

Joseph L., Omland K.E. Phylogeography: its development and impact in Australo-Papuan ornithology with special reference to paraphyly in Australian birds. *Emu-Austral Ornith.* 2009;109(1):1-23. doi 10.1071/MU08024

Kamchatka, Kurils and Komandor Islands. History of landscape development in Siberia and Far East. Moscow: Nauka Publ., 1974 (in Russian)

Kamp L., Pasinelli G., Milanesi P., Drovetski S.V., Kosiński Z., Kossenko S., Robles H., Schweizer M. Significant Asia-Europe divergence in the middle spotted woodpecker (Aves, Picidae). *Zool Scripta.* 2019;48(1):17-32. doi 10.1111/ZSC.12320

Kholodova M.V. Comparative phylogeography: molecular methods, ecological interpretation. *Mol Biol.* 2009;43(5):847-854. doi 10.1134/S002689330905015X]

Kim H.R., Park Y.C. Genetic isolation of Korean populations of *Apodemus peninsulae* (Rodentia: Muridae) from their neighboring populations. *Genes Genomics*. 2015;37:999-1005. doi 10.1007/s13258-015-0331-0

Kim S.-I., Park S.-K., Lee H., Oshida T., Kimura J., Kim Y.-J., Nguyen S.T., Sashika M., Min M.-S. Phylogeography of Korean raccoon dogs: implications of peripheral isolation of a forest mammal in East Asia. *J Zool.* 2013;290(3):225-235. doi 10.1111/jzo.12031

Klicka J., Zink R.M. The importance of recent ice ages in speciation: a failed paradigm. *Science*. 1997;277(5332):1666-1669. doi 10.1126/science.277.5332.1666

Krehenwinkel H., Graze M., Rödder D., Tanaka K., Baba Y.G., Muster C., Uhl G. A phylogeographical survey of a highly dispersive spider reveals eastern Asia as a major glacial refugium for Palaearctic fauna. *J Biogeogr.* 2016;43(8):1583-1594. doi 10.1111/jbi.12742

Krijgsman W. The Mediterranean: *Mare Nostrum* of Earth sciences. *Earth Planet Sci Lett*. 2002;205(1-2):1-12. doi 10.1016/S0012-821X(02)01008-7

Kryukov A.P. Phylogeography and hybridization of corvid birds in the Palearctic region. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov J Genet Breed*. 2019;23(2):232-238. doi 10.18699/VJ19.487

Kryukov A.P., Goroshko O.A. Breeding success of interspecies hybrids: reduced fertility in a hybrid magpie population (*Pica pica* × *Pica serica*, Aves). *Biol Bull Rev.* 2025;15(3):377-384. doi 10.1134/S2079086425700057

Kryukov A.P., Suzuki H. Phylogeography of carrion, hooded and jungle crows (Aves, Corvidae) inferred from partial sequencing of the mitochondrial Cytochrome *b* gene. *Russian Journal of Genetics.* 2000;36(8):922-929

Kryukov A., Iwasa M.A., Kakizawa R., Suzuki H., Pinsker W., Haring E. Synchronic east-west divergence in azure-winged magpies (*Cyanopica cyanus*) and magpies (*Pica pica*). *J Zool Syst Evol Res.* 2004;42:342-351. doi 10.1111/j.1439-0469.2004.00287.x

Kryukov A., Spiridonova L., Mori S., Arkhipov V., Redkin Ya., Goroshko O., Lobkov E., Haring E. Deep phylogeographic breaks in magpie *Pica pica* across the Holarctic: concordance with bioacoustics and phenotypes. *Zool Sci.* 2017;34(3):185-200. doi 10.2108/zs160119

Kryukov A.P., Spiridonova L.N., Tyunin A.P., Kryukov K.A., Dorda B.A. Complete mitochondrial genomes of five subspecies of the Eurasian magpie *Pica pica*, obtained with Oxford Nanopore MinION, and their interpretation regarding intraspecific taxonomy. *Mitochondrial DNA B.* 2020;5(3):3792-3793. doi 10.1080/23802359.2020.1838354

Kryukov A.P., Goroshko O.A., Arkhipov V.Y., Red'kin Y.A., Lee S.I., Dorda B.A., Kryukov K.A., Kapun M., Haring E. Introgression at the emerging secondary contact zone of magpie *Pica pica* subspecies (Aves: Corvidae): integrating data on nuclear and mitochondrial markers, vocalizations and field observations. *Org Divers Evol.* 2022;22:1037-1064. doi 10.1007/s13127-022-00568-6

Kryukov A.P., Kryukov K.A., Collier K., Fang B., Edwards S. Mitogenomics clarifies the position of the Nearctic magpies (*Pica hudsonia* and *Pica nuttalli*) within the Holarctic magpie radiation. *Curr Zool.* 2024;70(5):618-630. doi 10.1093/cz/zoad048

Kvist L., Martens J., Ahola A., Orell M. Phylogeography of a Palaearctic sedentary passerine, the willow tit (*Parus montanus*). *J Evol Biol.* 2001;14(6):930-941. doi 10.1046/j.1420-9101.2001.00354.x

Lee M.Y., Lissovsky A.A., Park S.K., Obolenskaya E.V., Dokuchaev N.E., Zhang Y.P., Yu L., Kim Y.J., Voloshina I., Myslenkov A., Choi T.Y., Min M.-S., Lee H. Mitochondrial cytochrome *b* sequence variation and population structure of Siberian chipmunk (*Tamias sibiricus*) in Northeastern Asia and population subdivision in South Korea. *Mol Cells.* 2008;26(6):566-575. doi 10.1016/S1016-8478(23)25237-1

Lee S., Parr C.S., Hwang Y., Mindell D.P., Choe J.C. Phylogeny of magpies (genus *Pica*) inferred from mtDNA data. *Mol Phylogenet Evol.* 2003;29:250-257. doi 10.1016/s1055-7903(03)00096-4

Li J.J., Shu Q., Zhou S.Z., Zhao Z.J., Zhang J.M. Review and prospects of Quaternary glaciation research in China. *J Glaciol Geocryol.* 2004;26(3):235-243. doi 10.7522/j.issn.1000-0240.2004.0041

Li S.H., Yeung C.K.L., Feinstein J., Han L., Le M.H., Wang C.X., Ding P. Sailing through the Late Pleistocene: unusual historical demography of an East Asian endemic, the Chinese Hwamei (*Leucodioptron canorum canorum*), during the last glacial period. *Mol Ecol.* 2009;18(4):622-633. doi 10.1111/j.1365-294X.2008.04028.x

Li X., Dong F., Lei F., Alström P., Zhang R., Ödeen A., Fjeldså J., Ericson P.G., Zou F., Yang X. Shaped by uneven Pleistocene climate: mitochondrial phylogeographic pattern and population history of white wagtail *Motacilla alba* (Aves: Passeriformes). *J Avian Biol.* 2016;47(2):263-274. doi 10.1111/jav.00826

Liu S., Wei C., Leader P.J., Carey G.J., Jia C., Fu Y., Alström P., Liu Y. Taxonomic revision of the Long-tailed Rosefinch *Carpodacus sibiricus* complex. *J Ornithol.* 2020;161:1061-1070. doi 10.1007/s10336-020-01801-9

Londei T. Association of *Pica* magpies with grazing ungulates: a clue to the genus' origins. *Rivista Italiana Ornitologia.* 2018;87(2):39-42. doi 10.4081/rio.2017.295

Madge S., Christie D.A., Kirwan G.M. Oriental Magpie (*Pica serica*), version 1.0. In: Billerman S.M., Keeney B.K., Rodewald P.G., Schulenberg T.S. (Eds) Birds of the World. Ithaca, NY, USA: Cornell Lab. Ornithology, 2020. doi 10.2173/bow.orimag1.01

Martínez J.G. Urraca – *Pica pica*. In: Salvador A., Morales M.B. (Eds) Enciclopedia Virtual de los Vertebrados Españoles. Madrid: Museo Nacional de Ciencias Naturales, 2016;1-23. doi 10.20350/digital CSIC/8709

Mayr E. Animal Species and Evolution. Cambridge, MA, USA: Belknap Press, 1963

McCormack J.E., Hird S.M., Zellmer A.J., Carstens B.C., Brumfield R.T. Applications of next-generation sequencing to phylogeography and phylogenetics. *Mol Phylogenet Evol.* 2013;66(2):526-538. doi 10.1016/j.ympev.2011.12.007

McKay B.D., Zink R.M. The causes of mitochondrial DNA gene tree paraphyly in birds. *Mol Phylogenet Evol.* 2010;54(2):647-650. doi 10.1016/j.ympev.2009.08.024

Mila B., McCormack J.E., Castaneda G., Wayne R.K., Smith T.B. Recent postglacial range expansion drives the rapid diversification of a songbird lineage in the genus *Junco*. *Proc Biol Sci.* 2007; 274(1626):2653-2660. doi 10.1098/rspb.2007.0852

Miraldo A., Hewitt G.M., Paulo O.S., Emerson B.C. Phylogeography and demographic history of *Lacerta lepida* in the Iberian Peninsula: multiple refugia, range expansions and secondary contact zones. *BMC Evol Biol.* 2011;11:170. doi 10.1186/1471-2148-11-170

Mori S., Hasegawa O., Eguchi K., Hayashi Y., Fujioka M., Kryukov A., Nishiumi I. The origin and trend of common magpie in Japan: microsatellite analysis of old and new introduced populations. *Ornithol Sci.* 2014;13(Suppl.):59. Available at: https://ioc26.ornithology.jp/ioc26_abst-all.pdf

Moura C.C.D.M., Bastian H.V., Bastian A., Wang E., Wang X., Wink M. Pliocene origin, ice ages and postglacial population expansion have influenced a panmictic phylogeography of the European Bee-Eater *Merops apiaster. Diversity*. 2019;11(1):12. doi 10.3390/d11010012

Nazarenko A.A. On faunistic cycles (extinction-expansion-extinction…) with special reference to the East Palearctic dendrophilous avifauna. *Zhurnal Obshchei Biologii = Journal of General Biology.* 1982;43(6):823-835 (in Russian)

Neto J.M., Arroyo J.L., Bargain B., Monros J.S., Matrai N., Prochazka P., Zehtindjiev P. Phylogeography of a habitat specialist with high dispersal capability: the Savi's warbler *Locustella luscinioides*. *PLoS One*. 2012;7(6):e38497. doi 10.1371/journal.pone.0038497

Omland K.E., Baker J.M., Peters J.L. Genetic signatures of intermediate divergence: population history of Old and New World Holarctic ravens (*Corvus corax*). *Mol Ecol.* 2006;15(3):795-808. doi 10.1111/j.1365-294X.2005.02827.x

Ottenburghs J., Lavretsky P., Peters J.L., Kawakami T., Kraus R.H. Population genomics and phylogeography. In: Kraus R.H.S. (Ed.) Avian Genomics in Ecology and Evolution. Springer Nature Switzerland AG, 2019;237-265. doi 10.1007/978-3-030-16477-5_8

Pârâu L.G., Wink M. Common patterns in the molecular phylogeography of western palearctic birds: a comprehensive review. *J Ornithol.* 2021;162(4):937-959. doi 10.1007/s10336-021-01893-x

Pârâu L.G., Frias-Soler R.C., Wink M. High genetic diversity among breeding Red-Backed Shrikes *Lanius collurio* in the Western Palearctic. *Diversity*. 2019;11(3):31. doi 10.3390/d11030031

Pavelková Řičánková V., Robovský J., Riegert J. Ecological structure of recent and last glacial mammalian Faunas in Northern Eurasia: the case of Altai-Sayan refugium. *PLoS One.* 2014;9(1):e85056. doi 10.1371/journal.pone.0085056

Pavlova A., Zink R.M., Rohwer S. Evolutionary history, population genetics, and gene flow in the common rosefinch (*Carpodacus erythrinus*). *Mol Phylogenet Evol.* 2005;36(3):669-681. doi 10.1016/j.ympev.2005.02.010

Pavlova A., Rohwer S., Drovetski S.V., Zink R.M. Different post-Pleistocene histories of Eurasian parids. *J Heredity*. 2006;97(4):389-402. doi 10.1093/jhered/esl011

Perktaş U., Quintero E. A wide geographical survey of mitochondrial DNA variation in the great spotted woodpecker complex, *Dendrocopos major* (Aves: Picidae). *Biol J Linn Soc*. 2013;108:173-188. doi 10.1111/j.1095-8312.2012.02003.x

Peters J.L., Zhuravlev Y., Fefelov I., Logie A., Omland K.E. Nuclear loci and coalescent methods support ancient hybridization as cause of mitochondrial paraphyly between gadwall and falcated duck (*Anas* spp.). *Evolution*. 2007;61(8):1992-2006. doi 10.1111/j.1558-5646.2007.00149.x

Pons J.-M., Olioso G., Cruaud C., Fuchs J. Phylogeography of the Eurasian green woodpecker (*Picus viridis*). *J Biogeogr.* 2011;38(2): 311-325. doi 10.1111/j.1365-2699.2010.02401.x

Pons J.-M., Masson C., Olioso G., Fuchs J. Gene flow and genetic admixture across a secondary contact zone between two divergent lineages of the Eurasian Green Woodpecker *Picus viridis*. *J Ornithol.* 2019;160:935-945. doi 10.1007/s10336-019-01675-6

Poschel J., Heltai B., Gracia E., Quintana M.F., Velo-Antón G., Arribas O., Valdeón A., Wink M., Fritz U., Vamberge M. Complex hybridization patterns in European pond turtles (*Emys orbicularis*) in the Pyrenean Region. *Sci Rep*. 2018;8(1):15925. doi 10.1038/s41598-018-34178-0

Qu Y., Zhang R., Quan Q., Song G., Li S.H., Lei F. Incomplete lineage sorting or secondary admixture: disentangling historical divergence from recent gene flow in the Vinous-throated parrotbill (*Paradoxornis webbianus*). *Mol Ecol*. 2012;21(24):6117-6133. doi 10.1111/mec.12080

Red'kin Y.A., Arkhipov V.Yu., Zhigir D.R. On subspecies taxonomy and nomenclature of Far Eastern subspecies of magpie *Pica pica* Linnaeus, 1758 of the group *"serica"*. *Russian Journal of Ornithology.* 2021;30(2053):1535-1544 (in Russian)

Reinegger R.D., Bhanda G. First sighting of Eurasian magpie (*Pica pica*) in Mauritius. *Bull Phaethon.* 2024;59:51-55

Remington C.L. Suture-zones of hybrid interaction between recently joined biotas. In: Dobzhansky T., Hecht M.K., Steere W.C. (Eds) Evolutionary Biology. Boston, MA: Springer, 1968;321-428. doi 10.1007/978-1-4684-8094-8_8

Rustamov A.K. Family Corvidae. In: Birds of the Soviet Union, Vol. 5. Moscow: Sovetskaya Nauka Publ., 1954;13-104 (in Russian)

Ryall C. Further records and updates of range expansion in House Crow *Corvus splendens*. *Bull Br Ornithol Club*. 2016;136(1):39-45

Saitoh T., Alström P., Nishiumi I., Shigeta Y., Williams D., Olsson U., Ueda K. Old divergences in a boreal bird supports long-term survival through the Ice Ages. *BMC Evol Biol.* 2010;10:35. doi 10.1186/1471-2148-10-35

Sakka H., Quere J.P., Kartavtseva I., Pavlenko M., Chelomina G., Atopkin D., Bogdanov A., Michaux J. Comparative phylogeography of four *Apodemus* species (Mammalia: Rodentia) in the Asian Far East: evidence of Quaternary climatic changes in their genetic structure. *Biol J Linn Soc.* 2010;100(4):797-821. doi 10.1111/j.1095-8312.2010.01477.x

Salinas P., Morinha F., Literak I., García J., Milá B., Blanco G. Genetic diversity, differentiation and historical origin of the isolated population of rooks *Corvus frugilegus* in Iberia. *J Avian Biol.* 2021; 52(3):e02689. doi 10.1111/jav.02689

Saunders M.A., Edwards S.V. Dynamics and phylogenetic implications of MtDNA control region sequences in New World Jays (Aves: Corvidae). *J Mol Evol.* 2000;51(2):97-109. doi 10.1007/s002390010070

Serizawa K., Suzuki H., Iwasa M., Tsuchiya K., Pavlenko M.V., Kartavtseva I.V., Chelomina G.N., Dokuchaev N.E., Han S.-H. A spatial aspect of mitochondrial DNA genealogy in *Apodemus peninsulae* from East Asia. *Biochem Genet.* 2002;40(5-6):149-161. doi 10.1023/a:1015841424598

Song G., Zhang R., DuBay S.G., Qu Y., Dong L., Wang W., Zhang Y., Lambert D.M., Lei F. East Asian allopatry and north Eurasian sympatry in Long-tailed Tit lineages despite similar population dynamics during the late Pleistocene. *Zool Scripta.* 2016;45(2):115-126. doi 10.1111/zsc.12148

Song G., Zhang R., Alström P., Irestedt M., Cai T., Qu Y., Ericson P.G.P., Fjeldså J., Lei F. Complete taxon sampling of the avian genus *Pica* (magpies) reveals ancient relictual populations and synchronous Late-Pleistocene demographic expansion across the Northern Hemisphere. *J Avian Biol.* 2018;49(2):e01612. doi 10.1111/jav.01612

Song G., Zhang R., Machado-Stredel F. Great journey of Great Tits (*Parus major* group): origin, diversification and historical demographics of a broadly distributed bird lineage. *J Biogeogr*. 2020;47: 1585-1598. doi 10.1111/jbi.13863

Stegmann B.K. Corvid Birds. Leningrad: USSR Acad. Sci. Publ., 1932 (in Russian)

Taberlet P., Fumagalli L., Wust-Saucy A.G., Cosson J.F. Comparative phylogeography and postglacial colonization routes in Europe. *Mol Ecol.* 1998;7(4):453-464. doi 10.1046/j.1365-294x.1998.00289.x

Toews D.P., Brelsford A. The biogeography of mitochondrial and nuclear discordance in animals. *Mol Ecol.* 2012;21(16):3907-3930. doi 10.1111/j.1365-294X.2012.05664.x

Toews D.P., Irwin D.E. Cryptic speciation in a Holarctic passerine revealed by genetic and bioacoustic analyses. *Mol Ecol.* 2008;17(11): 2691-2705. doi 10.1111/j.1365-294X.2008.03769.x

Wang E., Van Wijk R.E., Braun M.S., Wink M. Gene flow and genetic drift contribute to high genetic diversity with low phylogeographical structure in European hoopoes (*Upupa epops*). *Mol Phylogenet Evol.* 2017;113:113-125. doi 10.1016/j.ympev.2017.05.018

Webb W.C., Marzluff J.M., Omland K.E. Random interbreeding between cryptic lineages of the Common Raven: evidence for speciation in reverse. *Mol Ecol.* 2011;20(11):2390-2402. doi 10.1111/j.1365-294X.2011.05095.x

Winkler D.W., Billerman S.M., Lovette I.J. Crows, Jays, and Magpies (Corvidae), version 1.0. In: Billerman S.M., Keeney B.K., Rodewald P.G., Schulenberg T.S. (Eds) Birds of the World. Ithaca, NY, USA: Cornell Lab. of Ornithology, 2020. doi 10.2173/bow.corvid1.01

Yi S., Kim S.J. Vegetation changes in western central region of Korean Peninsula during the last glacial (ca. 21.1–26.1 cal kyr BP). *Geosciences J*. 2010;14:1-10. doi 10.1007/s12303-010-0001-9

Zhang H., Yan J., Zhang G., Zhou K. Phylogeography and demographic history of Chinese black-spotted frog populations (*Pelophylax nigromaculata*): evidence for independent refugia expansion and secondary contact. *BMC Evol Biol.* 2008;8:21. doi 10.1186/1471-2148-8-21

Zhang R., Song G., Qu Y., Alstrom P., Ramos R., Xing X., Ericson P., Fjeldsa J., Wang H., Yang X., Krustin A., Shestopalov A., Choe L.C., Fumin L. Comparative phylogeography of two widespread magpies: importance of habitat preference and breeding behavior on genetic structure in China. *Mol Phylogenet Evol.* 2012;65(2):562-572. doi 10.1016/j.ympev.2012.07.011

Zhao N., Dai C., Wang W., Zhang R., Qu Y., Song G., Chen K., Yang X., Zou F., Lei F. Pleistocene climate changes shaped the divergence and demography of Asian populations of the great tit *Parus major*: evidence from phylogeographic analysis and ecological niche models. *J Avian Biol.* 2012;43(4):297-310. doi 10.1111/j.1600-048X.2012.05474.x

Zink R.M. Comparative phylogeography in North American birds. *Evolution.* 1996;50(1):308-317. doi 10.1111/j.1558-5646.1996.tb04494.x

Zink R.M., Barrowclough G.F. Mitochondrial DNA under siege in avian phylogeography. *Mol Ecol.* 2008;17:2107-2121. doi 10.1111/j.1365-294X.2008.03737.x

Zink R.M., Pavlova A., Drovetski S., Rohwer S. Mitochondrial phylogeographies of five widespread Eurasian bird species. *J Ornithology.* 2008;149:399-413. doi 10.1007/s10336-008-0276-z

Zink R.M., Pavlova A., Drovetski S., Wink M., Rohwer S. Taxonomic status and evolutionary history of the *Saxicola torquata* complex. *Mol Phylogenet Evol.* 2009;52(3):769-773. doi 10.1016/j.ympev.2009.05.016

Conflict of interest. The author declares no conflict of interest.
Received December 13, 2024. Revised February 4, 2025. Accepted February 5, 2025.

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2025;29(4):594-599

doi 10.18699/vjgb-25-62

Genetic potential for biofilm formation of clinical strains of *Pseudomonas aeruginosa*

U.M. Nemchenko, N.L. Belkova, E.S. Klimenko, N.E. Smurova, R.E. Zugeeva, V.V. Sinkov, E.D. Savilov

Institute of Epidemiology and Microbiology, Scientific Center for Family Health and Human Reproduction Problems, Irkutsk, Russia
nlbelkova@gmail.com

Abstract. *Pseudomonas aeruginosa* is one of the leading causes of nosocomial respiratory tract infections and plays an important role in lower respiratory tract infection in patients with cystic fibrosis (CF). Biofilms, which are organized cell clusters, ensure the survival of microorganisms in unfavorable environmental conditions and contribute to the chronicity of infection and the formation of persistent forms. The aim of this study was to determine the phenotypic ability and genetic potential for biofilm formation in clinical strains of *P. aeruginosa* persisting in patients with CF against the background of constant intake of antimicrobial drugs. Bacteriological, genetic, and bioinformatic methods were used to characterize five *P. aeruginosa* strains obtained from patients with CF. Phenotypically, all strains were classified as moderately biofilm-forming, while the biofilm formation coefficient varied from 2.10 to 3.15. Analysis of draft genomes revealed differences in the representation of some genes or individual loci of three of the four known signaling pathways (cAMP/Vfr, Gac/Rsm, and c-di-GMP) that have been described in *P. aeruginosa* genomes and are related to the regulation of biofilm formation. In addition, differences in the representation of genes such as *frzE*, *tcpE*, and *rcsC* are shown. Of undoubted interest is the analysis of genes such as *pppA*, *icmF*, *clpV1*, *trpE*, *trpG*, and *stp1*, which are used for extended multilocus typing PubMLST and differed in the structure of loci in all analyzed strains. These genes can be used to identify clinical strains of *P. aeruginosa* and to characterize their biofilm-forming properties. Thus, genes potentially participating in both biofilm formation and regulation have been characterized in the genomes of clinical *P. aeruginosa* strains that persist for a long time in patients receiving continuous antibiotic therapy. Characterization of the genetic potential for biofilm formation makes it possible to search for reliable genetic markers of this process in order to monitor the evolution of the pathogen as a result of long-term persistence in the host organism.

Key words: *Pseudomonas aeruginosa*; cystic fibrosis; biofilms; whole genome sequencing; signaling pathways

For citation: Nemchenko U.M., Belkova N.L., Klimenko E.S., Smurova N.E., Zugeeva R.E., Sinkov V.V., Savilov E.D. Genetic potential for biofilm formation of clinical strains of *Pseudomonas aeruginosa*. *Vavilovskii Zhurnal Genetiki i Selektsii* = *Vavilov J Genet Breed*. 2025;29(4):594-599. doi 10.18699/vjgb-25-62

Funding. This work was carried out within the framework of technology No. 123051600027-1.

Генетический потенциал к образованию биопленок клинических штаммов *Pseudomonas aeruginosa*

У.М. Немченко, Н.Л. Белькова, Е.С. Клименко, Н.Е. Смурова, Р.Е. Зугеева, В.В. Синьков, Е.Д. Савилов

Институт эпидемиологии и микробиологии, Научный центр проблем здоровья семьи и репродукции человека, Иркутск, Россия
nlbelkova@gmail.com

Аннотация. Бактерии вида *Pseudomonas aeruginosa* являются одной из ведущих причин нозокомиальных инфекций дыхательного тракта и играют важную роль в инфицировании нижних дыхательных путей у больных муковисцидозом (МВ). Биопленки, представляя собой организованные скопления клеток, обеспечивают выживаемость микроорганизмов в неблагоприятных условиях окружающей среды и способствуют хронизации инфекции и формированию персистирующих форм. Целью работы стало определение фенотипической способности и генетического потенциала к биопленкообразованию у клинических штаммов *P. aeruginosa*, персистирующих у пациентов с МВ на фоне постоянного приема антимикробных препаратов. Для характеристики пяти штаммов *P. aeruginosa*, полученных от пациентов с МВ, в работе применены бактериологические, молекулярно-генетические и биоинформатические методы. Фенотипически все штаммы отнесены к умеренно-образующим биопленку, при этом коэффициент биопленкообразования варьировал от 2.10 до 3.15. Анализ драфт-геномов выявил различия в представленности некоторых генов или отдельных локусов трех из четырех известных сигнальных путей, цАМФ/Vfr, Gac/Rsm и c-di-GMP, которые описаны в геномах *P. aeruginosa* и имеют отношение к регуляции образования биопленок. Дополнительно показаны отличия в представленности таких генов, как *frzE*, *tcpE* и *rcsC*. Несомненно интересен анализ генов *pppA*, *icmF*, *clpV1*, *trpE*, *trpG* и *stp1*, которые используются для

© Nemchenko U.M., Belkova N.L., Klimenko E.S., Smurova N.E., Zugeeva R.E., Sinkov V.V., Savilov E.D., 2025
This work is licensed under a Creative Commons Attribution 4.0 License

расширенного мультилокусного типирования PubMLST и различаются по структуре локусов у всех проанализированных штаммов. Эти гены потенциально могут быть применены для типирования клинических штаммов *P. aeruginosa* с целью характеристики их биопленкообразующих свойств. Таким образом, в геномах клинических штаммов *P. aeruginosa*, длительно персистирующих у пациентов на фоне постоянного получения антибиотикотерапии, охарактеризованы гены, которые потенциально могут участвовать как в процессе биопленкообразования, так и в его регуляции. Характеристика генетического потенциала к образованию биопленок дает возможность поиска надежных генетических маркеров этого процесса для мониторинга эволюции возбудителя в результате длительной персистенции в организме хозяина.

Ключевые слова: *Pseudomonas aeruginosa*; муковисцидоз; биопленки; полногеномное секвенирование; сигнальные пути

Introduction

Pseudomonas aeruginosa is one of the leading causes of nosocomial respiratory tract infections and plays an important role in lower respiratory tract infections in patients with cystic fibrosis (CF) (Parkins et al., 2018; Shaginyan et al., 2019). Surface colonization and subsequent biofilm formation and development provide numerous advantages for infectious agents. Biofilms, which are organized clusters of cells enclosed in a polysaccharide matrix and protected from adverse environmental conditions, including antimicrobials, disinfectants, and antiseptics, ensure microorganism survival. Such structural organization contributes to the increased heterogeneity of the bacterial population and the selection of cells that counteract damaging effects by acquiring and accumulating genetic mutations. Therefore, biofilm-forming microorganisms significantly contribute to the chronicity of infection and the formation of persistent forms (Penesyan et al., 2021).

Considering the pathogenetic potential of *P. aeruginosa* in the biofilm state in patients with CF, it should be noted that bacterial cells are tolerant not only to antibiotics, disinfectants, and antiseptics, but also to factors of the innate and adaptive defense system of the body (Jurado-Martín et al., 2021; Fernández-Billón et al., 2023). In response to anaerobic conditions, competition for resources, high concentrations of antibiotics, and immune responses of the body, such as neutrophil attacks in the lungs in CF, *P. aeruginosa* undergoes microevolution, acquiring spontaneous mutations that lead to the selection of cells that better survive long-term colonization (Winstanley et al., 2016; Jurado-Martín et al., 2021). Experimental studies have shown the ability of *P. aeruginosa* to form biofilm structures in the sputum of patients with CF (Bjarnsholt et al., 2009), which is a decisive factor for survival and tolerance to antibiotics, and the inability to completely eliminate bacteria is directly associated with the chronicity of the infection (Elfadadny et al., 2024). Understanding the mechanisms of adaptation and evolution of the pathogen during chronic respiratory infections in patients with CF may help discover new treatment methods for *P. aeruginosa* infections.

Modern sequencing technologies allow us to analyze the complete genomes of opportunistic microorganisms not only for their typing but also for identifying molecular markers that are potentially significant for the infectious agent. Currently, a map of the main signaling pathways characterized in the *P. aeruginosa* genomes and related to the regulation of biofilm formation has been created on the KEGG PATHWAY Database platform (PATHWAY: ko02025; https://www.genome.jp/kegg-bin/show_pathway?ko02025). The ko02025 map contains information on 90 loci included in four main signaling pathways: the cAMP/Vfr pathway, the quorum sensing (QS) system, the Gac/Rsm pathway, and the c-di-GMP signaling pathway.

The aim of this study was to determine the phenotypic ability and genetic potential for biofilm formation in clinical strains of *P. aeruginosa* persisting in patients with CF against the background of constant use of antimicrobial drugs.

Materials and methods

The objects of the study were five clinical strains of *P. aeruginosa* from the working collection of the microbiome and microecology laboratory of the Institute of Epidemiology and Microbiology, Scientific Center for Family Health and Human Reproduction Problems. The strains were isolated from the sputum of patients with CF who were treated at Ivano-Matreninskaya City Children's Clinical Hospital (Irkutsk, Russia) and who received long-term antibiotic therapy. The strains were identified using morpho-biochemical tests and were confirmed using mass spectrometry (Nemchenko et al., 2022). Sensitivity to AMPs was determined according to EUCAST criteria (Nemchenko et al., 2024).

The ability of the strains to form biofilms (BF) was studied using the G.A. O'Toole plate method (O'Toole, 2011), with our own modifications (Nemchenko et al., 2020; Sitnikova et al., 2022). Briefly, the ability of cultures to form BFs was determined using a 96-well sterile flat-bottomed plastic immunological plate. Daily bacterial culture was standardized in sterile meat-peptone broth (MPB) to 1×10^6 CFU/ml. The culture suspension and control were inoculated with 150 µl per well of the plate into four replicates. Sterile MPB served as the background control. The plate was incubated in a dry-air thermostat for 18–20 h at 37 °C. Biofilms were stained using a modified G.A. O'Toole method: planktonic cells were removed by pipetting, the plate was washed three times with sterile saline, dried for 10–15 min without a lid, and the biofilms were stained with 1 % gentian violet, followed by alcohol extraction according to the method (O'Toole, 2011). The biomass of the formed films was estimated from the optical density (OD) of the gentian violet dye extracts at 492 nm (STAT FAX®4300 spectrophotometer, USA). The biofilm formation coefficient (BFC) was calculated as the A492exp/A492control ratio. Strains with BFC values ≤ 2 units were considered to be weakly BF-forming; those with a BFC of 2–3.99 had a moderate ability to form BFs (Nemchenko et al., 2020; Grigorova et al., 2021).

Whole-genome sequencing. Genomic DNA was isolated using a Quick-DNA Fungal/Bacterial Miniprep Kit (Zymo Research, USA). Whole-genome sequencing of strains was performed on Illumina NextSeq 550 equipment using the

Table 1. Brief results of genome assembly and annotation of *P. aeruginosa* strains

Indicator	IMB101	IMB105	IMB103	IMB100	IMB104
Genome assembly					
Number of reads	12,369,102	13,477,896	3,737,929	36,044,976	14,328,761
Number of contigs	121	274	278	258	487
N50	255,738	298,770	550,311	65,489	238,191
Coverage	279	686	592	1,500	179
Genome annotation					
GC, %	66	66	66	66	65
CDS number	5,895	5,874	5,910	5,548	5,935

Illumina® DNA Prep Tagmentation, IDT® for Illumina® DNA/RNA UD Indexes Set Tagmentation, and NextSeq 500/550 High Output Kit v.2.5 (300 Cycles) library preparation reagent kits according to the manufacturer's recommendations.

Bioinformatics analysis. Genome assembly was performed using the SPAde sv.3.11.1 program (Bankevich et al., 2012). Contig alignment and orientation correction were performed using MAUVE 2.4.0 (Rissman et al., 2009) and the *P. aeruginosa* PAO1 reference genome (GenBank AE004091.2) (Table 1). Functional annotation was performed using Prokka 1.14.6 (Seemann, 2014). The genes involved in biofilm formation were identified using the KEGG (Kanehisa et al., 2022) and PubMLST (Jolley et al., 2018) databases. The complete genomes of the strains obtained in this study were deposited in the NCBI database project PRJNA1026796.

The study was conducted using equipment from the Center for collective use "Center for the development of progressive personalized health technologies", and "Collection of human microbiota of the Irkutsk region" of the Institute of Epidemiology and Microbiology, Scientific Center for Family Health and Human Reproduction Problems (Irkutsk).

Results

Five clinical strains of *P. aeruginosa* isolated from patients with CF who had received long-term antibiotic therapy were analyzed in the study. Phenotypically, all strains were classified as moderately biofilm-forming, with the BFC varying from 2.10 to 3.15 (Table 2). The minimum BFC value was determined for strain IMB101, and the maximum values were 3.11 and 3.15 for IMB100 and IMB104, respectively.

Table 2. Phenotypic characteristics and genetic potential of *P. aeruginosa* strains for biofilm formation

Indicator	IMB101	IMB105	IMB103	IMB100	IMB104
Biofilm formation coefficient	2.10	2.50	2.91	3.11	3.15
cAMP/Vfr signaling pathway					
pil; twitching motility proteins	*pilT/Y/Q*	*pilT/Y/Q/A*	*pilT/Y/Q/A*	*pilT/Y/Q*	*pilT/Y/Q/A*
cyaA; adenylate cyclase	–	–	+	–	–
Gac/Rsm pathway					
hcpA; secreted protein Hcp	–	–	–	1–3	–
hcpC; secreted protein Hcp	1	–	–	–	–
c-di-GMP signaling pathway					
wbp; phosphomannose isomerase/ mannose-1-phosphate guanylyl transferase	*wbpABDEI*	*wbpA*	*wbpA*	–	–
Additional genes potentially involved in biofilm formation not included in the ko02025 map					
frzE; gliding motility regulatory protein	–	–	–	+	–
tcpE; toxin coregulated pilus biosynthesis protein E	+	–	+	–	+
rcsC; sensor histidine kinase RcsC	1–15	1–15	1–15	1–13	1–14

Note. Numbers indicate the number of variants of the *hcpA*, *hcpC*, and *rcsC* genes according to genome annotation using Prokka 1.14.6 (Seemann, 2014).

Table 3. Description of loci identified in the genomes of the studied *P. aeruginosa* strains for genes used for extended MLST typing (PubMLST) and included in the ko02025 map

Gene; synthesized protein	IMB101	IMB105	IMB103	IMB100	IMB104
Gac/Rsm pathway					
hcp1, protein secretion apparatus assembly protein	1	1	1	1	1
pppA, serine/threonine phosphatase	9	4	3	4	16
icmF1, type VI secretion protein IcmF	72	5	19	120	39
clpV1, secretion protein ClpV1	67	176	22	113	120
ppkA, serine/threonine protein kinase PpkA	44	187	113	–	138
fha1, Fha domain-containing protein	127	6	26	–	326
gacA, response regulator GacA	3	3	3	3	3
gacS, sensor/response regulator hybrid protein	59	487	18	–	24
retS, sensor histidine kinase MifS	59	–	–	108	154
QS pathway					
trpE, anthranilate synthase component I	58	5	14	5	116
trpG, anthranilate synthase component II	1	48	12	27	72
stp1, serine/threonine phosphoprotein phosphatase Stp1	34	82	4	13	70

Whole-genome sequencing was performed for all strains. The primary objective of this study was to identify genes that could participate in biofilm formation and its regulation. The ko02025 map (KEGG PATHWAY Database) was used for routine search. Additionally, we analyzed the loci used in typing *P. aeruginosa* strains on the PubMLST platform (Jolley et al., 2018) and manually searched for genes that had previously been shown to participate in the process of biofilm formation or its regulation (Thelin, Taylor, 1996; Kearns, Shimkets, 1998; Wall et al., 2018).

We showed that all tested genes are localized in chromosomal DNA. Differences between the genomes of the studied strains were found in the presence or absence of some genes or in the representation of loci of three signaling pathways: cAMP/Vfr, Gac/Rsm, and c-di-GMP (Table 2). An additional search for genes involved in biofilm formation revealed differences in the presence or absence of genes primarily involved in regulatory processes: *frzE* (gliding motility regulatory protein), *tcpE* (toxin coregulated pilus biosynthesis protein E), and *rcsC* (sensor histidine kinase RcsC) (Table 2). Of greatest interest are genes that not only participate in signaling pathways according to the ko02025 map (KEGG PATHWAY Database) but are also used for extended multilocus typing PubMLST (Table 3).

The *hcp1* and *gacA* genes were completely identical in the genomes of the studied strains. Genes such as *ppkA*, *fha1*, and *gacS* were not identified in the IMB100 strain, which showed a fairly high biofilm formation coefficient (3.11). Note that *retS* was absent in the genomes of the IMB103 and IMB105 strains, which had biofilm formation coefficients of 2.91 and 2.50, respectively. Of undoubted interest is the analysis of genes such as *pppA*, *icmF*, *clpV1*, *trpE*, *trpG*, and *stp1*, which differed in the structure of loci in all the analyzed strains. These genes can be used to identify clinical strains of *P. aeruginosa* and to characterize their biofilm-forming properties.

Discussion

P. aeruginosa is an opportunistic pathogen that causes infections in immunocompromised or CF patients. *P. aeruginosa* infection in CF patients occurs as a mild acute infection that subsequently progresses to chronic respiratory disease. It has been suggested that two distinct, mutually exclusive sets of virulence factors are associated with the two stages of infection (Brencic, Lory, 2009). The type III secretion system (T3SS) and type IV pili genes are thought to be associated with acute disease, whereas the type VI secretion system (T6SS) HSI-I and biofilm formation are important during chronic infection (Deretic et al., 1995; Brencic, Lory, 2009). There is currently an active search for genes, the expression of which differs according to the lifestyles of bacterial pathogens and which may be biomarkers of the transition from acute to chronic infection and vice versa (Cao et al., 2023).

In this study, we analyzed the phenotypic and genetic properties of *P. aeruginosa* strains isolated from patients with CF receiving continuous antibacterial therapy. All strains were defined as moderate biofilm-forming, but their genomes showed differences in the presence/absence of some genes or in the loci of signaling pathways that were characterized in the genomes of clinical *P. aeruginosa* strains and were related to both biofilm formation and regulation of this process. The identification of genetic markers of phenotypes associated with biofilm structure formation is of undoubted interest.

These genes may be used for extended multilocus typing PubMLST and may be responsible for individual stages of biofilm formation or regulation. In our studies, we identified 12 such genes, 9 of which belong to the Gac/Rsm signaling pathway, and 3 – to the QS pathway.

The two-component GacS/GacA system stimulates the expression of two small regulatory RNAs, RsmY and RsmZ, which in turn regulate the translational repressor RsmA. Members of the RsmA/CsrA family have been identified in the genomes of many Gram-negative bacteria, including *P. aeruginosa*, *P. fluorescens*, *Escherichia coli*, and some species of *Salmonella*, *Legionella*, *Proteus*, *Helicobacter*, and *Erwinia*, where they have been implicated in the regulation of phenotypes such as virulence, motility, QS systems, and stress response (Brencic, Lory, 2009).

QS systems are a form of bacterial intercellular communication used by many species to determine population density and coordinate gene expression (Coggan, Wolfgang, 2012). QS is achieved by producing autoinducer signaling molecules so that an increase in bacterial population density leads to their accumulation. Once a threshold concentration is reached, autoinducers bind their cognate receptors, which directly or indirectly activate gene expression. Three QS systems are encoded in the genomes of *P. aeruginosa*: two N-acyl-homoserine lactone (AHL)-based and a 2-alkyl-4-quinolone (AQ)-based signaling system. These three QS systems are involved in the regulation of virulence factor production, biofilm maturation, and motility phenotypes (Coggan, Wolfgang, 2012).

It should be noted that studies of transcriptional profiles of different clinical *P. aeruginosa* strains grown under planktonic and biofilm conditions showed that transcriptional profiles detected under planktonic growth conditions were quite similar, and more divergent transcriptional profiles were recorded when isolates were grown under biofilm conditions (Thöming et al., 2020). The model experiments showed that different groups of clinical isolates follow parallel evolutionary pathways and produce similar phenotypes. This convergence of organismal phenotypes was observed for a variety of traits, including the formation of different biofilm structures characterized by specific transcriptional signatures, as well as virulence and motility phenotypes (Thöming et al., 2020).

It can be assumed that, despite the different sequence types identified in patients with CF, the transition to a persistent form during chronic *P. aeruginosa* infection will not simply stimulate the expression of certain genes to create a certain pathogen phenotype but will also form the corresponding genotype, realizing the potential of genetic heterogeneity of the bacterial population. It should also be noted that genes that participate in biofilm formation or regulation of this process (according to the ko02025 map) are used for extended multilocus typing of PubMLST and can be used to type clinical strains of *P. aeruginosa* in order to characterize their biofilm-forming properties.

Conclusion

In the genomes of clinical strains of *P. aeruginosa* that persist for a long time in patients with CF against the background of constant antibiotic therapy, genes that can potentially participate both in the process of biofilm formation and in its regulation have been characterized. Characterization of the genetic potential for biofilm formation makes it possible to search for reliable genetic markers of this process to monitor the evolution of the pathogen as a result of long-term persistence in the host organism.

References

Bankevich A., Nurk S., Antipov D., Gurevich A.A., Dvorkin M., Kulikov A.S., Lesin V.M., … Sirotkin A.V., Vyahhi N., Tesler G., Alekseyev M.A., Pevzner P.A. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol*. 2012;19(5):455-477. doi 10.1089/cmb.2012.0021

Bjarnsholt T., Jensen P.Ø., Fiandaca M.J., Pedersen J., Hansen C.R., Andersen C.B., Pressler T., Givskov M., Høiby N. *Pseudomonas aeruginosa* biofilms in the respiratory tract of cystic fibrosis patients. *Pediatr Pulmonol*. 2009;44(6):547-558. doi 10.1002/ppul.21011

Brencic A., Lory S. Determination of the regulon and identification of novel mRNA targets of *Pseudomonas aeruginosa* RsmA. *Mol Microbiol*. 2009;72(3):612-632. doi 10.1111/j.1365-2958.2009.06670.x

Cao P., Fleming D., Moustafa D.A., Dolan S.K., Szymanik K.H., Redman W.K., Ramos A., Diggle F.L., Sullivan C.S., Goldberg J.B., Rumbaugh K.P., Whiteley M. A *Pseudomonas aeruginosa* small RNA regulates chronic and acute infection. *Nature*. 2023;618(7964):358-364. doi 10.1038/s41586-023-06111-7

Coggan K.A., Wolfgang M.C. Global regulatory pathways and cross-talk control *Pseudomonas aeruginosa* environmental lifestyle and virulence phenotype. *Curr Issues Mol Biol*. 2012;14(2):47-70. doi 10.21775/cimb.014.047

Deretic V., Schurr M.J., Yu H. *Pseudomonas aeruginosa*, mucoidy and the chronic infection phenotype in cystic fibrosis. *Trends Microbiol*. 1995;3(9):351-356. doi 10.1016/s0966-842x(00)88974-x

Elfadadny A., Ragab R.F., AlHarbi M., Badshah F., Ibáñez-Arancibia E., Farag A., Hendawy A.O., De Los Ríos-Escalante P.R., Aboubakr M., Zakai S.A., Nageeb W.M. Antimicrobial resistance of *Pseudomonas aeruginosa*: navigating clinical impacts, current resistance trends, and innovations in breaking therapies. *Front Microbiol*. 2024;15:1374466. doi 10.3389/fmicb.2024.1374466

Fernández-Billón M., Llambías-Cabot A.E., Jordana-Lluch E., Oliver A., Macià M.D. Mechanisms of antibiotic resistance in *Pseudomonas aeruginosa* biofilms. *Biofilm*. 2023;5:100129. doi 10.1016/j.bioflm.2023.100129

Grigorova E.V., Nemchenko U.M., Voropaeva N.M., Belkova N.L., Noskova O.A., Savilov E.D. Effect of disinfectants with different active ingredients on biofilm formation in *Pseudomonas aeruginosa*. *Bull Exp Biol Med*. 2021;171(6):745-749. doi 10.1007/s10517-021-05308-y

Jolley K.A., Bray J.E., Maiden M.C.J. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. *Wellcome Open Res*. 2018;3:124. doi 10.12688/wellcomeopenres.14826.1

Jurado-Martín I., Sainz-Mejías M., McClean S. *Pseudomonas aeruginosa*: an audacious pathogen with an adaptable arsenal of virulence factors. *Int J Mol Sci*. 2021;22(6):3128. doi 10.3390/ijms22063128

Kanehisa M., Sato Y., Kawashima M. KEGG mapping tools for uncovering hidden features in biological data. *Protein Sci*. 2022;31(1):47-53. doi 10.1002/pro.4172

Kearns D.B., Shimkets L.J. Chemotaxis in a gliding bacterium. *Proc Natl Acad Sci USA*. 1998;95(20):11957-11962. doi 10.1073/pnas.95.20.11957

Nemchenko U.M., Kungurtseva E.A., Grigorova E.V., Belkova N.L., Markova Yu.A., Noskova O.A., Chemezova N.N., Savilov E.D. Simulation of bacterial biofilms and estimation of the sensitivity of healthcare-associated infection pathogens to bactericide Sekusept active. *Klinicheskaya Laboratornaya Diagnostika = Russian Clini-*

cal Laboratory Diagnostics. 2020;65(10):652-658. doi 10.18821/0869-2084-2020-65-10-652-658 (in Russian)

Nemchenko U.M., Sitnikova K.O., Belkova N.L., Grigorova E.V., Voropaeva N.M., Sukhoreva M.V., Sukhareva E.S., Savilov E.D. Effects of antimicrobials on *Pseudomonas aeruginosa* biofilm formation. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov J. Genet. Breed*. 2022;26(5):495-501. doi 10.18699/VJGB-22-60

Nemchenko U.M., Belkova N.L., Klimenko E.S., Smurova N.E., Savilov E.D. Phenotypic and genetic variability of clinical isolates of *Pseudomonas aeruginosa* persisting in patients with cystic fibrosis. *Kliniceskaa Mikrobiologia i Antimikrobnaa Himioterapia = Clinical Microbiology and Antimicrobial Chemotherapy*. 2024;26(S1):42 (in Russian)

O'Toole G.A. Microtiter dish biofilm formation assay. *J Vis Exp*. 2011; 30(47):2437. doi 10.3791/2437

Parkins M.D., Somayaji R., Waters V.J. Epidemiology, biology, and impact of clonal *Pseudomonas aeruginosa* infections in cystic fibrosis. *Clin Microbiol Rev*. 2018;31(4):e00019-18. doi 10.1128/cmr.00019-18

Penesyan A., Paulsen I.T., Kjelleberg S. Three faces of biofilms: a microbial lifestyle, a nascent multicellular organism, and an incubator for diversity. *NPJ Biofilms Microbiomes*. 2021;7(1):80. doi 10.1038/s41522-021-00251-2

Rissman A.I., Mau B., Biehl B.S., Darling A.E., Glasner J.D., Perna N.T. Reordering contigs of draft genomes using the Mauve Aligner. *Bioinformatics*. 2009;25(16):2071-2073. doi 10.1093/bioinformatics/btp356

Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*. 2014;30(14):2068-2069. doi 10.1093/bioinformatics/btu153

Shaginyan I.A., Avetisyan L.R., Chernukha M.Yu., Siyanova E.A., Burmistrov E.M., Voronkova A.Yu., Kondratieva E.I., Chuchalin A.G., Gintzburg A.L. Epidemiological significance of genome variations in *Pseudomonas aeruginosa* causing chronic lung infection in patients with cystic fibrosis. *Kliniceskaya Mikrobiologiya i Antimikrobnaya Himioterapiya = Clinical Microbiology and Antimicrobial Chemotherapy*. 2019;21(4):340-351. doi 10.36488/cmac.2019.4.340-351 (in Russian)

Sitnikova K.O., Nemchenko U.M., Voropaeva N.M., Grigorova E.V., Savilov E.D., Markova Yu.A., Belkova N.L. The effectiveness of biofilm formation of daily cultures of clinically significant strains of opportunistic bacteria. *Acta Biomedica Scientifica*. 2022;7(5-1): 119-128. doi 10.29413/ABS.2022-7.5-1.13 (in Russian)

Thelin K.H., Taylor R.K. Toxin-coregulated pilus, but not mannose-sensitive hemagglutinin, is required for colonization by *Vibrio cholerae* O1 El Tor biotype and O139 strains. *Infect Immun*. 1996; 64(7):2853-2856. doi 10.1128/iai.64.7.2853-2856.1996

Thöming J.G., Tomasch J., Preusse M., Koska M., Grahl N., Pohl S., Willger S.D., Kaever V., Müsken M., Häussler S. Parallel evolutionary paths to produce more than one *Pseudomonas aeruginosa* biofilm phenotype. *NPJ Biofilms Microbiomes*. 2020;6:2. doi 10.1038/s41522-019-0113-6

Wall E., Majdalani N., Gottesman S. The complex Rcs regulatory cascade. *Annu Rev Microbiol*. 2018;72:111-139. doi 10.1146/annurev-micro-090817-062640

Winstanley C., O'Brien S., Brockhurst M.A. *Pseudomonas aeruginosa* evolutionary adaptation and diversification in cystic fibrosis chronic lung infections. *Trends Microbiol*. 2016;24:327-337. doi 10.1016/j.tim.2016.01.008

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

Received November 8, 2024. Revised January 17, 2025. Accepted February 14, 2025.

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2025;29(4):600-607

doi 10.18699/vjgb-25-63

Env-pseudoviruses based on the HIV-1 genetic variant circulating in Siberia

N.B. Rudometova, A.A. Fando, D.N. Shcherbakov, B.N. Zaitsev, A.P. Rudometov, L.I. Karpenko

State Research Center Center of Virology and Biotechnology "Vector" Rospotrebnadzor, Koltsovo, Novosibirsk region, Russia
nadenkaand100@mail.ru

Abstract. Despite numerous efforts of the global community, it is still not possible to stop the HIV/AIDS pandemic. To stop the spread of the virus, an effective preventive vaccine is needed, as well as the search for new antiviral agents. In order to be able to quickly and adequately evaluate the developed vaccine constructs, characterize HIV-specific antibodies and potential drugs, a reliable testing method is needed. In this regard, pseudotype neutralization assays using a panel of Env-pseudoviruses of different HIV-1 subtypes has proven itself well. Currently, separate panels of Env-pseudoviruses of the main genetic subtypes of HIV-1 (A, B, C and a number CRFs) have been created. These panels are necessary to obtain standardized data sets that can be used to rank the effectiveness of the vaccine and identify promising candidates for further study. Currently, the HIV-1 subtype A6 dominates in the European part of Russia, and the recombinant form CRF63_02A6, which has currently been detected in more than 80 % of new HIV-1 cases in Siberia, dominates in Siberia. The aim of this work was to expand and characterize the collection of Env-pseudoviruses obtained on the basis of the recombinant form CRF63_02A6 of HIV-1 circulating in Siberia. In this study, two new variants of Env-pseudoviruses based on CRF63_02A6 of HIV-1 were obtained, characterized, and included in our collection. At present, the collection includes 13 Env-pseudoviruses that are CCR5-tropic. Phylogenetic analysis of the full-length nucleotide sequences of the *env* gene confirmed that all 13 pseudoviruses cluster with the reference sequences of the recombinant form CRF63_02A6. The Env-pseudoviruses were characterized using broadly neutralizing antibodies (bnAbs) targeting different regions of vulnerability of HIV-1 located on the surface of Env glycoprotein complexes. It was shown that the Env-pseudoviruses are sensitive to neutralization by bnAbs VRC01 and 10E8; moderately sensitive to neutralization by bnAbs PG9 and PGT126; and resistant to neutralization by antibodies 2G12 and 2F5. The resulting collection is an important addition to the existing panels of pseudoviruses against other HIV-1 subtypes in the world.

Key words: HIV-1; CRF63_02A6; Env-pseudoviruses; bnAbs; virus neutralization assay

For citation: Rudometova N.B., Fando A.A., Shcherbakov D.N., Zaitsev B.N., Rudometov A.P., Karpenko L.I. Env-pseudoviruses based on the HIV-1 genetic variant circulating in Siberia. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov J Genet Breed*. 2025;29(4):600-607. doi 10.18699/vjgb-25-63

Funding. The study was funded by the state assignment of FBRI SRC VB "Vector" of Rospotrebnadzor.

Env-псевдовирусы на основе генетического варианта, циркулирующего на территории Сибири

Н.Б. Рудометова, А.А. Фандо, Д.Н. Щербаков, Б.Н. Зайцев, А.П. Рудометов, Л.И. Карпенко

Государственный научный центр вирусологии и биотехнологии «Вектор» Роспотребнадзора, р. п. Кольцово, Новосибирская область, Россия
nadenkaand100@mail.ru

Аннотация. Несмотря на многочисленные усилия мирового сообщества, остановить пандемию ВИЧ/СПИД пока не удается. Чтобы остановить распространение вируса, требуются эффективная профилактическая вакцина, а также поиск новых противовирусных агентов. Для того чтобы иметь возможность быстро и адекватно оценивать разрабатываемые вакцинные конструкции, характеризовать ВИЧ-специфические антитела и потенциальные лекарственные препараты, необходим надежный метод тестирования. В этом аспекте хорошо зарекомендовал себя анализ нейтрализации вирусного псевдотипа с использованием панели Env-псевдовирусов разных субтипов ВИЧ-1. В настоящее время созданы отдельные панели псевдовирусов основных генетических подтипов ВИЧ-1: A, B, C и ряда CRF. Эти панели необходимы для получения стандартизированных наборов данных, которые могут быть использованы для ранжирования эффективности вакцины и выявления перспективных кандидатов для дальнейшего исследования. На сегодняшний день в европейской части России доминирует субтип A6 ВИЧ-1, а в Сибири – рекомбинантная форма CRF63_02A6, которая обнаруживается более чем в 80 % новых случаев ВИЧ-1 в Сибири. Цель работы состояла в расширении и характеристике

© Rudometova N.B., Fando A.A., Shcherbakov D.N., Zaitsev B.N., Rudometov A.P., Karpenko L.I. , 2025
This work is licensed under a Creative Commons Attribution 4.0 License

коллекции Env-псевдовирусов, полученных на основе рекомбинантной формы CRF63_02A6 ВИЧ-1, циркулирующей в регионах Сибири. В настоящем исследовании получены и охарактеризованы два новых варианта Env-псевдовирусов на основе CRF63_02A6 ВИЧ-1, которые вошли в нашу коллекцию. Коллекция включает 13 Env-псевдовирусов, являющихся CCR5-тропными. Филогенетический анализ полноразмерных нуклеотидных последовательностей гена *env* подтвердил, что все 13 псевдовирусов кластеризуются с референсными последовательностями рекомбинантной формы CRF63_02A6. Env-псевдовирусы были охарактеризованы с помощью широко нейтрализующих антител (bnAb), нацеленных на разные регионы уязвимости ВИЧ-1, расположенных на поверхности гликопротеиновых комплексов Env. Показано, что Env-псевдовирусы чувствительны к нейтрализации bnAb VRC01 и 10E8; умеренно чувствительны к нейтрализации bnAb PG9 и PGT126; и устойчивы к нейтрализации антителами 2G12 и 2F5. Полученная коллекция – важное дополнение к существующим в мире панелям псевдовирусов против других подтипов ВИЧ-1.

Ключевые слова: ВИЧ-1; CRF63_02A6; Env-псевдовирусы; bnAb; анализ вирусной нейтрализации

Introduction

Human immunodeficiency virus type 1 (HIV-1) remains a major global public health problem. Despite the successes of antiretroviral therapy (ART), it is still not possible to stop the spread of HIV. According to WHO, about 1.2 million people are infected with HIV-1 annually[1]. This is largely due to the fact that the use of drugs to treat HIV infection is accompanied by the emergence of resistance mutations in the virus[2]. As a result, the development of more effective drugs is required. Vaccination of the population could provide reliable protection against HIV/AIDS, but, unfortunately, an effective vaccine has not yet been created (Levy, 2024). Developing a vaccine against HIV-1 is a very difficult task due to the high variability of the virus and its ability to integrate into the human genome. Research in this area is currently focused on the development of immunogens capable of inducing antibodies that neutralize a broad range of HIV-1 isolates (bnAbs) (Trkola, Moore, 2024).

In order to be able to rapidly and reliably evaluate vaccine constructs under development, characterize HIV-specific antibodies and search for potential drugs (entry inhibitors), a reliable testing method is needed. Env-pseudovirus technology has proven itself to be the best in this regard. Neutralizing activity assays using pseudoviruses have several advantages over traditional replicating virus systems in peripheral blood mononuclear cell cultures. First, pseudoviruses are incapable of replication and can be safely handled outside an expensive biosafety level 3 laboratory. Second, neutralization or inhibition assays can be performed using a continuous cell line, thereby reducing the need for primary donor cells. Together, these elements ensure the accuracy, reproducibility and standardization of pseudovirus-based assays (Rudometova et al., 2022a).

Currently, separate panels of pseudoviruses of the main HIV-1 genetic subtypes (A, B, C and a number of CRFs) have been created (Li et al., 2005; Hraber et al., 2017; Wang et al., 2018; Stefic et al., 2019). These panels are needed to obtain standardized data sets that can be used to rank vaccine efficacy and identify promising candidates for further study (de Camp et al., 2014).

It should be noted that in the Russian Federation, the genetic diversity of HIV-1 has regional specificity (Antonova et al., 2023). In the European part of Russia, the HIV-1 subtype A6 dominates (Antonova et al., 2023; Kuznetsova et al., 2023), and in the Siberian region, the HIV-1 recombinant form CRF63_02A6 (Maksimenko et al., 2020; Rudometova et al., 2021), which is currently detected in more than 80 % of new HIV-1 cases in Siberia (Sivay et al., 2022). Some data indicate that HIV-1 CRF63_02A6 is more infectious and may have higher replicative activity than other subtypes, which may lead to an increase in the number of people infected with this variant of the virus (Bogacheva et al., 2017). In connection with the above, there is a need to develop a separate collection of pseudoviruses for HIV-1 CRF63_02A6.

Previously, we described the production of a number of HIV-1 CRF63_02A6 Env-pseudoviruses (Rudometova et al., 2022b). However, to ensure the relevance of the collection, ongoing work is required to obtain new pseudoviruses based on circulating HIV-1 strains.

The aim of this work was to expand and characterize the collection of Env-pseudoviruses obtained based on the recombinant form CRF63_02A6 of HIV-1 circulating in the Siberian region of Russia.

Materials and methods

For the study, 63 serum samples taken from HIV-infected blood donors from the Kemerovo Region and the Altai Republic were selected. They were provided by regional AIDS centers in accordance with the decisions made by the ethical committees of the above-mentioned organizations. Each sample was assigned an anonymous number in accordance with the requirements of the ethical standards of the Russian Federation. Then, RNA was isolated from individual serum using the MAGNO-Sorb reagent kit (Amplisens, Russia) according to the manufacturer's recommendations. The isolated RNA samples were stored at –80 °C.

To obtain amplified variants of the HIV-1 *env* gene, a reverse transcription – PCR reaction (RT-PCR) was performed using the Env_S-1 and Env_02A_AS_1 primers (5′-TTGG GTGTCAACATAGCAGAATAGG-3′ and 5′-CCTGTGGC CTGACTGGAAAGC-3′) and the SuperScript™ IV One-Step RT-PCR System reagent kit (Invitrogen, USA) according to the manufacturer's recommendations. After RT-PCR, the amplification products were analyzed by electrophoresis in a 1 % agarose gel.

Amplified variants of the HIV-1 *env* gene were cloned into the commercial expression vector pcDNA3.1/V5-His TOPO TA (Invitrogen). Transformation of chemically com-

[1] HIV statistics, globally and by WHO region, 2024
[2] HIV drug resistance: brief report, 2024.

petent *Escherichia coli* strain Stbl3 cells (Thermo Fisher Scientific, USA) was performed using the heat shock method (Chang et al., 2017) and clones carrying a plasmid with the integrated *env* gene were selected.

To produce Env-pseudoviruses, HEK293 cell culture was transfected with the constructed recombinant pEnv plasmids together with the pSG3Δenv core plasmid (NIH Reagent Program) using Lipofectamine 3000 (Invitrogen) in a 6-well plate format. Forty-eight hours after transfection, pseudovirus particles were collected by filtering the culture medium through a 0.45-micron filter; concentrated in 20 % sucrose solution for 3 hours at 25,000 rpm; aliquots of 1 ml were made and stored at –80 °C.

Electron microscopic images were obtained by the negative contrast method using a JEM-1400 electron microscope with an accelerating voltage of 80 kV, at magnifications from 10,000 to 80,000x.

The functional activity of pseudoviruses was determined using the TZM-bl cell line (NIH Reagent Program, USA) according to the method described previously (Revilla et al., 2011). Pseudovirus clones were considered functional and suitable for further work if the luminescence level of TZM-bl cells carrying these clones was at least 50 times higher than the luminescence level of cells without the addition of pseudovirus particles.

The nucleotide sequences of the amplified variants of the HIV-1 *env* gene were determined by sequencing using the Sanger method (Genomics Collective Use Center, Institute of Chemical Biology and Fundamental Medicine, Siberian Branch of the Russian Academy of Sciences, Novosibirsk). The resulting sequencing chromatograms were analyzed using the BioEdit program.

The subtype of HIV-1 *env* gene variants was determined using the HIV BLAST online resource (https://www.hiv.lanl.gov/content/sequence/BASIC_BLAST/basic_blast.html). Phylogenetic tree construction was performed using IQ-TREE on the portal https://www.hiv.lanl.gov/content/sequence/IQTREE/iqtree.html. Reference sequences from the Los Alamos HIV Sequence Database (http://www.hiv.lanl.gov/) were used to construct phylogenetic trees. Tropism of the obtained HIV-1 Env pseudoviruses to the coreceptor (CCR5 or CXCR4) was determined by the nucleotide sequence of the V3 loop using the geno2pheno online program (http://coreceptor.geno2pheno.org/).

Neutralization assay was performed according to the method described previously (Revilla et al., 2011). In the work, monoclonal bnAbs to HIV-1 were used, targeting different regions of vulnerability on the surface of trimeric gp120/gp41 complexes – 2G12, VRC01, PG9, PGT126, 2F5, 10E8 (NIH HIV Reagent Program). Samples were tested in duplicate; the experiment was repeated twice. Statistical data processing and IC50 calculation were performed using the GraphPad Prism 9 software.

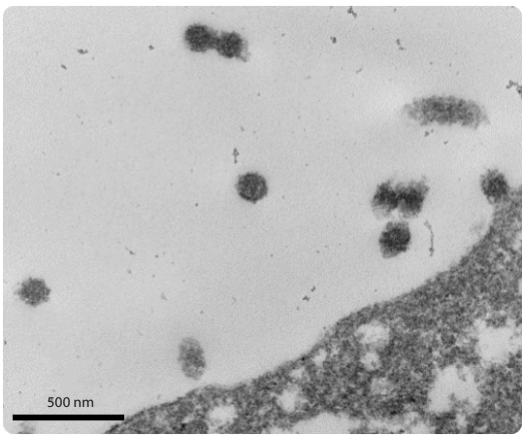
Results

Obtaining new HIV-1 Env-pseudoviruses variants

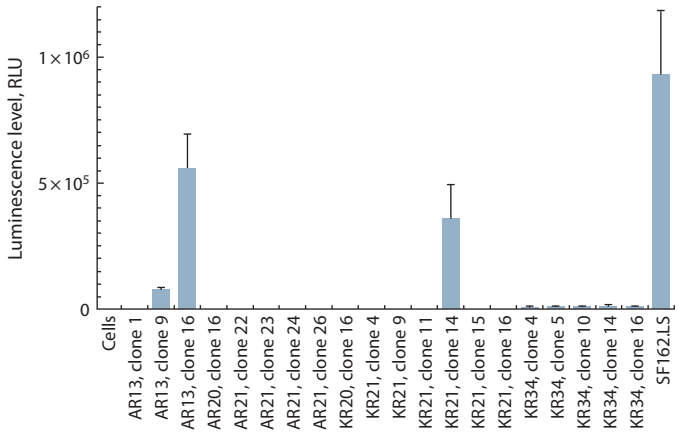
As a result of the study, five full-length variants of the *env* gene were amplified using RT-PCR and cloned into an expression plasmid vector. For each *env* gene variant, one to six genetic constructs with an insert were obtained and used for co-transfection.

Next, we produced Env-pseudoviruses and determined their functional activity. The formation of pseudovirus particles was recorded using electron microscopy (Fig. 1, *a*). Functional analysis of the obtained pseudovirus variants showed that the assembly of functionally active pseudovirus particles occurs for two pseudovirus variants: 22RUAR13 (clone 16) and

a



b



c

Sample code	Region	Phylogenetic affiliation	Tropism
22RUAR13, clone 16	Altai Republic	CRF63_02A6	R5
22RUKR21, clone 14	Kemerovo region	CRF63_02A6	R5

Fig. 1. Characteristics of HIV-1 Env-pseudoviruses.

a – electron micrograph of HEK293 cells after transfection with budded pseudoviral particles of the recombinant form CRF63_02A6; *b* – functional activity of the obtained variants of HIV-1 Env-pseudoviruses (data are presented as mean ± standard deviation); Env-pseudovirus SF162.LS subtype B from the international reference panel of pseudoviruses was used as a positive control; *c* – characteristics of functionally active HIV-1 Env-pseudoviruses.

Fig. 2. Maximum likelihood phylogenetic tree of the env gene of HIV-1 pseudoviruses (visualized using the iTOL online tool).

Green dots mark the env genes of clones 22RUAR13 (clone 16) and 22RUKR21 (clone 14). Purple dots mark the *env* genes of pseudoviruses obtained earlier (Rudometova et al., 2022b).

22RUKR21 (clone 14) (Fig. 1, *b*). The luminescence level of the cells carrying these clones was more than 150 times higher than the luminescence level of pure cells. Sequencing data and phylogenetic analysis confirmed that the functional Env-pseudoviruses belong to the recombinant form CRF63_02A6 and are CCR5-tropic (Fig. 1, *c*).

Characteristics of the HIV-1 CRF63_02A6 Env-pseudoviruses collection

Previously, we obtained 11 Env-pseudoviruses belonging to the recombinant form CRF63_02A6 (Rudometova et al., 2022b). In this work, we added two new pseudoviruses to the collection and decided to conduct a more detailed characterization of it. A phylogenetic analysis of the full-length nucleotide sequences of the env gene was carried out, which confirmed that all 13 pseudoviruses cluster with the reference sequences of the recombinant form CRF63_02A6 (Fig. 2). From the presented data (Fig. 2), it is evident that a certain genetic heterogeneity is observed due to the presence of differences in the nucleotide sequences of the *env* genes of the pseudoviruses included in the collection.

Neutralization phenotype of the HIV-1 CRF63_02A6 Env-pseudoviruses collection

Env-pseudoviruses with functional activity were tested in a pseudovirus neutralization assay with antibodies capable of neutralizing a broad range of HIV-1 isolates (bnAbs). bnAbs recognizing the main sites of HIV-1 vulnerability on

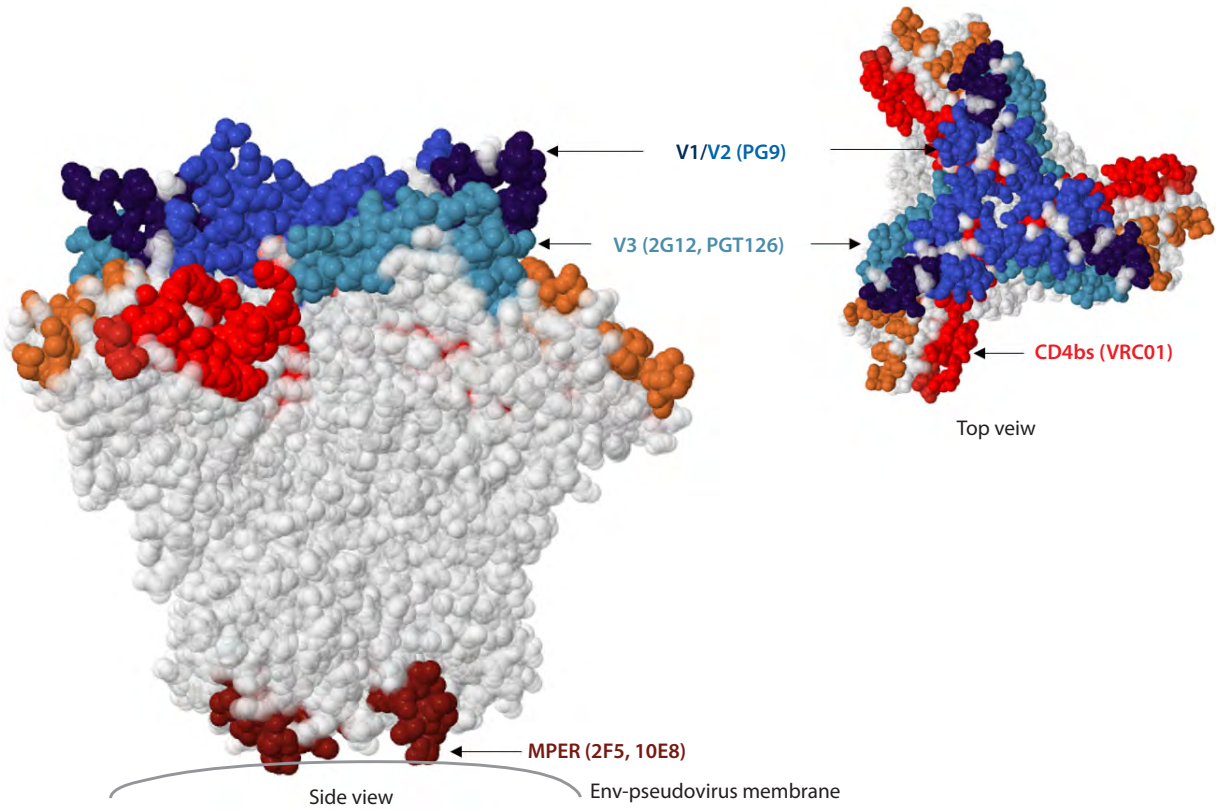


Fig. 3. Sites of vulnerability of broadly neutralizing antibodies in the Env trimer model (PDB: 4ZMJ).

The HIV-1 Env trimer model was visualized using the HIV 3D Structure Viewer program. The designations of broadly neutralizing antibodies (bnAbs) binding to the vulnerability sites of the virus and used in this work are given in brackets.

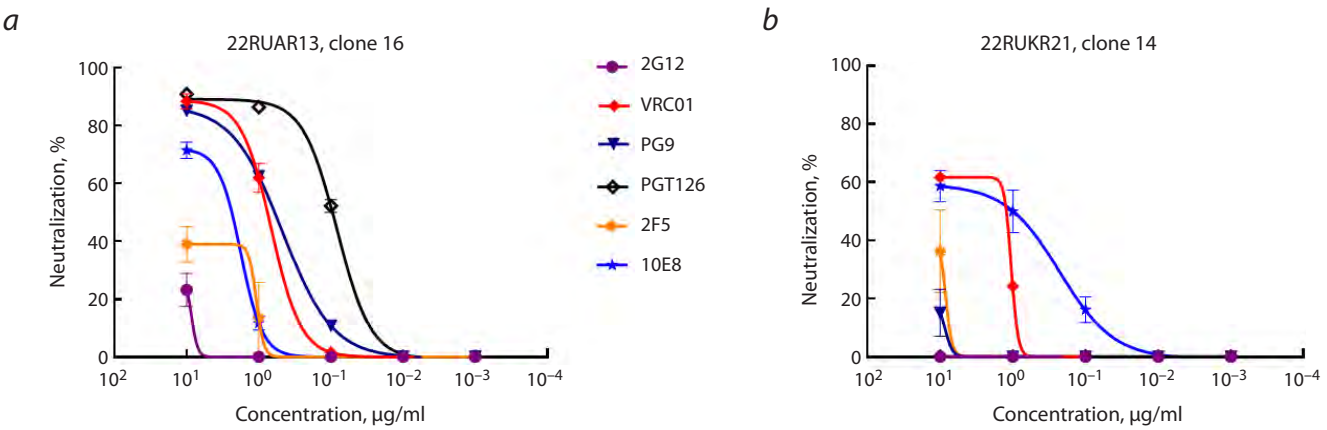


Fig. 4. Neutralizing activity of bnAbs against Env-pseudoviruses 22RUAR13, clone 16 (*a*) and 22RUKR21, clone 14 (*b*).

the gp120/gp41 trimeric complex were selected for analysis (Fig. 3): bnAbs VRC01 – the virus binding to the CD4 cellular receptor; bnAbs PG9 – the structure of the V1/V2 variable loops at the top of the trimer; bnAbs 2G12 and PGT126 – the V3 loop region in the glycan complex; bnAbs 2F5 and 10E8 – the membrane-proximal external region of gp41 (MPER) (Walsh, Seaman, 2021; Thavarajah et al., 2024).

As an example, Fig. 4 shows typical neutralization curves of Env-pseudoviruses by monoclonal broadly neutralizing antibodies for the new variants 22RUAR13 (clone 16) and 22RUKR21 (clone 14).

According to the obtained data (Fig. 4), Env-pseudovirus 22RUAR13 (clone 16) was highly sensitive to neutralization by bnAbs VRC01, PG9 and PGT126, showed moderate sensitivity to neutralization by antibody 10E8, and was resistant to neutralization by antibodies 2G12 and 2F5. At the same time, pseudovirus 22RUKR21 (clone 14) demonstrated moderate sensitivity to neutralization by bnAbs VRC01 and 10E8 and

IC_{50} values determined for bnAbs against HIV-1 CRF63_02A6 Env-pseudoviruses

Env-pseudovirus	IC_{50}, μg/ml					
	VRC01	PG9	2G12	PGT126	2F5	10E8
22RUAR13	0.75	0.62	>10	0.097	>10	2.41
22RUKR21	1.22	>10	>10	>10	>10	1.02
16RU12	0.11	>10	>10	>10	>10	0.02
16RU13	1.71	4.83	>10	0.11	>10	0.05
16RU19	2.63	1.61	>10	0.03	8.43	0.12
16RU20	0.11	0.22	>10	0.16	>10	0.12
16RU21	0.11	4.23	2.31	>10	>10	0.01
16RU22	2.93	6.72	>10	0.01	8.0	0.15
16RU23	3.81	>10	>10	>10	>10	0.11
16RU27	0.31	>10	1.71	0.19	>10	0.61
16RU28	0.31	1.0	1.42	0.01	4.51	0.32
16RU35	0.43	>10	>10	0.11	3.52	0.19
16RU36	0.32	0.22	>10	0.41	>10	0.26

was resistant to neutralization by antibodies 2G12, PG9, 2F5 and PGT126.

Table shows the IC_{50} values determined for bnAbs against our entire collection of HIV-1 CRF63_02A6 Env-pseudoviruses, including 22RUAR13 and 22RUKR21. Comparative analysis of IC_{50} showed that Env-pseudoviruses of this collection have rather heterogeneous sensitivity to neutralization by bnAbs (see the Table), which is probably due to their antigenic differences (Supplementary Material)[3]. As can be seen from Table, HIV-1 CRF63_02A6 Env-pseudoviruses demonstrate high sensitivity to neutralization by bnAbs VRC01 and 10E8.

These antibodies are known to have high neutralizing activity against other HIV-1 subtypes, including A, B, C, recombinant forms CRF02_AG and CRF01_AE (Wu et al., 2010; Huang et al., 2012; Hraber et al., 2017; Wang et al., 2018; Stefic et al., 2019; Wieczorek et al., 2024). In contrast, Env-pseudoviruses demonstrated moderate or even low sensitivity to neutralization against 2F5, PG9, 2G12 and PGT126 antibodies. This is likely due to the fact that bnAbs PG9, 2G12 and PGT126 recognize conformational epitopes in complex with glycans, which can interfere with efficient binding and, as a consequence, neutralization of the virus (Sanders et al., 2002; Doores, Burton, 2010; Krumm et al., 2016).

Discussion

Recombination and mutational variability of HIV-1 have a significant impact on the changes in circulating HIV-1 viruses in Russia. As a result, not only the emergence of new recombinant forms of HIV is observed, but also their spread with the formation of new, different phylogenetic clusters of HIV-1. Such changes in the genetic characteristics of modern HIV-1 must be considered by developers of antiretroviral drugs and vaccines (Rashid et al., 2022; Nair et al., 2024). This requires verification tools, which include collections of Env-pseudoviruses of different HIV-1 subtypes (deCamp et al., 2014).

In this study, two new variants of HIV-1 Env-pseudoviruses were obtained and characterized based on genetic variants circulating in the Siberian Federal District. Phylogenetic analysis showed that both variants belong to the recombinant form CRF63_02A6 and are CCR5-tropic (Fig. 1, *c*; Fig. 2). The obtained variants of HIV-1 Env-pseudoviruses expanded the existing collection. Thus, our collection currently includes 13 Env-pseudoviruses belonging to the recombinant form CRF63_02A6. Env-pseudoviruses were characterized using bnAbs targeting different regions of HIV-1 vulnerability located on the surface of Env glycoprotein complexes. The analysis results showed that the Env-pseudoviruses included in the collection exhibit high sensitivity to bnAbs VRC01 and 10E8; moderate sensitivity to bnAbs PG9 and PGT126 and resistance to antibodies 2G12 and 2F5 (see the Table).

The advantage of this collection is that it is relatively representative, as it includes pseudoviruses that differ in both genetic diversity and sensitivity to bnAbs. According to the P. Hraber et al. (2017), a collection of 12 pseudoviruses of a certain HIV-1 subtype may be sufficient for the initial assessment of the efficacy of vaccine candidates.

Due to the fact that the recombinant form of CRF63_02A6 HIV-1 currently accounts for about 80 % in the Siberia and continues to actively spread (Sivay et al., 2022), the developed collection is an important addition to existing collections in the world against other HIV-1 subtypes.

[3] Supplementary Material is available at:
https://vavilov.elpub.ru/jour/manager/files/Suppl_Rudomet_Engl_29_4.pdf

Conclusion

This work supplements and characterizes the collection of Env-pseudoviruses of HIV-1 recombinant form CRF63_02A6, which will allow to carry out a complex of scientific and applied works: to study the antiviral activity of created chemotherapeutic drugs and to evaluate the effectiveness of vaccines against HIV-1; to study the breadth and spectrum of neutralizing properties of monoclonal broadly neutralizing antibodies and to search for new broadly neutralizing antibodies by screening the neutralizing properties of blood sera of HIV-1 infected people, as well as to study the subtype-specific features of circulating genetic variants of HIV-1.

References

Antonova A.A., Kuznetsova A.I., Ozhmegova E.N., Lebedev A.V., Kazennova E.V., Kim K.V., Tumanov A.S., Glinkina L.N., Bobkova M.R. Genetic diversity of HIV-1 at the current stage of the epidemic in the Russian Federation: an increase in the prevalence of recombinant forms. *VICH-infektciya i Immunosupressii = HIV Infection and Immunosuppressive Disorders*. 2023;15(3):61-72. doi 10.22328/2077-9828-2023-15-3-61-72 (in Russian)

Bogacheva N.V., Blednyh N.A., Totmenin A.V., Mirdzhamalova F.O., Sokolov J.V., Gashnikova N.M. Creature a collection of up-to-date HIV-1 isolates including major Russian genetic variants of virus. *VICH-infektciya i Immunosupressii = HIV Infection and Immunosuppressive Disorders*. 2017;9(1):65-76. doi 10.22328/2077-9828-2017-9-1-65-76 (in Russian)

Chang A.Y., Chau V., Landas J.A., Pang Y. Preparation of calcium competent *Escherichia coli* and heat-shock transformation. *JEMI Methods*. 2017;1:22-25

deCamp A., Hraber P., Bailer R.T., Seaman M.S., Ochsenbauer C., Kappes J., Gottardo R., Edlefsen P., Self S., Tang H., Greene K., Gao H., Daniell X., Sarzotti-Kelsoe M., Gorny M.K., Zolla-Pazner S., LaBranche C.C., Mascola J.R., Korber B.T., Montefiori D.C. Global panel of HIV-1 Env reference strains for standardized assessments of vaccine-elicited neutralizing antibodies. *J Virol*. 2014; 88(5):2489-2507. doi 10.1128/JVI.02853-13

Doores K.J., Burton D.R. Variable loop glycan dependency of the broad and potent HIV-1-neutralizing antibodies PG9 and PG16. *J Virol*. 2010;84(20):10510-10521. doi 10.1128/JVI.00552-10

Hraber P., Rademeyer C., Williamson C., Seaman M.S., Gottardo R., Tang H., Greene K., Gao H., LaBranche C., Mascola J.R., Morris L., Montefiori D.C., Korber B. Panels of HIV-1 subtype C Env reference strains for standardized neutralization assessments. *J Virol*. 2017;91(19):e00991-17. doi 10.1128/JVI.00991-17

Huang J., Ofek G., Laub L., Louder M.K., Doria-Rose N.A., Longo N.S., Imamichi H., … Wyatt R., Haynes B.F., Kwong P.D., Mascola J.R., Connors M. Broad and potent neutralization of HIV-1 by a gp41-specific human antibody. *Nature*. 2012;491(7424):406-412. doi 10.1038/nature11544

Krumm S.A., Mohammed H., Le K.M., Crispin M., Wrin T., Poignard P., Doores K.J. Mechanisms of escape from the PGT128 family of anti-HIV broadly neutralizing antibodies. *Retrovirology*. 2016; 13:8. doi 10.1186/s12977-016-0241-5

Kuznetsova A.I., Munchak Ia.M., Lebedev A.V., Tumanov A.S., Kim K.V., Antonova A.A., Ozhmegova E.N., Pronin A.Yu., Drobyshevskaya E.V., Kazennova E.V., Bobkova M.R. Genetic diversity of capsid protein (p24) in human immunodeficiency virus type-1 (HIV-1) variants circulating in the Russian Federation. *Voprosy Virusologii = Problems of Virology*. 2023;68(1):66-78. doi 10.36233/0507-4088-161 (in Russian)

Levy Y. DNA and protein HIV vaccines: how should we mix it? *Lancet HIV*. 2024;11(5):e274-e275. doi 10.1016/S2352-3018(24)00092-4

Li M., Gao F., Mascola J.R., Stamatatos L., Polonis V.R., Koutsoukos M., Voss G., Goepfert P., Gilbert P., Greene K.M., Bilska M., Kothe D.L., Salazar-Gonzalez J.F., Wei X., Decker J.M., Hahn B.H., Montefiori D.C. Human immunodeficiency virus type 1 env clones from acute and early subtype B infections for standardized assessments of vaccine-elicited neutralizing antibodies. *J Virol*. 2005; 79(16):10108-10125. doi 10.1128/JVI.79.16.10108-10125.2005

Maksimenko L.V., Totmenin A.V., Gashnikova M.P., Astakhova E.M., Skudarnov S.E., Ostapova T.S., Yaschenko S.V., Meshkov I.O., Bocharov E.F., Maksyutov R.A., Gashnikova N.M. Genetic diversity of HIV-1 in Krasnoyarsk Krai: area with high levels of HIV-1 recombination in Russia. *BioMed Res Int*. 2020;2020:9057541. doi 10.1155/2020/9057541

Nair M., Gettins L., Fuller M., Kirtley S., Hemelaar J. Global and regional genetic diversity of HIV-1 in 2010-21: systematic review and analysis of prevalence. *Lancet Microbe*. 2024;5(11):100912. doi 10.1016/S2666-5247(24)00151-4

Rashid A., Li K., Feng Y., Ahmad T., Getaneh Y., Yu Y., Hu X., Abidi S.H., Shao Y. HIV-1 genetic diversity a challenge for AIDS vaccine development: a retrospective bibliometric analysis. *Hum Vaccin Immunother*. 2022;18(1):2014733. doi 10.1080/21645515.2021.2014733

Revilla A., Delgado E., Christian E.C., Dalrymple J., Vega Y., Carrera C., González-Galeano M., … Pérez-Álvarez L., Nájera R., Montefiori D.C., Seaman M.S., Thomson M.M. Construction and phenotypic characterization of HIV type 1 functional envelope clones of subtypes G and F. *AIDS Res Hum Retroviruses*. 2011;27(8):889-901. doi 10.1089/AID.2010.0177

Rudometova N.B., Shcherbakova N.S., Shcherbakov D.N., Mishenova E.V., Delgado E., Ilyichev A.A., Karpenko L.I., Thomson M.M. Genetic diversity and drug resistance mutations in reverse transcriptase and protease genes of HIV-1 isolates from Southwestern Siberia. *AIDS Res Hum Retroviruses*. 2021;37(9):716-723. doi 10.1089/AID.2020.0225

Rudometova N.B., Shcherbakov D.N., Rudometov A.P., Ilyichev A.A., Karpenko L.I. Model systems of human immunodeficiency virus (HIV-1) for *in vitro* efficacy assessment of candidate vaccines and drugs against HIV-1. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov J Genet Breed*. 2022a;26(2):214-221. doi 10.18699/VJGB-22-26

Rudometova N.B., Shcherbakova N.S., Shcherbakov D.N., Taranov O.S., Zaitsev B.N., Karpenko L.I. Construction and characterization of HIV-1 *env*-pseudoviruses of the recombinant form CRF63_02A and subtype A6. *Bull Exp Biol Med*. 2022b;172(6):729-733. doi 10.1007/s10517-022-05466-7

Sanders R.W., Venturi M., Schiffner L., Kalyanaraman R., Katinger H., Lloyd K.O., Kwong P.D., Moore J.P. The mannose-dependent epitope for neutralizing antibody 2G12 on human immunodeficiency virus type 1 glycoprotein gp120. *J Virol*. 2002;76(14):7293-7305. doi 10.1128/jvi.76.14.7293-7305.2002

Sivay M.V., Maksimenko L.V., Osipova I.P., Nefedova A.A., Gashnikova M.P., Zyryanova D.P., Ekushov V.E., Totmenin A.V., Nalimova T.M., Ivlev V.V., Kapustin D.V., Pozdnyakova L.L., Skudarnov S.E., Ostapova T.S., Yaschenko S.V., Nazarova O.I., Chernov A.S., Ismailova T.N., Maksutov R.A., Gashnikova N.M. Spatiotemporal dynamics of HIV-1 CRF63_02A6 sub-epidemic. *Front Microbiol*. 2022;13:946787. doi 10.3389/fmicb.2022.946787

Stefic K., Bouvin-Pley M., Essat A., Visdeloup C., Moreau A., Goujard C., Chaix M.L., Braibant M., Meyer L., Barin F. Sensitivity to broadly neutralizing antibodies of recently transmitted HIV-1 clade CRF02_AG viruses with a focus on evolution over time. *J Virol*. 2019;93(2):e01492-18. doi 10.1128/JVI.01492-18

Thavarajah J.J., Hønge B.L., Wejse C.M. The use of broadly neutralizing antibodies (bNAbs) in HIV-1 treatment and prevention. *Viruses*. 2024;16(6):911. doi 10.3390/v16060911

Trkola A., Moore P.L. Vaccinating people living with HIV: a fast track to preventive and therapeutic HIV vaccines. *Lancet Infect Dis*. 2024; 24(4):e252-e255. doi 10.1016/S1473-3099(23)00481-4

Walsh S.R., Seaman M.S. Broadly neutralizing antibodies for HIV-1 prevention. *Front Immunol*. 2021;12:712122. doi 10.3389/fimmu.2021.712122

Wang H., Yuan T., Li T., Li Y., Qian F., Zhu C., Liang S., Hoffmann D., Dittmer U., Sun B., Yang R. Evaluation of susceptibility of HIV-1 CRF01_AE variants to neutralization by a panel of broadly neutralizing antibodies. *Arch Virol*. 2018;163(12):3303-3315. doi 10.1007/s00705-018-4011-7

Wieczorek L., Sanders-Buell E., Zemil M., Lewitus E., Kavusak E., Heller J., Molnar S., … Ake J., Krebs S.J., Peel S., Tovanabutra S., Polonis V.R. Evolution of HIV-1 envelope towards reduced neutralization sensitivity, as demonstrated by contemporary HIV-1 subtype B from the United States. *PLoS Pathog*. 2023;19(12):e1011780. doi 10.1371/journal.ppat.1011780

Wieczorek L., Chang D., Sanders-Buell E., Zemil M., Martinez E., Schoen J., Chenine A.L., … Michael N.L., Crowell T.A., Ake J.A., Tovanabutra S., Polonis V.R. Differences in neutralizing antibody sensitivities and envelope characteristics indicate distinct antigenic properties of Nigerian HIV-1 subtype G and CRF02_AG. *Virol J*. 2024;21(1):148. doi 10.1186/s12985-024-02394-y

Wu X., Yang Z.Y., Li Y., Hogerkorp C.M., Schief W.R., Seaman M.S., Zhou T., … Kwong P.D., Roederer M., Wyatt R.T., Nabel G.J., Mascola J.R. Rational design of envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1. *Science*. 2010;329(5993):856-861. doi 10.1126/science.1187659

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

Received November 3, 2024. Revised March 9, 2025. Accepted March 10, 2025.

BIOINFORMATICS AND SYSTEMS BIOLOGY
Original article

Vavilovskii Zhurnal Genetiki i Selektsii
Vavilov Journal of Genetics and Breeding. 2025;29(4):608-614

doi 10.18699/vjgb-25-64

Counting touching wheat grains in images based on elliptical approximation

D.R. Avzalov[1, 2], E.G. Komyshev[1, 3], D.A. Afonnikov [1, 2, 3]

[1] Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
[2] Novosibirsk State University, Novosibirsk, Russia
[3] Kurchatov Genomic Center of ICG SB RAS, Novosibirsk, Russia
ada@bionet.nsc.ru

Abstract. The number of grains of a cereal plant characterizes its yield, while grain size and shape are closely related to its weight. To estimate the number of grains, their shape and size, digital image analysis is now generally used. The grains in such images may be completely separated, touching or densely packed. In the first case, the simplest binarization/segmentation algorithms, such as the watershed algorithm, can achieve high accuracy in segmentation and counting grains in an image. However, in the case of touching grains, simple machine vision algorithms may lead to inaccuracies in determining the contours of individual grains. Therefore, methods for accurately determining the contours of individual grains when they are in contact are relevant. One approach is based on the search for pixels of the grain contact area, in particular, by identification of concave points on the grain contour boundary. However, some grains may have chips, depressions and bulges, which leads to the identification of the corner points that do not correspond to the grain contact region. Additional data processing is required to avoid these errors. In this paper, we propose an algorithm for the identification of wheat grains in an image and determine their boundaries in the case when they are touching. The algorithm is based on using a modification of the concave point search algorithm and utilizes a method of assigning contour boundary pixels to a single grain based on approximation of grain contours by ellipses. We have shown that the proposed algorithm can identify grains in the image more accurately compared to the algorithm without such approximation and the watershed algorithm. However, the time cost for such an algorithm is significant and grows rapidly with increasing number of grains and contours including multiple grains.

Key words: wheat; grains; counting; digital images; segmentation; algorithm; concave points

For citation: Avzalov D.R., Komyshev E.G., Afonnikov D.A. Counting touching wheat grains in images based on elliptical approximation. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov J Genet Breed*. 2025;29(4):608-614. doi 10.18699/vjgb-25-64

Funding. The work was funded by the Kurchatov Genomic Center of ICG SB RAS, Agreement No. 075-15-2019-1662 with the Ministry of Science and Higher Education of the Russian Federation.

Определение числа соприкасающихся зерен пшеницы на изображениях на основе эллиптической аппроксимации

Д.Р. Авзалов[1, 2], Е.Г. Комышев[1, 3], Д.А. Афонников [1, 2, 3]

[1] Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
[2] Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия
[3] Курчатовский геномный центр ИЦиГ СО РАН, Новосибирск, Россия
ada@bionet.nsc.ru

Аннотация. Количество зерен растения напрямую характеризует его урожайность, а размер и форма тесно связаны с массой семян. Для оценки количества зерен, их формы и размеров в настоящее время, как правило, используют анализ цифровых изображений. Зерна на таких изображениях могут быть полностью разделены, соприкасаться или быть плотно упакованными. В случае разделенных зерен высокую точность выделения и подсчета зерен на изображении позволяют получить самые простые алгоритмы бинаризации/сегментации, например алгоритм водораздела. Но в случае соприкасающихся зерен простые алгоритмы машинного зрения могут приводить к неточностям в определении контуров отдельных зерен. В этой связи актуальными являются методы точного определения контуров индивидуальных зерен в случае их соприкосновения. Один из подходов основан на поиске пикселей области соприкосновения зерен, в частности с помощью поиска угловых точек на границе контура зерен. Однако зерна могут иметь сколы, впадины и выпуклости, что приводит к идентификации угловых точек, которые не соответствуют области контакта зерен. Это влечет за собой ошибки и для их устранения требует дополнительной обработки данных, фильтрации ложных угловых точек. В настоящей работе мы предлагаем алгоритм идентификации зерен пшеницы на изображении, который позволяет идентифицировать касающиеся зерна и определять их границы на изображении.

© Avzalov D.R., Komyshev E.G., Afonnikov D.A., 2025
This work is licensed under a Creative Commons Attribution 4.0 License

Он базируется на модификации алгоритма поиска угловых точек и использует метод отнесения пикселей границы контура к одному зерну на основе аппроксимации контуров зерен эллипсами. Мы показали на тестовых изображениях, что предложенный алгоритм позволяет более точно идентифицировать зерна на изображении по сравнению с алгоритмом без такой аппроксимации и алгоритмом водораздела. Однако временные затраты для такого алгоритма существенны и быстро растут с увеличением количества зерен и контуров, включающих несколько зерен.
Ключевые слова: пшеница; зерна; подсчет; цифровые изображения; сегментация; алгоритм; угловые точки

Introduction

One of the most important areas of cereal breeding and genetic research is identification of genes that control yield. The number of grains of a cereal plant directly characterizes its yield, and the grain size and shape are closely related to its weight (Zhang X. et al., 2014; Brinton, Uauy, 2019). Estimation of the number of grains, their shape and size is frequently performed using high-performance phenotyping (Afonnikov et al., 2016; Li et al., 2020; Kolhar, Jagtap, 2023), which is based on digital image analysis (Tanabata et al., 2012; Whan et al., 2014; Komyshev et al., 2017). These methods are characterized by high performance, low equipment costs and simplicity of image acquisition protocols. An additional advantage of these methods in comparison with manual counting is that it is possible to accurately determine not only the number of grains, but also their characteristics (size, shape, color) (Tanabata et al., 2012; Cervantes et al., 2016; Komyshev et al., 2020), which is difficult or impossible in the case of manual counting. Furthermore, the obtained images can be stored without changes, while grain characteristics may vary depending on the duration of storage (Afonikov et al., 2022).

Typical images for analysis are grains on a light background obtained using a digital camera, desktop scanner, or smartphone (Herridge et al., 2011; Tanabata et al., 2012; Whan et al., 2014; Komyshev et al., 2017). The grains in such images may be completely separated, touching, or densely packed. In case of separated grains, the simplest binarization/segmentation algorithms can be used to isolate grains in the image, for example, the watershed algorithm (Roerdink, Meijster, 2000). In this case, the size and shape of each grain can be estimated (Mebatsion et al., 2013). However, such a protocol requires a significant amount of time spent on careful placement of grains, which makes it difficult to analyze a large number of samples. In the case of dense packing, it is difficult to determine the contours of individual grains due to possible overlap, and an objective assessment can only be expected for the number of grains in the image, not for their size and shape. In the case of touching grains, simple machine vision algorithms can lead to inaccuracies in determining the contours of individual grains, but the shape and size characteristics of individual grains can be evaluated. This protocol does not require careful placement of grains on the surface, which reduces the time spent on analysis. In this regard, the development of methods for determining the contours of individual grains in the case of their contact in images is relevant.

To solve this problem, approaches based on machine vision (Wang, Paliwal, 2006; Qin et al., 2013) and deep learning (Yang et al., 2021) have been developed. Deep learning methods are currently being actively developed, but they require large samples of images marked manually for training, which is labor-intensive. Machine vision algorithms are less demanding on the size of training samples and their annotations, and are also being actively developed (Liang et al., 2022; Lin et al., 2023). To extract grains, they use binarization algorithms to detect grain contours and subsequent shape analysis for complex contours that include two or more grains. One of the approaches used for analyzing complex contours is based on finding pixels in the grain contact area. To search for such pixels, an algorithm is used to find corner points on the grain contour boundary (Gao et al., 2017; Liu et al., 2017; Tan et al., 2019; Liang et al., 2022; Zhang J. et al., 2022). This algorithm allows to quickly identify points where the contour boundary bends sharply. Pixels where the curvature of the contour line is greatest are considered potential points of grain contact. However, grains may have chips, cavities, and bumps, which leads to the identification of corner points that do not correspond to the grain contact area. This leads to errors, fixing which requires additional data processing to filter out false corner points.

In this paper, an algorithm is proposed for identification of wheat grains in the image, which allows us to identify touching grains and determine their boundaries in the image. It is based on a modification of the corner point search algorithm and uses the method of assigning pixels of the contour boundary to a single grain based on the approximation of grain contours by ellipses.

Materials and methods

Identification of grain contours in an image. The general scheme of the image processing method is shown in Figure 1. It includes the following steps:

- Image preprocessing (size reduction and Gaussian filter).
- Binarization (separation of the areas of grains and the background).
- Search for corner points in the grain contours.
- Post-processing of contour pixels based on ellipse approximation.

Each of the stages involved the use of various computer vision algorithms described below.

The high resolution of the original images (3,968×2,976 pixels) negatively affected the running time of computer vision algorithms. During the preprocessing step, the image resolution was reduced to 1,984×1,488 pixels. It was found empirically that downsizing reduces the running time of the algorithms by several times without significant loss in the grain counting accuracy. For this purpose, a method based on bilinear interpolation was used (Gonzalez, Woods, 2004). In the original image, non-overlapping windows with a size of 2×2 pixels were replaced by one pixel in the converted image. The intensity values of the red, green, and blue components of this pixel were calculated based on the corresponding pixel components of the input window.

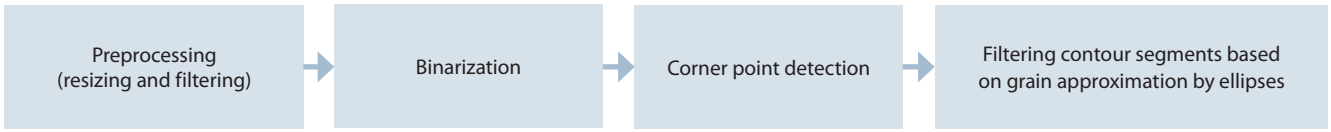


Fig. 1. The main stages of image processing for counting wheat grains.

Let (x_i, y_j), $i, j = 1, 2$, be 2×2 neighboring pixels with coordinates (x_i, y_j), $f(x_i, y_j)$ is the color component intensity for pixel i, j. The color intensity for the resulting pixel in the downsized image $f(x, y)$ was calculated using the following algorithm:
1. Linear interpolation in the X direction:

$$f(x, y_i) = f(x_1, y_i) + f(x_2, y_i). \quad (1)$$

2. Linear interpolation of the obtained values in the Y direction:

$$f(x, y) = f(x, y_1) + f(x, y_2). \quad (2)$$

The image obtained by (1, 2) was downsized by a factor of 2.

A Gaussian filter (Gedraite, Hadad, 2011) was applied to the downsized image to eliminate noise. The filtered image was converted to HSV color space. This transformation improves the identification of the differences between the grains and the background (Fisenko V.T., Fisenko T.Yu., 2008; Domasev, Gnatyuk, 2009).

Smoothing by the mean shift algorithm (Comaniciu, Meer, 1999) was performed at the next step of image preprocessing. The image was converted to grayscale and binarized using the Otsu algorithm (Otsu, 1979).

The results of preprocessing and binarization for a scaled image after Gaussian filtering are shown in Figure 2.

Contour analysis and corner pixel identification. Contours corresponding to the grain regions were identified in the binary masked image. A contour is a curve that connects all pixels of the edge of the grain area in a binarized image. Each contour was analyzed independently.

The contour boundary is traversed; for each pixel p of the boundary, the value of the corner response function, CRF, was calculated (Tan et al., 2019):

$$CRF(p) = \frac{n_p}{A}, \; A = \pi R^2, \quad (3)$$

where n_p is the number of pixels that belong to a contour inside a circle with radius R and center at pixel p, and A is the total number of pixels in this circle (Fig. 3*a*). The value $R = 7$ was selected empirically from the set [3, 4, …, 10] yielding the best accuracy values (at $R > 10$, the algorithm performance decreased substantially). The value of this function is close to ~0.5 on the "straight" part of the grain contour boundary. The large fraction of grain pixels inside the circle (large CRF value) indicates pixels belonging to the corner (Fig. 3*b*, *c*).

The pixel was determined as a corner point if $CRF > 0.6$ (Zhang J. et al., 2022). The number of grains within the contour could be estimated using the number of corner points $N_{corners}$ and the number of contour boundary segments between corner points R_{closed} as $N_{grains} = N_{corners}/2 - R_{closed} + 1$ (Liu et al., 2017). The disadvantage of the algorithm lies in the assumption of an elliptical grain contour. In real-world examples, false corner points can be detected due to chips and irregularities on the grain surface.

Approximation of grain contours by ellipses. Since grains have a shape close to elliptical, they can be approximated by ellipses, which helps breaking the contour of several grains into segments belonging to the same grain even if there are some false corners. The problem is to split the set of segments into subsets in such a way that each subset represents segments of the border of one grain (Fig. 4*a*, *b*).

The algorithm was based on finding the optimal partitioning of contour segments belonging to ellipses, minimizing the partitioning error. The partitioning error is the total error of pixel approximation of each of their ellipses. For some grains, several ellipses can be inscribed in its contour segments (Fig. 4*c*). Therefore, a penalty was added to the error value to choose from several inscribed ellipses the one that has the center position closest to the average center.

The contour was described by a set of pixels p_{ik}, $i = 1, \ldots, m_k$. The coordinates of ellipse pixels were represented by quadratic form $a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2b_1x + 2b_2y + 1 = 0$.

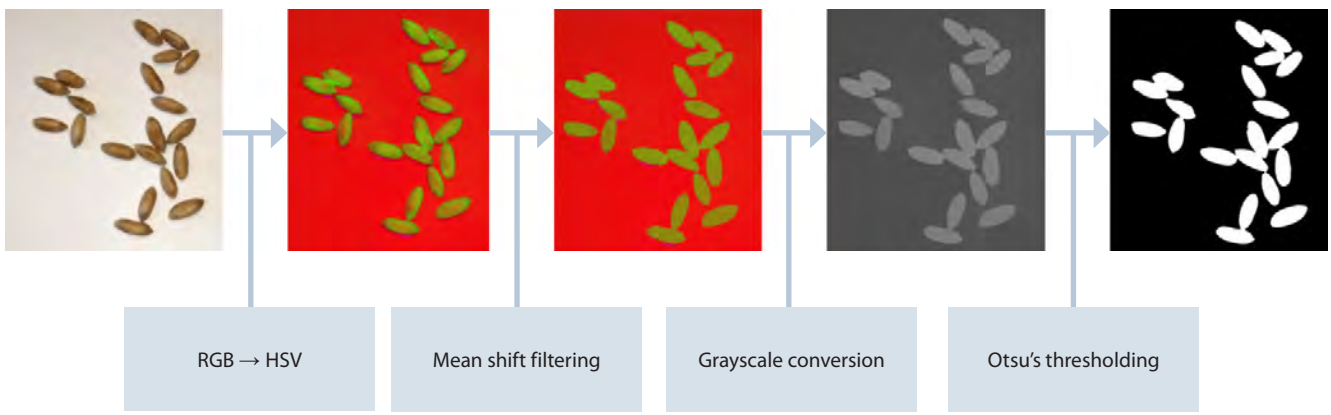


Fig. 2. Results of preprocessing and binarization of grain images.

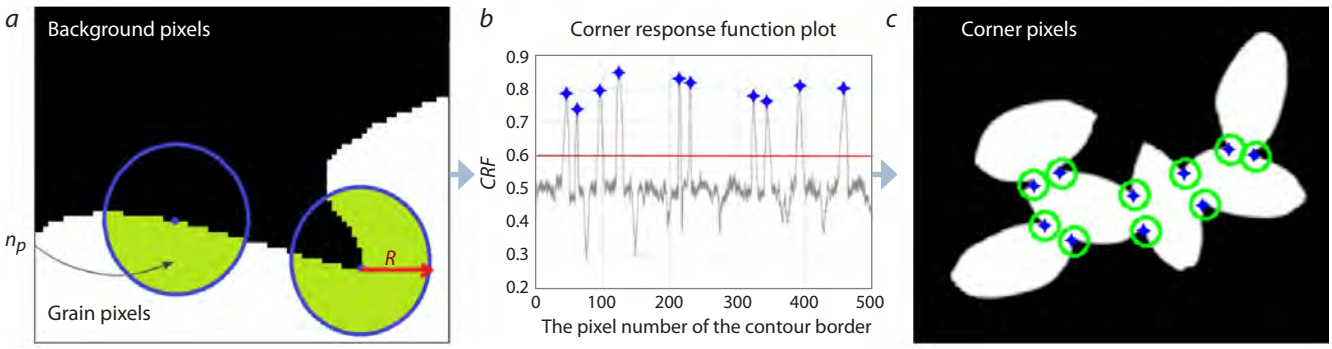


Fig. 3. Using the corner response function to identify the corner pixels of the contour.

a, Visualization of the *CRF* calculation: background pixels are displayed in black, grain pixels outside and inside the *CRF* circle are displayed in white and green, respectively; *b*, the *CRF* plot (*Y*-axis) for contour pixels (*X*-axis), the peaks corresponding to the corner points are shown by blue diamonds; *c*, corner pixels shown in the binarized image (blue diamonds inside green circles).

Fig. 4. The elliptical approximation algorithm used to identify contour boundary segments belonging to the same grain.

a, Pixels belonging to the same segment of the boundary between corners are shown in the same color; *b*, pixels of the contour belonging to the same grain are shown in the same color; *c*, alternate location of two ellipses for segments of the same grain, their centers c_i and c_j is shown as yellow and green dots, the axis of ellipse *B* is shown.

The coordinates of each pixel were substituted in the equation representing the equation system:

$$A = \begin{pmatrix} x_1^2 & 2x_1y_1 & y_1^2 & 2x_1 & 2y_1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ x_{m_k}^2 & 2x_{m_k}y_{m_k} & y_{m_k}^2 & 2x_{m_k} & 2y_{m_k} \end{pmatrix}, b = (-1, \ldots -1)^T, \quad (4)$$

$$||A\cdot\alpha - b||^2 \rightarrow \min_{\alpha}, \quad \alpha = (a_{11}, a_{12}, a_{22}, b_1, b_2). \quad (5)$$

The optimal solution can be found using the least squares method and SVD decomposition (Brinton, Uauy, 2019). The solution to α* can be found as:

$$\alpha^* = A^+b = UD^{-1}V^Tb. \quad (6)$$

In this case, the total partitioning error was calculated using the following formula:

$$E_{\text{partition}} = \sum_k E_k + P, \quad E_k = ||A\cdot\alpha_k^* - b||, \quad (7)$$

where

$$P = \sum_{i=1}^{N} \sum_{j=i+1}^{N} \frac{2}{B\cdot||\text{center}_i - \text{center}_j||} \quad (8)$$

is the penalty for "incorrect" partitioning, and *B* is the minimum axis length of the inscribed ellipses (Fig. 4*c*).

As a result of finding the partition of the contour with the lowest error, the number of grains is equal to the number of subsets in the partition. This approach solves the problem with extra corner points, which affects the accuracy of the method. An example of the result of such an algorithm when identifying grains is shown in Figure 5.

Figure 5 demonstrates the success of the algorithm application. Two touching grains in the lower right corner (the turquoise and green colors of the contours) were split correctly despite the extra corner pixel for the right grain.

The described algorithms were implemented in Python v.3.9, using the OpenCV v. 4.6.0 (Howse, 2013) and Numpy v. 1.21 libraries (https://numpy.org/).

Evaluation of the accuracy of grain identification in an image. For each image, the number of grains was found using the described algorithm. For each image, the accuracy of the algorithm was estimated using the following metric:

$$CR = 100 \cdot \left(1 - \frac{|N - N^*|}{N}\right), \quad (9)$$

where N^* is the number of grains determined using the proposed algorithm, and N is the true number of grains. Two accuracy measures were calculated for contour analysis: CR_{cp} for the algorithm without correction by ellipses and CR_{cpe} for the algorithm with correction by inscribed ellipses.

Additionally, we estimated the average accuracy for grain identification images for the erosion (CR_{e}) and watershed CR_{w} methods (Zhang J. et al., 2022). Additionally, the time

Fig. 5. The result of splitting the contours of several grain groups based on the ellipse approximation algorithm.

The identified corner pixels are shown as green circles. Contour pixels that belong to the same grain are shown in the same color.

Fig. 6. Example of an image of wheat grains in contact.

for image processing, *T*, was estimated for the algorithm with correction using inscribed ellipses. Calculations were performed on a laptop with an Intel i5 4 * 2.9 GHz processor and 6 GB of RAM running the Windows 10 operating system.

Results and discussion

A set of 9 images of wheat grains on a white sheet of paper was used for analysis. Images were obtained using a HUAWEI P20 smartphone with a Sony IMX380 camera and flash. Imaging was performed in auto mode, the high dynamic range (HDR) option was turned off. Grains in the image were placed loosely, but they had a significant number of contacts. An example of a grain image is shown in Figure 6.

Six images had 20 grains each; the sample also included images of 31, 46, and 51 grains.

The Table shows the results of counting the number of grains using two algorithms: using only corner points (cp) and using correction by the ellipse-based corner point algorithm (cpe). The Table shows estimates of the number of grains identified by the algorithm, as well as measures of accuracy and calculation time.

The average accuracy value for the cp algorithm was $CR_{cp} = 0.90$, for the cpe algorithm $CR_{cp} = 0.96$, for the watershed algorithm $CR_w = 0.77$, and for the erosion algorithm $CR_e = 0.93$. These data demonstrate that the ellipse-based grain number correction algorithm yields the most accurate results. However, the accuracy of the algorithm depends on the number of grains that form complex contours: the more grains in the contour, the lower the accuracy. The accuracy also depends on the number of grains, because for large numbers of grains, the probability of their contact is higher.

The correction of the seed detection algorithm by the ellipse-based method yielded performance comparable to some previously published methods based on corner points

Accuracy of grain counting methods based on corner points without and with correction based on the ellipse method for 9 test images

Image number	*N*	*M*	N_{cp}	CR_{cp}, %	N_{cpe}	CR_{cpe}, %	*T*, sec
1	20	2	21	95.0	20	100.0	13.57
2	20	3	22	90.0	20	100.0	20.01
3	20	3	20	100.0	20	100.0	19.54
4	20	4	21	95.0	20	100.0	31.33
5	20	5	21	95.0	18	90.0	80.09
6	20	5	24	80.0	21	95.0	82.87
7	31	7	33	93.5	32	96.8	270.11
8	46	4	53	84.8	45	97.8	43.50
9	53	9	66	75.5	45	85.0	1,121.30

Note. Columns represent the number of grains, *N*; the maximum number of grains in the contour, *M*; the number of grains determined by the cp algorithm, N_{cp}; the accuracy of the cp algorithm, CR_{cp}; the number of grains determined by the cpe algorithm, N_{cpe}; the accuracy of the cpe algorithm CR_{cpe}; the running time of the cpe algorithm, *T* (sec).

identification. Tan et al. (2019) used a corner point detection algorithm for counting rice grains in combination with a neural network with error back propagation for subsequent correction of segmentation results. On average, grain segmentation accuracy for different rice varieties was 94 %. Wang and Paliwal (2006) applied the watershed algorithm to seed images after segmentation and transformation depending on the distances between background and grain pixels. The algorithm was used to count grains in images for six types of plants (winter wheat, hard white-grain wheat, and hard amber wheat, barley, oats, and rye). The proportion of correctly identified grains ranged from 88.6 to 94.4 % for wheat and from 55.4 to 79.0 % for the other plants. Liu et al. (2017) suggested an algorithm based on detecting feature points in the image and estimating the correlation between their number and the number of grains in the image. The authors tested the algorithm for wheat and rice grains and showed that the error of their method was from 0 to 4.7 % (on average, 0.1 % for wheat and 1.5 % for rice), while applying the usual watershed algorithm led to an error of 14 to 40 % for wheat and of 20 to 50 % for rice seeds. Liang et al. (2022) proposed a comprehensive approach that identifies the number of non-touching grains by the K-means method, a layered watershed algorithm was used to identify and count rarely touching grains, and a dividing line algorithm was used for densely lying grains. The accuracy of the method was 99.65 %.

Thus, the corner point algorithm with ellipse-based correction in general yielded an accuracy comparable to existing algorithms (especially in the case when there are few contours with a large number of touching grains). However, the execution time of this algorithm increases quickly with increasing both the number of complex contours and the number of grains analyzed. For a number of grains more than 20 and a large number of contours with multiple grains, the time of execution becomes unacceptable for analysis. This occurs because with increasing number of contiguous grains in a single contour, the number of possible combinations of ellipse location subsets increases. Moreover, the number of contour pixels, the coordinates of which are used to compose the systems of linear equations (4), is growing. To reduce the execution time (in the case of the described implementation of the method), it is possible to parallelize the algorithm. Further optimization is possible by using estimates of the number of possible grains, for example, by the contour area. This will reduce the number of partitions.

Note also that the algorithm uses an approximation of the grain shape by ellipses, which may not be applicable to grains of a more complex shape, such as beans. However, in the general case, this algorithm allows to use an arbitrary shape of grain contours, which in the future can be implemented for grains of other plant species.

Conclusions

An algorithm is proposed for identifying and counting grains in digital images in the case of their touching. The algorithm is based on binarization of the image to select contours containing grains and further processing of these contours. Processing consists of finding corner points on the contour and then selecting them by assigning pixels of the contour border to a single grain based on the approximation of grains by ellipses. The post-processing step allowed to exclude false regions of grain contacts in the contour. Analysis of the test images showed that in the case when the number of touching grains is small, the algorithm allows us to obtain a high accuracy of grain counting (up to 100 %, and systematically better than without ellipse-based approximation). However, when contours include a large number of touching grains, the execution time of the algorithm increases, which makes it impractical in analysis.

References

Afonnikov D.A., Genaev M.A., Doroshkov A.V., Komyshev E.G., Pshenichnikova T.A. Methods of high-throughput plant phenotyping for large-scale breeding and genetic experiments. *Russ J Genet.* 2016;52(7):688-701. doi 10.1134/S1022795416070024

Afonnikov D.A., Komyshev E.G., Efimov V.M., Genaev M.A., Koval V.S., Gierke P.U., Börner A. Relationship between the characteristics of bread wheat grains, storage time and germination. *Plants.* 2022;11(1):35. doi 10.3390/plants11010035

Brinton J., Uauy C. A reductionist approach to dissecting grain weight and yield in wheat. *J Integr Plant Biol*. 2019;61(3):337-358. doi 10.1111/jipb.12741

Cervantes E., Martín J.J., Saadaoui E. Updated methods for seed shape analysis. *Scientifica*. 2016;2016(1):5691825. doi 10.1155/2016/5691825

Comaniciu D., Meer P. Mean shift analysis and applications. In: Proceedings of the Seventh IEEE International Conference on Computer Vision. Vol. 2. Kerkyra, Greece, 1999;1197-1203. doi 10.1109/ICCV.1999.790416

Domasev M.V., Gnatyuk S.P. Color, Color Management, Color Calculations and Measurements. St. Petersburg: Piter Publ., 2009 (in Russian)

Fisenko V.T., Fisenko T.Yu. Computer Processing and Image Recognition. St. Petersburg, 2008 (in Russian)

Gao L., Zhao C., Liu M. Segmentation of touching seeds based on shape feature and multiple concave point detection. In: 2017 IEEE International Conference on Imaging Systems and Techniques (IST), Beijing, 2017;1-5. doi 10.1109/IST.2017.8261448

Gedraite E.S., Hadad M. Investigation on the effect of a Gaussian Blur in image filtering and segmentation. In: Proceedings ELMAR-2011, Zadar, Croatia, 2011;393-396

Gonzalez R.C., Woods R.E. Digital Image Processing. CRC Press, Boca Raton, FL, 2004

Herridge R.P., Day R.C., Baldwin S., Macknight R.C. Rapid analysis of seed size in *Arabidopsis* for mutant and QTL discovery. *Plant Methods*. 2011;7(1):3. doi 10.1186/1746-4811-7-3

Howse J. OpenCV Computer Vision with Python. Birmingham: Packt Publishing, 2013

Kolhar S., Jagtap J. Plant trait estimation and classification studies in plant phenotyping using machine vision. A review. *Inf Process Agric*. 2023;10(1):114-135. doi 10.1016/j.inpa.2021.02.006

Komyshev E.G., Genaev M.A., Afonnikov D.A. Evaluation of the SeedCounter, a mobile application for grain phenotyping. *Front Plant Sci.* 2017;7:1990. doi 10.3389/fpls.2016.01990

Komyshev E.G., Genaev M.A., Afonnikov D.A. Analysis of color and texture characteristics of cereals on digital images. *Vavilovskii Zhurnal Genetiki i Selektsii = Vavilov J Genet Breed*. 2020;24(4):340-347. DOI 10.18699/VJ20.626

Li Z., Guo R., Li M., Chen Y., Li G. A review of computer vision technologies for plant phenotyping. *Comput Electron Agric*. 2020;176:105672. doi 10.1016/j.compag.2020.105672

Liang N., Sun S., Yu J., Taha M.F., He Y., Qiu Z. Novel segmentation method and measurement system for various grains with complex touching. *Comput Electron Agric*. 2022;202:107351. doi 10.1016/j.compag.2022.107351

Lin W., Ma D., Su Q., Liu S., Liao H., Yao H., Xu P. Image segmentation method for physically touching soybean seeds. *Software Impacts*. 2023;18:100591. doi 10.1016/j.simpa.2023.100591

Liu T., Chen W., Wang Y., Wu W., Sun C., Ding J., Guo W. Rice and wheat grain counting method and software development based on Android system. *Comput Electron Agric*. 2017;141:302-309. doi 10.1016/j.compag.2017.08.011

Mebatsion H.K., Paliwal J., Jayas D.S. Automatic classification of non-touching cereal grains in digital images using limited morphological and color features. *Comput Electron Agric*. 2013;90:99-105. doi 10.1016/j.compag.2012.09.007

Otsu N. A threshold selection method from gray-level histograms. In: IEEE Transactions on Systems, Man, and Cybernetics. 1979;9(1): 62-66. doi 10.1109/TSMC.1979.4310076

Qin Y., Wang W., Liu W., Yuan N. Extended-maxima transform watershed segmentation algorithm for touching corn kernels. *Adv Mech Eng*. 2013;5:268046. doi 10.1155/2013/268046

Roerdink J.B.T.M., Meijster A. The watershed transform: definitions, algorithms and parallelization strategies. *Fundam Inform*. 2000; 41(2):187-228

Tan S., Ma X., Mai Z., Qi L., Wang Y. Segmentation and counting algorithm for touching hybrid rice grains. *Comput Electron Agric*. 2019;162:493-504. doi 10.1016/j.compag.2019.04.030

Tanabata T., Shibaya T., Hori K., Ebana K., Yano M. SmartGrain: high-throughput phenotyping software for measuring seed shape through image analysis. *Plant Physiol*. 2012;160(4):1871-1880. doi 10.1104/pp.112.205120

Wang W., Paliwal J. Separation and identification of touching kernels and dockage components in digital images. *Can Biosyst Eng*. 2006;48:7

Whan A.P., Smith A.B., Cavanagh C.R., Ral J.P.F., Shaw L.M., Howitt C.A., Bischof L. GrainScan: a low cost, fast method for grain size and colour measurements. *Plant Methods*. 2014;10:23. doi 10.1186/1746-4811-10-23

Yang S., Zheng L., He P., Wu T., Sun S., Wang M. High-throughput soybean seeds phenotyping with convolutional neural networks and transfer learning. *Plant Methods*. 2021;17(1):50. doi 10.1186/s13007-021-00749-y

Zhang J., Liu S., Wu W., Zhong X., Liu T. Research on a rapid identification method for counting universal grain crops. *PLoS One*. 2022;17(9):e0273785. doi 10.1371/journal.pone.0273785

Zhang X., Deng Z., Wang Y., Li J., Tian J. Unconditional and conditional QTL analysis of kernel weight related traits in wheat (*Triticum aestivum* L.) in multiple genetic backgrounds. *Genetica*. 2014; 142(4):371-379. doi 10.1007/s10709-014-9781-6

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

Received December 28, 2024. Revised January 22, 2025. Accepted February 6, 2025.

Прием статей через электронную редакцию на сайте http://vavilov.elpub.ru/index.php/jour
Предварительно нужно зарегистрироваться как автору, затем в правом верхнем углу страницы выбрать «Отправить рукопись». После завершения загрузки материалов обязательно выбрать опцию «Отправить письмо», в этом случае редакция автоматически будет уведомлена о получении новой рукописи.

«Вавиловский журнал генетики и селекции (Vavilov Journal of Genetics and Breeding)» до 2011 г. выходил под названием «Информационный вестник ВОГиС»/ "The Herald of Vavilov Society for Geneticists and Breeding Scientists".

Сетевое издание «Вавиловский журнал генетики и селекции (Vavilov Journal of Genetics and Breeding)» – реестровая запись СМИ Эл № ФС77-85772, зарегистрировано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций 14 августа 2023 г.

Издание включено ВАК Минобрнауки России в Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук, Russian Science Citation Index, Российский индекс научного цитирования, ВИНИТИ, Web of Science CC, Scopus, PubMed Central, DOAJ, ROAD, Ulrich's Periodicals Directory, Google Scholar.

Открытый доступ к полным текстам:
русскоязычная версия – на сайте https://vavilovj-icg.ru/
и платформе Научной электронной библиотеки, elibrary.ru/title_about.asp?id=32440

англоязычная версия – на сайте vavilov.elpub.ru/index.php/jour
и платформе PubMed Central, https://www.ncbi.nlm.nih.gov/pmc/journals/3805/

При перепечатке материалов ссылка обязательна.

✉ email: vavilov_journal@bionet.nsc.ru

Издатель: Федеральное государственное бюджетное научное учреждение
«Федеральный исследовательский центр Институт цитологии и генетики
Сибирского отделения Российской академии наук»,
проспект Академика Лаврентьева, 10, Новосибирск, 630090.
Адрес редакции: проспект Академика Лаврентьева, 10, Новосибирск, 630090.
Секретарь по организационным вопросам С.В. Зубова. Тел.: (383)3634977.
Издание подготовлено информационно-издательским отделом ИЦиГ СО РАН. Тел.: (383)3634963*5218.
Начальник отдела: Т.Ф. Чалкова. Редакторы: В.Д. Ахметова, И.Ю. Ануфриева. Дизайн: А.В. Харкевич.
Компьютерная графика и верстка: Т.Б. Коняхина, О.Н. Савватеева.

Дата публикации 30.06.2025. Формат 60 × 84 $^{1}/_{8}$. Уч.-изд. л. 19.2.