

Сетевое издание

ВАВИЛОВСКИЙ ЖУРНАЛ ГЕНЕТИКИ И СЕЛЕКЦИИ

Основан в 1997 г.

Периодичность 8 выпусков в год

DOI 10.18699/vjgb-24-15

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 (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 – канд. 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 – 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. Розозин – канд. биол. наук (США)

А.О. Рувинский – д-р биол. наук, проф. (Австралия)

Е.Ю. Рыкова – д-р биол. наук (Россия)

Е.А. Салина – 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 – д-р биол. наук, проф. РАН (Россия)

Э.К. Хуснутдинова – д-р биол. наук, проф. (Россия)

М. Чен – д-р биол. наук (Китайская Народная Республика)

Ю.Н. Шавруков – д-р биол. наук (Австралия)

Р.И. Шейко – чл.-кор. НАНБ, д-р с.-х. наук (Беларусь)

С.В. Шестаков – академик РАН, д-р биол. наук (Россия)

Н.К. Янковский – академик РАН, д-р биол. наук (Россия)

Online edition

VAVILOV JOURNAL OF GENETICS AND BREEDING

VAVILOVSKII ZHURNAL GENETIKI I SELEKTSII

*Founded in 1997**Published 8 times annually*

DOI 10.18699/vjgb-24-15

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), 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, Cand. 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. Beshpalova*, 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*, 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*, 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*, 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

Молекулярная и клеточная биология

- 131 **ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ**
Транскрипционный фактор DREF регулирует экспрессию гена микроРНК *bantam Drosophila melanogaster*.
М.Б. Шварц, М.М. Прудникова, О.В. Андреевков, Е.И. Волкова, И.Ф. Жимулев, О.В. Антоненко, С.А. Демаков
- 138 **ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ**
Оценка клеточных линий с индуцируемой деплецией компонентов когезина и конденсинов посредством анализа морфологии метафазных хромосом. А.М. Юнусова, А.В. Смирнов, И.Е. Пристяжнюк, Т.А. Шнайдер, Е.К. Мальцева, С.Д. Афонникова, О.А. Гусев, Н.Р. Баттулин (на англ. языке)

Генетика и селекция растений

- 148 **ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ**
Изменчивость плюсовых деревьев сосны обыкновенной (*Pinus sylvestris* L.) в Среднем и Верхнем Поволжье по ISSR-маркерам. О.В. Шейкина, Е.М. Романов
- 155 **ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ**
Многомерный анализ многолетних климатических данных в связи с урожайностью, скороспелостью и проблемой глобального потепления. В.М. Ефимов, Д.В. Речкин, Н.П. Гончаров

Иммунитет растений к болезням

- 166 **ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ**
Изучение генетической коллекции земляники (*Fragaria* L.) по устойчивости к мучнистой росе. А.С. Лыжин, И.В. Лукьянчук
- 175 **ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ**
In silico поиск R-генов у примитивных культурных видов картофеля. А.А. Гурина, М.С. Ганчева, Н.В. Алпатьева, Е.В. Рогозина

Генетика животных

- 185 **ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ**
Различия в плодовитости между двумя линиями *Drosophila melanogaster* дикого типа коррелируют с различиями в экспрессии гена *Jeh1*, кодирующего фермент деградации ювенильного гормона. О.В. Андреевкова, Н.В. Адоньева, В.М. Ефимов, Н.Е. Грунтенко

190

ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ

Влияние породы и среды на длину теломер лейкоцитов у крупного рогатого скота. Н.С. Юдин, А.В. Игошин, Г.А. Ромашов, А.А. Мартынов, Д.М. Ларкин (на англ. языке)

Медицинская генетика

198

ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ

Аберрантное метилирование генов развития плаценты в ворсинах хориона спонтанных абортусов с трисомией 16. О.Ю. Васильева, Е.Н. Толмачева, А.Э. Дмитриев, Я.А. Даркова, Е.А. Саженова, Т.В. Никитина, И.Н. Лебедев, С.А. Васильев

204

ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ

Состав бактериального микробиома мокроты пациентов с разными патоморфологическими формами немелкоклеточного рака легкого. В.Г. Дружинин, Е.Д. Баранова, П.С. Деменков, Л.В. Мацкова, А.В. Ларионов

215

ОБЗОР

Роль сиртуинов в эпигенетической регуляции и контроле старения. Е.М. Самойлова, С.Е. Романов, Д.А. Чудакова, П.П. Лактионов

228

ОБЗОР

Влияние транспозонов на развитие болезни Альцгеймера. Р.Н. Мустафин, Э.К. Хуснутдинова

Актуальные технологии

239

ОБЗОР

Невирусные системы внутриклеточной доставки инструментов редактирования генома. И.Х. Шайхутдинов, П.В. Ильясов, О.В. Грибкова, Л.В. Лимарева

249

ОБЗОР

Проблемы создания фаговых библиотек антител и пути их решения. В.С. Арипов, Н.В. Волкова, А.А. Ильичев, Д.Н. Щербаков

Molecular and cell biology

- 131 ORIGINAL ARTICLE
Transcription factor DREF regulates expression of the microRNA gene *bantam* in *Drosophila melanogaster*. M.B. Schwartz, M.M. Prudnikova, O.V. Andreenkov, E.I. Volkova, I.F. Zhimulev, O.V. Antonenko, S.A. Demakov
- 138 ORIGINAL ARTICLE
Assessing cell lines with inducible depletion of cohesin and condensins components through analysis of metaphase chromosome morphology. A.M. Yunusova, A.V. Smirnov, I.E. Prist'yazhnuk, T.A. Shnaider, E.K. Maltseva, S.D. Afonnikova, O.A. Gusev, N.R. Battulin

Plant genetics and breeding

- 148 ORIGINAL ARTICLE
Variability of Scots pine (*Pinus sylvestris* L.) plus trees in the Middle and Upper Volga Region with the use of ISSR markers. O.V. Sheikina, E.M. Romanov
- 155 ORIGINAL ARTICLE
Multivariate analysis of long-term climate data in connection with yield, earliness and the problem of global warming. V.M. Efimov, D.V. Rechkin, N.P. Goncharov

Plant immunity

- 166 ORIGINAL ARTICLE
Study of a genetic collection of strawberry (*Fragaria* L.) for resistance to powdery mildew. A.S. Lyzhin, I.V. Luk'yanchuk
- 175 ORIGINAL ARTICLE
In silico search for and analysis of *R* gene variation in primitive cultivated potato species. A.A. Gurina, M.S. Gancheva, N.V. Alpatieva, E.V. Rogozina

Animal genetics

- 185 ORIGINAL ARTICLE
Fertility differences between two wild-type *Drosophila melanogaster* lines correlate with differences in the expression of the *Jheh1* gene, which codes for an enzyme degrading juvenile hormone. O.V. Andreenkova, N.V. Adonyeva, V.M. Efimov, N.E. Gruntenko

- 190 ORIGINAL ARTICLE
Influence of breed and environment on leukocyte telomere length in cattle. N.S. Yudin, A.V. Igoshin, G.A. Romashov, A.A. Martynov, D.M. Larkin

Medical genetics

- 198 ORIGINAL ARTICLE
Aberrant methylation of placental development genes in chorionic villi of spontaneous abortions with trisomy 16. O.Yu. Vasilyeva, E.N. Tolmacheva, A.E. Dmitriev, Ya.A. Darkova, E.A. Sazhenova, T.V. Nikitina, I.N. Lebedev, S.A. Vasilyev
- 204 ORIGINAL ARTICLE
Composition of the sputum bacterial microbiome of patients with different pathomorphological forms of non-small-cell lung cancer. V.G. Druzhinin, E.D. Baranova, P.S. Demenkov, L.V. Matskova, A.V. Larionov
- 215 REVIEW
Role of sirtuins in epigenetic regulation and aging control. E.M. Samoilova, S.E. Romanov, D.A. Chudakova, P.P. Laktionov
- 228 REVIEW
Involvement of transposable elements in Alzheimer's disease pathogenesis. R.N. Mustafin, E.K. Khusnutdinova

Mainstream technologies

- 239 REVIEW
Non-viral systems for intracellular delivery of genome editing tools. I.H. Shaikhutdinov, P.V. Ilyasov, O.V. Gribkova, L.V. Limareva
- 249 REVIEW
Problems of creating antibody phage libraries and their solutions. V.S. Aripov, N.V. Volkova, A.A. Ilyichev, D.N. Shcherbakov

DOI 10.18699/vjgb-24-20

Transcription factor DREF regulates expression of the microRNA gene *bantam* in *Drosophila melanogaster*

M.B. Schwartz, M.M. Prudnikova, O.V. Andreenkov, E.I. Volkova, I.F. Zhimulev, O.V. Antonenko, S.A. Demakov  

Institute of Molecular and Cellular Biology of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
 demakov@mcb.nsc.ru

Abstract. The *bantam* gene encodes a vital microRNA and has a complex expression pattern in various tissues at different stages of *Drosophila* development. This microRNA is involved in the control of normal development of the ocular and wing imaginal discs, the central nervous system, and also in maintaining the undifferentiated state of stem cells in the ovaries of adult females. At the cellular level, *bantam* stimulates cell proliferation and prevents apoptosis. The *bantam* gene is a target of several conserved signaling cascades, in particular, Hippo. At the moment, at least ten proteins are known to directly regulate the expression of this gene in different tissues of *Drosophila*. In this study, we found that the *bantam* regulatory region contains motifs characteristic of binding sites for DREF, a transcription factor that regulates the expression of Hippo cascade genes. Using transgenic lines containing a full-length *bantam* lethality-rescuing deletion fragment and a fragment with a disrupted DREF binding site, we show that these motifs are functionally significant because their disruption at the *bantam* locus reduces expression levels in the larvae and ovaries of homozygous flies, which correlates with reduced vitality and fertility. The effect of DREF binding to the promoter region of the *bantam* gene on its expression level suggests an additional level of complexity in the regulation of expression of this microRNA. A decrease in the number of eggs laid and a shortening of the reproductive period in females when the DREF binding site in the regulatory region of the *bantam* gene is disrupted suggests that, through *bantam*, DREF is also involved in the regulation of *Drosophila* oogenesis.

Key words: microRNA; genetic regulation; mutagenesis; transcription; transcription factors.

For citation: Schwartz M.B., Prudnikova M.M., Andreenkov O.V., Volkova E.I., Zhimulev I.F., Antonenko O.V., Demakov S.A. Transcription factor DREF regulates expression of the microRNA gene *bantam* in *Drosophila melanogaster*. *Vavilovskii Zhurnal Genetiki i Seleksii* = *Vavilov Journal of Genetics and Breeding*. 2024;28(2):131-137. DOI 10.18699/vjgb-24-20

Транскрипционный фактор DREF регулирует экспрессию гена микроРНК *bantam* *Drosophila melanogaster*

М.Б. Шварц, М.М. Прудникова, О.В. Андреевков, Е.И. Волкова, И.Ф. Жимулев, О.В. Антоненко, С.А. Демаков  

Институт молекулярной и клеточной биологии Сибирского отделения Российской академии наук, Новосибирск, Россия
 demakov@mcb.nsc.ru

Аннотация. Ген *bantam* кодирует жизненно важную микроРНК и имеет сложный паттерн экспрессии в различных тканях на разных стадиях развития дрозофилы. Эта микроРНК обеспечивает нормальное развитие глазных и крыловых имажинальных дисков, центральной нервной системы, а также участвует в поддержании недифференцированного состояния стволовых клеток в яичниках взрослых самок. На клеточном уровне *bantam* стимулирует пролиферацию клеток и препятствует апоптозу. Ген *bantam* является мишенью нескольких консервативных сигнальных каскадов, в частности Hippo. На сегодняшний день известно не менее 10 белков, напрямую регулирующих экспрессию этого гена в разных тканях дрозофилы. В настоящей работе мы обнаружили, что регуляторная область *bantam* содержит мотивы, характерные для сайтов связывания DREF – транскрипционного фактора, который регулирует экспрессию генов каскада Hippo. Используя трансгенные линии, содержащие полноразмерный фрагмент, спасающий летальность делеции *bantam*, и фрагмент с нарушенным сайтом связывания DREF, мы показали, что эти мотивы имеют функциональное значение, поскольку их нарушение в локусе *bantam* снижает уровень экспрессии в личинках и яичниках гомозиготных мух, что коррелирует со сниженной жизнеспособностью и фертильностью. Влияние связывания DREF с промоторной областью гена *bantam* на уровень его экспрессии предполагает дополнительный уровень сложности регуляции экспрессии этой микроРНК. Снижение количества откладываемых яиц и сокращение репродуктивного периода у самок при нарушении сайта связывания DREF в регуляторной области гена *bantam* позволяют предполагать, что через *bantam* DREF также участвует в регуляции оогенеза дрозофилы.

Ключевые слова: микроРНК; генетическая регуляция; мутагенез; транскрипция; факторы транскрипции.

Introduction

The *bantam* gene encodes a vital microRNA that is expressed in many tissues throughout the *Drosophila* life cycle. At the cellular level, *bantam* stimulates cell proliferation and prevents apoptosis (Brennecke et al., 2003). This microRNA provides normal development of the ocular and wing imaginal discs, the central nervous system, and it is also involved in maintaining the undifferentiated state of stem cells in the ovaries of adult females (Shcherbata et al., 2007; Peng et al., 2009; Reddy, Irvine, 2011; Slattery et al., 2013; Weng, Cohen, 2015).

Expression of the *bantam* gene is controlled by a wide range of transcription factors that form different assemblies in different tissues at different stages of development. Such conserved morphogens as *Notch*, *Wingless*, and *Dpp* are involved in the regulation of the *bantam* gene expression in different tissues (Herranz et al., 2008; Oh et al., 2010; Ku, Sun, 2017). Currently, about ten proteins that directly regulate *bantam* expression in different tissues are known. Dysregulation of *bantam* leads to abnormal development of many imaginal organs, and problems with vitality and fertility (Hipfner et al., 2002; Brennecke et al., 2003; Shcherbata et al., 2007).

One of the key regulators of *bantam* expression is the transcriptional coactivator Yorkie (Yki), which provides the necessary level of *bantam* expression by binding to tissue-specific transcription factors (Peng et al., 2009; Slattery et al., 2013; Nagata et al., 2022). Yki is part of the highly conserved Hippo signaling cascade (Oh et al., 2010). In *Drosophila*, this cascade suppresses the process of cell division, inducing cell cycle arrest and apoptosis. The cascade begins with the kinase Hippo (Hpo), which triggers the sequential phosphorylation of a number of proteins and, finally, leads to the inactivation of Yki (Huang et al., 2005; Reddy, Irvine, 2011). The *hpo* gene promoter has been shown to contain DRE motifs (DREF responding element). The promoter of the *warts* (*wts*) gene, the second key kinase of the Hippo cascade, also contains DRE motifs. The transcription factor DREF binds to these motifs, enhancing the expression of *hpo* and *wts* in the ocular imaginal discs (Fujiwara et al., 2012; Vo et al., 2014). Thus, DREF positively regulates the Hippo cascade and, accordingly, reduces the expression level of *bantam*.

In our work, we found that the *bantam* regulatory region contains motifs characteristic of DREF binding sites and tested the consequences of their disruption at the organism level. We examined how disruption of DREF binding sites at the *bantam* locus affected fly viability and fertility.

Materials and methods

Obtaining a mutation of the DREF binding site. To obtain the mutation, we used a DNA fragment from the *bantam* locus of 4709 bp in length (3L:637635-642343, release=r6.23), including DRE and DRE-like motifs. The DRE and DRE-like motifs were isolated from each other using the unique EcoRV restriction site located between them. To mutate each of these motifs, restriction with endonuclease ClaI (AT|CG|AT) was performed, and then the 5'-overhangs were end-filled using a Klenow fragment. This involved the insertion of two nucleotides (CG) into the key part of each motif. Then both parts of the 4709 bp DNA fragment with mutated DRE and DRE-like motifs were ligated and the resulting DNA fragment DREF was inserted at the KpnI and NotI restriction sites into the

previously obtained pUni-mod vector containing the attB recombination site (Andreenkov et al., 2016). To obtain the transgenic line “DREF” we used the attP/attB system for specific integration, and a line containing an attP site in the region 10A1-2 of the X chromosome (Andreenkov et al., 2016).

Fly stocks. The “4.7” transgenic line’s genotype was y^1 , $Df(1)w^{67c23}$, 10A1-2-“4.7”; *ban^{Δ1}/TM6B*. It contained a transgene insertion with a full-length DNA fragment from the *bantam* locus with a length of 4709 bp into the 10A1-2 region (Schwartz et al., 2019).

The transgenic line “DREF” had the following genotype: y^1 , $Df(1)w^{67c23}$, 10A1-2-“DREs_mut”; *ban^{Δ1}/TM6B*, and it contained the insertion of a transgene with a modified fragment 4.7 in the 10A1-2 region. The modification of the 4.7 fragment consisted in mutations in potential binding sites for the DREF protein. In all experiments performed, homozygotes for transgene insertion were studied.

The following lines were used as control: “yw” with the genotype y^1 $Df(1)w^{67c23}$; “ban+” with the genotype y^1 , $Df(1)w^{67c23}$; +/TM6B, where TM6B is a balancer chromosome with *Tb* phenotype – short body of larvae and adults; “*Δban*” with the genotype y^1 , $Df(1)w^{67c23}$; *ban^{Δ1}/TM6B*.

Lines of flies were kept at +23 °C on standard food with addition of dry yeast.

Immunostaining. Indirect immunofluorescence staining of polytene chromosomes was performed according to the previously described protocol (Kolesnikova et al., 2013) using mouse monoclonal antibodies to the DREF protein (kindly provided by C.M. Hart, USA), diluted in a ratio of 1:200, with subsequent coloring by goat-anti-mouse-Alexa 488 antibodies (Thermo Fisher Scientific, # A28175), diluted in a ratio of 1:600.

Fly viability assessment. To determine the viability of transgenic line flies, 5 females and 5 males of the same line were placed in one glass vial. Once every 5 days, the flies were transferred to fresh food, and the experiment continued for a month. Flies of the “TM6” line were used as a control. Based on the results of 3 repetitions during the experiment, each line had on average the following number of descendants: “4.7” – 454 ± 92; “DREF” – 287 ± 112; “TM6” – 756 ± 289.

The viability of homozygotes for the *ban^{Δ1}* deletion in transgenic lines was determined as the ratio of the number of flies with a body of normal length (*Tb*+) to the total number of hatched offspring. Since the *Tb* phenotype is identified at the stages of larvae, pupae and adults, it is possible to distinguish flies lacking the balancer, homozygous for the *ban^{Δ1}* deletion, from heterozygous flies at different stages of development. The “ban+” line, without a transgene and without a deletion, was used as a control. The viability of flies at different developmental stages in transgenic lines was compared with the viability of control flies using Student’s *t* test, with a preliminary check for normal distribution using the Shapiro–Wilk test.

Female fertility assessment. To determine the fertility of females of transgenic lines, they were crossed with males of the “yw” line. Females of the “yw” line were used as control. In each cross we used 5 females and 5 males. Every day the flies were transferred to fresh food and the number of eggs laid was counted. The experiment continued until the death of the last transgenic female in a vial. The number of eggs during the entire experiment was normalized to the number

of females. The experiment was repeated three times. When plotting the fertility dynamics curve, the number of eggs laid by females each day was normalized to the current number of living females. Differences in fertility levels between lines were assessed using Student's *t* test, as well as the χ^2 test.

Determination of the expression level of mature microRNA *bantam*. The expression level of mature microRNA *bantam* was determined by quantitative PCR combined with a reverse transcription reaction (qRT-PCR), adapted for the study of microRNAs through the use of an extended stem-loop primer (Chen et al., 2005; Kramer, 2011). We used *U6* snRNA as a reference gene (Zhang et al., 2017). To obtain cDNA, we used 5 μ g of total RNA, M-MuLV-PH revertase, and accompanying reagents according to the manufacturer's instructions (Biolabmix). Relative *bantam* gene expression was determined using the $\Delta\Delta C_t$ method. qRT-PCR was carried out using a thermal cycler BioRad C-1000 (USA).

The experiment was done in two biological replicates. We used 30 μ l of the following reaction mixtures: for *bantam* detection – 3 μ l of 5 μ M ban-F and ban-R primers, 3 μ l of 2.5 μ M TaqMan-ban probe, 3 μ l of 10xAS buffer, 3 μ l of 4 μ M dNTP, 1 u. a. of Taq polymerase; for *U6* detection – 3 μ l of 10 μ M U6-F and U6-R primers, 3 μ l of 2.5 μ M TaqMan *U6* probe, 3 μ l of 10xAS buffer, 3 μ l of 4 μ M dNTP, 1 u. a. of Taq polymerase, 3 μ l of 10 mM MgCl₂ to the final concentration of Mg²⁺ equal to 2.5 mM for a reaction. The nucleotide sequences of the primers and probes used in the experiments are given in 5'→3' orientation: ban-SL – gtcgtatccagtgcagggtccgaggtattcgcactg gatacacaatcag, ban-F – cgccggcagatgagatcattttg, ban-R – cagt gcagggtccgaggt, TaqMan-ban – cgcaactggatacacaatcagcttt, U6-SL – gtcgtatccagtgcagggtccgaggtattcgcactggatacagcggc catgc, U6-F – gccgcatacagagaagatta, U6-R – agtgcagggtc cgaggta, TaqMan-U6 – ttcgcaactggatacagcggcctatgc.

Results and discussion

The *bantam* regulatory region contains DREF binding sites

To study the role of DREF in the regulation of *bantam* gene expression, we used transgenic fly lines. The “4.7” fly line contained a transgene insertion with a 4709 bp long DNA fragment from the *bantam* locus (Fig. 1, a) into the 10A1-2 region of the X chromosome. This fragment, hereafter referred to as fragment 4.7 (see Fig. 1, b), contains a sequence encoding the *bantam* miRNA hairpin as well as two putative promoters of the *bantam* gene (Brennecke et al., 2003; Qian et al., 2011). It has been shown previously that the 4.7 fragment rescues the lethal ban^{Δ1} deletion, removing approximately 21 kb from the *bantam* locus (Schwartz et al., 2019). In the fragment 4.7, 1.2 kb upstream of the *bantam* hairpin, we found TATCGATA and TATCGATG motifs corresponding to DRE and DRE-like elements, respectively (Ohler et al., 2002). Both motifs are characteristic of binding sites for the transcription factor DREF (see Fig. 1, c).

The “DREF” fly line contained a transgene insertion with the 4.7 fragment with disrupted DREF binding sites. To disrupt the DREF binding sites, we introduced mutations in the DRE (DREF-responding element) and DRE-like motifs (see Fig. 1, c). It should be noted that DRE and DRE-like include CGATA motifs that form the binding site for the insulator protein BEAF-32. The mutations we introduced did not disrupt the CGATA motifs and, accordingly, did not destroy the BEAF-32 binding site. Immunolocalization on polytene chromosomes of *Drosophila* larvae showed that in line “4.7” in the distal part of the 10A1-2 band at the site of transposon insertion there was an additional DREF localization signal (Fig. 2, b, red arrow). At the same time, this additional signal was absent

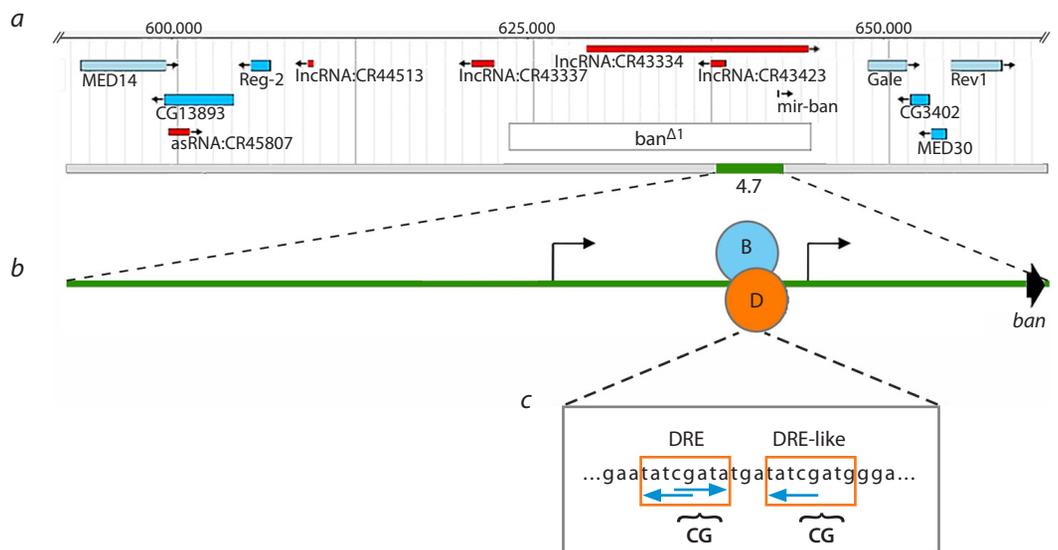


Fig. 1. Molecular and genetic organization of the *bantam* locus.

a – DNA fragment 4.7 containing the *bantam* gene (green rectangle), ban^{Δ1} deletion (white rectangle). The location of other genes in the region (blue and red rectangles); b – the 4.7 DNA fragment organization scheme. Binding sites for DREF and BEAF-32 proteins (ovals D and B, respectively). Curved arrows correspond to the positions of putative promoters of the *bantam* gene (Brennecke et al., 2003; Qian et al., 2011). The *bantam* hairpin position (black arrow); c – nucleotide sequence with binding sites for DREF and BEAF-32 proteins. Motifs characteristic of the DREF (orange rectangles) and BEAF-32 (blue arrows) binding sites. Double nucleotide insertions that disrupt DREF binding sites are indicated by curly brackets.

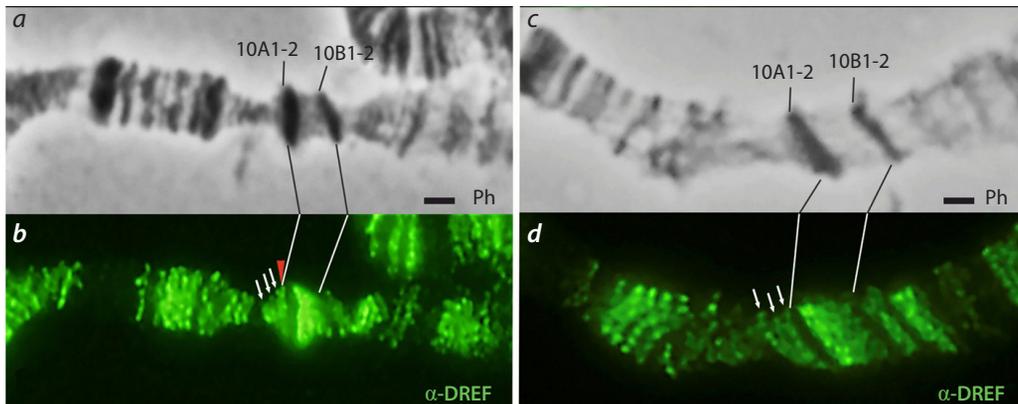


Fig. 2. Immunolocalization of the DREF protein in the 10A1-2 region of the X chromosome transgenic fly lines.

Fly lines: “4.7” (a, b) and “DREF” (c, d). Microphotographs of polytene X chromosomes of larval salivary glands in phase contrast (Ph) mode and after staining with antibodies against DREF (green). White arrows indicate endogenous DREF localization signals in the 9F region. The red triangle marks an additional localization signal of DREF in the distal part of the 10A1-2 band, corresponding to the localization of the transgene in the “4.7” line. There is no additional signal in the “DREF” line (d). Scale bar 1 μ m.

in the “DREF” line (see Fig. 2, d). It can be assumed that the DRE and DRE-like motifs we found correspond to the binding site of the DREF protein, and the mutations of these motifs we introduced lead to disruption of DREF protein binding.

Disruption of the DREF binding site in the *bantam* regulatory region affects fly viability

We found that flies of the “DREF” line survived against the background of the *ban* ^{Δ 1} deletion, but the viability of such flies was significantly reduced compared to flies of the “4.7” line (Fig. 3), as well as compared to control “TM6” flies, containing neither the transgene nor the *ban* ^{Δ 1} deletion. Moreover, the death of flies homozygous for *ban* ^{Δ 1} in the “DREF” line mainly occurred at the late pupal stage, which coincides with the characteristic lethality of the *ban* ^{Δ 1} deletion (Brennecke et al., 2003). The use of control containing TM6B balancer allowed taking into account the influence of the balancer itself on the fly viability.

In order to determine whether the decrease in viability of transgenic “DREF” flies was associated with *bantam* gene expression, we examined the expression level of the mature *bantam* microRNA in the larvae of transgenic and control flies. The studies were carried out using Real-time PCR adapted for microRNA (Chen et al., 2005; Kramer, 2011). The “yw” line with a normal *bantam* locus served as a positive control. We used the “*Aban*” line with the *ban* ^{Δ 1} deletion as negative control. Larvae homozygous for the insertions were selected. Since individuals homozygous for the *ban* ^{Δ 1} deletion die at the pupal stage, studies were carried out on larvae. As expected, in the “*Aban*” line, the mature microRNA *bantam* was not detected in larvae homozygous for *ban* ^{Δ 1} (Fig. 4). The expression of *bantam* was reduced in both transgenic lines, “4.7” and “DREF”, compared to the control “yw”. A significant decrease in the expression level of mature microRNA in the “4.7” line seems surprising given that the viability of flies in the “4.7” line did not differ from the control. This may be explained by the fact that although in the “4.7” line, the level of *bantam* microRNA expression was significantly reduced, it remained at a sufficient level in all the tissues where it was necessary

for fly survival. And in the “DREF” line, the *bantam* expression level was ubiquitously at a low, threshold level, which significantly affected viability. It is also possible that in the “DREF” line, *bantam* expression was reduced only in certain tissues critical for fly survival. The data obtained suggest that although the “DREF” transgene rescues the *ban* ^{Δ 1} deletion, it does not contain all the regulatory elements required for full *bantam* expression.

The fundamental ability of the mutant transgene in the “DREF” line to “rescue” the *ban* ^{Δ 1} deletion indirectly confirms that the introduced mutations did not destroy the binding site of the BEAF-32 protein. We have previously shown that destruction of the BEAF-32 protein binding site in the *bantam* regulatory region leads to death at the late pupal stage (Schwartz et al., 2019).

Interestingly, in the “DREF” transgenic line, the viability of adult flies homozygous for *ban* ^{Δ 1} was sex-dependent. The proportion of males homozygous for *ban* ^{Δ 1} was only 30 % of

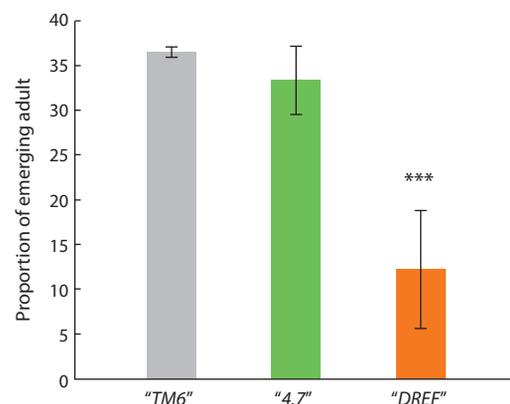


Fig. 3. Effect of DREF binding to the *bantam* regulatory region on adult fly viability in transgenic fly lines.

The proportion of adults homozygous for the *ban* ^{Δ 1} deletion in transgenic fly lines “4.7” (green column) and “DREF” (orange column). The proportion of adults in the control fly line “TM6” (gray column), containing only native wild-type *bantam* locus. *** $p < 0.001$.

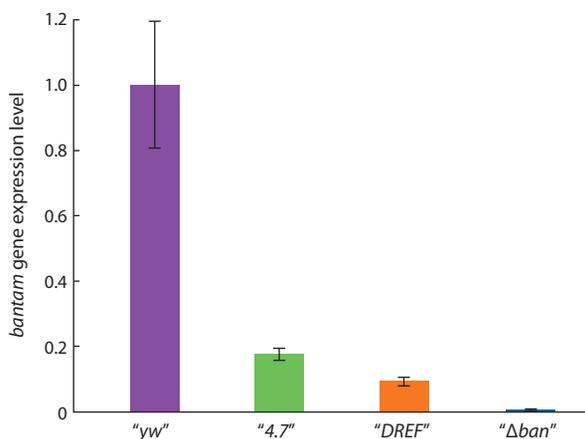


Fig. 4. The expression of *bantam* microRNA in the 3rd instar larvae.

The amount of mature microRNA *bantam* in the transgenic lines "4.7" and "DREF" and the control line "Δban", homozygous for the *ban*^{Δ1} deletion. The expression of *bantam* in the control line "yw" was conditionally taken as one. The *bantam* expression levels were normalized to the expression level of the U6 snRNA reference gene.

all adult flies, which is significantly lower than in the "4.7" line and in the control ($p < 0.01$). This may be explained by the fact that disruption of the DREF binding site in the *bantam* regulatory region has a greater effect on the viability of males than females. Another explanation could be the different levels of transgene activity on the X chromosome associated with dosage compensation. Despite the fact that males of the "DREF" line, homozygous for *ban*^{Δ1}, were less viable than females, they did not have problems with fertility.

Disruption of the DREF binding site in the *bantam* promoter region significantly reduces female fertility

We studied the fertility of females of transgenic lines, estimating the average number of eggs laid per female (see Materials and methods). In the "4.7" line, the fertility of females homozygous for *ban*^{Δ1} was significantly reduced compared to females of the "yw" control line and amounted to 32.7 % of the fertility of control "yw" females, taken as 100 % (Fig. 5). The fertility of females homozygous for *ban*^{Δ1} in the "DREF" line was only 10.5 %. These data indicate that the 4.7 fragment does not contain all the regulatory elements necessary for the normal progression of oogenesis, and the mutation of the binding site in the "DREF" line disrupts this process even more crucially.

However, it should be taken into account that the life expectancy of females in transgenic lines was significantly reduced compared to control females of the "yw" line ($p < 0.05$). Thus, in the "yw" line, 50 % of the females died on average on the 22nd day of the experiment, in the "4.7" line – on the 16th day, and in the "DREF" line – on the 10th day.

At the same time, a reduction in life expectancy was not the only explanation for the decrease in the number of eggs laid by females of transgenic lines. Analysis of the dynamics of egg laying showed that females of transgenic lines not only laid fewer eggs on each day of the experiment than females of the control "yw" line ($p < 0.001$), but also finished laying eggs much earlier (Fig. 6). The reproductive period in females

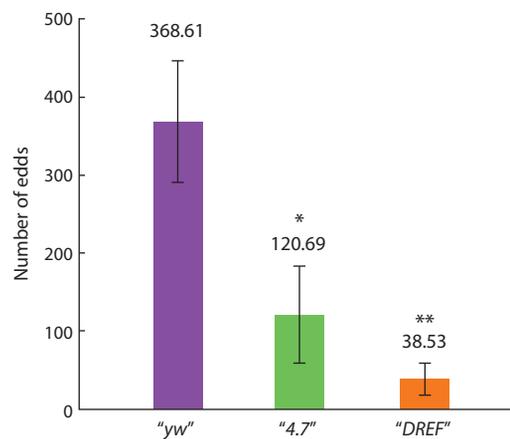


Fig. 5. Fertility of transgenic fly line females.

Average number of eggs per female from the control line "yw" and transgenic lines "4.7" and "DREF", homozygous for the transgenes and for the *ban*^{Δ1} deletion. ** $p < 0.01$; * $p < 0.05$.

of the "4.7" line lasted on average 16.3 days ($p < 0.01$), and in "DREF" females – 8.3 days ($p < 0.001$), while the reproductive period in control "yw" females was 24.7 days.

Disruption of the DREF binding site reduces *bantam* miRNA expression in adult fly ovaries

The early completion of egg laying in transgenic lines is similar to the previously described situation with inactivation of *bantam* microRNA in germline stem cells in the ovaries of adult flies (Shcherbata et al., 2007). According to the authors, in such females, about 14 % of germline stem cells left their niche per day. This can lead to both a general decrease in fertility and a shortening of the reproductive period.

We decided to test whether DREF binding to the *bantam* regulatory region actually affects the expression level of mature microRNA. Using qRT-PCR, we showed that the expression of *bantam* microRNA in the ovaries of females of the transgenic lines "4.7" and "DREF" was lower than in females of the control "yw" line (Fig. 7).

The results obtained are consistent with the fact that in both transgenic lines the fertility of females was reduced compared to control "yw" females. Moreover, in the "DREF" line, *bantam* expression was reduced not only compared to "yw", but also compared to "4.7". This indicates that DREF binding to the regulatory region of the *bantam* gene plays an important role in its expression in the ovaries, and it is the disruption of this binding that may explain the significant decrease in fertility in females of the "DREF" line.

Conclusion

In this work, we tested the functionality of a potential binding site for the transcription factor DREF, found in the regulatory region of the *bantam* gene. Disruption of this DREF binding site has a significant impact on fly viability and female fertility. This is accompanied by a significant decrease in the expression of mature *bantam* microRNA in both whole larvae and the ovaries of adult flies. As was previously shown, DREF

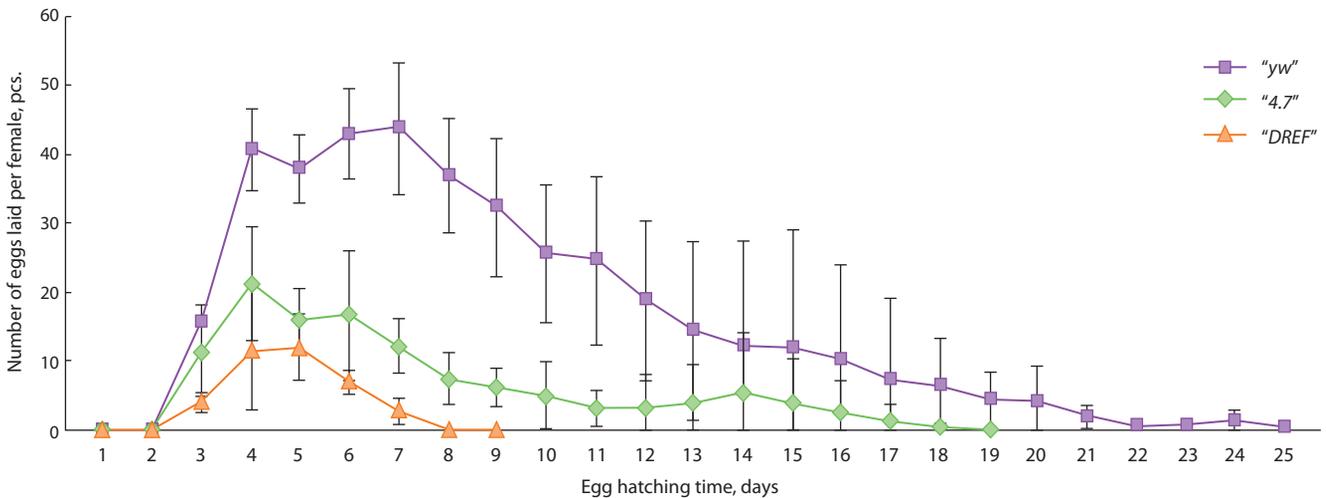


Fig. 6. Temporal dynamics of female fertility.

Average number of eggs laid per day per female in the control "yw" line and in the transgenic lines "4.7" and "DREF", homozygous for transgenes and for the *ban^{Δ1}* deletion.

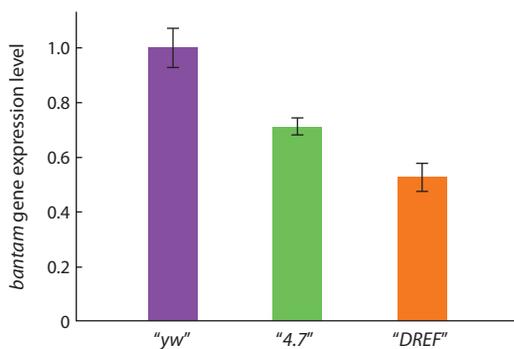


Fig. 7. The expression of mature *bantam* microRNA in the ovaries of transgenic females.

The amount of mature microRNA *bantam* in the transgenic lines "4.7" and "DREF", homozygous for the *ban^{Δ1}* deletion. The expression of *bantam* in the control "yw" line was conditionally taken as one. The *bantam* expression levels were normalized to the expression level of the *U6* snRNA reference gene.

positively influences the activity of the Hippo cascade, thereby indirectly limiting the expression of *bantam* (Fujiwara et al., 2012; Vo et al., 2014). The effect of DREF binding to the promoter region of the *bantam* gene on its expression level suggests an additional level of complexity in the regulation of the expression of this microRNA.

A decrease in the number of eggs laid and a shortening of the reproductive period in females when the DREF binding site in the regulatory region of the *bantam* gene is disrupted suggests that DREF is also involved in the regulation of *Drosophila* oogenesis through *bantam*.

References

Andreenkov O.V., Andreenkova N.G., Volkova E.I., Georgiev P.G., Goncharova A.A., Pokholkova G.V., Demakov S.A. Ectopic tethering of the Chromator protein in UAS>DBD(GAL4) system as approach for studying of the insulator proteins in *Drosophila mel-*

nogaster polytene chromosomes. *Tsitologiya = Cell and Tissue Biology*. 2016;58(6):493-497 (in Russian)

Brennecke J., Hipfner D.R., Stark A., Russell R.B., Cohen S.M. *bantam* encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene *hid* in *Drosophila*. *Cell*. 2003;113(1):25-36. DOI 10.1016/s0092-8674(03)00231-9

Chen C., Ridzon D.A., Broomer A.J., Zhou Z., Lee D.H., Nguyen J.T., Barbisin M., Xu N.L., Mahuvakar V.R., Andersen M.R., Lao K.Q., Livak K.J., Guegler K.J. Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Res*. 2005;33(20):e179. DOI 10.1093/nar/gni178

Fujiwara S., Ida H., Yoshioka Y., Yoshida H., Yamaguchi M. The warts gene as a novel target of the *Drosophila* DRE/DREF transcription pathway. *Am. J. Cancer Res*. 2012;2(1):36-44

Herranz H., Pérez L., Martín F.A., Milán M. A Wingless and Notch double-repression mechanism regulates G1-S transition in the *Drosophila* wing. *EMBO J*. 2008;27(11):1633-1645. DOI 10.1038/emboj.2008.84

Hipfner D.R., Weigmann K., Cohen S.M. The *bantam* gene regulates *Drosophila* growth. *Genetics*. 2002;161(4):1527-1537. DOI 10.1093/genetics/161.4.1527

Huang J., Wu S., Barrera J., Matthews K., Pan D. The Hippo signaling pathway coordinately regulates cell proliferation and apoptosis by inactivating Yorkie, the *Drosophila* Homolog of YAP. *Cell*. 2005;122(3):421-434. DOI 10.1016/j.cell.2005.06.007

Kolesnikova T.D., Posukh O.V., Andreyeva E.N., Bebyakina D.S., Ivankin A.V., Zhimulev I.F. *Drosophila* SUUR protein associates with PCNA and binds chromatin in a cell cycle-dependent manner. *Chromosoma*. 2013;122(1-2):55-66. DOI 10.1007/s00412-012-0390-9

Kramer M.F. Stem-loop RT-qPCR for miRNAs. *Curr. Protoc. Mol. Biol*. 2011;95(1):15.10.1-15.10.15. DOI 10.1002/0471142727.mb1510s95

Ku H.Y., Sun Y.H. Notch-dependent epithelial fold determines boundary formation between developmental fields in the *Drosophila* antenna. *PLoS Genet*. 2017;13(7):e1006898. DOI 10.1371/journal.pgen.1006898

Nagata R., Akai N., Kondo S., Saito K., Ohsawa S., Igaki T. Yorkie drives supercompetition by non-autonomous induction of autophagy via *bantam* microRNA in *Drosophila*. *Curr. Biol*. 2022;32(5):1064-1076.e4. DOI 10.1016/j.cub.2022.01.016

- Oh H., Irvine K.D. Yorkie: the final destination of Hippo signaling. *Trends Cell Biol.* 2010;20(7):410-417. DOI 10.1016/j.tcb.2010.04.005
- Ohler U., Liao G.C., Niemann H., Rubin G.M. Computational analysis of core promoters in the *Drosophila* genome. *Genome Biol.* 2002;3:research0087.1. DOI 10.1186/gb-2002-3-12-research0087
- Peng H.W., Slattery M., Mann R.S. Transcription factor choice in the Hippo signaling pathway: *Homothorax* and *yorkie* regulation of the microRNA *bantam* in the progenitor domain of the *Drosophila* eye imaginal disc. *Genes Dev.* 2009;23(19):2307-2319. DOI 10.1101/gad.1820009
- Qian J., Zhang Z., Liang J., Ge Q., Duan X., Ma F., Li F. The full-length transcripts and promoter analysis of intergenic microRNAs in *Drosophila melanogaster*. *Genomics.* 2011;97(5):294-303. DOI 10.1016/j.ygeno.2011.02.004
- Reddy B.V.V.G., Irvine K.D. Regulation of *Drosophila* glial cell proliferation by Merlin–Hippo signaling. *Development.* 2011;138(23):5201-5212. DOI 10.1242/dev.069385
- Schwartz (Berkaeva) M.B., Pankova T.E., Demakov S.A. ADF1 and BEAF-32 chromatin proteins affect nucleosome positioning and DNA decompaction in 61C7/C8 interband region of *Drosophila melanogaster* polytene chromosomes. *Vavilov Journal of Genetics and Breeding.* 2019;23(2):154-159. DOI 10.18699/VJ19.475
- Shcherbata H.R., Ward E.J., Fischer K.A., Yu J.Y., Reynolds S.H., Chen C.H., Xu P., Hay B.A., Ruohola-Baker H. Stage-specific differences in the requirements for germline stem cell maintenance in the *Drosophila* ovary. *Cell Stem Cell.* 2007;1(6):698-709. DOI 10.1016/j.stem.2007.11.007
- Slattery M., Voutev R., Ma L., Nègre N., White K.P., Mann R.S. Divergent transcriptional regulatory logic at the intersection of tissue growth and developmental patterning. *PLoS Genet.* 2013;9(9):e1003753. DOI 10.1371/journal.pgen.1003753
- Vo N., Horii T., Yanai H., Yoshida H., Yamaguchi M. The Hippo pathway as a target of the *Drosophila* DRE/DREF transcriptional regulatory pathway. *Sci. Rep.* 2014;4:7196. DOI 10.1038/srep07196
- Weng R., Cohen S.M. Control of *Drosophila* Type I and Type II central brain neuroblast proliferation by *bantam* microRNA. *Development.* 2015;142(21):3713-3720. DOI 10.1242/dev.127209
- Zhang X., Aksoy E., Girke T., Raikhel A.S., Karginov F.V. Transcriptome-wide microRNA and target dynamics in the fat body during the gonadotrophic cycle of *Aedes aegypti*. *Proc. Natl. Acad. Sci. USA.* 2017;114(10):E1895-E1903. DOI 10.1073/pnas.1701474114

Funding. The work was funded by the Russian Science Foundation grant No. 19-14-00051П. The authors gratefully acknowledge the resources provided by the "Molecular and Cellular Biology" core facility of the Institute of Molecular and Cellular Biology of the Siberian Branch of the Russian Academy of Sciences supported by the fundamental scientific research program on the project FWGZ-2021-0014.

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

Received November 7, 2023. Revised November 29, 2023. Accepted November 29, 2023.

DOI 10.18699/vjgb-24-16

Assessing cell lines with inducible depletion of cohesin and condensins components through analysis of metaphase chromosome morphology

A.M. Yunusova¹, A.V. Smirnov¹, I.E. Pristyazhnuk ¹, T.A. Shnaider ¹, E.K. Maltseva², S.D. Afonnikova ¹,
O.A. Gusev ^{3, 4, 5}, N.R. Battulin ^{1, 2} 

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 Life Improvement by Future Technologies (LIFT) Center, Moscow, Russia

⁴ Kazan Federal University, Kazan, Russia

⁵ Endocrinology Research Center, Moscow, Russia

 battulin@gmail.com

Abstract. One of the most productive strategies for finding the functions of proteins is to study the consequences of loss of protein function. For this purpose, cells or organisms with a knockout of the gene encoding the protein of interest are obtained. However, many proteins perform important functions and cells or organisms could suddenly lose fitness when the function of a protein is lost. For such proteins, the most productive strategy is to use inducible protein degradation systems. A system of auxin-dependent protein degradation is often implemented. To use this system, it is sufficient to introduce a transgene encoding a plant-derived auxin-dependent ubiquitin ligase into mammalian cells and insert a sequence encoding a degron domain into the gene of interest. A crucial aspect of development of cell lines engineered for inducible protein depletion is the selection of cell clones with efficient auxin-dependent degradation of the protein of interest. To select clones induced by depletion of the architectural chromatin proteins RAD21 (a component of the cohesin complex) and SMC2 (a component of the condensin complex), we propose to use the morphology of metaphase chromosomes as a convenient functional test. In this work, we obtained a series of clones of human HAP1 cells carrying the necessary genetic constructs for inducible depletion of RAD21 and SMC2. The degradation efficiency of the protein of interest was assessed by flow cytometry, Western blotting and metaphase chromosome morphology test. Based on our tests, we showed that the clones we established with the SMC2 degron effectively and completely lose protein function when induced by auxin. However, none of the HAP1 clones we created with the RAD21 degron showed complete loss of RAD21 function upon induction of degradation by auxin. In addition, some clones showed evidence of loss of RAD21 function even in the absence of induction. The chromosome morphology test turned out to be a convenient and informative method for clone selection. The results of this test are in good agreement with flow cytometry analysis and Western blotting data.

Key words: SMC proteins; degron; chromosome condensation.

For citation: Yunusova A.M., Smirnov A.V., Pristyazhnuk I.E., Shnaider T.A., Maltseva E.K., Afonnikova S.D., Gusev O.A., Battulin N.R. Assessing cell lines with inducible depletion of cohesin and condensins components through analysis of metaphase chromosome morphology. *Vavilovskii Zhurnal Genetiki i Seleksii* = *Vavilov Journal of Genetics and Breeding*. 2024;28(2):138-147. DOI 10.18699/vjgb-24-16

Оценка клеточных линий с индуцируемой деплецией компонентов когезина и конденсинов посредством анализа морфологии метафазных хромосом

A.M. Юнусова¹, A.B. Смирнов¹, И.Е. Пристяжнюк ¹, Т.А. Шнайдер ¹, Е.К. Мальцева², С.Д. Афонникова ¹,
О.А. Гусев ^{3, 4, 5}, Н.Р. Баттулин ^{1, 2} 

¹ Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия

² Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия

³ Life Improvement by Future Technologies (LIFT) Center, Москва, Россия

⁴ Казанский федеральный университет, Казань, Россия

⁵ Национальный медицинский исследовательский центр эндокринологии» Министерства здравоохранения Российской Федерации, Москва, Россия

 battulin@gmail.com

Аннотация. Одна из самых продуктивных стратегий поиска функций различных белков – исследование последствий потери функции белка. Часто для этого получают клетки или организмы с нокаутом гена, кодирующего белок интереса. Однако многие белки выполняют настолько важные функции, что клетка или организм резко теряют жизнеспособность при потере функции такого белка. Для этих белков наиболее продуктивной стратегией является применение систем индуцируемой деградации белка. Часто используют систему ауксин-

зависимой деградации белков. Для применения этой системы достаточно ввести в клетки млекопитающих трансген, кодирующий растительную ауксин-зависимую убиквитин лигазу, и включить в ген интереса последовательность, кодирующую дегроновый домен. Важный этап создания клеток, способных к индуцируемой деплеции белка, – отбор клеточных клонов с эффективной ауксин-зависимой деградацией белка интереса. Для отбора клонов с индуцируемой деплецией архитектурных белков хроматина RAD21 (компонент когезинового комплекса) и SMC2 (компонент конденсинового комплекса) мы предлагаем использовать морфологию метафазных хромосом как удобный функциональный тест. В данной работе мы получили серию клонов клеток человека HAP1, несущих необходимые генетические конструкции для индуцируемой деплеции RAD21 и SMC2. Эффективность деградации белка интереса была оценена с помощью проточной цитофлуориметрии, Вестерн-блоттинга и теста на морфологию метафазных хромосом. На основе проведенных тестов мы продемонстрировали, что созданные нами клоны с дегроном SMC2 эффективно и полно теряют функцию белка при индукции ауксином. При этом ни один из созданных нами клонов HAP1 с дегроном RAD21 не показал полной потери функции RAD21 при индукции деградации ауксином. Кроме того, некоторые клоны имели признаки потери функции RAD21 даже в отсутствие индукции. Использованный нами тест на морфологию хромосом оказался удобным и информативным для отбора клонов. Результаты этого теста хорошо согласуются с данными проточной цитофлуориметрии анализа и Вестерн-блоттинга.

Ключевые слова: SMC белки; дегрон; конденсация хромосом.

Introduction

Chromatin architectural proteins play a crucial role in maintaining the three-dimensional structure of the genome (Kabirova et al., 2023). Among them, special attention is drawn to the cohesin and condensin complexes belonging to the SMC (structural maintenance of chromosomes) family of proteins. Cohesin has many different functions: it ensures cohesion of sister chromatids after replication (Losada et al., 1998), forms loops and TADs (topologically associating domains) via the loop extrusion mechanism (Nuebler et al., 2018), and is involved in the repair of DNA breaks (Litwin et al., 2018). Condensins, on the other hand, organize the loops of metaphase chromosomes during cell division (Gibcus et al., 2018). These functions can be considered critical for maintaining cell life; therefore, homozygous loss-of-function mutations in the genes encoding cohesin subunits are lethal for cells.

The inability to obtain dividing cells without cohesin makes it difficult to study the cohesin complex. To characterize the consequences of cohesin loss, researchers use various tricks. For example, V.C. Seitan et al. utilized a conditional knockout approach to examine the effects of cohesin loss in postmitotic thymocytes *in vivo* (Seitan et al., 2011). The absence of cell division in mature thymocytes makes them more tolerant to the severe consequences of cohesin loss, such as disruption of mitotic mechanics due to loss of chromatid cohesion. However, for use in cell cultures, conditional knockout of proteins crucial for cell division is almost inapplicable, since cells actively proliferate in culture, and DNA excision of a gene fragment using Cre recombinase in a specific cell rarely occurs. Therefore, it takes a long time (several days or weeks) for a knockout to occur in a significant part of the cells.

The study of chromatin architectural proteins has greatly advanced with the development of inducible protein degradation methods. A comprehensive overview of these technologies can be found in the article by E. de Wit and E.P. Nora (2023). Among the various systems used for degrading proteins, the auxin-dependent protein depletion system is currently the most widely employed (Phanindhar, Mishra, 2023). While this system offers great potential for solving scientific problems, its application comes with several challenges that are not adequately addressed in the existing literature.

One of these problems is the selection of clones that carry all the necessary genetic modifications and are truly capable of

inducible depletion of the protein of interest. The difficulty is that the addition of the degron tag can negatively affect protein function, which reduces cell fitness. Under such conditions, a selection advantage is given to cells that have somehow blocked the functioning of the introduced system of auxin-dependent protein degradation, for example, due to epigenetic silencing of exogenous plant ubiquitin ligase (Yunusova et al., 2021). Therefore, clone selection based on Western blotting or loss of protein function tests is of particular importance for the application of auxin degron technology.

Here, we used the involvement of cohesin and condensins in the formation of the metaphase chromosome to assess the completeness of the loss of protein function in cell clones with inducible depletion of these complexes. We showed that assessment of chromosome morphology is a convenient functional test that allows screening of clones.

Materials and methods

Cell culture and cell lines. The human HAP1 cell line (a near-haploid cell line derived from the KBM-7 cell line) was purchased from Horizon Discovery. HCT116-based cell lines with auxin-inducible degron-tagged *RAD21* and *SMC2* genes were kindly provided by Dr. Masato Kanemaki. Cells were maintained at 37 °C in a humidified atmosphere with 5 % CO₂ in growth medium that consisted of IMDM supplemented with 10 % FBS (vol/vol), 2 mM GlutaMAX (all from Thermo Fisher Scientific, USA) and 50 U/ml penicillin/50 mg/ml streptomycin (Capricorn Scientific GmbH, Germany). After reaching 70–80 % confluency, cells were detached with 0.05 % trypsin/EDTA and replated at a 1:3 ratio into new cell culture dishes. Cells subjected to flow cytometric analysis were resuspended in PBS. Flow cytometric analysis was performed on a BD FACSAria (BD Biosciences).

Auxin treatment. For inducing degradation of proteins fused with miniIAA7, 500 μM Indole 3 acetic Acid (IAA, I2886, Sigma-Aldrich, USA) was added directly to the culture medium. HCT116 cell lines were treated with 1 μM 5-Ph-IAA at the appropriate time intervals.

Plasmids and constructs. Donor vectors with homology arms, degron tag (miniIAA7-eGFP) and selection cassette were assembled in one reaction using the Gibson Assembly (NEBuilder HiFi DNA Assembly Master Mix, NEB, USA) in the pMK290 backbone (Addgene, 72828). miniIAA7-eGFP

List of used primers and sgRNAs

Primers	Target	Sequence (5'–3')
Homology arms	<i>RAD21</i>	RAD21 HAR_F: gcgaattggagctccccgggatccactttggcctttaccctcttga RAD21 HAR_R: ttggcgctgcaccggatcctataatggaaccttggtccagg RAD21 HAL_F: acgcacggtgtgggtcggtggagctagaagcattatagctagtg RAD21 HAL_R: aacaaaagctgggtaccggatcccactgaagtctgagttcaaaagtg
Homology arms	<i>SMC2</i>	SMC2 HAR_F: gcgaattggagctccccgggatcctaggagctggggctgcaaaa SMC2 HAR_R: ttggcgctgcaccggatccaacttcgacgtgtgctccttt SMC2 HAL_F: tccctgaagaggttactaactactaaaagtatttctcatcttg SMC2 HAL_R: ggaacaaaagctgggtaccggatcctcatggtgtctgtagtgca
sgRNA	<i>RAD21</i>	CCAAGGTTCCATATTATATA
sgRNA	<i>SMC2</i>	ACCACCCAAAGGAGCACATG
sgRNA	<i>AAVS1</i>	GGGGCCACTAGGGACAGGAT
Genotyping	<i>RAD21</i>	hRAD21_F: CAGCGTGCTCTTGCTAAACT hRAD21_R: AAGATTGCCAGTGTACTGATGGAA hRAD21_F1: CACAGGGAGTGATTGATAAGGGA hRAD21_R1: TGGGGGCAATTTGTAAGCAC hRAD21_F2: GCTGACACAGGAAGAACCCTA hRAD21_R2: TCAAGAGGGTGACCATTGTTGT
Genotyping	<i>SMC2</i>	hSMC2_F: CAAGCAGTCAACCACCAGGA hSMC2_R: TCACAACCACAATAATTGGACCAT hSMC2_out_F: CATGTGACACTTGATGGGGGA hSMC2_out_R: GTACGGCCATATATCAGGGGA
Genotyping	<i>eGFP</i>	GFP_R: CAGCTCGACCAGGATGGG

fragment was amplified from the vector pSH-EFIRE5-B-Scipin-miniIAA7-mEGFP (Addgene, 129719). Homology arms for *RAD21* (NCBI Entrez Gene ID: 5885) and *SMC2* (NCBI Entrez Gene ID: 10592) were PCR-amplified from human genomic DNA with Q5 polymerase (NEB, USA). For auxin receptor F-box protein overexpression AtAFB2 plasmids were used (pSH-EFIRE5-P-AtAFB2-mCherry-weak NLS vector (Addgene, 129717)). SgRNA targeting the last codon of the *RAD21* and *SMC2* genes were cloned into gRNA_Cloning Vector (Addgene, 41824). For CRISPR/Cas9 gene targeting the human codon-optimized Cas9 expression plasmid (Addgene, 41815) was used. The list of primers and gRNA sequences is shown in Table.

Generation of HAP1 cell lines with an auxin-inducible degron system. Generation of degron cell lines was done essentially as previously described (Yunusova et al., 2021). Briefly, HAP1 cells were electroporated at conditions of 1200 V, 30 ms, 1 pulse, using the Neon Transfection system (Thermo Fisher Scientific), according to the manufacturer's instructions with minor modifications. Per electroporation, 250,000 cells were resuspended in 10 µL of DPBS containing 1 µg of the plasmids with the ratio 1:1:2 (gRNA:Cas9:Donor vector for recombination, accordingly). Immediately following pulsation, cells were transferred into pre-warmed cell media without antibiotics. The next day cells were split into 10 cm dishes at 1:4 and 1:10 dilutions and placed under Hygromycin B selection (0.8 mg/ml) or puromycin selection (1 µg/ml). The medium was replaced every three days. After 10 to 14 days of selection, single-cell clones were visible, and a subset of clones was handpicked with pipette tips under microscope.

Then part of the cells were lysed in PBND lysis buffer (0.2 mg/ml proteinase K, 10 mM Tris-HCl pH 8, 50 mM KCl, 2.5 mM MgCl₂, 0.45 % (v/v) NP-40, and 0.45 % (v/v) Tween 20) or DNA extraction buffer (0.2 mg/ml proteinase K, 10 mM TrisHCl pH 8, 100 mM NaCl, 25 mM EDTA-Na₂, 0.5 % SDS) for 1 h at 55 °C followed by proteinase K inactivation for 10 min at 95 °C. The target regions were amplified by PCR with HS-Taq DNA Polymerase (Biolabmix, Russia). The parameters were as follows: 95 °C for 30 s, then 34 cycles of 95 °C for 10 s, 60 °C (unless otherwise stated) for 30 s, 72 °C for 1 min/kb, and a final step at 72 °C for 5 min. The amplified products were analyzed by agarose gel electrophoresis.

Protein detection. The cells were washed twice with PBS and scraped from the surface in the presence of RIPA buffer (50 mM Tris-HCl pH 8, 150 mM NaCl, 1 % Triton X-100, 0.5 % sodium deoxycholate, and 0.1 % SDS) containing the protease inhibitor cocktail (1x Complete ULTRA, 1x PhosSTOP (both from Roche, Switzerland), 5 mM NaF (Sigma-Aldrich)). After that, the cells were sonicated by three 10 s pulses at 33–35 % power settings with UW 2070 (Bandelin Electronics, Germany). Lysates were centrifuged at 14,000 g for 20 min at 2 °C, frozen, and stored at –80 °C. The protein concentrations in cell lysates were quantified using Pierce BCA Protein Assay Kit (Thermo Fisher Scientific). Equal amounts (20 µg) of total protein were separated on 10 % SDS-PAGE and then transferred onto the Immun-Blot PVDF membrane (Bio-Rad, USA). After blocking in 5 % milk/TBST for 2 h, the membrane was incubated with primary antibodies against RAD21 and SMC2 (#12673/#8720, Cell Signaling Technology, USA) at 4 °C overnight. On the following day, membranes were incubated with horseradish peroxidase-con-

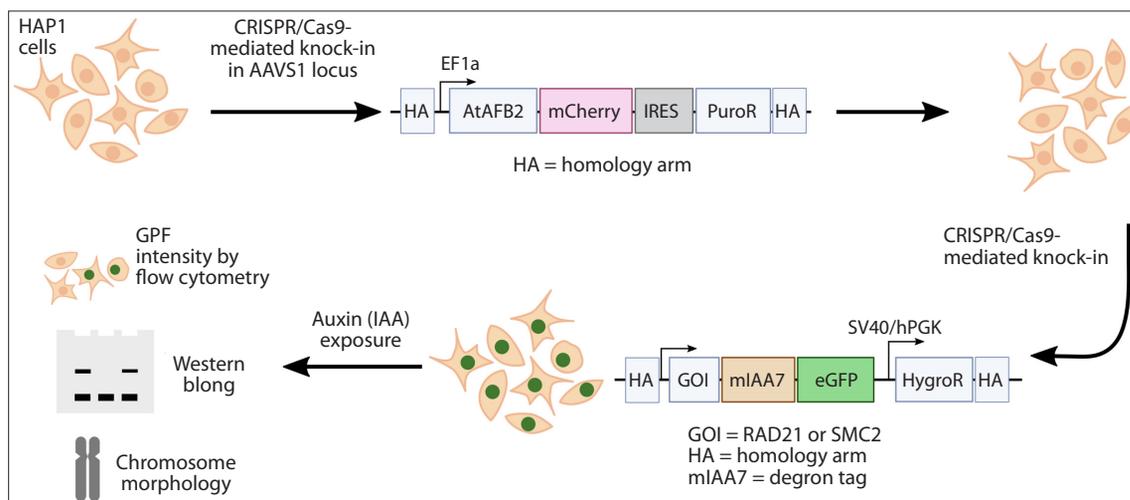


Fig. 1. Overview of experimental design.

jugated secondary antibodies (#7074, Cell Signaling Technology) for 2 h at room temperature. Detection was performed with Clarity™ (Bio-RAD) and detected iBright™ FL1500 (Thermo Fisher Scientific).

Chromosome spread. Chromosome preparations were made according to the previously described protocol with minor modifications (Kruglova et al., 2008). Briefly, human cell cultures were exposed to 50 ng/ml Colcemid (Merck KGaA, Darmstadt, Germany) for 3 h followed by auxin treatment for 2 hours. Afterwards, cells were detached with 0.05 % Trypsin-EDTA solution (Capricorn Scientific GmbH, Germany) and resuspended in hypotonic solution (0.38 M KCl) for 15 min at 37 °C. Then, cells were fixed with Carnoy fixative (3:1 methanol: glacial acetic acid), dropped onto cold wet glass slides, and stained with 1 µg/ml 4',6-diamidino-2-phenylindole (DAPI) (Sigma-Aldrich). The samples were analyzed using a Carl Zeiss Axioscop 2 fluorescence microscope at the Center for Collective Use of Microscopy of the Institute of Cytology and Genetics SB RAS (Novosibirsk, Russia). Image processing was carried out using ISIS software (MetaSystems GmbH, Germany). At least 50 metaphase plates were analyzed for each experimental group. Three categories of metaphase plates were distinguished: metaphases with separated chromatids, metaphases with non-separated chromatids, and an intermediate category of metaphase plates in which the chromatids either lay parallel, in close proximity, but were not in contact, or some of the chromosomes were with non-separated chromatids.

Results

Introduction of modifications into the genome of cells

In this study, we implemented a clone selection system based on chromosome morphology to obtain HAP1 cells capable of inducible depletion of RAD21 (a component of the cohesin complex) and SMC2 (a component of the condensin I and II complexes).

To generate these cells, we employed the methodology described in the study by (Yunusova et al., 2021). The process of obtaining cell lines with auxin-dependent degradation of the

protein of interest involved two rounds of genome modification, as illustrated in Fig. 1. Firstly, we performed targeted integration of an exogenous construct that encodes a degron fused with the *eGFP* gene and a selectable marker in front of the stop codon of the gene of interest. Secondly, we introduced the *AtAFB2* gene, a component of plant ubiquitin ligase and an auxin receptor, into the cell genome through either random or targeted integration. In fact, modification of the gene of interest and integration of ubiquitin ligase *AtAFB2* can be carried out in the reverse order, since effective operation of the system occurs only in cells that have both modifications.

In the presence of auxin, the *AtAFB2* protein interacts with the degron domain, leading to polyubiquitination and subsequent degradation of the chimeric protein in the proteasome. This system allows for the controlled depletion of the protein of interest upon auxin induction.

PCR genotyping of cell clones with a degron

It is crucial to modify all alleles of the gene of interest to prevent residual wild-type alleles from maintaining protein function. For this study, we selected the human cell line HAP1 as the experimental cell line. The near haploid karyotype of HAP1 cells simplifies the process of obtaining modified sub-clones, as modification of a single allele of the gene of interest is sufficient. However, the presence of pseudogenes poses a significant challenge in clone selection, as they can complicate the identification of clones with the desired modification.

One such pseudogene, *RAD21P1*, located on the X chromosome, shares a high degree of homology (over 93 %) with the C-terminal fragment of the *RAD21* gene (Supplementary Material 1)¹. It is in this region that we inserted an exogenous construct containing a degron tag, *eGFP*, and a selectable marker (see Fig. 1). Therefore, the selection of appropriate primers for PCR genotyping is a critical step in the workflow. We carefully selected and tested several pairs of primers, as well as their combinations (see the Table), using DNA isolated from the HCT116 *RAD21_mAC* cell line as a positive control sample. These cells were obtained and intensively charac-

¹ Supplementary Materials 1 and 2 are available at: <https://vavilovj-icg.ru/download/pict-2024-28/appx7.pdf>

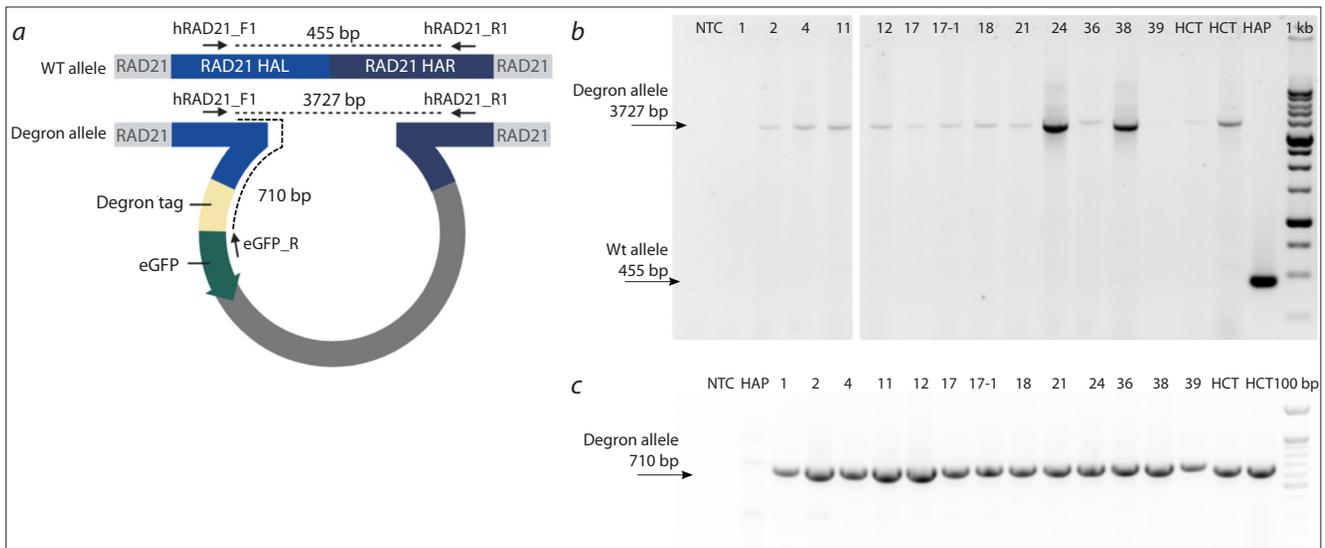


Fig. 2. PCR genotyping of knockin in the *RAD21* gene.

a – scheme of primer annealing for the wild-type allele and the knockin allele; *b* – results of genotyping with primers hRAD21_F1-hRAD21_R1. It can be seen that none of the 13 cell clones tested contain the wild-type allele; *c* – genotyping results with primers hRAD21_F1-GFP_R. It can be seen that all 13 cell clones tested have the knockin allele. 1, 2, 4, 11, 12, 17, 17-1, 18, 21, 24, 36-1, 38, 39 – numbers of the tested clones; HCT – HCT116 *RAD21*_mAC; HAP – HAP1 without modifications; NTC – no template control; 1 kb and 100 bp ladders.

terized by another group (Yesbolatova et al., 2020), but since the design of the modification of the endogenous *RAD21* gene was similar to what we used, these cells can be used as a reference for interpreting our results. Intact wild-type human DNA and DNA from one of the clones with targeted modification of the *RAD21* locus, obtained in our laboratory, were also used as test samples. Indeed, some primer combinations amplified a nonspecific PCR product from the pseudogene even in the absence of a wild-type allele. Therefore, to genotype the selected clones, we used the hRAD21_F1/hRAD21_R1 primer combination (Fig. 2, *a*), which amplified only a specific product. PCR genotyping of clones with modification of the *SMC2* gene did not reveal amplification of nonspecific fragments.

At the first stage, the *AtAFB2* gene was integrated into the *AAVS1* safe-harbor locus, and puromycin-resistant cells were selected. These cells were then used for the second round of modification – insertion – of a construct with a degron at the end of the *RAD21* gene. After selection on hygromycin, we selected more than a hundred colonies and performed PCR genotyping with primers hRAD21_F1/hRAD21_R1. Using the selected primers (see Fig. 2, *a*), we genotyped cell clones that had successfully passed selection for resistance to the antibiotic hygromycin. Figure 2, *b*, *c* shows the results of genotyping of 13 selected cell clones carrying the degron tag modification in the *RAD21* gene.

The same genotyping strategy was used to establish HAP1 cell lines with the *SMC2* degron (Supplementary Material 2).

Assessment of the degree of protein depletion upon induction of degradation by auxin

The degree of degradation of the chimeric protein RAD21_miniIAA7_eGFP was assessed by the number of GFP-positive cells on a flow cytometer after 2 hours of exposure to auxin. Most of the clones demonstrated low efficiency of degradation of the target protein (Fig. 3, *a*). We assume that this is due to

low expression in cells of the auxin receptor *AtAFB2*, which is integrated into the *AAVS1* locus. Since we used *AtAFB2* fused to the fluorescent protein mCherry (see Fig. 1), its expression level can be detected using a flow cytometer. Indeed, the level of mCherry fluorescence was higher in clones with more efficient degradation of RAD21_miniIAA7_eGFP (data not shown). We believe that integration of *AtAFB2* transgene into a random location in the genome rather than into the *AAVS1* locus is a better strategy, since in this case it is possible to select for clones in which the insertion provides strong, persistent expression of *AtAFB2*. This might happen due to the selection of clones with a high copy number of the insertion or with a successful epigenetic landscape at the site of integration. Based on this analysis, we continued to functionally characterize the depletion efficiency of only those clones that showed a degree of degradation comparable to the HCT116 *RAD21*_mAC cell line (Yesbolatova et al., 2020).

A similar analysis was carried out for subclones with the chimeric protein *SMC2*_miniIAA7_eGFP. In this case, both subclones tested had a degree of chimeric protein degradation comparable to the control sample (*SMC2*_mAID_Clover) (see Fig. 3, *b*). *SMC2*_mAID_Clover cells were previously obtained using a similar strategy by tagging the *SMC2* gene in (Yesbolatova et al., 2020).

Selected cell clones were assessed for loss of protein function by analyzing metaphase chromosome morphology.

Assessment of loss of function of cohesin and condensins based on chromosome morphology analysis

Since the architectural proteins of chromatin – the cohesin and condensin complexes – have a well known function in the formation of mitotic chromosomes, changes in chromosome morphology can serve as a convenient criterion for assessing the function of these proteins. Therefore, we assessed the proportion of cells with abnormal chromosome morphology in

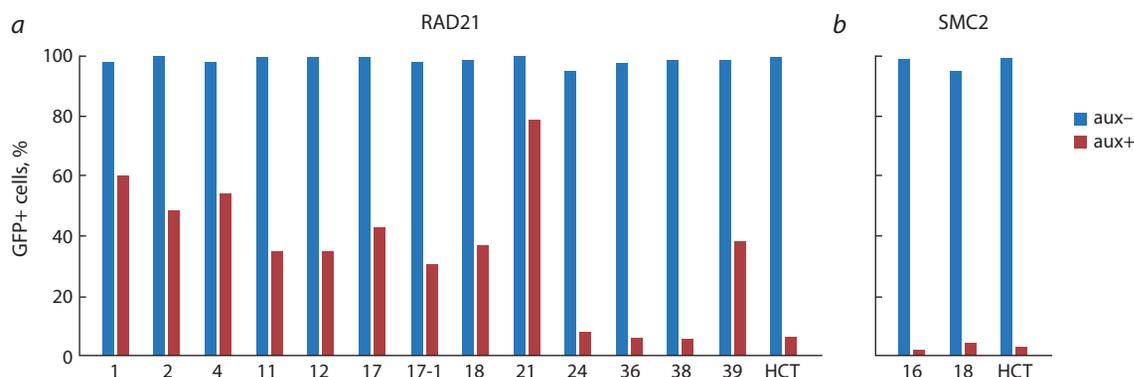


Fig. 3. Evaluation of degradation efficiency of a protein of interest using flow cytometry.

a – the proportion of cells that have not degraded (GFP+) fusion protein RAD21_{miniIAA7_eGFP}. 1, 2, 4, 11, 12, 17, 17-1, 18, 21, 24, 36, 38, 39 – numbers of tested clones; HCT – positive control, cells with efficient degradation of the protein of interest HCT116 RAD21_{mAC}; *b* – the proportion of cells that have not degraded (GFP+) fusion protein SMC2_{miniIAA7_eGFP}. 16, 18 – numbers of tested clones; HCT – positive control, cells with efficient degradation of the protein of interest HCT116 SMC2_{mAC}.

clones before and after depletion of the target protein induced by auxin exposure. Cohesin ensures cohesion of sister chromatids; therefore, in clones with RAD21_{miniIAA7_eGFP}, we counted the number of metaphase plates with unconnected chromatids (Fig. 4). Condensins are responsible for the compaction of metaphase chromosomes and the formation of their rod-shaped morphology; therefore, for clones with SMC2_{miniIAA7_eGFP}, we counted the number of metaphase plates with non-compacted chromosomes (Fig. 5).

It is clearly seen from Fig. 4 that in unmodified cells, cohesion of sister chromatids is observed in the vast majority of metaphase plates (see Fig. 4, *a*). We use HCT116 RAD21_{mAC} cells as a positive control, since these cells demonstrate effective depletion of RAD21 (Yesbolatova et al., 2020). For these cells, without auxin exposure, all plates have chromosomes with sister chromatid cohesion. And when depletion of RAD21 is induced by auxin, no plates with normal chromosome morphology remain, and the vast majority of plates contain separate chromatids (see Fig. 4, *b*). This shows that the chromosome morphology test allows us to assess cohesin function both before and after induction of RAD21 depletion by auxin. None of the three HAP1 cell clones (see Fig. 4, *c–e*) showed loss of cohesin function comparable to the positive control. In clones 24 and 38, there is no loss of chromatid cohesion in most laminae. And in clone 36, even in the absence of auxin induction, 25 percent of the plates did not have chromatid cohesion, that is, cohesin function was impaired without induction.

In the case of the SMC2 degren, both selected clones showed the same complete loss of condensin function after induction as the control and no evidence of protein function deficiency without auxin induction (see Fig 5).

Evaluation of cohesin and condensins depletion based on Western blotting

Assessing the efficiency of degradation of a protein of interest based on flow cytometry is indirect, since in this way only the protein fused with degren and fluorescent domains will be included in the analysis. It is possible to propose several scenarios in which, despite the degradation of the fusion protein,

functional molecules of the protein of interest remain in the cells (we discuss these scenarios in Discussion). Therefore, it is important to complement the analysis of the morphology of metaphase chromosomes with the Western blotting methods.

Figure 6, *a* shows the results of Western blotting of the RAD21 protein. It is clearly seen that the amount of RAD21 protein in unmodified cells is greater than in modified cells. In positive control cells HCT116 RAD21_{mAC} after induction with auxin, RAD21 is not detected at all. In clones 24, 36, 38, two forms of the protein are detected, a heavier one corresponding to the fusion protein with GFP and a lighter one corresponding to the wild-type protein. After exposure to auxin, fusion protein decreases (clones 24, 38) or completely disappears (clone 36). However, in clones 24, 36, 38, the wild-type form of the RAD21 protein does not decrease after auxin induction.

Western blotting of the SMC2 protein carried out for clone 16 also showed that knockin of the degren domain into the SMC2 gene leads to a decrease in the amount of protein compared to unmodified cells. After induction of protein degradation by auxin, SMC2 is not detected (see Fig. 6, *b*).

Discussion

Protein function studies are often based on the loss of function strategy. However, for many proteins, loss of function will have a dramatic impact on viability. The use of inducible depletion systems makes it possible to apply a loss of function approach to such proteins. However, like any complex technology, inducible protein depletion has its own limitations. In this article, we described an approach based on the use of chromosome morphology as a convenient functional test for selecting clones with inducible depletion of the architectural chromatin proteins cohesin and condensins.

The difficulty of obtaining cell clones capable of inducible protein depletion appears to depend greatly on the properties of the particular protein of interest. In our work, clones capable of effective loss of SMC2 function were obtained easily, without any trouble at the stage of genotyping the clones or at subsequent stages of assessing the completeness of degradation. At the same time, the attempt to obtain a clone capable of

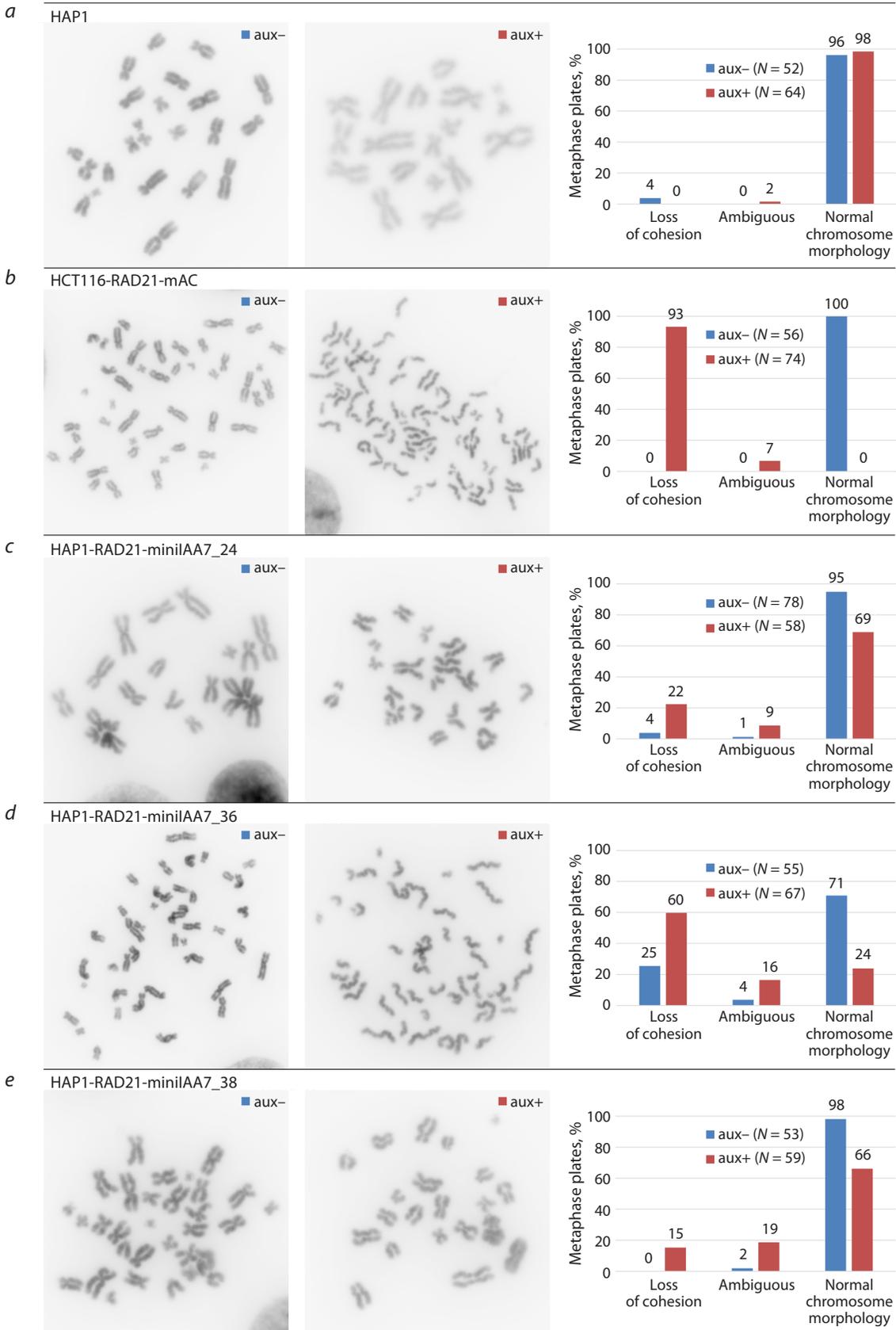


Fig. 4. Assessment of changes in chromosome morphology after induction of RAD21 depletion.

For each cell clone, one metaphase plate is presented, illustrating the most common chromosome morphology without induction (aux-) and after induction (aux+). Also presented is a diagram summarizing the proportion of metaphase plates with loss of chromatid cohesion, normal chromosome morphology and ambiguous. For each condition, the number of analyzed metaphase plates is indicated – N.

a – HAP1 cells without modifications – negative control; b – HCT116 RAD21_mAC, cells with effective RAD21 depletion – positive control; c–e – cell clones tested.

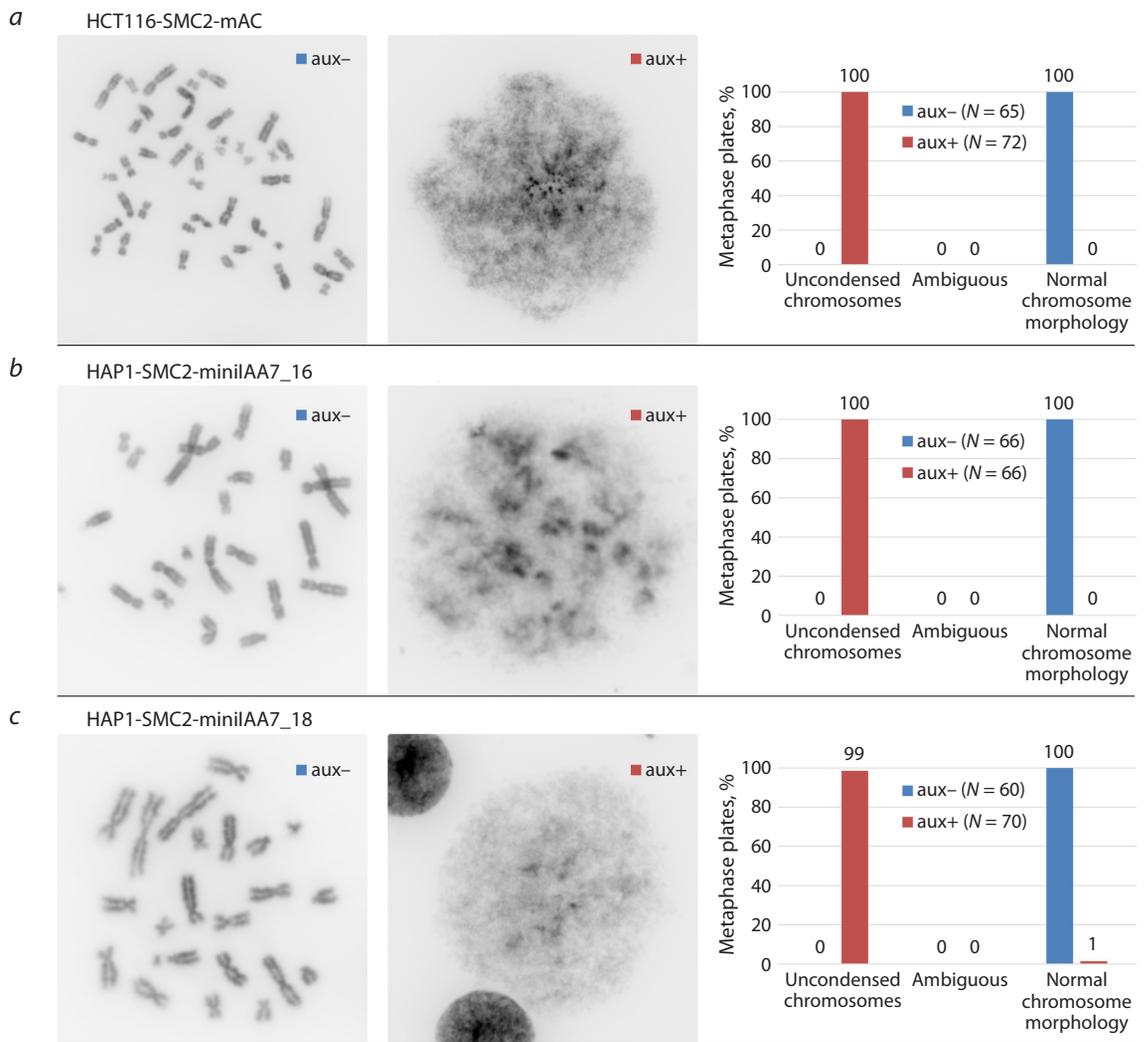


Fig. 5. Assessment of changes in chromosome morphology after induction of SMC2 depletion.

For each cell clone, one metaphase plate is presented, illustrating the most common chromosome morphology without induction (aux-) and after induction (aux+). Also presented is a diagram summarizing the proportion of metaphase plates with uncondensed chromosomes, normal chromosome morphology and ambiguous. For each condition, the number of analyzed metaphase plates is indicated – N. a – HCT116-SMC2-mAC; cells with effective SMC2 depletion – positive control; b, c – cell clones tested.

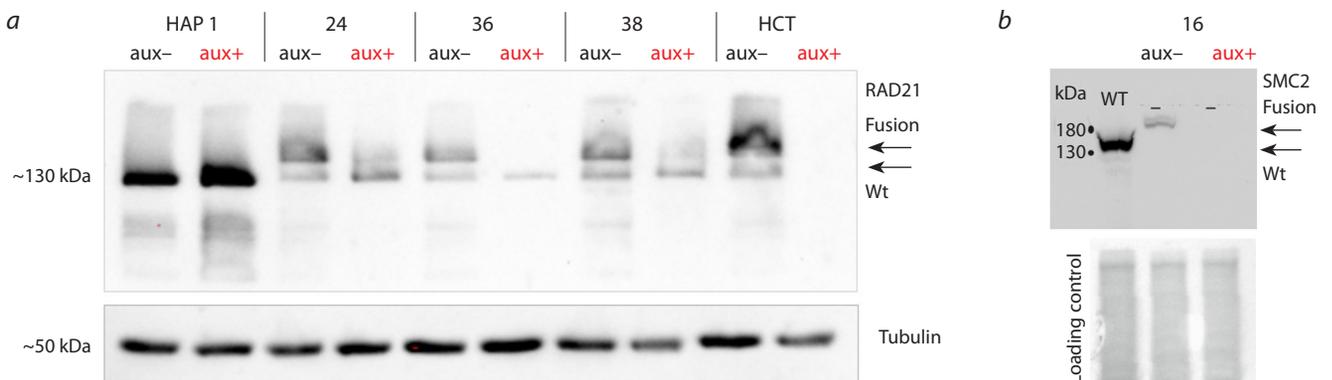


Fig. 6. Identification of proteins of interest using Western blotting.

a – Western blotting of RAD21 in cells without modification (HAP), in cells with effective depletion of RAD21 (HCT) and three cell clones tested (24, 36, 38) without induction (aux-) and after induction of depletion (aux+). For RAD21, the sizes of the fusion protein and the wild-type protein are indicated. Tubulin detection was used as an internal control; b – Western blotting of SMC2 in cells without modification (wt), and cell clone tested (16) without induction (aux-) and after induction of depletion (aux+). Total protein staining (loading control).

efficient degradation of the RAD21 protein was not successful: the clones we obtained suffer from loss of protein function without auxin induction and incomplete protein degradation after induction. This is clearly visible when comparing these parameters with HCT116 RAD21_mAC cells, which we chose as a positive control (these cells are capable of effective auxin-induced depletion of RAD21) (Yesbolatova et al., 2020).

We believe that in the case of RAD21, the main problem is the decrease in the amount of RAD21 in cells even in the absence of auxin induction (“basal degradation”). This decrease may have two independent causes: 1) instability of the transcript/protein due to fusion with the degron tag and GFP. Theoretically, introduction of a genetic construct could reduce the knockin transcription efficiency of *RAD21*; cause a decrease in transcript stability; reduce the efficiency of translation of the knockin transcript; cause a decrease in the stability of the fusion protein (Yu et al., 2015); 2) background ubiquitin ligase activity in the absence of auxin. Such activity will lead to polyubiquitinylation and degradation of the protein carrying the degron tag (Li et al., 2019). Apparently, both occur in our case. To address the first cause, one can do little except make a fusion not with the C-terminus, but with the N-terminus of the peptide. But the second cause depends entirely on the properties of the peptide used and the auxin-dependent ubiquitin ligase. We used the ubiquitin ligase AtAFB2, which is thought to have lower basal activity compared to OsTIR1 (Li et al., 2019). However, a system with even lower activity has recently emerged. It was created based on OsTIR1 by replacing the ligand with a synthetic auxin analogue 5-Ph-IAA and rationally designing the active site of the enzyme (Yesbolatova et al., 2020). Our positive control cells are designed with just such a system. The Western blot in Fig. 6 clearly shows that the amount of the heavy fusion form of the protein in the positive control is significantly greater than in clones 24, 36, 38 that we created based on AtAFB2. This may be explained by the higher basal activity of AtAFB2. In addition, the use of alternative depletion systems, such as dTAG, may be a productive strategy for achieving degradation of architectural chromatin proteins (Nabet et al., 2018).

As we noted above, the key stage in obtaining cells capable of inducible depletion is the complete modification of all alleles of the gene of interest. All cell clones used in this work did not have wild-type alleles when analyzed by PCR (see Fig. 2). However, Western blotting reveals a band corresponding in size to unmodified RAD21. Since there are no wild-type alleles during genotyping, this eliminates the possibility of contamination of samples with unmodified cells. Theoretically, repair of the break introduced by Cas9 can lead to the loss of the primer annealing site (we previously reported this issue (Korablev et al., 2020)); such damaged alleles will not be detected by PCR, however, they can produce a functional transcript. But for haploid HAP1 cells, this explanation does not apply since each clone contains exactly one modified *RAD21* allele. Alternative explanations suggest that wild-type *RAD21* appears in cells containing the correct modification of *RAD21*. For example, a transcript from a modified allele can be spliced and, as a result, a fusion peptide is not formed. During translation, a peptide bond may also not be formed, for example, like in 2A viral peptides. One can speculate

that something similar may form between the fusion parts of the peptide. However, both of these explanations in our case remain at the level of speculation.

For some genes, pseudogenes could be the source of a functional peptide (Zhang et al., 2023); however, in the case of *RAD21*, the pseudogene is highly mutated and contains many stop codons and therefore does not have a functional open reading frame.

Of course, the most direct way to assess depletion of a protein of interest is Western blotting, but this method is highly dependent on the quality and specificity of the antibodies. It is known that the specificity of many antibodies is questioned (Baker, 2015).

Conclusion

Therefore, for many proteins, the use of Western blotting can be problematic. Therefore, a convenient functional test is an excellent way to characterize cells capable of induced depletion of a protein of interest. The chromosome morphology test used in this work proved to be very informative.

References

- Baker M. Reproducibility crisis: Blame it on the antibodies. *Nature*. 2015;521(7552):274-276. DOI 10.1038/521274a
- de Wit E., Nora E.P. New insights into genome folding by loop extrusion from inducible degron technologies. *Nat. Rev. Genet.* 2023; 24(2):73-85. DOI 10.1038/s41576-022-00530-4
- Gibcus J.H., Samejima K., Goloborodko A., Samejima I., Naumova N., Nuebler J., Kanemaki M.T., Xie L., Paulson J.R., Earnshaw W.C., Mirny L.A., Dekker J. A pathway for mitotic chromosome formation. *Science*. 2018;359(6376):eaao6135. DOI 10.1126/science.aao6135
- Kabirova E., Nurislamov A., Shadskiy A., Smirnov A., Popov A., Salnikov P., Battulin N., Fishman V. Function and evolution of the loop extrusion machinery in animals. *Int. J. Mol. Sci.* 2023;24(5):5017. DOI 10.3390/ijms24055017
- Korablev A., Lukyanchikova V., Serova I., Battulin N. On-target CRISPR/Cas9 activity can cause undesigned large deletion in mouse zygotes. *Int. J. Mol. Sci.* 2020;21(10):3604. DOI 10.3390/ijms21103604
- Kruglova A.A., Kizilova E.A., Zhelezova A.I., Gridina M.M., Golubitsa A.N., Serov O.L. Embryonic stem cell/fibroblast hybrid cells with near-tetraploid karyotype provide high yield of chimeras. *Cell Tissue Res*. 2008;334(3):371-380. DOI 10.1007/s00441-008-0702-9
- Li S., Prasanna X., Salo V.T., Vattulainen I., Ikonen E. An efficient auxin-inducible degron system with low basal degradation in human cells. *Nat. Methods*. 2019;16(9):866-869. DOI 10.1038/s41592-019-0512-x
- Litwin I., Pilarczyk E., Wysocki R. The emerging role of cohesin in the DNA damage response. *Genes (Basel)*. 2018;9(12):581. DOI 10.3390/genes9120581
- Losada A., Hirano M., Hirano T. Identification of *Xenopus* SMC protein complexes required for sister chromatid cohesion. *Genes Dev*. 1998;12(13):1986-1997. DOI 10.1101/gad.12.13.1986
- Nabet B., Roberts J.M., Buckley D.L., Paulk J., Dastjerdi S., Yang A., Leggett A.L., Erb M.A., Lawlor M.A., Souza A., Scott T.G., Vittori S., Perry J.A., Qi J., Winter G.E., Wong K.-K., Gray N.S., Bradner J.E. The dTAG system for immediate and target-specific protein degradation. *Nat. Chem. Biol.* 2018;14:431-441. DOI 10.1038/s41589-018-0021-8
- Nuebler J., Fudenberg G., Imakaev M., Abdennur N., Mirny L.A. Chromatin organization by an interplay of loop extrusion and compartmental segregation. *Proc. Natl. Acad. Sci. USA*. 2018;115(29):E6697-E6706. DOI 10.1073/pnas.1717730115

- Phanindhar K., Mishra R.K. Auxin-inducible degron system: an efficient protein degradation tool to study protein function. *Biotechniques*. 2023;74(4):186-198. DOI 10.2144/btn-2022-0108
- Seitan V.C., Hao B., Tachibana-Konwalski K., Lavagnoli T., Mirabontenbal H., Brown K.E., Teng G., Carroll T., Terry A., Horan K., Marks H., Adams D.J., Schatz D.G., Aragon L., Fisher A.G., Krangel M.S., Nasmyth K., Merckenschlager M. A role for cohesin in T-cell-receptor rearrangement and thymocyte differentiation. *Nature*. 2011;476(7361):467-471. DOI 10.1038/nature10312
- Yesbolatova A., Saito Y., Kitamoto N., Makino-Itou H., Ajima R., Nakano R., Nakaoka H., Fukui K., Gamo K., Tominari Y., Takeuchi H., Saga Y., Hayashi K., Kanemaki M.T. The auxin-inducible degron 2 technology provides sharp degradation control in yeast, mammalian cells, and mice. *Nat. Commun.* 2020;11(1):5701. DOI 10.1038/s41467-020-19532-z
- Yu K., Liu C., Kim B.-G., Lee D.-Y. Synthetic fusion protein design and applications. *Biotechnol. Adv.* 2015;33(1):155-164. DOI 10.1016/j.biotechadv.2014.11.005
- Yunusova A., Smirnov A., Shnaider T., Lukyanchikova V., Afonnikova S., Battulin N. Evaluation of the OsTIR1 and AtAFB2 AID systems for genome architectural protein degradation in mammalian cells. *Front. Mol. Biosci.* 2021;8:757394. DOI 10.3389/fmolb.2021.757394
- Zhang M., Zhao Y., Liu X., Ruan X., Wang P., Liu L., Wang D., Dong W., Yang C., Xue Y. Pseudogene MAPK6P4-encoded functional peptide promotes glioblastoma vasculogenic mimicry development. *Commun. Biol.* 2023;6(1):1059. DOI 10.1038/s42003-023-05438-1

Funding. We thank Masato Kanemaki for sharing the HCT-116 RAD21-mAC and SMC2-mAC cell lines. This work was supported by Russian Science Foundation grant No. 23-74-00055. Cell culture was performed at the Collective Center of ICG SB RAS "Collection of Pluripotent Human and Mammalian Cell Cultures for Biological and Biomedical Research", project number FWNR-2022=0019 (<https://ckp.icgen.ru/cells/>; http://www.biores.cytogen.ru/brc_cells/collections/ICG_SB_RAS_CELL). Cryoarchiving and primary analysis of control cell lines HCT-116 RAD21-mAC and SMC2-mAC was supported by the Ministry of Science and Higher Education of the Russian Federation, grant 075-15-2021-1344.

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

Received November 14, 2023. Revised November 29, 2023. Accepted December 4, 2023.

DOI 10.18699/vjgb-24-17

Variability of Scots pine (*Pinus sylvestris* L.) plus trees in the Middle and Upper Volga Region with the use of ISSR markers

O.V. Sheikina , E.M. Romanov 

Volga State University of Technology, Yoshkar-Ola, Russia
 shejkinaov@volgatech.net

Abstract. One of the serious issues in forest breeding is how to reduce the variability level in breeding populations of forest tree species that is a set of selected plus trees. The problem is that variability is jeopardized by the risk of losing the genetic diversity of future artificial forests, as well as emerging inbreeding depression in the seed plus trees progeny. DNA markers are an effective tool to study variability, identify features of the genetic structure and degree of plant differentiation. The research focuses on assessing the level of the genetic diversity and the degree of differentiation of plus trees of various geographic origin with the use of ISSR markers. We used six ISSR primers to study 270 plus trees grown in the Penza region, the Chuvash Republic, the Republic of Tatarstan and the Mari El Republic. The samples of plus trees under study were characterized by different levels of genetic diversity. Two hundred fifteen PCR fragments were identified for six ISSR primers in total, while the number of amplified fragments varied from 186 to 201 in different plus trees samples. The genetic variability varied within the following limits: 95.7–96.9 %, polymorphic loci; 1.96–1.97, the number of alleles per locus; 1.31–1.48, the number of effective alleles per locus; finally, 0.291–0.429, Shannon's index; 0.205–0.298, the expected heterozygosity. According to the analysis of molecular variance (AMOVA), 82 % of the variability of ISSR markers is typical for the plus tree samples, while only 18 % is variability among the compared groups of trees from different geographical zones. The dendrogram generated by UPGMA showed that the plus trees grown in the Penza region, the Chuvash Republic and the Republic of Tatarstan are similar in term of the genetic structure of plus trees, while the plus gene pool of Scots pine from the Mari El Republic stands alone. The results of the research prove that the level of genetic diversity, the structure of genetic variability, and the nature of differentiation of plus trees are consistent with those previously elicited for natural populations of Scots pine in the Middle and Upper Volga region.

Key words: *Pinus sylvestris* L.; plus trees; genetic diversity; differentiation; ISSR markers.

For citation: Sheikina O.V., Romanov E.M. Variability of Scots pine (*Pinus sylvestris* L.) plus trees in the Middle and Upper Volga Region with the use of ISSR markers. *Vavilovskii Zhurnal Genetiki i Seleksii* = *Vavilov Journal of Genetics and Breeding*. 2024;28(2):148-154. DOI 10.18699/vjgb-24-17

Изменчивость плюсовых деревьев сосны обыкновенной (*Pinus sylvestris* L.) в Среднем и Верхнем Поволжье по ISSR-маркерам

О.В. Шейкина , Е.М. Романов 

Поволжский государственный технологический университет, Йошкар-Ола, Россия
 shejkinaov@volgatech.net

Аннотация. Снижение уровня изменчивости селекционных популяций лесных древесных видов, представляющих собой совокупность отобранных плюсовых деревьев, считается одной из ключевых проблем в лесной селекции. Она связана с опасностью потери генетического разнообразия будущих искусственно созданных лесов, а также с риском возникновения инбредной депрессии семенного потомства плюсовых деревьев. Эффективным инструментом для изучения изменчивости, определения особенностей генетической структуры и степени дифференциации растений являются ДНК-маркеры. Наше исследование направлено на оценку уровня генетического разнообразия и степени дифференциации плюсовых деревьев разного географического происхождения с применением ISSR-маркеров. С использованием шести ISSR-праймеров изучено 270 плюсовых деревьев из Пензенской области, Чувашской Республики, Республик Татарстан и Марий Эл. Сравнимые выборки характеризовались разным уровнем генетического разнообразия. Всего для шести ISSR-праймеров обнаружено 215 ПЦР-фрагментов, при этом у разных выборок число амплифицированных фрагментов варьировало от 186 до 201. Основные показатели генетической изменчивости находились в следующих пределах: доля полиморфных локусов 95.7–96.9 %, число аллелей на locus 1.96–1.97, число эффективных аллелей 1.31–1.48, индекс Шеннона 0.291–0.429, ожидаемая гетерозиготность 0.205–0.298. По результатам анализа молекулярной дисперсии (AMOVA) установлено, что 82 % варибельности ISSR-локусов обнаруживается внутри выборок плюсовых деревьев и только 18 % приходится на изменчивость между

сравниваемыми группами деревьев из разных географических районов. Построение UPGMA-дендрограммы показало близость генетической структуры плюсовых деревьев из Пензенской области, Чувашской Республики и Республики Татарстан и обособленность плюсового генофонда сосны обыкновенной из Республики Марий Эл. Результаты исследований указывают на то, что уровень генетического разнообразия, структура генетической изменчивости и характер дифференциации плюсовых деревьев соответствуют ранее выявленным для природных популяций сосны обыкновенной в Среднем и Верхнем Поволжье.

Ключевые слова: *Pinus sylvestris* L.; плюсовые деревья; генетическое разнообразие; дифференциация; ISSR-маркеры.

Introduction

Mass selection of plus trees constitutes the Russian selective seed production of the main forest-forming species (Tarakanov et al., 2021). A number of phenotypic characteristics, such as height, diameter, trunk quality, disease resistance, etc., are critical for plus trees selection in natural plantations. First-order tree gene banks, an integral part of the forest-seed establishment, serve for the mass production of seeds of forest tree species by the vegetative offspring of plus trees (Tsarev et al., 2021). One of the concerns, while introducing forest seed programs for Scots pine based on the principles of plus selection, is that the genetic diversity of the plus gene pool declines. This happens due to the selection of a limited number of plus trees, as well as the risk of inbreeding depression of seed offspring that are in proximity to related clones in tree gene banks (Koelewijn et al., 1999; Hosius et al., 2006). Therefore, further studies are necessary to research the diversity of the selected plus gene pool and identify the nature of its differentiation using both morphometric characters (Tarakanov, Kalchenko, 2015; Besschetnova, Besschetnov, 2017) and molecular markers (Shigapov, 1995; Milyutina et al., 2013; Ilinov, Raevsky, 2021).

Molecular markers have become an effective tool that solves a wide range of issues in the field of forest selection and seed production, as well as estimates the genetic diversity of plus trees (Sheikina, 2022b). To assess the variability of Scots pine plus trees and tree gene banks established by their offspring, different researchers used isoenzymes (Shigapov, 1995), ISSR markers (Milyutina et al., 2013; Khanova et al., 2020) and microsatellites (Ilinov, Raevsky, 2021; Kamalov et al., 2022). The results of comparative studies of the genetic diversity of the plus gene pool of tree species and natural populations showed contradictory results. A number of works note that plus trees can be characterized by a level of genetic variability comparable to natural populations (Bergman, Ruetz, 1991; Ilyinov, Raevsky, 2023). On the other hand, we may witness a decrease in allelic diversity in samples of plus trees growing in tree gene banks (Shigapov, 1995; Ilyinov, Raevsky, 2017).

Until now, in the Middle and Upper Volga regions, there have been studies of ISSR loci polymorphism only for a small sampling of 36 plus trees in the Republic of Mari El (Milyutina et al., 2013). However, no assessment of the genetic diversity of Scots pine plus trees in other parts of the Middle and Upper Volga region has been made. Meanwhile, ISSR markers are widely employed to study the characteristics of the population genetic structure of Scots pine in China (Hui-yu et al., 2005), Portugal (Cipriano et al., 2013), on the East European Plain and in the Urals (Vidyakin et al., 2015; Vasilyeva et al., 2021; Chertov et al., 2022; Sboeva et al., 2022), in the Perm

Territory (Prishnivskaya et al., 2019) and in the Volga region (Sheikina, 2022a).

The objective of this paper is to study the genetic variability and differentiation of Scots pine plus trees from the Middle Volga region based on the analysis of ISSR markers. We assumed that the level of genetic diversity, the structure of genetic variability and the nature of differentiation of plus trees selected as a result of breeding is comparable to those previously identified for natural populations of Scots pine in the Middle and Upper Volga region.

Materials and methods

The object of the study was plus trees of Scots pine (*Pinus sylvestris* L.) or their clones from four regions of the Middle and Upper Volga region. Samples for molecular genetic research in the Republic of Tatarstan were stored from plus trees growing on the territory of the Zelenodolsk forestry. The remaining samples were stored from clones of plus trees growing at forest seed production facilities: in the Chuvash Republic from a first-order tree gene bank in the Ibresinsky forestry, in the Penza Region from a first-order tree gene bank in the Chaadayevsky forestry, in the Mari El Republic from a collection and uterine plot in the Sernursky forestry. In total, the authors studied 270 trees.

The material for DNA extraction was dried pine needles. The CTAB technique (Doyle J.J., Doyle J.L., 1987) was employed for DNA preparations. Six ISSR primers were used for PCR: (CA)₆AGCT, (CA)₆AG, (CA)₆GT, (CA)₆AC, (AG)₈T and (AG)₈GCT (Hui-yu et al., 2005). PCR was carried out in a MJ MiniTM Gradient Thermal Cycler (Bio-Rad, USA) according to the following program: 94 °C – 5 min; 35 cycles: 94 °C – 45 s, 60 °C – 45 s, 72 °C – 45 s; 72 °C – 7 min. To perform PCR, we used the components of the commercial set Encyclo Plus PCR kit (Evrogen, Russia) with the following concentration: 10 × PCR buffer – 1 μl; dNTPs – 0.2 μl (10 mM); primer – 0.1 μl (100 μM); DNA preparation – 1 μl (20 ng); Taq polymerase – 0.1 μl (2 units/μl); water – 7.6 μl. PCR for each sample was performed in triplicate to check the repeatability of the DNA fingerprints obtained. PCR results were visualized with the use of electrophoresis in a 1.5 % agarose gel in 1 × TBE buffer at an electric field voltage of 80 V and staining with an ethidium bromide solution. Gel images were obtained using the GelDoc 2000 gel documentation system (Bio-Rad, USA) and the Quantity One® Version 4.6.3 software package. The '100 bp+3.0 kb DNA Ladder' marker (Evrogen, Russia) was used to calculate the lengths of PCR fragments.

Interpretation of the results of molecular genetic analysis was based on compiling a binary matrix, in which PCR fragments present in the electropherogram were designated

as ‘1’, and those absent, as ‘0’. Calculation of genetic diversity indicators, analysis of molecular variance (AMOVA) and principal coordinate analysis (PCoA) were done in the GenAlEx program (Peakall, Smouse, 2012). The statistical significance of differences between the average values of genetic diversity indicators of samples of plus trees was assessed using single-factor analysis of variance. A dendrogram illustrating the genetic relationship of samples of plus trees was drawn based on the frequency of occurrence of ISSR loci in the POPTREEW program (Takezaki et al., 2014) using the Unweighted Pair-Group Method with Arithmetic Mean (UPGMA) with bootstrap support for 10,000 replications.

Results

The authors identified 215 amplified DNA fragments for six ISSR primers, 99.5 % of which turned out to be polymorphic (Table 1). For samples of plus trees of different geographical origins, the number of PCR fragments varied from 186 in the Republic of Mari El to 201 in the Penza Region, and the percentage of polymorphic loci ranged from 95.7 to 96.9. The number of rare PCR fragments with an occurrence frequency of less than 5 % in different samples varied from 1 to 23, and the number of unique ones, from 0 to 2.

Shannon information index and expected heterozygosity were different for the studied sample of plus trees. Plus trees from the Mari El Republic proved to have the lowest

values of indicators ($I = 0.291$, $He = 0.205$). While pine from the Penza Region showed the maximum values of genetic variability ($I = 0.429$, $He = 0.298$). The differences between the samples are significant ($p = 0.01$). In terms of the number of alleles per locus and the number of effective alleles, plus trees of different geographical origins did not differ ($Na = 1.96-1.97$, $Ne = 1.31-1.48$) at $p = 0.01$. In total, for all the trees studied, the number of alleles per locus was 1.99, the number of effective alleles was 1.37, the Shannon index was 0.363, and the expected heterozygosity was 0.230. Figure 1 exemplifies the spectra of PCR fragments.

The different ISSR primers used in PCR allowed us to analyze from 27 to 40 loci, 80.6–93.5 % of which are polymorphic (Table 2). The high level of polymorphism suggests that the studied set of markers can be a useful and informative tool in assessing the genetic variability of both natural populations of an economically valuable species, as well as forest crops and objects of a genetic breeding complex, including plus trees. Other indicators of genetic diversity for different ISSR primers varied in the following ranges: the number of alleles per locus, from 1.62 to 1.90; the number of effective alleles, from 1.31 to 1.41; the Shannon index, from 0.331 to 0.393; the expected heterozygosity, from 0.206 to 0.252.

Analysis of molecular variance proved that 82 % of genetic variability is distributed within samples of plus trees from different geographical areas of the Middle Volga region

Table 1. Indicators of genetic diversity of Scots pine plus trees

Geographical origin	Number of trees	N (N_{05} , R)	P, %	Na	Ne	I	He
Penza Region	63	201 (1, 1)	96.5	1.96 ± 0.013	1.48 ± 0.021	0.429 ± 0.014	0.298 ± 0.010
Republic of Tatarstan	66	199 (23, 1)	96.9	1.97 ± 0.021	1.36 ± 0.021	0.343 ± 0.015	0.232 ± 0.011
Chuvash Republic	70	194 (6, 2)	95.7	1.96 ± 0.020	1.44 ± 0.022	0.391 ± 0.016	0.273 ± 0.010
Mari El Republic	71	186 (4, 0)	95.7	1.96 ± 0.020	1.31 ± 0.020	0.291 ± 0.015	0.205 ± 0.010
Total	270	215	99.5	1.99 ± 0.020	1.37 ± 0.011	0.363 ± 0.008	0.230 ± 0.006
Fisher's F test ($F_{0.01} = 4.94$)	–	–	–	0.07	0.99	14.87	13.92

Note. Mean \pm standard error. N – number of PCR fragments; N_{05} – number of PCR fragments with a frequency <5 %; R – number of unique PCR fragments; P – percentage of polymorphic loci; Na – number of alleles per locus; Ne – number of effective alleles; I – Shannon index; He – expected heterozygosity.

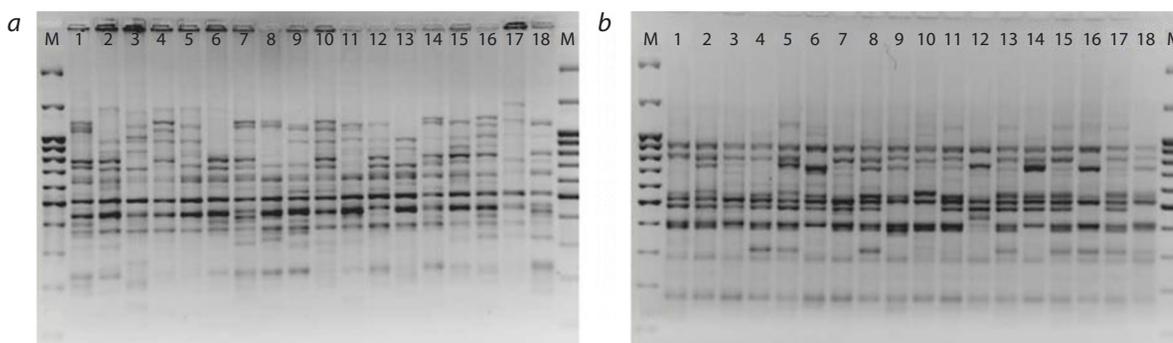


Fig. 1. DNA profiles showing polymorphism of Scots pine plus trees obtained with ISSR primers $(CA)_6AGCT$ (a) and $(AG)_6T$ (b).

1–18 – DNA sample numbers, M – DNA length marker 100 bp + 3.0 kb DNA Ladder.

Table 2. Indicators of genetic diversity of ISSR primers

ISSR primer	N (N ₀₅)	P, %	Na	Ne	I	He
(CA) ₆ AGCT	34 (2)	81.6	1.65 ± 0.064	1.36 ± 0.026	0.359 ± 0.020	0.228 ± 0.014
(CA) ₆ AG	40 (1)	86.3	1.77 ± 0.048	1.31 ± 0.023	0.331 ± 0.017	0.206 ± 0.012
(CA) ₆ GT	35 (0)	92.9	1.90 ± 0.033	1.41 ± 0.029	0.338 ± 0.019	0.248 ± 0.014
(CA) ₆ AC	39 (1)	91.0	1.88 ± 0.033	1.35 ± 0.024	0.363 ± 0.017	0.227 ± 0.012
(AG) ₈ T	27 (0)	93.5	1.89 ± 0.042	1.41 ± 0.033	0.393 ± 0.021	0.252 ± 0.016
(AG) ₈ GCT	40 (3)	80.6	1.62 ± 0.062	1.37 ± 0.025	0.359 ± 0.019	0.231 ± 0.013

Note. Mean ± standard error. N – number of PCR fragments; N₀₅ – number of PCR fragments with a frequency <5 %; P – percentage of polymorphic loci; Na – number of alleles per locus; Ne – number of effective alleles; I – Shannon index; He – expected heterozygosity.

Table 3. Distribution of intra- and interbreeding population genetic variability of Scots pine plus trees according to the analysis results of molecular variance

Source of variability	df	SS	MS	V	Total variability, %
For all samplings					
Among breeding populations	3	1409.3	469.7	6.5	18.0
Within breeding populations	299	8040.9	30.2	30.2	82.0
Penza Region and Republic of Tatarstan					
Among breeding populations	1	428.9	428.9	6.1	16.0
Within breeding populations	127	4148.7	32.7	32.7	84.0
Penza Region and Chuvash Republic					
Among breeding populations	1	405.3	405.3	5.6	14.0
Within breeding populations	131	4516.9	34.5	34.5	86.0
Penza Region and Mari El Republic					
Among breeding populations	1	642.3	642.3	9.2	24.0
Within breeding populations	132	3924.4	29.7	29.7	76.0
Republic of Tatarstan and Chuvash Republic					
Among breeding populations	1	395.4	395.4	5.4	15.0
Within breeding populations	134	4116.5	30.7	30.7	85.0
Republic of Tatarstan and Mari El Republic					
Among breeding populations	1	332.8	332.8	4.5	15.0
Within breeding populations	135	3523.9	26.1	26.1	85.0
Chuvash Republic and Mari El Republic					
Among breeding populations	1	608.7	608.7	8.2	23.0
Within breeding populations	139	3892.2	28.0	28.0	77.0

Note. df – number of degrees of freedom; SS – sum of squares; MS – standard deviation; V – dispersion.

(Table 3). Interbreeding population variation accounts for 18 % of genetic diversity. Pairwise comparisons of plus trees from different geographic areas showed that interbreeding population variation could account for 14 to 24 %. The greatest genetic subdivision is characterized by samples from the Penza Region and the Mari El Republic (24 %), as well as from the Chuvash Republic and the Mari El Republic (23 %). The share of interpopulation variability was 14–16 % in the remaining cases. In all cases, the significance level was $p < 0.001$.

Samples of trees from the Penza Region, Republic of Tatarstan and Chuvash Republic were included in one cluster on the UPGMA dendrogram with a high bootstrap value (100) (Fig. 2). The sample of plus trees from the Mari El Republic was assigned to a separate cluster.

The authors analyzed the principal coordinates for individual Scots pine trees (Fig. 3, a) and samples of plus trees (Fig. 3, b) based on pairwise Nei's genetic distances. Analysis of the principal coordinates for individual Scots pine trees

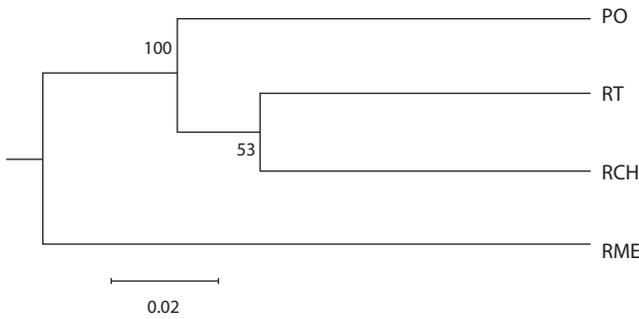


Fig. 2. UPGMA dendrogram drawn with Nei's genetic distance between plus trees of *P. sylvestris* L.

PO – Penza Region, RT – Republic of Tatarstan, RCH – Chuvash Republic, RME – Mari El Republic.

showed that the three principal axes account for 17.03 % of the polymorphism of ISSR loci, with the first coordinate accounting for 8.45 % and the second for 4.96 % of the total variability. At the same time, 81.02 % of the total diversity occurs in the first and second coordinates at the level of plus tree samples. The authors did not identify any geographic gradients along the axes. However, one can note the similarity in the distribution of samples on the first axis with the location of the sampling areas of plus trees in relation to the river Volga along the first axis. Plus trees from the Mari El Republic and Republic of Tatarstan grow on the left bank, while those from the Penza Region and Chuvash Republic grow on the right one.

Discussion

The paper discusses the genetic variability and differentiation of the plus gene pool of Scots pine from different regions of the Middle and Upper Volga region. To preserve the genetic diversity of a species in the process of artificial regeneration, it seems to be crucial that breeding populations are highly variable. Literature review showed that the percentage of polymorphic ISSR loci in Scots pine populations may vary from 42 to 100 % (Hui-yu et al., 2005; Cipriano et al., 2013; Vidyakin et al., 2015; Prishnivskaya et al., 2019). The percentage varied between 95.7–96.9 % and averaged 99.5 % for the studied samplings of plus trees and the selected

markers, which aligns with the results of the previous studies. The value the authors got for the proportion of polymorphic loci of plus trees was comparable to the data acquired for 12 natural populations of Scots pine (96.7 %) from the Upper and Middle Volga region, studied with the same set of ISSR markers (Sheikina, 2022a).

Other indicators of genetic diversity identified among the studied samples of plus trees were not inferior to the values typical for natural populations. Thus, the values of the number of effective alleles and expected heterozygosity for plus trees were 1.31–1.48 and 0.205–0.298, respectively, while for natural populations these were 1.27–1.39 and 0.174–0.241 (Sheikina, 2022a). A similar value of the expected heterozygosity of ISSR loci ($H_e = 0.239$) was identified for plus Scots pine trees from the Republic of Bashkortostan (Khanova et al., 2020). Lower values of expected heterozygosity were determined for Scots pine populations on the Russian Plain ($H_e = 0.046–0.239$) (Vidyakin et al., 2015; Prishnivskaya et al., 2019; Vasilyeva et al., 2021; Sboeva et al., 2022) and in the Urals ($H_e = 0.149–0.185$) (Chertov et al., 2022). High values of expected heterozygosity ($H_e = 0.447–0.488$) were typical for Portuguese populations (Cipriano et al., 2013), 1.5–2.4 times higher than the values described above.

For the samplings of plus trees under study, the Shannon index varied from 0.331 to 0.393. In other studies, the Shannon index identified for populations from different parts of Russia was 0.087–0.357 (Vasilyeva et al., 2021; Chertov et al., 2022; Sboeva et al., 2022). Higher values of the Shannon index ($I = 0.636–0.681$) were determined for Scots pine populations from Portugal (Cipriano et al., 2013). Differences in levels of genetic diversity may be explained by both geographic variability and the fact that studies have used different sets and numbers of ISSR markers.

Based on the analysis of the molecular dispersion of plus pine trees of different geographical origins, the authors found that that 18 % of the variability of ISSR loci accounts for the interpopulation component. These data comply with the information previously acquired for natural populations of Scots pine in the Upper and Middle Volga region (14 %) (Sheikina, 2022a). The value of this parameter was 37–48 % (Vasilyeva et al., 2021; Chertov et al., 2022; Sboeva et al., 2022) for the populations from the East European Plain and

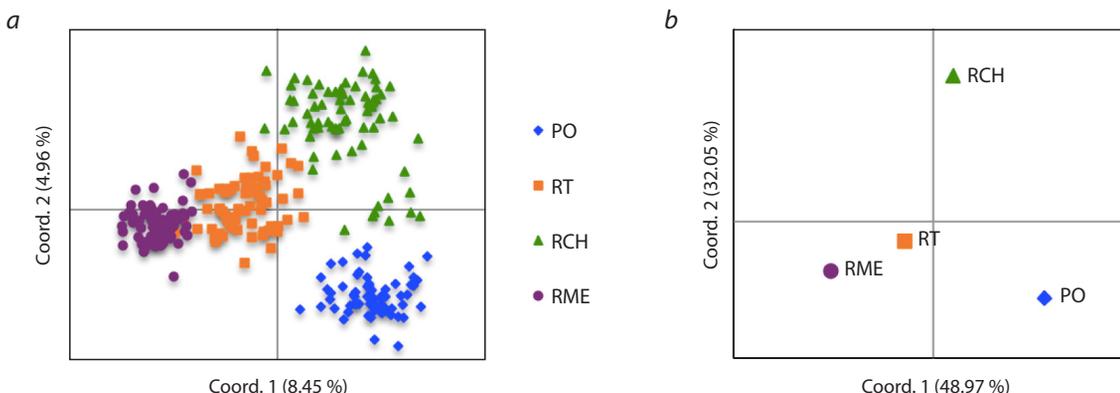


Fig. 3. Spatial location of the principal coordinate (PCoA) of Scots pine plus trees (a) and geographic origins (b).

PO – Penza Region, RT – Republic of Tatarstan, RCH – Chuvash Republic, RME – Mari El Republic.

the Urals. Assessment of the differentiation of Scots pine populations from various parts by measuring the indicator of genetic subdivision (Gst) proved that the interpopulation component of the variability of ISSR loci can account for from 5.8 to 55.8 % (Hui-yu et al., 2005; Cipriano et al., 2013; Vidyakin et al., 2015; Vasilyeva et al., 2021; Chertov et al., 2022; Sboeva et al., 2022; Sheikina, 2022a). Relatively low values of the genetic subdivision indicator were found for populations from the Middle Urals (Gst = 0.155) (Sboeva et al., 2022) and from Portugal (Gst = 0.058) (Cipriano et al., 2013). Higher values of the genetic subdivision indicator were shown for populations from China (Gst = 0.396) (Hui-yu et al., 2005), from the East European Plain (Gst = 0.439–0.558) (Vidyakin et al., 2015; Vasilyeva et al., 2021; Sboeva et al., 2022) and from the Urals (Gst = 0.362) (Chertov et al., 2022). The indicator of genetic subdivision of natural populations from the Upper and Middle Volga region was 0.161 (Sheikina, 2022a). Thus, the data on the structure of genetic variability in samples of plus trees acquired in this study do not contradict previously described results for natural populations of Scots pine.

Clustering of plus trees samples with the UPGMA method showed the isolation of the plus gene pool of Scots pine from the Mari El Republic from three other groups of trees. Tree samples from the Penza Region, Chuvash Republic and Republic of Tatarstan constitute a single cluster with a similar genetic structure. While assessing the population structure of pine forests in the Middle and Upper Volga regions, the authors also traced the differences between populations growing in the Mari El Republic, on the right Volga riverbank, from left-bank populations growing in the Chuvash Republic and the Penza Region (Sheikina, 2022a). The identified differentiation of populations and plus gene pools of Scots pine of different geographical origins may be the result of the intersection of the species' migration routes in the post-glacial period. Specifically, with the allozyme analysis, the authors discovered that five different Pleistocene refugia could have participated in creating the gene pool of Scots pine populations on the East European Plain (Sannikov et al., 2020).

Conclusion

The studied plus trees samples taken from various parts of the Middle and Upper Volga regions differ in the level of polymorphism of ISSR loci. A level of genetic diversity of the plus gene pool of Scots pine selected during breeding is comparable to natural populations in the region under the study. The structure of genetic variability and the nature of differentiation of samples of plus trees of various geographical origin also correspond to the population genetic structure of natural populations.

The authors' results proved that the ISSR markers described by A.I. Vidyakin et al. (Vidyakin et al., 2015) are feasible for the study of the population genetic structure of Scots pine. Moreover, the high level of variability of the selected loci (80.6–93.5 %) allows recommending this set for assessing the genetic variability of natural populations, forest crops and objects of the unified genetic breeding pool (UGBP). Further studies with the use of other types of molecular markers are necessary to improve the reliability of genetic diversity assessment and differentiation of plus trees.

References

- Bergman F., Ruetz W. Isozyme genetic variation and heterozygosity in random tree samples and selected orchard clones from the same Norway spruce populations. *For. Ecol. Manag.* 1991;46(1-2):39-47. DOI 10.1016/0378-1127(91)90243-0
- Besschetnova N.N., Besschetnov V.P. Variability of morphometrical characteristics of needles at a clonal plantation of plus trees of scots pine (*Pinus sylvestris* L.). *Vavilovskii Zhurnal Genetiki i Seleksii = Vavilov Journal of Genetics and Breeding.* 2017;21(2):198-206. DOI 10.18699/VJ17.237 (in Russian)
- Chertov N., Nechaeva Y., Zhulanov A., Pystogova N., Danilova M., Boronnikova S., Kalendar R. Genetic structure of *Pinus* populations in the Urals. *Forests.* 2022;13(8):1278. DOI 10.3390/f13081278
- Cipriano J., Carvalho A., Fernandes C., Gaspar M.J., Pires J., Bento J., Roxo L., Louzada J., Lima-Brito J. Evaluation of genetic diversity of Portuguese *Pinus sylvestris* L. populations based on molecular data and inferences about the future use of this germplasm. *J. Genet.* 2013;92(2):e41-e48. DOI 10.1007/s12041-013-0241-3
- Doyle J.J., Doyle J.L. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* 1987;19:11-15
- Hosius B., Leinemann L., Konnerth M., Bergmann F. Genetic aspects of forestry in the Central Europe. *Eur. J. For. Res.* 2006;125(4):407-417. DOI 10.1007/s10342-006-0136-4
- Hui-yu L., Jing J., Gui-feng L., Xu-jun M., Jing-xiang D., Shi-lie L. Genetic variation and division of *Pinus sylvestris* provenances by ISSR markers. *J. For. Res.* 2005;16(3):216-218. DOI 10.1007/BF02856818
- Ilinov A.A., Raevskiy B.V. Comparative evaluation of the genetic diversity of natural populations and clonal seed orchards of *Pinus sylvestris* L. and *Picea × fennica* (Regel) Kom. in Karelia. *Russ. J. Genet. Appl. Res.* 2017;7(6):607-616. DOI 10.1134/S2079059717060065
- Ilinov A.A., Raevsky B.V. Genetic diversity of scots pine trees of different selection categories in plus stands of Karelia. *Ekologicheskaya Genetika = Ecological Genetics.* 2021;19(1):23-35. DOI 10.17816/ecogen50176 (in Russian)
- Ilinov A.A., Raevsky B.V. Genetic evaluation by microsatellite loci of *Pinus sylvestris* L. plus trees. *Lesnoy Zhurnal = Russian Forestry Journal.* 2023;3:48-68. DOI 10.37482/0536-1036-2023-3-48-68 (in Russian)
- Kamalov R.M., Peturenko M.Yu., Degtyareva A.P. Variability of indicators of molecular markers in clones of plus trees *Pinus sylvestris* L. *Trudy Sankt-Peterburgskogo Nauchno-Issledovatel'skogo Instituta Lesnogo Khozyaystva = Proceedings of the Saint Petersburg Forestry Research Institute.* 2022;3:4-14. DOI 10.21178/2079-6080.2022.3.4 (in Russian)
- Khanova E., Konovalov V., Timeryanov A., Isyanyulova R., Rafikova D. Genetic and selection assessment of the scots pine (*Pinus sylvestris* L.) in forest seed orchards. *Wood Res.* 2020;65(2):283-292. DOI 10.37763/wr.1336-4561/65.2.283292
- Koelwijn H.P., Koski V., Savolainen O. Magnitude and timing of inbreeding depression in Scots pine (*Pinus sylvestris* L.). *Evolution.* 1999;53(3):758-768. DOI 10.1111/j.1558-5646.1999.tb05370.x
- Milyutina T.N., Sheikina O.V., Novikov P.S. Molecular-genetic research of variation in clonal progeny of *Pinus sylvestris* plus trees using ISSR markers. *Khvoynye Boreal'noy Zony = Conifers of the Boreal Area.* 2013;31(1-2):102-105 (in Russian)
- Peakall R., Smouse P.E. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research an update. *Bioinformatics.* 2012;28(19):2537-2539. DOI 10.1093/bioinformatics/bts460
- Prishnivskaya Ya.V., Nassonova E.S., Chertov N.V., Zhulanov A.A., Vasileva Yu.S., Boronnikova S.V., Kalendar R.N. Genetic diversity within species of two species woody plants populations in Perm Krai. *Bulleten Nauki i Praktiki = Bulletin of Science and Practice.* 2019;5(4):58-68. DOI 10.33619/2414-2948/41/06 (in Russian)

- Sannikov S.N., Petrova I.V., Egorov E.V., Sannikova N.S. Searching for and revealing the system of pleistocene refugia for the species *Pinus sylvestris* L. *Russ. J. Ecol.* 2020;51(3):215-223. DOI 10.1134/S1067413620030133
- Sboeva Y., Chertov N., Nechaeva Y., Valeeva A., Boronnikova S., Kalendar R. Genetic diversity, structure, and differentiation of *Pinus sylvestris* L. populations in the East European Plain and the Middle Urals. *Forests.* 2022;13(11):1798. DOI 10.3390/f13111798
- Sheikina O.V. Genetic structure and differentiation of Scots pine (*Pinus sylvestris* L.) populations in the Middle and Upper Volga Regions. *Ekologicheskaya Genetika = Ecological Genetics.* 2022a; 20(4):261-270. DOI 10.17816/ecogen110866 (in Russian)
- Sheikina O.V. Application of molecular markers in forest breeding and seed production in Russia: experience and prospects (review). *Vestnik Povolzhskogo Gosudarstvennogo Tekhnologicheskogo Universiteta. Seriya: Les. Ekologiya. Privodopolzovanie = Vestnik of Volga State University of Technology. Series: Forest. Ecology. Nature Management.* 2022b;2(54):64-79. DOI 10.25686/2306-2827.2022.2.64 (in Russian)
- Shigapov Z.H. Comparative genetic analysis of forest seed plantations and natural populations of Scots pine. *Lesovedenie = Russian Journal of Forest Science.* 1995;3:19-24 (in Russian)
- Takezaki N., Nei M., Tamura K. POPTREEW: web version of POPTREE for constructing population trees from allele frequency data and computing some other quantities. *Mol. Biol. Evol.* 2014; 31(6):1622-1624. DOI 10.1093/molbev/msu093
- Tarakanov V.V., Kalchenko L.I. Phenetic Analysis of Clonal and Natural Populations of *Pinus sylvestris* L. in the Altai Territory. Novosibirsk: Acad. Publ. House "Geo", 2015 (in Russian)
- Tarakanov V.V., Palenova M.M., Parkina O.V., Rogovtsev R.V., Tretyakova R.A. Forest tree breeding in Russia: achievements, challenges, priorities (review). *Lesokhozyastvennaya Informatsiya = Forestry Information.* 2021;1:100-143. DOI 10.24419/LHI.2304-3083.2021.1.09 (in Russian)
- Tsarev A.P., Laur N.V., Tsarev V.A., Tsareva R.P. The current state of forest breeding in the Russian Federation: the trend of recent decades. *Lesnoy Zhurnal = Russian Forestry Journal.* 2021;6:38-55. DOI 10.37482/0536-1036-2021-6-38-55 (in Russian)
- Vasilyeva Y., Chertov N., Nechaeva Y., Sboeva Y., Pystogova N., Boronnikova S., Kalendar R. Genetic structure, differentiation and originality of *Pinus sylvestris* L. populations in the East of the East European Plain. *Forests.* 2021;12(8):999. DOI 10.3390/f12080999
- Vidyakin A.I., Boronnikova S.V., Nechayeva Y.S., Pryshnivskaya Y.V., Boboshina I.V. Genetic variation, population structure, and differentiation in Scots pine (*Pinus sylvestris* L.) from the northeast of the Russian plain as inferred from the molecular genetic analysis data. *Russ. J. Genet.* 2015;51(12):1213-1220. DOI 10.1134/S1022795415120133

Funding. The research work was supported by the Russian Science Foundation (RSF, No. 23-16-00220), <https://rscf.ru/en/project/23-16-00220/> using equipment of the Core Facility Centre "Ecology, biotechnologies and processes for obtaining environmentally friendly energy carriers" of Volga State University of Technology, Yoshkar-Ola.

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

Received February 17, 2023. Revised September 29, 2023. Accepted September 29, 2023.

DOI 10.18699/vjgb-24-18

Multivariate analysis of long-term climate data in connection with yield, earliness and the problem of global warming

V.M. Efimov , D.V. Rechkin, N.P. Goncharov 

Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
 gonch@bionet.nsc.ru

Abstract. Climate change is the key challenge to agriculture in the XXI century. Future agricultural techniques in the Russian Federation should involve the optimization of land utilization. This optimization should apply algorithms for smart farming and take into consideration possible climate variations. Due to timely risk assessment, this approach would increase profitability and production sustainability of agricultural products without extra expenditures. Also, we should ground farming optimization not on available empirical data encompassing limited time intervals (month, year) or human personal evaluations but on the integral analysis of long-term information bodies using artificial intelligence. This article presents the results of a multivariate analysis of meteorological extremes which caused crop failures in Eastern and Western Europe in last 2600 years according to chronicle data and paleoreconstructions as well as reconstructions of heliophysical data for the last 9000 years. This information leads us to the conclusion that the current global warming will last for some time. However, subsequent climate changes may go in any direction. And cooling is more likely than warming; thus, we should be prepared to any scenario. Plant breeding can play a key role in solving food security problems connected with climate changes. Possible measures to adapt plant industry to the ongoing and expected climate changes are discussed. It is concluded that future breeding should be based on the use of highly adapted crops that have already been produced in pre-breeding programs, ready to meet future challenges caused by potential climate change. Key words: climate; global warming; models; next generation breeding; adaptability; earliness.

For citation: Efimov V.M., Rechkin D.V., Goncharov N.P. Multivariate analysis of long-term climate data in connection with yield, earliness and the problem of global warming. *Vavilovskii Zhurnal Genetiki i Seleksii = Vavilov Journal of Genetics and Breeding*. 2024;28(2):155-165. DOI 10.18699/vjgb-24-18

Многомерный анализ многолетних климатических данных в связи с урожайностью, скороспелостью и проблемой глобального потепления

В.М. Ефимов , Д.В. Речкин, Н.П. Гончаров 

Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
 gonch@bionet.nsc.ru

Аннотация. Изменение климата – определяющая проблема растениеводства XXI в. Агротехнологии будущего в Российской Федерации должны быть связаны с оптимизацией землепользования, которая будет осуществляться как с применением алгоритмов «умного» сельского хозяйства, так и на основе прогноза потенциальных климатических изменений. Это позволит повысить рентабельность и устойчивость производства сельскохозяйственной продукции без дополнительных затрат за счет своевременного учета возможных рисков. При этом оптимизация системы возделывания культур должна базироваться не на имеющихся эмпирических показателях, полученных по ограниченному во времени точкам учета (месяц, год), и субъективной человеческой оценке, а на комплексном анализе массивов многолетней информации. В настоящей работе приведены результаты многомерного анализа метеорологических экстремумов и связанных с ними неурожаев в Восточной и Западной Европе за последние 2600 лет по летописным данным и палеореконструкциям, а также реконструкциям гелиофизических данных за последние 9000 лет. Отмечается, что идущее глобальное потепление продлится еще некоторое время. Однако последующие изменения климата могут быть направлены в любую сторону, причем, скорее, в сторону похолодания, а не потепления, поэтому надо быть готовыми к любым сценариям будущего. Для нивелирования последствий этих изменений селекция может сыграть ключевую роль в решении проблем продовольственной безопасности. Обсуждаются перспективы разработки мер адаптации растениеводства к происходящим и ожидаемым изменениям климата, и сделан вывод, что селекция будущего должна базироваться на использовании уже отработанных в программах предварительной селекции (pre-breeding) высокоадаптированных сельскохозяйственных культур, потенциально отвечающих будущим вызовам, обусловленным потенциальным изменением климата. Ключевые слова: климат; глобальное потепление; модели; селекция будущего; адаптивность; скороспелость.

Introduction

Geologic record indicates that dramatic climate changes directly affected humanity throughout its history (Gupta, 2004). The cooling in the Palearctic, sometimes reaching extreme values, started in the late Pleistocene, about 27 ka BP, and ended about 14 ka BP (Prentice, 2009). Archeological data show that humans inhabiting the Palearctic either successfully adapted to the changes or migrated to areas with better conditions. In VIII–X centuries BC, the latter group proceeded to productive economy (Shnirelman, 2012), that is, farming based on plant domestication and, later, animal domestication (Goncharov, 2013). At present, it is obvious that Toynbee's (1954) 'challenge and response' theory of the history of civilisations, which presumes that the transition to agriculture was the response of hunters and gatherers to abrupt aridization caused by the melting of Late Pleistocene glaciers, has found no evidence (Trifonov, Karakhanyan, 2004). Moreover, agriculture in West Asia started against the background of wears relative humidization.

Local and global climatic changes constitute the primary concern for the XXI century, particularly, its first decades, when measures for mitigation of severe consequences for humanity and the global agroecosystem are urgent (Eckardt et al., 2023). The oncoming climate change may exert numerous adverse effects on crop production throughout the globe. It will demand significant extension of biodiversity (germ plasm pool), required for the inclusion of new characters and traits and completely new previously uncultivated plant species in breeding. The collection, preservation, and subsequent effective and intelligent use of the biodiversity of crops and their wild relatives in the changing climate are critical issues (Eastwood et al., 2022).

Many relatives of commercially important crops have considerable polymorphisms for seasonal adaptation (Goncharov, Chikida, 1995; Leigh et al., 2022; Liang, Tian, 2023); thus, they can be useful for improving the general adaptability of crops to local and global climate changes.

The consequences of climate changes cause anxiety to experts in various fields (Kattsov et al., 2011; Ruddiman et al., 2016), including biologists and agrarians, who deal with a wide variety of objects (Baltzoi et al., 2015; Gurova, Osipova, 2018; Morgounov et al., 2018; Eastwood et al., 2022; and others). It is beyond doubt that current climatic trends adversely influence the performance of many extensively grown crops. These trends are a considerable commercial risk to global agriculture and other farm industries (Lobell, Gourdj, 2012). It is predicted that climate changes will affect not only the production but also the quality of food (Atkinson et al., 2008), thereby jeopardizing food security. The broadening of polymorphism for many characters (Trifonova et al., 2021) with ensuing breeding for optimal duration of vegetative period (earliness) of crops are becoming more and more relevant (Kamran et al., 2014; Smolenskaya et al., 2022).

The potential climate changes demand modification of breeding programs for new generation cultivars and breeds. The requirements for higher adaptability to future environmental condition changes should be put at the forefront, which will apparently differ from the present. Neither the scale nor the directions of the changes can be unambiguously assessed. Present forecasts concern mainly risks associated with the broad (standard) use of conventional crop architectonics (Jatayev et al., 2020; Liu et al., 2022). This problem is significant, because about 70 % of dwarf common wheat commercial cultivars bear only two alleles of the *Rht* genes – *Rht-B1b* and/or *Rht-D1b* (Sukhikh et al., 2021).

There is no consensus among investigators as to whether the recorded climate changes are caused by anthropogenic or natural factors (Ruddiman et al., 2016; Lobkovsky et al., 2022; and others). At last time, the Sun is shown to operate in distinct modes – a main general mode, a Grand minimum mode corresponding to an inactive Sun, and a possible Grand maximum mode corresponding to an unusually active Sun (Solanki et al., 2004; Usoskin et al., 2014). It states that the heat flux from the Sun to Earth, the so-called solar constant, is in fact not constant, at least, on the millennial scale. This flux demonstrates variations unpredictable with the current state of knowledge.

It is clear that the presently observed global warming started long before the industrial boom. We just live in a time of changes, when the solar heat flux on our planet starts its change once again to induce a climatic upheaval (Usoskin et al., 2014; Biswas et al., 2023). In view of all this, the analysis of historical and modern data to assess the possible limits of natural climatic variations appears to be essential for choosing land use strategies and for successful farming in the future.

The goal of this study is to analyze the limits of climate variability, meteorological extremes, and crop failures in Eastern and Western Europe over 2600 years from chronicles (Barash, 1989), paleoreconstructions of air temperatures (Sleptsov, Klimenko, 2005), and composite solar physics data across millennia based on proxy methods (Clette et al., 2014; Wu et al., 2018). We applied multivariate analysis approaches, in particular, the principal component analysis (PCA), for more comprehensive coverage and deeper analysis of the processes under study.

Materials and methods

The following data were invoked:

- 1) chronicle data on years with meteorological extremes and crop failures in Western and Eastern Europe over 2600 years, from X century BC to XVI century AD from monograph of S.I. Barash (1989);
- 2) climate reconstructions for Eastern Europe (East European Plain) over the last 2000 years according to paleoclimatic data reported by A.M. Sleptsov and V.V. Klimenko (2005);

- 3) Solar activity reconstruction over the last 9000 years according to indirect data (Wu et al., 2018);
- 4) Wolf sunspot numbers SN(v2.0) (1700–2022) from (WDC-SILSO, Royal Observatory of Belgium, Brussels).

We added some attributes to chronicle data presented in S.I. Barash (1989) to compensate insufficient accuracy in the characterization of years. Specifically, we introduced an integral attribute bulking crop failures in general in addition to attributes reflecting data on failures caused by particular factors: drought, overflowing, etc. In this case, the formation of this characteristic was carried out exclusively according to the data of S.I. Barash (Fig. 1).

Long-term data were processed by the principal component analysis (PCA) for time series, PCA-TS (Karhunen, 1947; Loève, 1948). By this method, any time series can be expanded into principal components that reflect the trend, quasicyclic fluctuations, and noise (Efimov et al., 1988). The modern PCA-TS version (Efimov et al., 2021) converts a unidimensional time series into a trajectory matrix (Takens, 1981). The matrix of Euclidean distances between its rows is calculated, and principal components (PCs) are extracted from the latter by the master coordinate procedure (Gower, 1966). Time series intervals uniform in variability patterns are detected from phase images drawn from principal components.

It should be mentioned that the opinion that statistical independence of principal components presumes their functional independence and, therefore, no principal component can be another component's derivative in the strict mathematical sense, because their correlation is zero, is false. A counterexample is the pair of time series, $\sin(t)$ and $\cos(t)$. The derivative of sine is cosine, their phase image is a circle, and their correlation is zero.

Similar situations often arise in the processing of real series by PCA. When one of the components is interpreted as a derivative of another, this fact can be used for predictions. Components often appear in pairs with close variances and frequencies, their phase images are close to circles, and which of them is the derivative of the other can be determined from the contributions of attributes to the components or from their shift with reference to each other. In another frequent case, contributions to one of the components are of the same sign and are similar in amplitude, forming a trend, whereas the other component is constituted by two sequential intervals of opposite signs, characterizing trend changes.

As each line has two directions, the researcher chooses the orientation at his discretion. Any principal component can be multiplied by “-1”. This will invert signs of contributions of all attributes. It is recommended that component orientation be chosen so that the signs of all contributions be positive in case of trend, whereas in case of opposite directions, negative contributions should be first and then positive ones. With this choice, the phase trajectory rotates mainly clockwise, a positive value of the derivative points

to the growth of the principal component, and a negative, to its drop.

The calculation of pairwise cross-correlations between the manifestations of the listed attributes of years involved the cutoff of noise related to less significant (minor) principal components (Supplementary Material 1)¹. We confined our studies to the effects of the detected modulation on events mentioned in chronicles; therefore, the diagonal matrix elements in the table presented in Supplementary Material 2 are always below unity. The mean error of the correlation coefficient for the specified number of objects ($n = 2600$) and the mean correlation coefficient value ($r = 0.500$) is $s_r = (1 - r^2) / \sqrt{n}$ no more than 0.014. Thus, pairwise correlation coefficients exceeding the triple mean error (0.044) were considered significant.

Results

We analyzed the collected by S.I. Barash's (1989) data transformed into matrix (see Fig. 1), in which objects are years and attributes in cells indicate favorable (good yield) or unfavorable (all other) events.

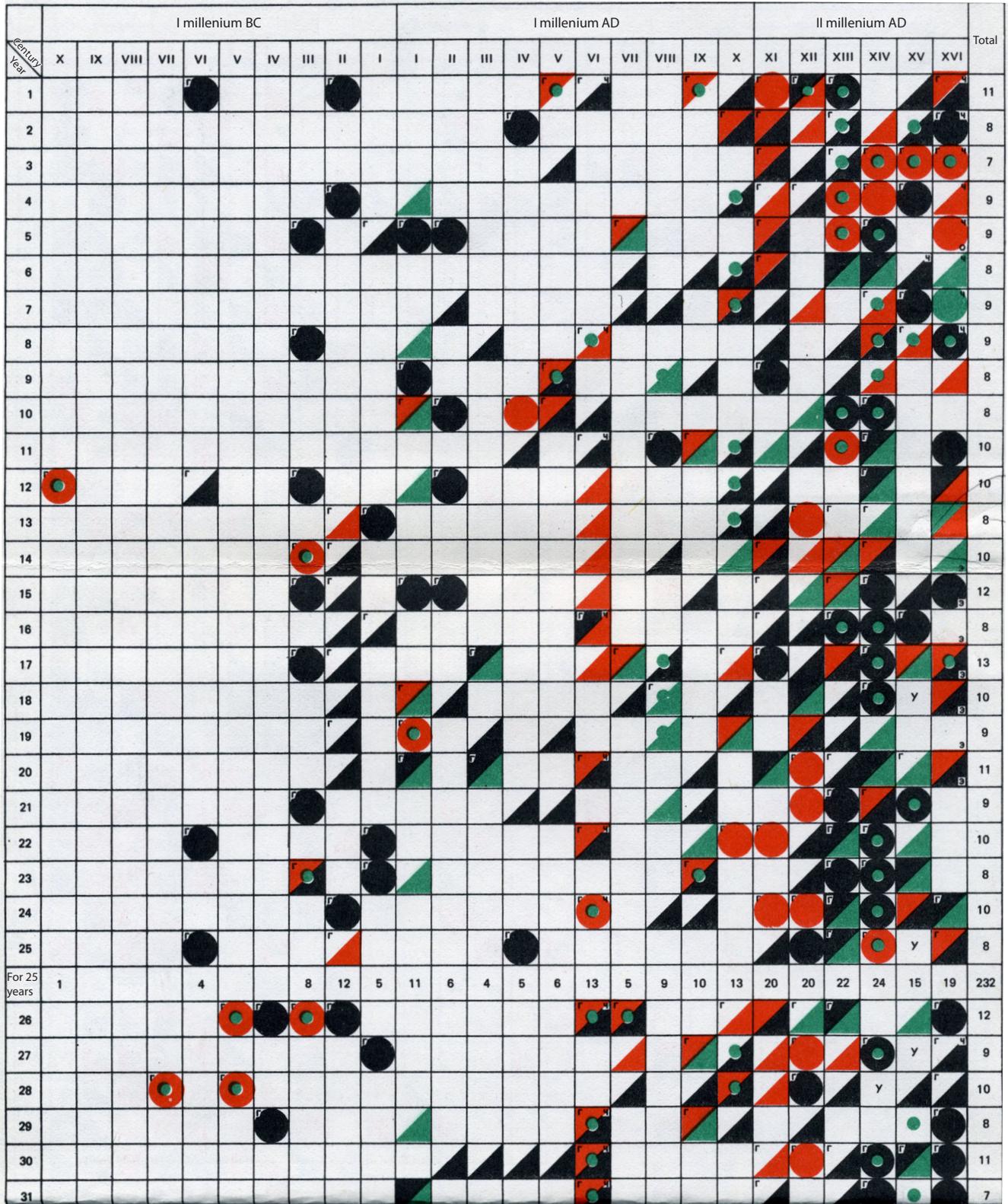
Designations in what follows: Eastern Europe: Dr-EE, drought and local poor crops; DrCF-EE, severe drought, extreme widespread crop failure; R-EE, rainy summer, local poor crops; RCF-EE, wet summer and extreme widespread crop failure; W-EE, severe winter; WPC-EE, extreme winter and local poor crops because of frost-killing of winter crops; F-EE, famine; L-EE, locust plague; Epi-EE, epidemics; P-EE, plague; Sp-EE, smallpox; SF-EE, spot fever; M-EE, murrain (epizootics); Y-EE, good yield; CF-EE, widespread crop failure. Western Europe: the same with the WE after the hyphen.

We introduced the CF-EE and CF-WE attributes. In S.I. Barash (1989), lean years are not marked by the CF-EE and CF-WE, but are “coded” in a hidden way in the Dr, DrCF, R, RCF, and WPC. We resigned out their effect in the two CF-EE and CF-WE attributes and, after that, we assessed the influence of all the listed natural phenomena. We found that crop failures are affected only by the DrCF and RCF, associated with widespread crop failures, rather than by local poor crops.

The data presented by S.I. Barash (1989) lack any information on typhus epidemics in Western Europe; thus, the SF-WE is null. For this reason, the overall matrix shows 15 natural phenomena for Eastern Europe and 14 ones for Western, 29 in total.

The data were processed by the PCA (Fig. 2, Supplementary Materials 1–3). The first (PC1) and second (PC2) principal components account for 22 % of the total sample variance (see Supplementary Material 1). Eigenvalues are conventionally arranged in decreasing order. They reflect information redistribution and the concentration of the most significant factors in the principal components (see Supplementary Material 3).

¹ Supplementary Materials 1–3 are available at:
https://vavilov.elpub.ru/jour/manager/files/Suppl_Efimov_Engl_28_2.pdf



 Droughts and local poor crops
  Rainy summers and local poor crops
  Extreme winters
 Severe droughts and extraordinary widespread crop failure
  Extremal rainy summers and extreme widespread crop failure
  Extreme winters and local crop failures because of frost-killing of winter crops
 r – famine; c – locust plague; e – epidemics; u – plague; o – smallpox; m – murrain (epizootics); y – good yield.

Fig. 1. Meteorological extremes and crop failures in Western (WE) and Eastern (EE) Europe over 2600 years, from X century BC to XV century AD, fragment from (Barash, 1989).

The sum of all eigenvalues equals the dimensionality of the correlation matrix (the number of attributes in the original sample). Therefore, when no regularities in the interactions of attributes can be recognized, each eigenvalue should be unity in case of a correlation matrix calculated from the original data matrix with centered and normalized attributes. Therefore, value “1” can be considered the threshold whose crossing by a principal component reflects a factor essential for sample description. Correspondingly, principal components whose eigenvalues fall short of this threshold should be considered insignificant. We will follow the terminology used by experts in computer and radio sciences (Oppenheim, Schafer, 1975) and name the informative set of PCs signal and other, insignificant components, statistical noise (see Supplementary Material 1). Here we regard the major principal components PC1 to PC12, accounting for about 66 % of the overall variance of the sample in question, as signal. Their eigenvectors are shown in Supplementary Material 3.

The first principal component (PC1; 14.5 % of the overall variance) reflects the influence of climatic (natural) factors on widespread crop failures and, consequently, on famine in Eastern and Western Europe. The greatest contribution amplitudes are made by the attributes DrCF-EE, RCF-EE, WPC-EE, DrCF-WE, RCF-WE, and WPC-WE. Consequently, the contributions of associated attributes CF-EE, F-EE, CF-WE, and F-EE are also high. The nature of the factor can be defined as yield vs. famine.

The second principal component (PC2; 7.5 % of the overall variance) reflects the influence of the factor determined by the set of attributes Dr-EE, R-EE, RCF-EE, and WPC-EE and their counterparts Dr-WE, R-WE, RCF-WE, W-EE, and CF-WE. In Western Europe, the signs of contributions of Dr-WE, R-WE, and W-EE are opposite to those of RCF-EE and CF-EE. Obviously, the second-rank factor in Western Europe shows that excess moisture in summer is the main cause of crop failures. In contrast, such climatic conditions in Eastern Europe do not result in crop failures, although they exert a certain influence on agriculture in general. The nature of the factor responsible for PC2 can be defined as differences in moisture regimes between Eastern and Western Europe.

By analyzing chronicle data on meteorological extremes and crop failures in Western and Eastern Europe for 2600 years, we assessed the similarities between the series of these events from the location of attributes in the phase space of principal components (see Fig. 2; Supplementary Material 2). For clarity, attributes similar in Eastern and Western Europe are connected with lines (see Fig. 2).

The third principal component (PC3; 5.8 % of the overall variance) shows that the effect of widespread crop failures in Western Europe manifests itself as a threat of famine, whereas in Eastern Europe famines are not so great threat as epidemics, first of all, plague. This fact stems from seeking food by the population of steppe and forest-steppe regions and contacts of humans with small animals inhabiting

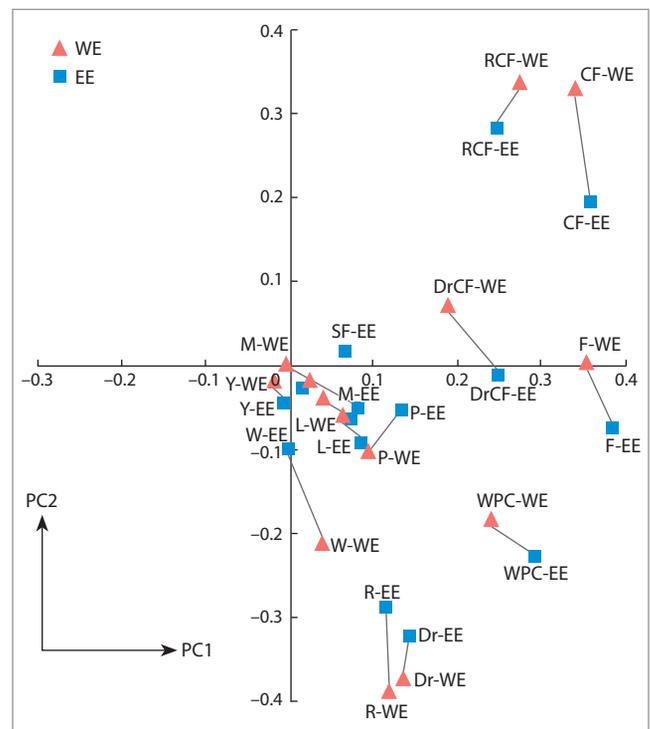


Fig. 2. The similarity in the manifestation of year natural phenomena in Western (WE) and Eastern (EE) Europe with respect to PC1 and PC2.

steppes and conveying plague: marmots, ground squirrels, tarbagans, and such.

The association of epidemics and epizootics with lean years can be explained by the poorer disease resistance in humans and livestock caused by undernourishment. Nature behaves in a different way in outbreaks of locust populations. The gregarious behavior of locusts is enhanced in generations produced by undernourished parents.

We assessed climate changes on the grounds of data reported by A.M. Sleptsov and V.V. Klimenko (2005), who attempted to reconstruct climate in Eastern Europe (East European Plain) from four kinds of sources: instrumental measurements, historical evidence, palynology, and dendrochronology. They reconstructed the variation in the annual average air temperatures in the East European Plain for the last 2000 years (Fig. 3).

A.M. Sleptsov and V.V. Klimenko note a negative trend in annual average temperatures, most pronounced in the last millennium; more precisely, from year 1200 to the second half of the XX century. They extrapolate these data to the 50 years to come and conclude that the so-called “global warming” is in fact of anthropogenic nature and that it rescued humanity from a “global cooling”, which would have been much more disastrous for our civilization, as shown by the history of the XIV–XVIII centuries. Also, they reveal a clear climatic rhythm of about 200 years, closely associated with solar activity variation (Sleptsov, Klimenko, 2005).

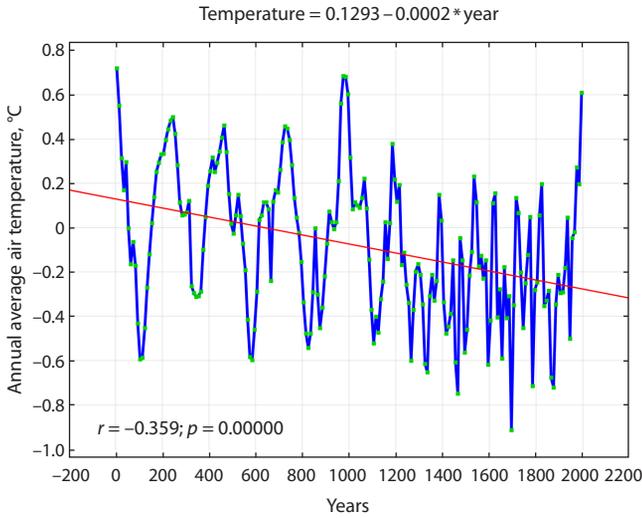


Fig. 3. Annual temperature deviations from the present values in the East European Plain (averaged over decades). Data from (Sleptsov, Klimenko, 2005; Fig. 3).

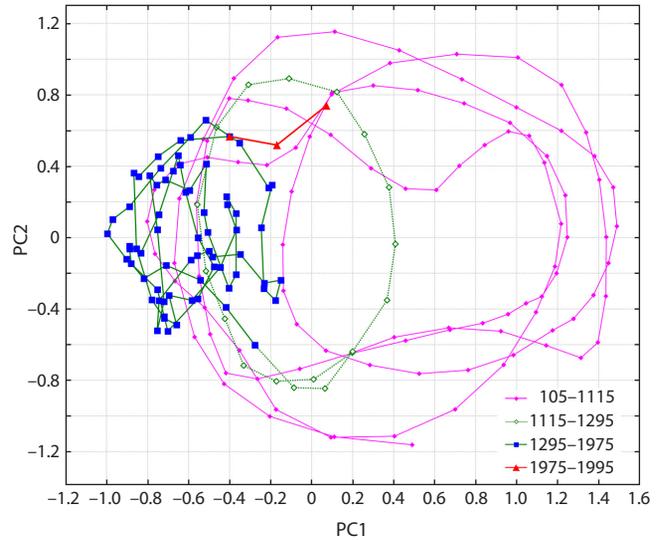


Fig. 4. Phase image of the variation of annual average temperature in the East Siberian Plain on the plane of two major principal components, PC1 and PC2.

A relatively beneficial time span lasted from the I to the XII century. After the XII century, it gave way to a cooling, which lasted nearly till present. However, additional information can be extracted from the data illustrated in Fig. 3.

The processing of this time series by the PCA (Figs. 4–5) clearly demonstrates its nonuniformity. It can be concluded from the coefficients of correlation of the first two principal components (PC1 and PC2) with the annual average tem-

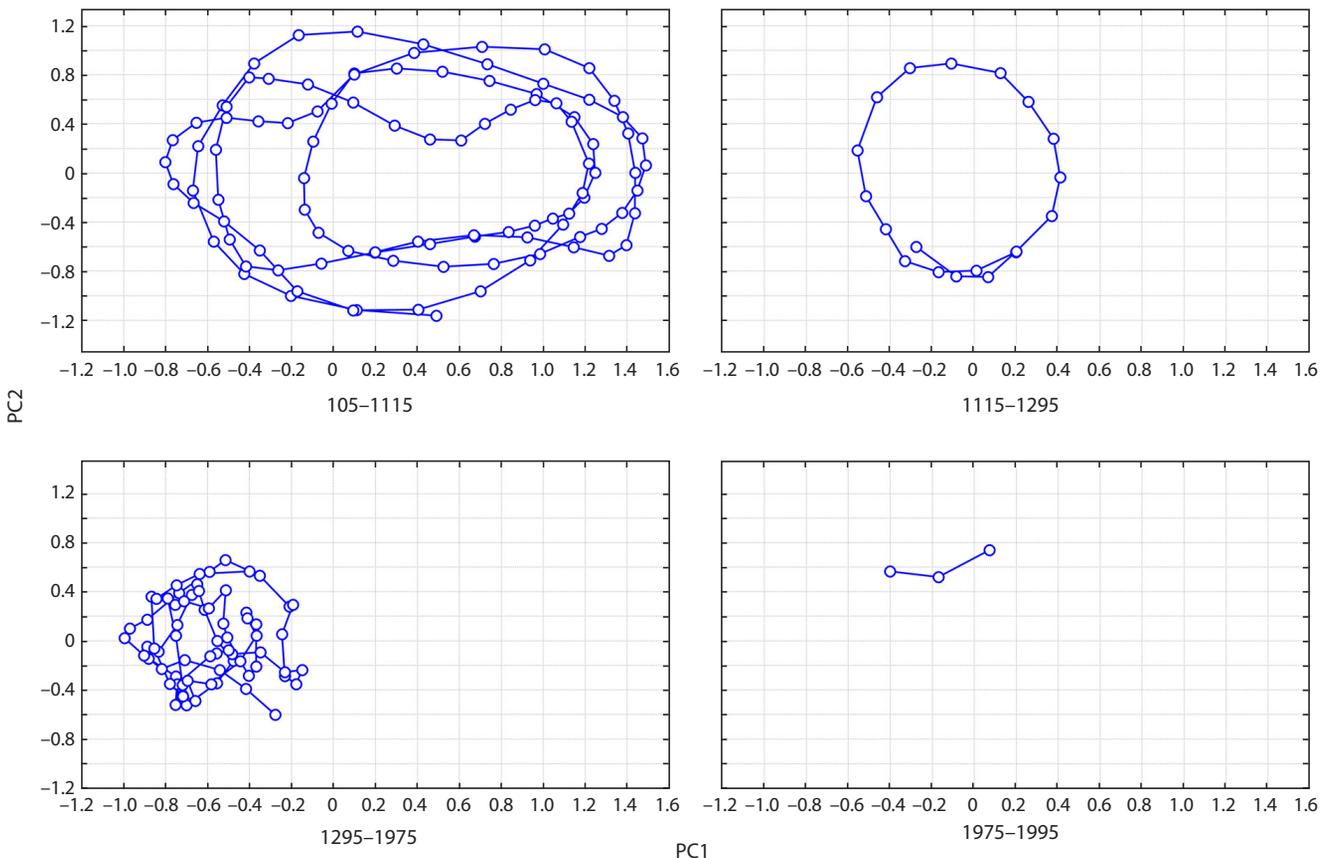


Fig. 5. Phase images of the variation of annual average temperature in the East Siberian Plain on the plane of PC1 and PC2, separately for each temperature phase.

Correlation coefficients ($\times 1000$) of two major principal components with the annual average temperature in the East Siberian Plain with different lags

Lag	1	2	3	4	5	6	7	8	9	10
PC1	492	621	723	795	825	820	777	701	592	459
PC2	-638	-640	-518	-314	-106	132	348	542	653	646

Note. Colors indicate: light-red and light-green – $p < 0.001$; red and green – $p < 10^{-4}$.

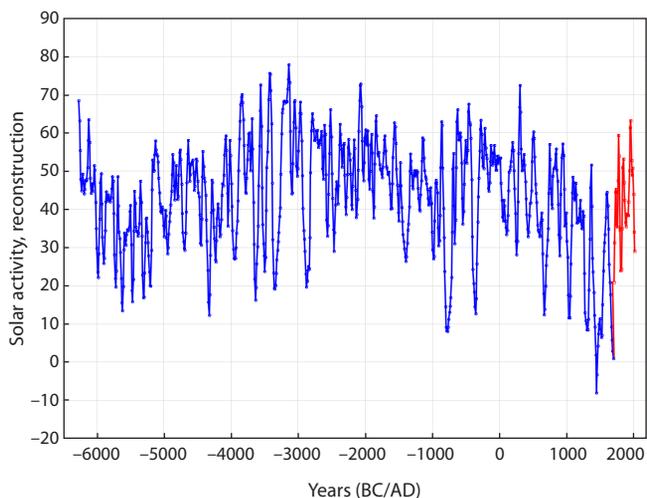


Fig. 6. Solar activity reconstruction for the last 9000 years (Wu et al., 2018).

perature with various lags that PC1 (47.7 % of variance) is responsible for air warming, and PC2 (24.9 % of variance), for its derivative (see the Table). This means that, when the trajectory of the series is above zero for PC2, it is bound to

move right, towards higher temperatures, until it falls below zero and turns back. This regularity shows no significant deviations (Figs. 6, 7). Four phases with different regimes are recognized in the time span analyzed: cyclic oscillations with period about 200 years (years 105–1115), transitional phase (1115–1295); quasichaotic variation (1295–1975), and warming (1975–present), see Figs. 4 and 5.

It follows from the phase images in Figs. 6, 7 that (1) the trajectory of the time series under consideration has exceeded the limits that confined it in the last seven centuries, (2) it has not exceeded the limits in which it stayed in the entire I millennium and the beginning of the II millennium, and (3) it is not inconceivable that the cyclic regime characteristic of the I millennium is returning. If this conclusion is true, further temperature increase should be expected in the next 50–60 years for natural causes, not related to human activity.

The results of time series processing (see Figs. 4–6) confirm the inferences from our earlier analysis (Efimov, Goncharov, 2013). Specifically, climate in Western and Eastern Europe experiences centuries-long oscillations for presently unknown reasons, abruptly turning from one climate regime to another. The most notable transitions occurred in the

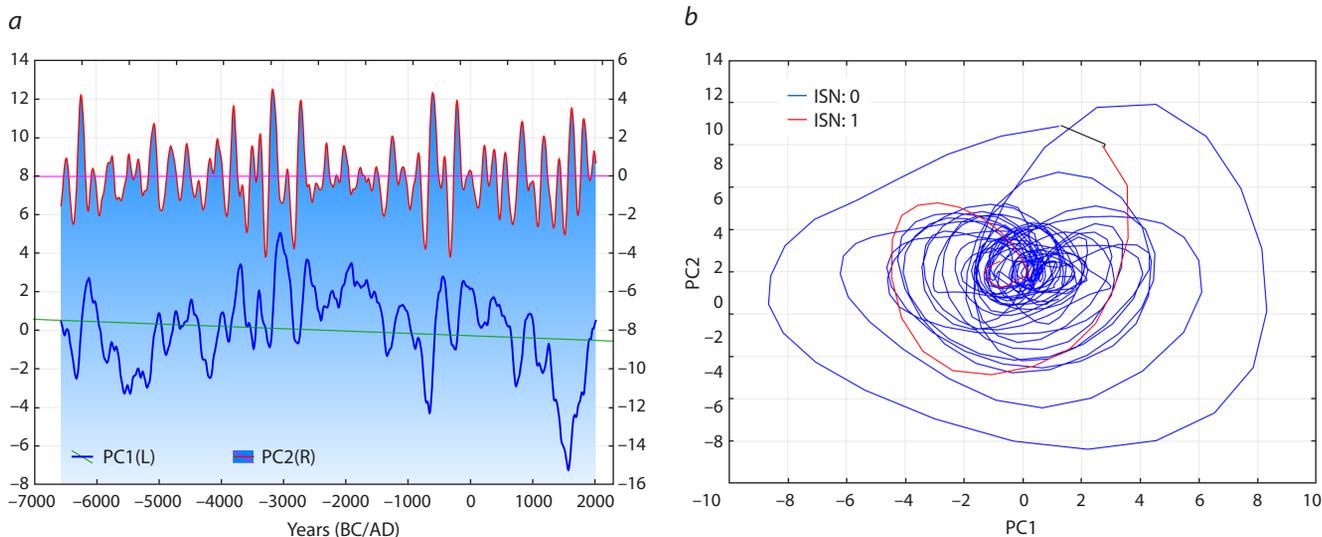


Fig. 7. The first two principal components (PC1 and PC2) of solar activity over the last 9000 years (a), their phase image (b). Variances PC1 = 36.3 %, PC2 = 28.5 %.

I and II millennia AD and in the XIV–XVIII centuries. The climate leaps observed now (see Figs. 6, 7) may be indicative of another transition from one climate phase to the next, whose closest analogue is the regime of the I millennium, warmer and more arid than the present.

In addition, multivariate analysis allows recognition of clearly different time ranges of the influence of heat flow onto the Earth (see Fig. 7). The reconstruction of solar activity for the last 9000 years (see Fig. 7, *a*) brings us to the suggestion that it is undergoing a profound change in the form of heat flow onto the Earth (Wu et al., 2018). The start of this change should be dated back to the XVI–XVII centuries, thereby rejecting the hypothesis of industrial activity as the main cause of the present climate change. Figure 7, *a* shows that the I millennium AD was warmer than the preceding I millennium BC or the subsequent II millennium AD.

Discussion

Our analysis of data reported by S.I. Barash (1989) indicates that lean years in Eastern Europe are those with droughts or excessive rains covering areas commensurable with the entire subcontinent (see Fig. 2, Supplementary Material 2). The same is observed in Western Europe, where the correlation between famine and excessive precipitation is more pronounced. Severe winters in Eastern Europe cause famine more often than in Western Europe. These differences stem from geographic features. Western Europe lies southwest and northeast and forms a natural barrier in the way of the North Atlantic Current, the main source of additional heat and moisture from the Atlantic (Hendry, 1982; Hogg, 1992; Hogg, Johns, 1995). The influence of Arctic air masses in Western Europe is weaker than in Eastern. In contrast, the effect of the North Atlantic Current on Eastern Europe is much weaker, and cold Arctic air masses greatly influence plant growth. Winter crops occupy a notable portion of arable areas in Eastern Europe, and their overwintering is of greater importance there.

The correlation between plague epidemics and crop failures in Eastern Europe is nearly two times closer than in Western (see Supplementary Material 2). This fact may be related to the predominance of arid and steppe-like agrolandscapes in Eastern Europe. They are home to populations of steppe rodents, which are plague seeders. Just these animals closely approach human dwellings in times of food shortage. The situation with smallpox is different: it shows poor correlation in both Eastern and Western Europe. Smallpox virus is transmitted from human to human without animal vectors. Therefore, only intense human traveling matters in this case.

For unknown natural causes, climate in Western and Eastern Europe undergoes centuries-long changes with abrupt transitions from one regime to another (see Fig. 3). The most likely cause of these changes is millennia-long variations in heat flow from the Sun to the Earth. The most notable transitions were noted in the onsets of the

I and II millennia (AD) and in the XIV–XVIII centuries. The climatic change observed at present may be a transition to another, unknown by now, climatic regime or the continuation of the cold climatic phase that started half a millennium ago. In fact, the current global warming began in the middle of the XX century, but it was considered the return to normal climatic conditions after the extraordinary cooling, the Little Ice Age of the XIV–XVIII centuries. It was not until recently that the notion appeared that this warming, if continued, would bring about catastrophic consequences, and humanity should be prepared for them in advance. As the main cause of this change was claimed to be human activity, its natural consequence was the illusion that it was the power of humanity to modify climate.

Although climate regime variations have long been studied and are of practical significance, their primary cause is still debatable. Some scientists state that they are caused by industrial activity (and proponents of this viewpoint succeeded in getting three Nobel awards: Peace Prize (Solomon et al., 2007), Prize in Economic Sciences (Nordhaus, 2019), and Prize in Physics (Manabe, 2019, 2023)). Others interpret the changes as a regular round of natural climatic fluctuations (Usoskin et al., 2014; Lobkovsky et al., 2022; and others).

The fact that the Earth receives nearly all heat from the Sun poses the question of regularities in the variation of this heat flow and predictability of changes. Naturally, attention is focused primarily on the trend and cyclic mode of the variation. However, as seen from analyses of solar activity, the results clearly depend on the scale of consideration. If we confine ourselves the epoch of regular direct solar activity observations over the last 300 years, the commonly known 11-year cycles are most pronounced. If we smooth them, the ascending trend is beyond dispute, and only unlimited increase can be forecasted, as is the present case.

When we increase the scale to the last millennium, we see a Middle Age dip in the middle of the II millennium AD, out of which we are just coming. The only prediction in this case is further rise. In covering three millennia, we see that such dips happened before, but they were not as deep as the current; therefore, the temperature after the end of such a dip slowly drifts to cooling, being accompanied by minor fluctuations. The prediction will be a short-time rise followed by a gentle trend to cooling. If we analyze the information on the longest attainable time span (see Fig. 6), we see that big dips happened even during the greatest rises, e. g., on the cusp of the IV and III millennia BC. Such a change may happen in XXI century. The cause of such dips is unknown, and no reliable statistical regularities have been revealed.

Conclusion

Scientists have long been discussing the prospects of using climatic models in the development of measures for the adaptation of various human activities to the current and expected climate changes (Kattsov et al., 2011, and others).

Climate is fully responsible for what lives and grows in a certain biome. Lately, the effect of climate changes on farming has been extensively investigated (Rauner, 1981; Sirotenko, 2001; Zolotokrylin et al., 2020; Cooper, Messina, 2023; and others). However, assessments of agricultural response in various regions are diverse. The main cause of this fact is differences in source data, methods for data processing, and methods for evaluating the influence.

The paradigm shift determined by the warming prediction is most often discussed in the context of probable aridization of huge areas (Trifonov, Karakhanyan, 2004), the resulting necessity of raising drought resistance of various crops (Zotova et al., 2020; Cooper, Messina, 2023), and search for new drought-resistant plant species applicable for cultivation (Baltzoi et al., 2015). Strategies of adaptation to climate changes may include better fitness of plant phenology to moisture availability (Ceccarelli et al., 2010), broader access to varieties with different duration of vegetative period (earliness) (Smolenskaya, Goncharov, 2023) in order to avoid stress at critical stages of their life cycles, water use improvement, and switch to breeding new-generation varieties for mitigating the rising unpredictability (Ceccarelli et al., 2010). Anyhow, breeders should take into consideration the high probability of climate changes in the decades to come, even if the formerly recorded extreme levels, which may aridize broad areas and shift agricultural zones from south to north, are not reached. In this case, the development of early varieties as a precautionary measure for improving agrocenosis adaptivity is an urgent task.

References

- Atkinson M.D., Kettlewell P.S., Poulton P.R., Hollins P.D. Grain quality in the broadbalk wheat experiment and the winter north atlantic oscillation. *J. Agric. Sci.* 2008;146(5):541-549. DOI 10.1017/S0021859608007958
- Baltzoi P., Fotia K., Kyrkas D., Nikolaou K., Paraskevopoulou A.T., Accogli A.R., Karras G. Low water-demand plants for landscaping and agricultural cultivations – A review regarding local species of Epirus/Greece and Apulia/Italy. *Agric. Agric. Sci. Proceedings*. 2015;4:250-260. DOI 10.1016/j.aaspro.2015.03.029
- Barash S.I. History of Crop Failures and Weather in Europe. Leningrad: Gidrometeoizdat Publ., 1989 (in Russian)
- Biswas A., Karak B.B., Usoskin I., Weissshaar E. Long-term modulation of solar cycles. *Space Sci. Rev.* 2023;219(3):19. DOI 10.1007/s11214-023-00968-w
- Ceccarelli S., Grando S., Maatougui M., Michael M., Slash M., Haghparsar R., Rahmanian M., Taheri A., Al-Yassin A., Benbelkacem A., Labdi M., Mimoun H., Nachit M. Plant breeding and climate changes. *J. Agric. Sci.* 2010;148(6):627-637. DOI 10.1017/S0021859610000651
- Clette F., Svalgaard L., Vaquero J.M., Cliver E.W. Revisiting the sunspot number: A 400-year perspective on the solar cycle. *Space Sci. Rev.* 2014;186:35-103. DOI 10.1007/s11214-014-0074-2
- Cooper M., Messina C.D. Breeding crops for drought-affected environments and improved climate resilience. *Plant Cell*. 2023;35(1):162-186. DOI 10.1093/plcell/koac321
- Eastwood R.J., Tambam B.B., Aboagye L.M., Akparov Z.I., Aladele S.E., Allen R., Amri A., ... Tapia Toll J., Vu D.T., Vu T.D., Way M.J., Yazbek M., Zorrilla C., Kilian B. Adapting agriculture to climate change: A synopsis of coordinated National Crop Wild Relative Seed Collecting Programs across five continents. *Plants*. 2022;11(14):1840. DOI 10.3390/plants11141840
- Eckardt N.A., Ainsworth E.A., Bahuguna R.N., Broadley M.R., Busch W., Carpita N.C., ... Rim E.Y., Ronald P.C., Salt D.E., Shigenaga A.M., Wang E., Wolfe M., Zhang X. Climate change challenges, plant science solutions. *Plant Cell*. 2023;35(1):24-66. DOI 10.1093/plcell/koac303
- Efimov V.M., Goncharov N.P. Weather extremes and crop failures in Europe. In: Abstracts from the Tenth Siberian Meeting on Climate-Ecological Monitoring. Tomsk, October 14-17, 2013. Tomsk: Agraf-Press Publ., 2013;53 (in Russian)
- Efimov V.M., Galaktionov Yu.K., Shushpanova N.F. Analysis and Forecast of Time Series by the Principal Component Method. Novosibirsk: Nauka Publ., 1988 (in Russian)
- Efimov V.M., Efimov K.V., Polunin D.A., Kovaleva V.Y. New possibilities of the PCA-Seq method in the analysis of time series (on the example of solar activity). *J. Phys. Conf. Ser.* 2021;2099(1):012034. DOI 10.1088/1742-6596/2099/1/012034
- Goncharov N.P. Plant domestication. *Vavilovskii Zhurnal Genetiki i Selekcii = Vavilov Journal of Genetics and Breeding*. 2013;17(4/2):884-899 (in Russian)
- Goncharov N.P., Chikida N.N. Genetics of the growth habit in *Aegilops squarrosa* L. *Genetika (Moscow)*. 1995;31(3):343-346
- Gower J.C. Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika*. 1966;53(3-4):325-338. DOI 10.1093/biomet/53.3-4.325
- Gupta A. Origin of agriculture and domestication of plants and animals linked to early Holocene climate amelioration. *Curr. Sci.* 2004;87(1):54-59
- Gurova T.A., Osipova G.M. The problem of combined stress resistance of plants under climate change in Siberia. *Sibirskiy Vestnik Sel'skokhozyaystvennoy Nauki = Siberian Herald of Agricultural Science*. 2018;48(2):81-92. DOI 10.26898/0370-8799-2018-2-11 (in Russian)
- Hendry R.M. On the structure of the deep Gulf Stream. *J. Mar. Res.* 1982;40(1):119-142
- Hogg N.G. On the transport of the Gulf Stream between Cape Hatteras and the Grand Banks. *Deep-Sea Res.* 1992;39(7-8):1231-1246. DOI 10.1016/0198-0149(92)90066-3
- Hogg N.G., Johns W.E. Western boundary currents. *Rev. Geophys.* 1995;33(S2):1311-1334. DOI 10.1029/95RG00491
- Jatayev S., Sukhikh I., Vavilova V., Smolenskaya S.E., Goncharov N.P., Kurishbayev A., Zotova L., Absattarova A., Serikbay D., Hu Y.-G., Borisjuk N., Gupta N.P., Jacobs B., de Groot S., Koekemoer F., Alharthi B., Lethola K., Cu D., Schramm C., Anderson P., Jenkins C., Soole K.L., Shavrukov Y., Langridge P. Green revolution 'stumbles' in a dry environment: Dwarf wheat plants with *Rht* genes fail to produce higher yield than taller genotypes under drought. *Plant Cell Environ.* 2020;43(10):2355-2364. DOI 10.1111/pce.13819

- Kamran A., Iqbal M., Spaner D. Flowering time in wheat (*Triticum aestivum* L.): A key factor for global adaptability. *Euphytica*. 2014; 197:1-26. DOI 10.1007/s10681-014-1075-7
- Karhunen K. Über lineare Methoden in der Wahrscheinlichkeitsrechnung. *Ann. Acad. Sci. Fennicae. Ser. A. I. Math.-Phys.* 1947; 37:1-79.
- Kattsov V.M., Meleshko V.P., Khlebnikova E.I., Shkolnik I.M. Assessment of climate impacts on agriculture in Russia over the first half of the XXI century: current opportunities provided by numerical modeling. *Agrofizika = Agrophysica*. 2011;(3):22-30 (in Russian)
- Leigh F.J., Wright T.I., Horsnell R.A., Dyer S., Bentley A.R. Progenitor species hold untapped diversity for potential climate-responsive traits for use in wheat breeding and crop improvement. *Heredity*. 2022;128(5):291-303. DOI 10.1038/s41437-022-00527-z
- Liang Y., Tian F. Plant genetics: Mechanisms of wild soybean adaptation. *Curr. Biol.* 2023;33(2):R82-R84. DOI 10.1016/j.cub.2022.12.009
- Liu Z., Hu Z., Lai X., Cao J., Zhang J., Ma X., Zhang X., Wang X., Ji W., Xu S. Multi-environmental population phenotyping suggests the higher risks of wheat *Rht-B1b* and *Rht-D1b* cultivars in global warming scenarios. *bioRxiv*. 2022;07:500398. DOI 10.1101/2022.07.18.500398
- Lobell D.B., Gourdji S.M. The influence of climate change on global crop productivity. *Plant Physiol.* 2012;160(4):1686-1697. DOI 10.1104/pp.112.208298
- Lobkovsky L.I., Baranov A.A., Ramazanov M.M., Vladimirova I.S., Gabsatarov Y.V., Semiletoev I.P., Alekseev D.A. Trigger mechanisms of gas hydrate decomposition, methane emissions, and glacier breakups in polar regions as a result of tectonic wave deformation. *Geosciences*. 2022;12(10):372. DOI 10.3390/geosciences12100372
- Loève M. Fonctions Aléatoires de Second Ordre. In: Lévy P. (Ed.). *Processus Stochastique et Mouvement Brownien*. Paris: Gauthier-Villars, 1948;366-420
- Manabe S. Role of greenhouse gas in climate change. *Tellus A: Dyn. Meteorol. Oceanogr.* 2019;71(1):1620078. DOI 10.1080/16000870.2019.1620078
- Manabe S. Nobel Lecture: Physical modeling of Earth's climate. *Rev. Mod. Phys.* 2023;95(1):010501. DOI 10.1103/RevModPhys.95.010501
- Morgounov A., Sonder K., Abugalieva A., Bhadauria V., Cuthbert R.D., Shamanin V., Zelenskiy Yu., DePauw R.M. Effect of climate change on spring wheat yields in North America and Eurasia in 1981-2015 and implications for breeding. *PLoS One*. 2018;13(10):e0204932. DOI 10.1371/journal.pone.0204932
- Nordhaus W. Climate change: The ultimate challenge for economics. *Am. Econ. Rev.* 2019;109(6):1991-2014. DOI 10.1257/aer.109.6.1991
- Oppenheim A.V., Schaffer R.W. *Digital Signal Processing*. New Jersey: Pearson, 1975
- Pearson K.L. III. On lines and planes of closest fit to systems of points in space. *Philos. Mag.* 1901;2(11):559-572
- Prentice R. Cultural responses to climate change in the Holocene. *Anthos*. 2009;1(1):3. DOI 10.15760/anthos.2009.41
- Rauner Yu.L. *Climate and Crop Productivity*. Moscow: Nauka Publ., 1981 (in Russian)
- Ruddiman W.F., Fuller D.Q., Kutzbach J.E., Tzedakis P.C., Kaplan J.O., Ellis E.C., Vavrus S., Roberts J., Fyfe C.N., He R.F., Lemmen C., Woodbridge J. Late Holocene climate: Natural or anthropogenic? *Rev. Geophys.* 2016;54(1):93-118. DOI 10.1002/2015RG000503
- Shnirelman V.A. *The Emergence of a Productive Economy: Centers of Ancient Agriculture*. Moscow: Librokom Publ., 2012 (in Russian)
- Sirotenko O.D. Crop modeling: Advances and problems. *Agron. J.* 2001;93(3):650-653. DOI 10.2134/agronj2001.933650ax
- Sleptsov A.M., Klimenko V.V. Generalization of paleoclimatic data and reconstruction of the climate of Eastern Europe for the last 2000 years. *Istoriya i Sovremennost' = History and Modernity*. 2005;(1):118-135 (in Russian)
- Smolenskaya S.E., Goncharov N.P. Allelic diversity of the *Vrn* genes and the control of growth habit and earliness in wheat. *Vavilovskii Zhurnal Genetiki i Seleksii = Vavilov Journal of Genetics and Breeding*. 2023;27(8):933-946. DOI 10.18699/VJGB-23-108 (in Russian)
- Smolenskaya S.E., Efimov V.M., Kruchinina Yu.V., Nemtsev B.F., Chepurnov G.Yu., Ovchinnikova E.S., Belan I.A., Zuev E.V., Zhou Chenxi, Piskarev V.V., Goncharov N.P. Earliness and morphotypes of common wheat cultivars of Western and Eastern Siberia. *Vavilovskii Zhurnal Genetiki i Seleksii = Vavilov Journal of Genetics and Breeding*. 2022;26(7):662-674. DOI 10.18699/VJGB-22-81 (in Russian)
- Solanki S.K., Usoskin I.G., Kromer B., Schüssler M., Beer J. An unusually active Sun during recent decades compared to the previous 11,000 years. *Nature*. 2004;431(7012):1084-1087. DOI 10.1038/nature02995
- Solomon S., Qin D., Manning M., Chen Z., Marquis M., Averyt K., Tignor M., Miller H. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge: Cambridge University Press, 2007
- Sukhikh I.S., Vavilova V.Y., Blinov A.G., Goncharov N.P. Diversity and phenotypical effect of the allelic variants of *Rht* dwarfing genes in wheat. *Russ. J. Genet.* 2021;57(2):127-138. DOI 10.1134/S1022795421020101
- Takens F. Detecting strange attractors in turbulence. In: Rand D., Young L.S. (Eds). *Dynamical Systems and Turbulence*. Warwick 1980. Lecture Notes in Mathematics. Vol. 898. Berlin: Springer, 1981;366-381. DOI 10.1007/BFb0091924
- Toynbee A.J. *A Study of History*. London: Oxford University Press, 1954
- Trifonov V.G., Karakhanyan A.S. *Geodynamics and History of Civilizations*. Moscow: Nauka Publ., 2004 (in Russian)
- Trifonova A.A., Dedova L.V., Zuev E.V., Goncharov N.P., Kudryavtsev A.M. Comparative analysis of the gene pool structure of *Triticum aestiopicum* wheat accessions conserved *ex situ* and recollected in field after 85 year later. *Biodivers. Conserv.* 2021; 30(2):329-342. DOI 10.1007/s10531-020-02091-6
- Usoskin I.G., Hulot G., Gallet Y., Roth R., Licht A., Joos F., Kovaltsov G.A., Thebault E., Khokhlov A. Evidence for distinct modes of solar activity. *Astron. Astrophys.* 2014;562(1):L10. DOI 10.1051/0004-6361/201423391
- Wu C.J., Usoskin I.G., Krivova N., Kovaltsov G.A., Baroni M., Bard E., Solanki S.K. Solar activity over nine millennia: A con-

- sistent multi-proxy reconstruction. *Astron. Astrophys.* 2018; 615: A93. DOI 10.1051/0004-6361/201731892
- Zolotokrylin A.N., Cherenkova E.A., Titkova T.B. Aridization of drylands in the European part of Russia: Secular trends and links to droughts. *Izvestiya Rossiiskoi Akademii Nauk. Seriya Geograficheskaya = Bulletin of the RAS. Geographic Series.* 2020;84(2): 207-217. DOI 10.31857/S258755662002017X (in Russian)
- Zotova L., Shamambaeva N., Lethola K., Alharthi B., Vavilova V., Smolenskaya S.E., Goncharov N.P., Jatayev S., Kurishbayev A., Gupta N.K., Gupta S., Schramm C., Anderson P., Jenkins C.L.D., Soole K.L., Shavrukov Yu. *TaDrAp1* and *TaDrAp2*, Partner genes of a transcription repressor, coordinate plant development and drought tolerance in spelt and bread wheat. *Int. J. Mol. Sci.* 2020;21(21):8296. DOI 10.3390/ijms21218296

Funding. This study was supported by the Russian Science Foundation (project No. 22-16-20026) and the Government of the Novosibirsk Region.

Financial transparency. Authors have no financial incentive in the presented materials or methods.

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

Received March 21, 2023. Revised November 13, 2023. Accepted November 18, 2023.

DOI 10.18699/vjgb-24-19

Study of a genetic collection of strawberry (*Fragaria* L.) for resistance to powdery mildew

A.S. Lyzhin , I.V. Luk'yanchuk 

I.V. Michurin Federal Scientific Center, Michurinsk, Russia
 Ranenburzhetc@yandex.ru

Abstract. Powdery mildew (*Sphaerotheca macularis* Mag. (syn. *Podosphaera aphanis* Wallr.)) is a dangerous disease of strawberry (*Fragaria* L.). The resistance of strawberry to powdery mildew is controlled polygenically. Several genetic loci with a large contribution to disease resistance have been identified in various strawberry varieties. Diagnostic DNA markers have been developed for QTL 08 *To-f*. They showed a high level of reliable gene detection in mapping populations. The purpose of this study was assessment of a strawberry genetic collection for resistance to powdery mildew and identification of promising strawberry forms for breeding for resistance to *S. macularis*. The objects of the study were wild species of the genus *Fragaria* L., varieties and selected seedlings of strawberry (*Fragaria* × *ananassa* Duch.) created in the I.V. Michurin Federal Scientific Center, and strawberry varieties introduced from various ecological and geographical regions. To identify QTL 08 *To-f*, DNA markers IB535110 and IB533828 were used. Locus 08 *To-f* was detected in 23.2 % of the analyzed strawberry genotypes, including wild species *F. moschata* and *F. orientalis*, strawberry varieties of Russian breeding (Bylinnaya and Sudarushka) and foreign breeding (Florence, Korona, Malwina, Ostara, Polka and Red Gauntlet). The correlation between the presence of markers IB535110 and IB533828 and phenotypic resistance (powdery mildew effect on strawberry plants is absent) was 0.649. The determination coefficient (R^2) showing the contribution of the studied locus to the manifestation of the trait was 0.421, that is, in 42.1 % of cases resistance was explained by the presence of QTL 08 *To-f*, and in 57.9 % of cases, by other genetic factors. All strawberry genotypes with locus 08 *To-f* were characterized by high field resistance to *S. macularis* in the conditions of Michurinsk, Tambov region. Thus, locus 08 *To-f* is promising for conferring resistance on local powdery mildew races, and markers IB535110 and IB533828 can be used in marker-assisted breeding programs to create powdery mildew-resistant strawberry genotypes.

Key words: strawberry; powdery mildew; resistance; molecular markers; QTL.

For citation: Lyzhin A.S., Luk'yanchuk I.V. Study of a genetic collection of strawberry (*Fragaria* L.) for resistance to powdery mildew. *Vavilovskii Zhurnal Genetiki i Seleksii* = *Vavilov Journal of Genetics and Breeding*. 2024;28(2):166-174. DOI 10.18699/vjgb-24-19

Изучение генетической коллекции земляники (*Fragaria* L.) по устойчивости к мучнистой росе

А.С. Лъжин , И.В. Лукъянчук 

Федеральный научный центр им. И.В. Мичурина, Мичуринск, Россия
 Ranenburzhetc@yandex.ru

Аннотация. Мучнистая роса (*Sphaerotheca macularis* Mag. (син. *Podosphaera aphanis* Wallr.)) – опасное заболевание земляники (*Fragaria* L.). Устойчивость земляники к мучнистой росе контролируется полигенно. У различных сортов земляники садовой идентифицировано несколько генетических локусов (QTL) с большим вкладом в устойчивость к болезни. Для QTL 08 *To-f* разработаны диагностические ДНК-маркеры, показавшие высокий уровень надежности выявления гена в картирующих популяциях. Цель настоящего исследования – изучение генетической коллекции земляники по устойчивости к мучнистой росе и идентификация перспективных для селекции на устойчивость к *S. macularis* форм. Объектами исследования были дикорастущие виды рода *Fragaria* L., сорта и отборные сеянцы земляники садовой (*F.* × *ananassa* Duch.) селекции Федерального научного центра им. И.В. Мичурина, а также формы, интродуцированные из различных эколого-географических регионов. Для идентификации QTL 08 *To-f* использовали маркеры IB535110 и IB533828. Локус 08 *To-f* обнаружен у 23.2 % генотипов земляники, в том числе у дикорастущих видов *F. moschata* и *F. orientalis*, сортов земляники садовой отечественной (Былинная, Сударушка) и зарубежной (Florence, Korona, Malwina, Ostara, Polka, Red Gauntlet) селекции. Степень корреляции между наличием маркеров IB535110 и IB533828 и фенотипической устойчивостью (признаки поражения мучнистой росой отсутствуют) составила 0.649. Коэффициент детерминации (R^2), показывающий вклад изучаемого локуса в формирование признака, равен 0.421, т.е. в 42.1 % случаев устойчивость определяется наличием QTL 08 *To-f*, тогда как в 57.9 % случаев влияние оказывают факторы внешней среды. Все генотипы земляники с локусом 08 *To-f* характеризуются высокой полевой устойчивостью к *S. macularis* в условиях г. Мичуринска Тамбовской области. Таким образом, локус 08 *To-f* является перспективным для при-

дания устойчивости к местным расам мучнистой росы, а маркеры IB535110 и IB533828 могут быть использованы в программах маркер-опосредованной селекции по созданию устойчивых к мучнистой росе генотипов земляники.

Ключевые слова: земляника; мучнистая роса; устойчивость; молекулярные маркеры; QTL.

Introduction

Powdery mildew is a dangerous disease of strawberry. The causative agent of powdery mildew is the obligate biotrophic fungus *Sphaerotheca macularis* Mag. (syn. *Podospaera aphanis* Wallr.). The greatest harmful effect to plantings is caused by the conidial stage of the pathogen – *Oidium fragariae* Harz. (Holod, Semenova, 2014; Tapia et al., 2021). Powdery mildew affects all above-ground plant organs. The infection manifests itself in the form of a white powdery coating of mycelium and conidia of the fungus. Severely affected strawberry leaves curl upward in the shape of a boat, infected peduncles form deformed fruits, tendrils and young rosettes are stunted in growth and subsequently die (Kennedy et al., 2013; Stolnikova, Kolesnikova, 2017). Strawberry yield losses from powdery mildew can exceed 60 % (Nelson et al., 1995; Lifshitz et al., 2007).

Powdery mildew especially affects strawberry plantings in protected soil (greenhouses, hotbeds, tunnels) due to favorable conditions for the development of the pathogen – elevated temperature and humidity (Sylla et al., 2013; Tapia et al., 2021). The development of the pathogen in strawberry plantings is facilitated by warm weather (temperature 18–24 °C) and high air humidity (about 100 %). The decline of the disease is observed when the air is excessively dry or there is an abundance of precipitation (it washes away the pathogen spores and improves the condition of the plants), as well as at temperatures below 15 °C and above 30 °C (Zubov, 1990, 2004). An earlier manifestation of the disease was noted in the spring after warm and snowy winters. In frosty and snowless winters, the main supply of infection dies, and late and weak development of the pathogen is observed (Govorova, Govorov, 2004).

Control of the spread of *S. macularis* in strawberry plantations is ensured primarily by the use of contact (sulfur) and systemic (captan, benomyl) fungicides (Bajpai et al., 2019; Palmer, Holmes, 2021). However, the active use of chemical plant protection products contradicts the global trend in the development of agriculture – its biologization and ecologization (Zhuchenko, 2009; Gorgitano, Pirilli, 2016). In addition, *S. macularis* is characterized by a high ability to develop resistance to fungicides (Carisse, Bouchard, 2010; Sombardier et al., 2010). In this regard, a promising direction for increasing resistance to powdery mildew is to identify from the existing assortment and create new strawberry varieties with genetically determined resistance to *S. macularis*.

Resistance of strawberry varieties to powdery mildew is controlled polygenically. The formation of the “powdery mildew resistance” trait depending on the genotype is influenced by both additive (the summed influence of alleles of one gene or several non-allelic genes expressed equally) and non-additive (interaction of alleles of a gene or non-allelic genes according to the type of dominance, overdominance and epistasis due to different levels of gene expression, and as a result one allele of a gene or gene suppressing another) gene effects. The

heritability of the “powdery mildew resistance” trait in the strawberry hybrid offspring, according to estimates by various authors, ranges from medium to high ($H^2 = 0.44–0.94$). This indicates prospects for increasing strawberry resistance to powdery mildew using breeding (Kennedy et al., 2014). Analysis of the inheritance of resistance to *S. macularis* in strawberry hybrid combinations shows continuous variability of hybrids from resistant to susceptible forms. In a number of strawberry combinations, transgression (additive genetic effects) may occur, leading to the appearance of seedlings that are superior in powdery mildew resistance to the parent forms. Additive effects, as reported by a number of authors (Zubov, 2004; Kennedy et al., 2014), play a major role in the formation of strawberry resistance to powdery mildew.

Some initial strawberry forms (wild species *F. orientalis*, *F. moschata* and *F. ovalis*, and interspecific hybrids 298-22-19-21 (FB₂ *F. orientalis*, *F. moschata*, *F. × ananassa*), 297-22-124, 297-28-84 (FB₁ *F. orientalis*, *F. × ananassa*) and 778-7 (FB₂ *F. ovalis*, *F. × ananassa*)) transfer a high level of powdery mildew resistance to a large number of hybrid forms, regardless of the combination of crossing (non-additive genetic effects). The predominance of non-additive gene effects makes it possible to identify donors of strawberry resistance to powdery mildew (Zubov, 2004; Davik, Honne, 2005).

In recent years, several major quantitative trait loci (QTLs) for strawberry resistance to powdery mildew have also been identified. However, they were characteristic only of specific crossing combinations and their manifestation varied depending on prevailing weather conditions. Thus, six QTLs were identified in the hybrid combination Emily × Fenella, and five QTLs were identified in the hybrid combination Red Gauntlet × Hapil. The most stable QTLs include the *FarPa1C* (Emily × Fenella) and *FarPa6D2* (Red Gauntlet × Hapil) loci. At the same time, validation of the identified QTLs in a genetically diverse sample of strawberry varieties and forms showed the high uniqueness of the identified powdery mildew resistance loci and their almost complete absence in other genotypes, which limits the possibilities of their use in strawberry breeding (Cockerton et al., 2018).

In the hybrid combination Sonata × Babett, three loci of resistance to powdery mildew (*FxaPMR5b*, *FxaPMR7A* and *FxaPMR7X2*) were identified. Of these, one QTL (*FxaPMR7A*) was identified when strawberry plants were cultivated in a greenhouse, and two QTLs (*FxaPMR5b* and *FxaPMR7X2*) were identified in strawberry plants of open ground (Sargent et al., 2019). However, diagnostic DNA markers for these loci have not been developed, which prevents their use in breeding practice for identifying strawberry forms resistant to *S. macularis*. In 2020, H. Koishihara and co-authors, based on an analysis of the strawberry hybrid combinations Miyazaki Natsu Haruka × 08 To-f, Miyazaki Natsu Haruka × Ohkimi and 09s E-b45e × Miyazaki Natsu Haruka, identified another QTL (08 To-f) with a high contribution to the manifestation of resistance to powdery mildew (15.7 %) (Koishihara et al.,

2020). The *08 To-f* locus size was 6.83 cM. To identify the *08 To-f* locus in the strawberry germplasm, diagnostic DNA markers IB535110 and IB533828 were developed. The reliability of identification of strawberry genotypes resistant to powdery mildew using the IB535110 and IB533828 markers in the analyzed crossing combinations was 98.5 %. These markers, according to the authors' recommendations (Koishihara et al, 2020), can be used for marker-assisted screening of strawberry forms resistant to powdery mildew.

The *MLO* locus (Mildew Resistance Locus O), encoding genes that affect susceptibility to the pathogen, can also make a certain contribution to the formation of strawberry resistance to *S. macularis*. Blocking the expression of alleles of these genes or transferring them to a recessive state contributes to the manifestation of strawberry resistance to powdery mildew. 68 *MLO* sequences have been identified in octoploid strawberry. The most important of them are the *FaMLO10*, *FaMLO17* and *FaMLO20* loci (Tapia et al., 2021). In addition, strawberry resistance to powdery mildew is influenced by the *TGA* family transcription factors (involved in the metabolism of salicylic acid). 11 *FaTGA* genes were identified in strawberry varieties. The genes *FaTGA1*, *FaTGA2*, *FaTGA5*, *FaTGA7*, *FaTGA8* and *FaTGA10* are characterized by the greatest specificity to powdery mildew infection (Feng et al., 2020).

Applied research is also being conducted to develop diagnostic DNA markers to identify powdery mildew-resistant strawberry genotypes. In particular, the scientific group of L.J. Cheng, based on SSR analysis of hybrid seedlings from crossing the varieties Darselect (susceptible to powdery mildew) and Sweet Charlie (resistant to powdery mildew), identified SSR markers FSS50 and FSS121, which showed a close relationship with the presence of phenotypic resistance to *S. macularis* (Liu et al., 2012). H.-J. Je with co-authors, based on the analysis of the hybrid combination Akihime (susceptible) × Seolhyang (resistant), developed the CAPS marker SP1-Eae I, which allows identifying strawberry genotypes resistant to powdery mildew (Je et al., 2015).

The development of diagnostic DNA markers suitable for use in marker-assisted breeding programs is an important step in increasing the efficiency of selection of pathogen-resistant genotypes and creating new varieties (Whitaker et al., 2012). However, data on the suitability of the identified molecular markers for the analysis of genetically diverse strawberry varieties or wild species of the genus *Fragaria L.*, and the prevalence of resistance loci in the strawberry germplasm, are not provided.

The purpose of this research was to study the strawberry genetic collection according to powdery mildew resistance and identify promising forms for breeding for resistance to *S. macularis*.

Materials and methods

In this work, we used 43 samples, consisting of wild species of the genus *Fragaria L.*, varieties and selected seedlings of garden strawberry (*F. × ananassa* Duch.) created in the FSSI "I.V. Michurin FSC", and strawberry forms originating from different ecological and geographical regions (Table 1).

Weather and climatic conditions of the growing seasons during research (2018–2022) differed from the long-term average. In 2018, average monthly air temperatures exceeded

long-term values by 0.7–2.4 °C. 2019 and 2020 have both cooler (July and August 2019, April and May 2020) and hotter months (April, May and June 2019, June, July and September 2020). In 2021 and 2022, the average monthly air temperature exceeded long-term values by 0.6–4.2 °C and 0.8–4.5 °C. Precipitation amount in 2018–2020 was lower than the long-term average values by 13.9–28.7 % and its distribution across the months was quite uniform. In 2021 and 2022, precipitation amount exceeded the long-term average by 2.3 and 16.1 %, respectively. At the same time, precipitation was uneven: there were periods of high humidity (for example, in the 2nd decade of July 2022, 58.8 mm of precipitation fell) and periods of insufficient and weak water availability (August 2021 – 31.3 mm of precipitation, August 2022 – 40.1 mm of precipitation.).

Phenotypic assessment of the powdery mildew resistance of strawberry genotypes was carried out in field conditions against a natural infectious background on a scale from 0 to 5 points, where 0 – no plants affected, 5 – all vegetative organs of the plant are severely affected (Zubov, 1990).

Total genomic DNA was isolated from fresh leaves; extraction was carried out using the CTAB method with modifications described by I.V. Luk'yanchuk et al. (2018).

The QTL *08 To-f* was identified with the dominant markers IB535110 and IB533828. DNA marker IB535110 was represented by a 500 bp amplicon. DNA marker IB533828 was represented by an amplicon of about 120 bp. These products were amplified only if the strawberry genotype had the *08 To-f* locus (Koishihara et al., 2020).

PCR reactions were performed in 15 µl final volume containing: 20 ng of genomic DNA, 0.2 mM of each dNTP, 2.5 mM MgCl₂, 0.2 µM of each primer, 0.2 U of Taq DNA polymerase and 1.5 µl of PCR-buffer (+ (NH₄)₂SO₄, –KCl). All components were produced by Thermo Fisher Scientific (USA).

Amplification was performed in a T100 Thermal Cycler (Bio-Rad, USA). PCR conditions were as follows: 94 °C for 1 min, followed by 35 cycles of 94 °C for 30 s, T_m for 30 s, 72 °C for 1 min, and with a final extension of 72 for 5 min. T_m – primers annealing temperature: 35110_v1F/35110_v1R – 60 °C; 22828_v6F/22828_v6R – 58 °C.

The amplification products were separated on a 2 % agarose gel and visualized by ethidium bromide staining. Gene Ruler 100 bp DNA Ladder (Thermo Fisher Scientific) was used as a molecular weight marker.

Experimental data were processed by methods of mathematical statistics using Microsoft Excel 2016 and STATISTICA 6.0. A comparison of the frequency of occurrence of the "powdery mildew resistance" trait in samples of Russian and foreign strawberry varieties was carried out using Student's *t* test. The reliability of the results of phytopathological assessment of strawberry genotypes was assessed using two-factor ANOVA. The contribution of a genetic determinant to the formation of a trait was assessed using the heritability coefficient (H^2), calculated using the following formula: $H^2 = V_G / V_P$, where V_G – genotypic variance, V_P – total phenotypic variance. The identification of a statistical relationship between the presence of DNA markers for the *08 To-f* locus of powdery mildew resistance and its phenotypic manifestation was carried out by regression analysis using the *F*-test.

Table 1. List of wild strawberry species, varieties and selected forms

Genotype	Crossing combination	Originator/origin
<i>F. orientalis</i> Los.	Wild species	Primorye Territory, Russia
<i>F. moschata</i> Duch.		European part of Russia
<i>F. ovalis</i> (Lehm.) Rydb.		British Columbia, Canada
<i>F. virginiana</i> subsp. <i>platypetala</i> (Rydb.) Staudt		
Izbrannitsa	Festivalnaya × Senga Sengana	I.V. Michurin Federal Scientific Center, Russia
Lastochka	922-67 × Privlekatelnaya	
Rubinovy kaskad		
Privlekatelnaya	Rubinovy kulon × Allbritton	
Urozhaynaya CGL	Senga Sengana × Redcoat	
Feyerverk		
Flora		
298-19-9-43	FB ₂ <i>F. orientalis</i> Los., <i>F. moschata</i> Duch., <i>F.</i> × <i>ananassa</i> Duch.	
28-11	Lakomaya × Maryshka	
922-67	FB ₃ <i>F. ovalis</i> (Lehm.) Rydb.	
Bereginya	Solovushka × Unduka	Federal Horticultural Research Center for Breeding, Agrotechnology and Nursery, Russia
Borovitskaya	Nadezhda × Red Gauntlet	
Vityaz	Syurpriz olimpiade × Festivalnaya romashka	
Kokinskaya zarya	Slavutich × 157-7	
Kubata	Kubenskaya × Holiday	
Sudarushka	Festivalnaya × Roxana	
Tsaritsa	Venta × Red Gauntlet	
Tsarskoselskaya	Pavlovchanka × Holiday	Institute for Engineering and Environmental problems in Agricultural Production, Russia
Festivalnaya	Obilnaya × Premier	Federal Research Center the N.I. Vavilov All-Russian Institute of Plant Genetic Resources, Russia
Bylinnaya	Persikovaya × Seyanets VIR-228613	Krymsk Experimental Breeding Station of Federal Research Center the N.I. Vavilov All-Russian Institute of Plant Genetic Resources, Russia
Clery	Sweet Charlie × Onebor	Consorzio Italiano Vivaisti (CIV), Italy
Joly	T2-6 × A20-17	
Murano	R6R1-26 × A030-12	
Limalexia	E0011 × E0021	Limgroup, Netherlands
Driscoll Jubilee	50C130 × 19A331	Driscoll's, United Kingdom
Korona	Tamella × Unduka	Plant Research International – WUR, Netherlands
Ostara	Red Gauntlet × Masherahs Daurente	
Polka	Unduka × Sivetta	
Sonata	Elsanta × Polka	
Albion	Diamante × Cal 94.16-1	University of California Davis, USA
Monterey	Cal. 27-85.06 × Albion	
Vima Tarda	Vima Zanta × Vicoda	Vissers International BV, Netherlands
Malwina	Sophie × clone Schimmelpfeng, Weihenstefan	Peter Stoppel, Germany
Florence	[Tioga × (Red Gauntlet × (Wiltguard × Gorella))] × (Providence × self)	MEIOSIS Ltd, United Kingdom
Symphony	Rhapsody × Holiday	MyInfield Research Services Ltd, United Kingdom
Kimberly	Gorella × Chandler	Gebr. Vissers, Netherlands
Flamenco	Evita × EMR77	East Malling Research Station, United Kingdom
Red Gauntlet	(New Jersey 1051 × Climax) × (Climax × New Jersey 1051)	Scotland
Barlidaun	MDUS 2359 × MDUS 2713	USA

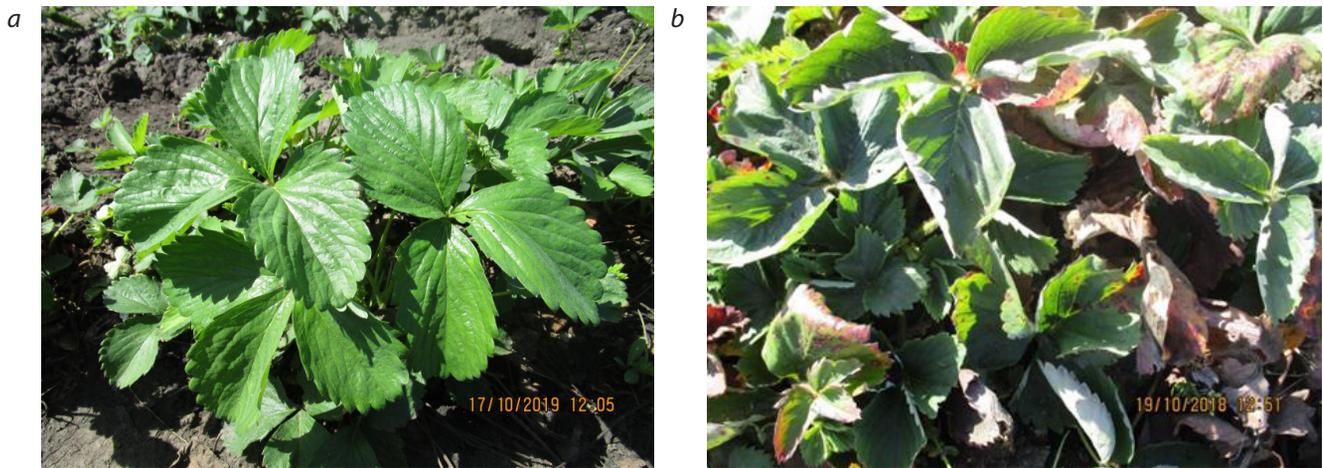


Fig. 1. Strawberry plants affected by powdery mildew.
a – the Sudarushka variety, no effect; b – the Festivalnaya variety, degree of effect – 3.0 points.

Results and discussion

During the research (2018–2022), in the conditions of Michurinsk, Tambov region, there were both relatively favorable years for the *S. macularis* (2018, 2019 and 2020) and unfavorable years for the pathogen (2021 and 2022). Under favorable conditions, the degree of strawberry genotypes affected by powdery mildew varied in the range from 0 to 4 points; under conditions unfavorable for the pathogen, plants affected did not exceed 1 point (Fig. 1, Table 2).

Analysis of variance of the obtained results showed a statistically significant influence on the manifestation of the “powdery mildew resistance” trait of both the genotype and the prevailing weather conditions (Table 3).

At the same time, the predominance of the influence of environmental conditions over the genotype in the formation of the “powdery mildew resistance” trait was noted. The heritability coefficient (H^2) was 34.8%. The low contribution of genetic variance to the total phenotypic variation of a trait is explained by weather and climatic conditions: out of five years of research, two years (2021 and 2022) were unfavorable for *S. macularis*, and as a result phenotypic differences between resistant and susceptible strawberry varieties did not appear. Taking into account only the years that were relatively favorable for the pathogen (2018, 2019 and 2020), when the differences between strawberry varieties were more contrasting, the contribution of the genotype to the formation of the trait was 50.7%.

Variation in the heritability coefficient of the powdery mildew resistance locus depending on the degree of infectious load is also described by other researchers. In particular, it is noted that low phenotypic variability of a trait reduces the calculated heritability coefficient (Kennedy et al., 2014).

Most of the studied strawberry collection (88.4% forms), including all analyzed wild species of the genus *Fragaria*, were characterized by high-level resistance to *S. macularis* – the degree of plants affected did not exceed 1.0 points. It should be noted that in the strawberry variety Sonata, which has resistance loci *FxaPMR5b*, *FxaPMR7A*, *FxaPMR7X2*

(Sargent et al., 2019), the degree of effect over the years of research was on average 0.2 points, and in some years – 1.0 points. In addition, the strawberry variety Clery, obtained by hybridization of the powdery mildew-resistant variety Sweet Charlie (Liu et al., 2012), in some years was characterized by leaf effect of 1 point (the average affected score for 2018–2022 was 0.4).

Over the years of research, 18 out of 43 analyzed strawberry genotypes (41.9% of the total number of forms) were characterized by the absence of powdery mildew effect. Among the strawberry varieties, the absence of *S. macularis* effect was detected in 35.9% genotypes. Among the strawberry varieties of Russian breeding, the absence of *S. macularis* effect was detected in 29.4% genotypes, among the foreign strawberry varieties – in 47.4% forms. At the same time, the differences in the distribution of the “powdery mildew resistance” trait in the samples of Russian and foreign strawberry varieties were statistically insignificant (at a significance level of $p \leq 0.05$ $t_{\text{fact}} = 0.4 \leq t_{\text{st}} = 4.3$).

The wild species *F. moschata*, *F. orientalis*, *F. ovalis* and *F. virginiana* subsp. *platypetala*, strawberry varieties Borovitskaya, Bylinnaya, Kubata, Sudarushka and Flora (Russian breeding), and Florence, Korona, Limalexia, Malwina, Murano, Ostara, Polka, Red Gauntlet and Vima Tarda (foreign breeding) are resistant to *S. macularis* (no plants affected over the years of research). The presence of populations with a high level of powdery mildew resistance in many wild strawberry species (*F. ovalis*, *F. virginiana*, *F. chiloensis*, etc.) is confirmed by literature data. At the same time, the ecological and geographical disunity of the habitats of these species suggests the presence of different mechanisms of plant resistance to *S. macularis* (Kennedy et al., 2013).

It should also be noted that strawberry genotypes that are resistant to powdery mildew in the Tambov region can be affected by the pathogen in other regions. For example, in the conditions of the Altai Territory (Western Siberia), strawberry varieties Korona and Polka were affected by 1.0–1.5 points, and the variety Bylinnaya – by up to 2.5 points (Stolnikova,

Table 2. Powdery mildew effect on the studied strawberry genotypes in the conditions of Michurinsk, Tambov region (2018–2022), and the presence of DNA markers for the *08 To-f* resistance locus

Genotype	Degree of effect, points						Phenotypic resistance*	QTL <i>08 To-f</i>	
	2018	2019	2020	2021	2022	Average		IB535110	IB533828
<i>F. moschata</i> Duch.	0.0	0.0	0.0	0.0	0.0	0.0	R	1	1
<i>F. orientalis</i> Los.	0.0	0.0	0.0	0.0	0.0	0.0	R	1	1
<i>F. ovalis</i> (Lehm.) Rydb.	0.0	0.0	0.0	0.0	0.0	0.0	R	0	0
<i>F. virginiana</i> subsp. <i>platypetala</i> (Rydb.) Staudt	0.0	0.0	0.0	0.0	0.0	0.0	R	0	0
Bereginya	0.0	1.0	0.0	0.0	1.0	0.4	S	0	0
Borovitskaya	0.0	0.0	0.0	0.0	0.0	0.0	R	0	0
Bylinnaya	0.0	0.0	0.0	0.0	0.0	0.0	R	1	1
Vityaz	1.0	1.0	0.0	0.0	0.0	0.4	S	0	0
Izbrannitsa	1.0	0.0	0.0	0.0	0.0	0.2	S	0	0
Kokinskaya zarya	1.0	0.0	1.0	0.0	0.0	0.4	S	0	0
Kubata	0.0	0.0	0.0	0.0	0.0	0.0	R	0	0
Lastochka	1.0	0.0	0.0	0.0	0.0	0.2	S	0	0
Privlekatelnaya	2.0	2.0	1.0	0.0	1.0	1.2	S	0	0
Rubinovy kaskad	0.0	1.0	0.0	0.0	0.0	0.2	S	0	0
Sudarushka	0.0	0.0	0.0	0.0	0.0	0.0	R	1	1
Urozhaynaya CGL	1.0	1.0	0.0	0.0	0.0	0.4	S	0	0
Feyerverk	1.0	0.0	0.0	0.0	0.0	0.2	S	0	0
Festivalnaya	3.0	2.0	2.0	1.0	1.0	1.8	S	0	0
Flora	0.0	0.0	0.0	0.0	0.0	0.0	R	0	0
Tsarskoselskaya	1.0	1.0	1.0	0.0	0.0	0.6	S	0	0
Tsaritsa	1.0	0.0	0.0	0.0	0.0	0.2	S	0	0
Albion	1.0	1.0	1.0	0.0	0.0	0.6	S	0	0
Barlidaun	3.0	2.0	2.0	1.0	1.0	1.8	S	0	0
Clery	1.0	0.0	1.0	0.0	0.0	0.4	S	0	0
Driscoll Jubilee	1.0	0.0	0.0	0.0	0.0	0.2	S	0	0
Flamenco	0.0	1.0	0.0	0.0	0.0	0.2	S	0	0
Florence	0.0	0.0	0.0	0.0	0.0	0.0	R	1	1
Joly	1.0	0.0	1.0	0.0	0.0	0.4	S	0	0
Kimberly	1.0	1.0	1.0	0.0	1.0	0.8	S	0	0
Korona	0.0	0.0	0.0	0.0	0.0	0.0	R	1	1
Limalexia	0.0	0.0	0.0	0.0	0.0	0.0	R	0	0
Malwina	0.0	0.0	0.0	0.0	0.0	0.0	R	1	1
Monterey	1.0	1.0	0.0	0.0	0.0	0.4	S	0	0
Murano	0.0	0.0	0.0	0.0	0.0	0.0	R	0	0
Ostara	0.0	0.0	0.0	0.0	0.0	0.0	R	1	1
Polka	0.0	0.0	0.0	0.0	0.0	0.0	R	1	1
Red Gauntlet	0.0	0.0	0.0	0.0	0.0	0.0	R	1	1
Sonata	1.0	0.0	0.0	0.0	0.0	0.2	S	0	0
Symphony	1.0	0.0	0.0	0.0	0.0	0.2	S	0	0
Vima Tarda	0.0	0.0	0.0	0.0	0.0	0.0	R	0	0
298-19-9-43	2.0	2.0	2.0	1.0	1.0	1.6	S	0	0
922-67	1.0	0.0	0.0	0.0	0.0	0.2	S	0	0
28-11	4.0	3.0	2.0	1.0	1.0	2.2	S	0	0
HCP _{0.5}						0.1			

* R – resistance (no plants affected); S – susceptibility (affected plants were noted); 1 – target amplicon of DNA marker is present; 0 – target amplicon of DNA marker is absent.

Table 3. ANOVA test of powdery mildew effect on strawberry genotypes over the years of research

Source of variation	SS	df	MS	F	p-value	Fcrit.
Genotype	63.42326	42	1.510078	10.07462	2.23E-28	1.457365
Weather conditions	10.8186	4	2.704651	18.04433	2.4E-12	2.425453
Error	25.1814	168	0.149889			
Total	99.42326	214				

Note. SS – sum of squared deviations; df – degrees of freedom; MS – mean squares; F – F-test (fact. value); p-value – significance of the results; Fcrit. – F-test (crit. value).

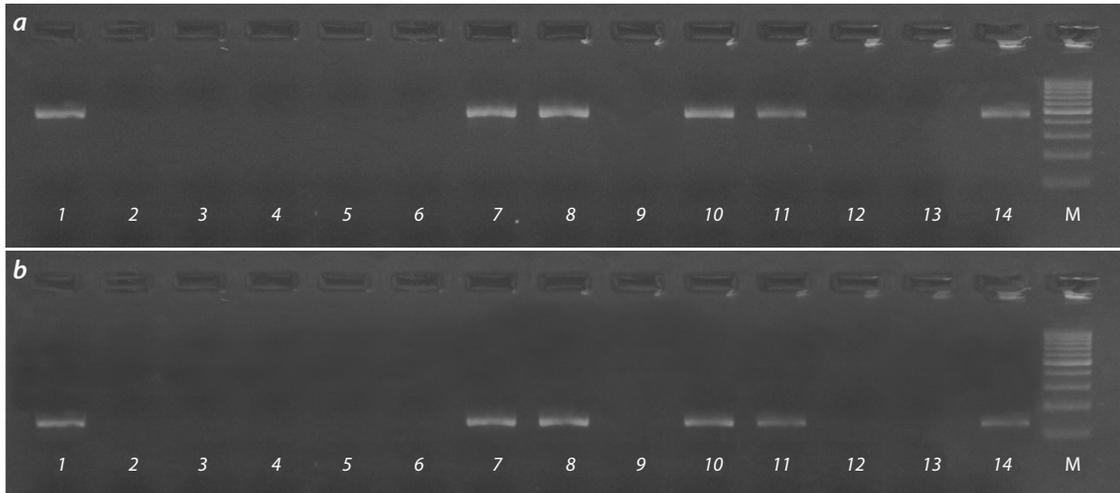


Fig. 2. Electrophoresis profile of markers IB535110 (a) and IB533828 (b) of 14 out of 43 analyzed strawberry genotypes.

1 – Red Gauntlet; 2 – Symphony; 3 – Bereginya; 4 – Izbrannitsa; 5 – Limalexia; 6 – Monterey; 7 – Bylinnaya; 8 – Korona; 9 – Barlidaun; 10 – Sudarushka; 11 – Polka; 12 – *F. ovalis*; 13 – *F. virginiana* subsp. *Platypetala*; 14 – *F. moschata*. M – molecular weight marker.

Kolesnikova, 2017). The results obtained are explained by the presence of physiological races of *S. macularis* that are specific to different regions.

To identify genetic factors of resistance, molecular genetic screening of the analyzed collection of strawberry genotypes was carried out using DNA markers IB535110 and IB533828, linked to the *08 To-f* powdery mildew resistance locus. Markers IB535110 and IB533828 were identified in 10 forms out of 43, which is 23.2 %. The results obtained for markers IB535110 and IB533828 are characterized by 100 % agreement with the phenotype assessment. An example of identification is shown in Fig. 2; the results are shown in Table 2.

Among the analyzed wild strawberries, the *08 To-f* locus was identified in *F. moschata* and *F. orientalis*. At the same time, the selected strawberry form 298-19-9-43 (three-species hybrid, second backcross generation from hybridization of *F. moschata* and *F. orientalis*) QTL *08 To-f* apparently did not inherit it from the original species, since the average degree of powdery mildew affect over the years of research was 1.6 points, the maximum was 2.0 points. Among the strawberry varieties, the proportion of genotypes with an identified powdery mildew resistance locus was 22.2 % (8 varieties out of 36). Among the analyzed Russian strawberry varieties, QTL *08 To-f* is present in two forms (Bylinnaya and Sudarushka) out of 17, which is 11.7 %. Among the analyzed foreign strawberry varieties, QTL *08 To-f* is present in six genotypes

(Florence, Korona, Malwina, Ostara, Polka and Red Gauntlet) out of 19, which is 31.6 %. It should also be noted that in the strawberry variety Sonata, in which the *FxaPMR5b*, *FxaPMR7A* and *FxaPMR7X2* loci were identified (Sargent et al., 2019), the *08 To-f* locus was not identified (markers IB535110 and IB533828 are absent). The obtained results confirm that the powdery mildew resistance of the Sonata variety is determined by QTL identified earlier.

Based on a BLAST analysis of the sequences of the used markers, it was suggested that the loci *FxaPMR7A* and *08 To-f* are orthologous (Sargent et al., 2019). For the strawberry variety Red Gauntlet, other QTLs were also mapped on chromosomes 2A, 4B, 6D, 7C and 7D (Cockerton et al., 2018). Therefore, in the Red Gauntlet variety, phenotypic resistance can be determined by the cumulative effect of several QTLs, and additional studies are necessary to compare the results obtained and clarify the number of identified QTLs and their location.

Analysis of the origin of the strawberry varieties, for which amplification fragments with markers IB535110 and IB533828 were obtained, indicates that the Florence and Ostara varieties were obtained using the Red Gauntlet variety, which, according to the data obtained, is the source of QTL *08 To-f*. The strawberry variety Korona was obtained in the crossing combination Tamella × Unduka, the Polka variety was obtained in the hybrid combination Unduka × Sivetta,

Table 4. Results of regression analysis of the dependence of strawberry phenotypic resistance to powdery mildew on the presence of the *08 To-f* locus

Index	df	SS	MS	F	F-value
Regression	1	4.40451	4.40451	29.79651	2.54E-06
Residual	41	6.060606	0.14782		
Total	42	10.46512			

Note. df – degrees of freedom; SS – sum of squared deviations; MS – mean squares; F – F-test (fact. value); F-value – significance of the results.

and therefore the source of QTL *08 To-f* for these forms is presumably the Unduka variety. To clarify, it is necessary to analyze the original parental strawberry forms for the presence of diagnostic markers IB535110 and IB533828.

It should also be noted that using the strawberry varieties Red Gauntlet and Unduka, varieties Borovitskaya (Nadezhda × Red Gauntlet), Tsaritsa (Venta × Red Gauntlet) and Bereginia (Solovushka × Unduka) were obtained, which, according to the results of a molecular genetic analysis, do not have the *08 To-f* powdery mildew resistance locus. The strawberry variety Bylinnaya was obtained in the crossing combination Persikovaya × Seyanets VIR-228613; for the original parental forms, there is no data on the presence of QTL *08 To-f*. For the strawberry variety Sudarushka, the initial parental forms are Festivalnaya and Roxana. The strawberry variety Festivalnaya was affected by powdery mildew over the years of research by an average of 1.8, with the effect varying over the years from 1.0 points to 3.0 points; the variety does not have markers IB535110 and IB533828. Therefore, the source of QTL *08 To-f* for the strawberry variety Sudarushka is presumably the Roxana variety.

A comparison of the results of molecular genetic analysis and phenotypic assessment of resistance to powdery mildew showed that all strawberry genotypes with identified markers IB535110 and IB533828 in the conditions of Michurinsk, Tambov region, are characterized by field resistance to *S. macularis* (no plants affected over the years of research). Thus, QTL *08 To-f* is a promising component of the genetic determinant of resistance to local races of *S. macularis*, and the strawberry varieties Bylinnaya, Sudarushka, Florence, Korona, Malwina, Ostara, Polka and Red Gauntlet as well as the wild species *F. moschata* and *F. orientalis* are valuable initial forms that can be used in breeding programs to create pathogen-resistant strawberry genotypes. At the same time, to reduce research time, financial and labor resources, it is permissible to use one of the two DNA markers. The disadvantage of the diagnostic markers used is the presence of only one target amplicon, and therefore it is possible to obtain false negative results due to PCR inhibition. To exclude false negative results, it is necessary to conduct a preliminary assessment of the quality of the extracted total strawberry DNA.

In addition, according to previous studies, the strawberry variety Bylinnaya is characterized by the presence of the *Rpfl* red stele root rot resistance gene (the causative agent is *Phytophthora fragariae* var. *fragariae* Hickman) (Lyzhin, Luk'yanchuk, 2020), and the strawberry variety Sudarushka is characterized by the presence of the *Rca2* anthracnose resistance gene (the causative agent is *Colletotrichum acuta-*

tum J.H. Simmonds) (Lyzhin et al., 2019). Therefore, these strawberry varieties are complex sources of resistance alleles to fungal pathogens.

It should also be noted that some of the studied strawberry genotypes (wild species *F. ovalis* and *F. virginiana* subsp. *platypetala*; strawberry varieties Borovitskaya, Kubata, Flora, Limalexia, Murano and Vima Tarda), in which QTL *08 To-f* is absent, were not affected by powdery mildew over the years of research.

The dependence of phenotypic resistance to *S. macularis* on the presence of QTL *08 To-f* in the genotype in the analyzed strawberry forms is described by the regression equation $y = 0.758x + 0.242$. Testing the significance of the regression model using Fisher's *F* test showed that at a significance level of 0.05, the null hypothesis about the absence of dependence between the variables is refuted (Table 4).

The degree of correlation between the presence of markers IB535110 and IB533828 and phenotypic resistance (no powdery mildew infection) was 0.649, which on the Chaddock scale corresponds to a perceptible communication between the traits ($0.5 < r_{xy} < 0.7$). The coefficient of determination (R^2), showing the contribution of the studied locus to the trait manifestation, is equal to 0.420, that is, in 42 % of cases, phenotypic resistance is determined by the presence of QTL *08 To-f*, and in 58 % of cases, it is affected by other factors. According to literature data, the coefficient of determination (R^2) for some resistance loci identified for powdery mildew varied from 0.16 to 0.57 (Cockerton et al., 2018).

The obtained results indicate the presence of additional genetic determinants of resistance to *S. macularis* in these strawberry forms. In this regard, identification of powdery mildew-resistant genotypes in the strawberry hybrid progeny using diagnostic DNA markers IB535110 and IB533828 is possible if the parental forms have QTL *08 To-f*. In other cases, strawberry resistance to *S. macularis* may be controlled by other genetic factors and, therefore, the use of markers IB535110 and IB533828 is inappropriate.

Conclusion

Thus, diagnostic DNA markers IB535110 and IB533828 make it possible to reliably identify the *08 To-f* powdery mildew resistance locus in the strawberry genoplasm and can be used in marker-assisted strawberry assortment improvement programs. Promising sources of resistance to *S. macularis*, according to the results of molecular genetic analysis, are the wild species *F. moschata* and *F. orientalis*, and strawberry varieties Bylinnaya and Sudarushka (Russian breeding), Florence, Korona, Malwina, Ostara, Polka and Red Gauntlet (foreign breeding).

References

- Bajpai S., Shukla P.S., Asiedu S., Pruski K., Prithiviraj B. A biostimulant preparation of brown seaweed *Ascophyllum nodosum* suppresses powdery mildew of strawberry. *Plant Pathol. J.* 2019;35(5): 406-416. DOI 10.5423/PPJ.OA.03.2019.0066
- Carisse O., Bouchard J. Age-related susceptibility of strawberry leaves and berries to infection by *Podosphaera aphanis*. *Crop Prot.* 2010;29(9):969-978. DOI 10.1016/j.cropro.2010.03.008
- Cockerton H.M., Vickerstaff R.J., Karlström A., Wilson F., Sobczyk M., He J.Q., Sargent D.J., Passey A.J., McLeary K.J., Pakozdi K., Harrison N., Lumberras-Martinez M., Antanaviciute L., Simpson D.W., Harrison R.J. Identification of powdery mildew resistance QTL in strawberry (*Fragaria × ananassa*). *Theor. Appl. Genet.* 2018;131(9):1995-2007. DOI 10.1007/s00122-018-3128-0
- Davik J., Honne B.I. Genetic variance and breeding values for resistance to a wind-borne disease [*Sphaerotheca macularis* (Wall. ex Fr.)] in strawberry (*Fragaria × ananassa* Duch.) estimated by exploring mixed and spatial models and pedigree information. *Theor. Appl. Genet.* 2005;111(2):256-264. DOI 10.1007/s00122-005-2019-3
- Feng J., Cheng Y., Zheng C. Expression patterns of octoploid strawberry *TGA* genes reveal a potential role in response to *Podosphaera aphanis* infection. *Plant Biotechnol. Rep.* 2020;14:55-67. DOI 10.1007/s11816-019-00582-9
- Gorgitano M.T., Pirilli M. Life cycle economic and environmental assessment for a greening agriculture. *Qual. – Access Success.* 2016; 17(S1):181-185
- Govorova G.F., Govorov D.N. Strawberry: Past, Present, Future. Moscow: Rosinformagrotekh Publ., 2004 (in Russian)
- Holod N., Semenova L. Susceptibility of strawberry varieties to powdery mildew. *Plodovodstvo i Vinogradarstvo Yuga Rossii = Fruit Growing and Viticulture of South Russia.* 2014;25(01):111-115 (in Russian)
- Je H.J., Ahn J.W., Yoon H.S., Kim M.K., Ryu J.S., Hong K.P., Lee S.D., Park Y.H. Development of cleaved amplified polymorphic sequence (CAPS) marker for selecting powdery mildew-resistance line in strawberry (*Fragaria × ananassa* Duchesne). *Korean J. Hortic. Sci. Technol.* 2015;33(5):722-729. DOI 10.7235/hort.2015.14133
- Kennedy C., Hasing T.N., Peres N.A., Whitaker V.M. Evaluation of strawberry species and cultivars for powdery mildew resistance in open-field and high tunnel production systems. *HortScience.* 2013; 48(9):1125-1129. DOI 10.21273/HORTSCI.48.9.1125
- Kennedy C., Osorio L.F., Peres N.A., Whitaker V.M. Additive genetic effects for resistance to foliar powdery mildew in strawberry revealed through divergent selection. *J. Am. Soc. Hortic. Sci.* 2014; 139(3):310-316. DOI 10.21273/JASHS.139.3.310
- Koishihara H., Enoki H., Muramatsu M., Nishimura S., Susumu Y.U.I., Honjo M. Marker associated with powdery mildew resistance in plant of genus *Fragaria* and use thereof. U.S. Patent No. 10,724,093. Washington, DC: U.S. Patent and Trademark Office. 2020
- Lifshitz C., David N., Shalit N., Slotzky S., Tanami Z., Elad Y., Dai N. Inheritance of powdery mildew resistance in strawberry lines from the Israeli germplasm collection. In: Proceedings of the NASS/NASGA conference (February 9-12, 2007, Ventura, California). Los Angeles, 2007;74-76
- Liu J., Duan K., Zhang Q., Ye Z., Gao Q. Genetic mapping and preliminary analysis of SSR marker for powdery mildew resistance in strawberry. *Acta Agriculturae Jiangxi.* 2012;24(11):49-52
- Luk'yanchuk I.V., Lyzhin A.S., Kozlova I.I. Analysis of strawberry genetic collection (*Fragaria L.*) for *Rca2* and *Rpfl* genes with molecular markers. *Vavilovskii Zhurnal Genetiki i Seleksii = Vavilov Journal of Genetics and Breeding.* 2018;22(7):795-799. DOI 10.18699/VJ18.423 (in Russian)
- Lyzhin A.S., Luk'yanchuk I.V. Analysis of polymorphism of strawberry genotypes (*Fragaria L.*) according to the strawberry red root spot resistance gene *Rpfl* for identification of strawberry forms promising for breeding and horticulture. *Vestsi Natsyyanal'nay Akademii Navuk Belarusi. Seryya Agrarnykh Navuk = Proceedings of the National Academy of Sciences of Belarus. Agrarian Series.* 2020; 58(3):311-320. DOI 10.29235/1817-7204-2020-58-3-311-320 (in Russian)
- Lyzhin A.S., Luk'yanchuk I.V., Zhanova E.V. Polymorphism of the *Rca2* anthracnose resistance gene in strawberry cultivars (*Fragaria × ananassa*). *Trudy po Prikladnoy Botanike, Genetike i Seleksii = Proceedings on Applied Botany, Genetics, and Breeding.* 2019; 180(1):73-77. DOI 10.30901/2227-8834-2019-1-73-77 (in Russian)
- Nelson M.D., Gubler W.D., Shaw D.V. Inheritance of powdery mildew resistance in greenhouse-grown versus field-grown California strawberry progenies. *Phytopathology.* 1995;85(4):421-424
- Palmer M.G., Holmes G.J. Fungicide sensitivity in strawberry powdery mildew caused by *Podosphaera aphanis* in California. *Plant Dis.* 2021;105(9):2601-2605. DOI 10.1094/PDIS-12-20-2604-RE
- Sargent D.J., Buti M., Šurbanovski N., Brurberg M.B., Alsheikh M., Kent M.P., Davik J. Identification of QTLs for powdery mildew (*Podosphaera aphanis*; syn. *Sphaerotheca macularis* f. sp. *fragariae*) susceptibility in cultivated strawberry (*Fragaria × ananassa*). *PLoS One.* 2019;14(9):e0222829. DOI 10.1371/journal.pone.0222829
- Sombardier A., Dufour M.-C., Blancard D., Corio-Costet M.-F. Sensitivity of *Podosphaera aphanis* isolates to DMI fungicides: distribution and reduced cross-sensitivity. *Pest Manag. Sci.* 2010;66:35-43. DOI 10.1002/ps.1827
- Stolnikova N.P., Kolesnikova A.V. Variety resistance of strawberry to powdery mildew in conditions of the south of Western Siberia. *Sadovodstvo i Vinogradarstvo = Horticulture and Viticulture.* 2017;5: 49-51. DOI 10.18454/VSTISP.2017.5.7593 (in Russian)
- Sylla J., Alsanian B.W., Krüger E., Becker D., Wohanka W. *In vitro* compatibility of microbial agents for simultaneous application to control strawberry powdery mildew (*Podosphaera aphanis*). *Crop Prot.* 2013;51:40-47. DOI 10.1016/j.cropro.2013.04.011
- Tapia R.R., Barbey C.R., Chandra S., Folta K.M., Whitaker V.M., Lee S. Evolution of the MLO gene families in octoploid strawberry (*Fragaria × ananassa*) and progenitor diploid species identified potential genes for strawberry powdery mildew resistance. *Hortic. Res.* 2021;8(1):153. DOI 10.1038/s41438-021-00587-y
- Whitaker V.M., Osorio L.F., Hasing T., Gezan S. Estimation of genetic parameters for 12 fruit and vegetative traits in the University of Florida strawberry breeding population. *J. Am. Soc. Hortic. Sci.* 2012;137(5):316-324. DOI 10.21273/JASHS.137.5.316
- Zhuchenko A.A. Biologization and ecologization of intensification processes in agriculture. *Vestnik OrelGAU = Bulletin of Orel State Agrarian University.* 2009;3:8-12 (in Russian)
- Zubov A.A. Genetic Features and Breeding of Strawberry. Michurinsk, 1990 (in Russian)
- Zubov A.A. Theoretical Foundations of Strawberry Breeding. Michurinsk: Michurin All-Russia Institute for the Genetics and Breeding of Fruit Plants, 2004 (in Russian)

Funding. This work was supported by the Ministry of Science and Higher Education of the Russian Federation under Agreement No. 075-15-2021-1050 of Sept. 28, 2021 within the framework of the project entitled "National network collection of plant genetic resources for effective scientific and technical development of the Russian Federation in the sphere of genetic technologies".

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

Received March 6, 2023. Revised October 12, 2023. Accepted October 12, 2023.

DOI 10.18699/vjgb-24-21

In silico search for and analysis of *R* gene variation in primitive cultivated potato species

A.A. Gurina , M.S. Gancheva ², N.V. Alpatieva ¹, E.V. Rogozina ¹

¹ Federal Research Center the N.I. Vavilov All-Russian Institute of Plant Genetic Resources (VIR), St. Petersburg, Russia

² St. Petersburg State University, St. Petersburg, Russia

 a.gurina@vir.nw.ru

Abstract. Pathogen recognition receptors encoded by *R* genes play a key role in plant protection. Nowadays, *R* genes are a basis for breeding many crops, including potato. Many potato *R* genes have been discovered and found suitable for breeding thanks to the studies of a wide variety of wild potato species. The use of primitive cultivated potato species (PCPS) as representatives of the primary gene pool can also be promising in this respect. PCPS are the closest to the early domesticated forms of potato; therefore, their investigation could help understand the evolution of *R* genes. The present study was aimed at identifying and analyzing *R* genes in PCPS listed in the open database of NCBI and Solomics DB. In total, the study involved 27 accessions belonging to three species: *Solanum phureja* Juz. & Bukasov, *S. stenotomum* Juz. & Bukasov and *S. goniocalyx* Juz. & Bukasov. Materials for the analysis were the sequencing data for the said three species from the PRJNA394943 and PRJCA006011 projects. An *in silico* search was carried out for sequences homologous to 26 *R* genes identified in potato species differing in phylogenetic distance from PCPS, namely nightshade (*S. americanum*), North- (*S. bulbocastanum*, *S. demissum*) and South-American (*S. venturii*, *S. berthaultii*) wild potato species, as well as the cultivated potato species *S. tuberosum* and *S. andigenum*. Homologs of all investigated protein-coding sequences were discovered in PCPS with a relatively high degree of similarity (85–100 %). Homologs of the *Rpi-R3b*, *Rpi-amr3* and *Rpi-ber1* genes have been identified in PCPS for the first time. An analysis of polymorphism of nucleotide and amino acid sequences has been carried out for 15 *R* genes. The differences in frequencies of substitutions in PCPS have been demonstrated by analysis of *R* genes, the reference sequences of which have been identified in different species. For all the studied NBS-LRR genes, the proportion of substituted amino acids in the LRR domain exceeds this figure for the NBS domain. The potential prospects of using PCPS as sources of resistance to *Verticillium* wilt have been shown.

Key words: *R* genes; NBS-LRR; polymorphism; *Solanum phureja*; *S. stenotomum*.

For citation: Gurina A.A., Gancheva M.S., Alpatieva N.V., Rogozina E.V. *In silico* search for and analysis of *R* gene variation in primitive cultivated potato species. *Vavilovskii Zhurnal Genetiki i Seleksii* = *Vavilov Journal of Genetics and Breeding*. 2024;28(2):175-184. DOI 10.18699/vjgb-24-21

In silico поиск *R*-генов у примитивных культурных видов картофеля

A.A. Гурина , M.C. Ганчева ², Н.В. Алпатьева ¹, Е.В. Рогозина ¹

¹ Федеральный исследовательский центр Всероссийский институт генетических ресурсов растений им. Н.И. Вавилова (ВИР), Санкт-Петербург, Россия

² Санкт-Петербургский государственный университет, Санкт-Петербург, Россия

 a.gurina@vir.nw.ru

Аннотация. Ключевую роль в защите растений от патогенов играют рецепторы, кодируемые *R*-генами. Они являются генетической основой для селекции многих сельскохозяйственных культур, в том числе картофеля. Множество генов устойчивости у картофеля стало известно и было вовлечено в селекцию благодаря изучению широкого разнообразия диких сородичей картофеля. Использование примитивных культурных видов (ПКВ), относящихся к первичному генофонду картофеля, также перспективно. Как наиболее близкие к ранним доместцированным формам картофеля, ПКВ представляют особый интерес для исследования эволюции генов устойчивости. Целью настоящего исследования стали поиск и анализ *R*-генов у ПКВ картофеля, геномы которых с различным качеством сборки представлены в базе данных NCBI. Исследовано 27 образцов, относящихся к трем видам: *Solanum phureja* Juz. & Bukasov, *S. stenotomum* Juz. & Bukasov и *S. goniocalyx* Juz. & Bukasov. Проведен *in silico* поиск последовательностей, гомологичных 26 *R*-генам, идентифицированных у различных по филогенетической отдаленности от ПКВ картофеля видов: паслёна (*S. americanum* Mill.), североамериканских (*S. bulbocastanum* Dunal., *S. demissum* Lindl.) и южноамериканских (*S. venturii* Hawkes & Hjert., *S. berthaultii* Hawkes) диких видов, а также видов культурного картофеля (*S. tuberosum* L., *S. andigenum* Juz. & Bukasov). Гомологи кодирующих последовательностей всех исследованных генов обнаружены у ПКВ картофеля с относительно высокой

степенью сходства (85–100 %). Впервые у примитивных культурных видов картофеля найдены гомологи генов *R3b*, *Rpi-amr3* и *Rpi-ber1*. Для 15 *R*-генов проведен анализ полиморфизма нуклеотидных и аминокислотных последовательностей. Приведены отличия в частоте замен у ПКВ картофеля при анализе *R*-генов, референсные последовательности которых идентифицированы у разных видов. Для всех изученных NBS-LRR генов доля замещенных аминокислот в LRR-домене превосходит этот показатель для NBS-домена. Показана потенциальная перспективность использования ПКВ картофеля в качестве источников устойчивости к вертициллезному увяданию. Ключевые слова: *R*-гены; NBS-LRR; полиморфизм; *Solanum phureja*; *S. stenotomum*.

Introduction

The main factor in plant evolution is the adaptation to unfavorable external conditions, including microorganisms and pests (Fang et al., 2022). Recently, significant progress has been made in understanding the molecular mechanisms of plant-pathogen interaction. It has been established that plants have a multi-level system of protection against pests, including the stages of recognition, signaling and initiation of a protective response (Zhang et al., 2019). Receptors localized on the surface or inside the cell have been shown to play a leading role in the activation of immunity, and *R* genes (resistance genes) encoding receptors are the genetic basis for breeding many crops for disease resistance (Deng et al., 2020).

In general, plants have two immune systems: PAMP, the pathogen-associated molecular patterns immunity, which includes the recognition of elicitors (conservative non-race-specific signals, such as polysaccharides, chitin, etc.), and ETI, the effector-triggered immunity, which includes the recognition of race-specific effectors. It is with the recognition of effectors that the action of *R* genes is associated: they directly or indirectly (mediated by other proteins) interact with effectors and trigger the plant immune response, which often includes the programmed cell death that blocks the further spread of the pathogen (Kourelis et al., 2018).

Potato is the most important non-grain crop. In terms of production volume, it ranks fourth among all agricultural crops; according to the FAO data, global potato production in 2020 was over 350 million tons (FAO, 2020). Despite its high adaptive potential, potato is affected by a variety of diseases and pests, 27 of which cause economically significant damage worldwide (Bradshaw, 2021). According to the FAO, potato diseases cause annual losses of about 11.6 % of the gross yield (FAO, 2010). Cultivation of resistant potato cultivars is necessary for stable agricultural production, optimized use of means of chemical control, and for obtaining high-quality products. Introgression of resistance genes from wild and cultivated potato relatives (species of the section *Petota* Dumort. of the genus *Solanum* L.) allows the creation of resistant varieties and breeding lines. The search for genes for resistance to pathogens and pests in representatives of different groups of the potato gene pool is a topical trend in research conducted by research centers in Europe, the USA, India and China (Bradshaw, 2021).

The breeding value of tuber-bearing *Solanum* species depends on their compatibility with cultivated potatoes and the nature of inheritance of the target trait (Rogozina, Khavkin, 2017). Primitive cultivated species, including *S. phureja* Juz. & Bukasov, *S. stenotomum* Juz. & Bukasov, *S. goniocalyx* Juz. & Bukasov (according to J. Hawkes (1990), *S. stenotomum* subsp. *goniocalyx* (Juz. & Bukasov) Hawkes), belong to the primary gene pool, representatives of which easily cross with

potato cultivars (Bradeen, Kole, 2011). In this regard, the use of primitive cultivated species as source material for breeding is of particular interest.

In 2011, the first article was published describing the genome sequence of the artificially created DM 1-3 516 R44 (DM1-3) doubled monoploid from the group Phureja (Potato Genome Sequencing Consortium (PGSC), 2011). The published sequence represented 86 % of the genome of an artificially created homozygous clone, and was obtained by integrating two assemblies of genomic sequences, i.e. the diploid heterozygous clone RH89-039-16 and the clone DM 1-3 516 R44 (CIP 801092) (The Potato Genome Sequencing Initiative, https://www.hutton.ac.uk/sites/default/files/documents/posters/sharma/Sharma_Potato_Sequencing_Initiative.pdf, accessed May 2, 2023).

By 2022, improvements in sequencing technologies made it possible to create assemblies of genomic sequences of representatives of all groups of the potato gene pool, i.e. wild and cultivated species, and tetraploid potato cultivars (Usadel, 2022). Genetic information on potato and related *Solanum* spp. presented in the database of NCBI, Spud and Solomics DBs contributes to a better understanding of their genetic differences, the process of species evolution, and helps in the search for genes that determine the traits of importance for breeding.

The objective of the present work was the search for and structural analysis of *R* genes using whole-genome sequencing data, including short-read data (sequence raw archive (SRA)), and genomes partially assembled (to the contig level) for PCPS.

Materials and methods

Material. The search for resistance genes used the potato reference genome sequence DM1-3 v4.3 (PGSC, 2011), short-read data resulting from the whole-genome sequencing in the PRJNA394943 Project (Li et al., 2018), as well as genome assemblies from the PRJCA006011 Project (Tang et al., 2022) for such PCPS as *Solanum phureja*, *S. stenotomum*, and *S. goniocalyx* (Table 1).

Twenty-six *R* genes were taken as objects for analysis (the complete list is presented in Supplementary Material 2)¹, the reference sequences of which are contained in the database of NCBI. These sequences were isolated from the genetic material of wild and cultivated potato species, as well as American black nightshade *S. americanum*.

***In silico* search for and analysis of *R* genes in genomes of PCPS.** At the first stage of *in silico* analysis, a search for *R* gene homologs was carried out in the whole-genome sequencing data and assemblies of genomes of PCPS.

¹ Supplementary Materials 1 and 2 are available at: https://vavilov.elpub.ru/jour/manager/files/Suppl_Gurina_Engl_28_2.pdf

Table 1. Whole-genome sequences and SRA data included in the analysis

Species	Designation	Number of samples (SRA data)	Genome assemblies
<i>S. phureja</i>	phu	9	DM1-3; PG6169; PG6225
<i>S. stenotomum</i>	stn	8	PG6029; PG6055
<i>S. goniocalyx</i>	gon	4	PG6148

Note. Genome assembly identification numbers are given according to the source (Tang et al., 2022). A full description of the material is presented in Supplementary Material 1.

Search for and analysis of R gene homologs in SRA data. After assessing the quality of sequencing (FASTQC v0.11.9) (Wingett, Andrews, 2018), low-quality reads were filtered (Trimmomatic 0.39). The process included the removal of PCR duplicates, as well as reads containing more than 20 % of nucleotides with Phred quality <5, or more than 10 % of unidentified nucleotides (Bolger et al., 2014). Further on, the reads of each sample were aligned to the reference *R* genes using the local multiple alignment algorithm of the bowtie2 v.2.3.5.1 program (Langmead, Salzberg, 2012). The results of alignment were processed using the samtools 1.10 and bedtools 2.30.0 programs; the same programs were used to assess the sufficiency and homogeneity of gene coverage (Quinlan, Hall, 2010). The search for variants was carried out using VarScan 2.4.4 with a minimum coverage of 5 (Koboldt et al., 2009). The results were processed using R 4.2.2 and Python 3.8.2.

Search for and analysis of R genes in genome assemblies for PCPS. The presence of gene sequences in several genome assemblies for PCPS from the PRJCA006011 Project (Tang et al., 2022) was checked using the local blast 2.5.0+ algorithm (Ladunga, 2017). Selection of sequences was based on their similarity to the reference ones (at least 80 %) and at least 80 % coverage of the latter. These sequences could represent both the complete sequences of *R* genes and their individual parts, i.e. UTRs, exons, and introns. Using ClustalW local alignment, sequences with a fully represented coding region of the gene were selected.

At the second stage, resistance genes identified in genomes of PCPS were analyzed for the presence of polymorphism in nucleotide and amino acid sequences. ClustalW local alignment in Mega X software was used to evaluate changes in amino acid sequences (Kumar et al., 2018). InterPro (Paysan-Lafosse et al., 2023) was used to calculate domain organization and key positions in the amino acid sequences of *R* genes.

Results

Homologs of all 26 original reference *R* genes were identified and their significant similarity (over 80 %) in the coding regions was revealed in the whole-genome sequencing data and assemblies of genomes of *S. phureja*, *S. stenotomum*, and *S. goniocalyx*. The non-coding sequences (UTRs and/or introns) in most putative homologs are very different from those in the reference genes, and the degree of their similarity, as a rule, does not exceed 60 %.

The quality and uniformity of SRA data coverage for each of the 26 *R* genes was assessed, and a search for complete coding sequences in genome assemblies for PCPS was performed. Further analysis excluded genes, the exact and complete

protein-coding sequences of which were unknown, as well as genes, the initial search for which demonstrated discrepancies between the data obtained from the study of genome assemblies and SRA data (for instance, the *Rpi-mch1* gene had extremely low coverage when SRA data were analyzed, although it was present in the assemblies). Taking into account the published data on the high degree of similarity of genes in clusters on chromosomes IV, VIII and IX, one reference gene was selected from each cluster for analyzing a group of homologs: *Rpi-R2-like* for the *Rpi-R2*, *Rpi-R2-like*, *Rpi-abpt*, and *Rpi-blb3* group of homologs; *Rpi-sto1* for the *Rpi-sto1*, *Rpi-blb1*, and *Rpi-bt1* group; and *Rpi-vnt1.3* among different allelic variants of the *Rpi-vnt1* gene.

Based on the results of the first stage of analysis, 15 *R* genes were selected. A number of these genes provide potato resistance to late blight; these are *Rpi-R1* (Ballvora et al., 2002), *Rpi-R2-like* (Lokossou et al., 2009), *Rpi-R3a* (Huang S. et al., 2005), *Rpi-R3b* (Li G. et al., 2011), *Rpi-sto1* (Vleeshouwers et al., 2008), *Rpi-blb2* (van der Vossen et al., 2005), *Rpi-vnt1.3* (Foster et al., 2009), *Rpi-ber1* (Monino-Lopez et al., 2021), *Rpi-R8* (Vossen et al., 2016), and *Rpi-amr3* in nightshade (Witek et al., 2021). Also, there was the *Rx* gene of resistance to potato virus X (Bendahmane et al., 1999), *Gpa2* – to pathotype Pa2 of the pale cyst nematode (van der Vossen et al., 2000), *Ve1*, *Ve2* – to verticillium wilt (Song et al., 2017), and *Tm2-ToMV* – to tomato and tobacco mosaic viruses (Wu X. et al., unpublished). All of these genes are representatives of two families: the *Ve1* and *Ve2* genes belong to the RLP/RKL (receptor-like proteins/receptor-like kinases) family, the remaining 13 genes are from the CC-NBS-LRR (Coiled-Coil Nucleotide Binding Site Leucine Rich Repeats) family.

Genome assemblies for PCPS were found to contain different numbers of copies of the analyzed *R* genes (Table 2). The *Rpi-R3a* gene is represented by a single sequence in most genomes. The *Rpi-amr3*, *Rpi-R1*, *Rpi-R3b*, *Rpi-blb2*, *Rpi-ber1*, *Ve1*, *Ve2*, *Rx*, *Gpa2*, and *Tm2-ToMV* genes have from one to 10 copies. More than a dozen copies were found for the *Rpi-sto1*, *Rpi-vnt1.3*, and *Rpi-R8* genes. The number of copies of the *Rpi-R2-like* gene in different assemblies can reach 45. No noticeable differences in the number of copies of each gene were observed between genome assemblies; the pairwise correlation coefficient between the number of copies in different assemblies exceeds 80 % (see Table 2).

Unlike other assemblies in the reference genome DM1-3, several genes (*Rpi-R1*, *Rpi-R3b*, and *Rpi-blb2*) do not have homologs capable of producing similar proteins (all homologous sequences contain premature stop codons) (see Table 2). We attribute this to the synthetic origin of the doubled monoplloid *S. phureja*.

Table 2. Number of gene copies in assemblies for PCPS, indicating the number of copies that do not contain premature stop codons (in parentheses)

Gene	DM1-3	PG6029	PG6225	PG6169	PG6148	PG6055
<i>Rpi-amr3</i>	6 (1)	11 (3)	7 (6)	6 (2)	9 (6)	8 (4)
<i>Rpi-R1</i>	3 (0)	8 (6)	7 (3)	4 (3)	7 (5)	9 (5)
<i>Rpi-R2-like</i>	26 (7)	34 (22)	35 (21)	24 (17)	45 (26)	40 (23)
<i>Rpi-R3a</i>	7 (1)	1 (1)	1 (1)	1 (1)	2 (2)	1 (1)
<i>Rpi-R3b</i>	2 (0)	5 (1)	3 (2)	5 (2)	7 (6)	4 (2)
<i>Rpi-sto1</i>	5 (3)	16 (12)	13 (8)	7 (5)	20 (13)	8 (6)
<i>Rpi-blb2</i>	1 (0)	3 (1)	4 (4)	2 (2)	4 (3)	2 (1)
<i>Rpi-ber1</i>	3 (1)	6 (1)	9 (2)	4 (1)	5 (1)	3 (1)
<i>Rpi-vnt1.3</i>	6 (5)	6 (6)	11 (6)	10 (9)	17 (12)	13 (7)
<i>Ve1</i>	3 (2)	4 (2)	6 (4)	7 (6)	5 (3)	5 (3)
<i>Ve2</i>	3 (1)	3 (2)	3 (1)	4 (1)	4 (1)	4 (1)
<i>Rpi-R8</i>	10 (1)	18 (12)	20 (9)	21 (15)	12 (7)	15 (9)
<i>Rx</i>	2 (1)	4 (3)	6 (4)	3 (1)	4 (4)	2 (2)
<i>Gpa2</i>	2 (2)	3 (2)	7 (4)	2 (2)	6 (5)	2 (2)
<i>Tm2-ToMV</i>	4 (3)	6 (5)	11 (8)	7 (7)	10 (7)	9 (6)

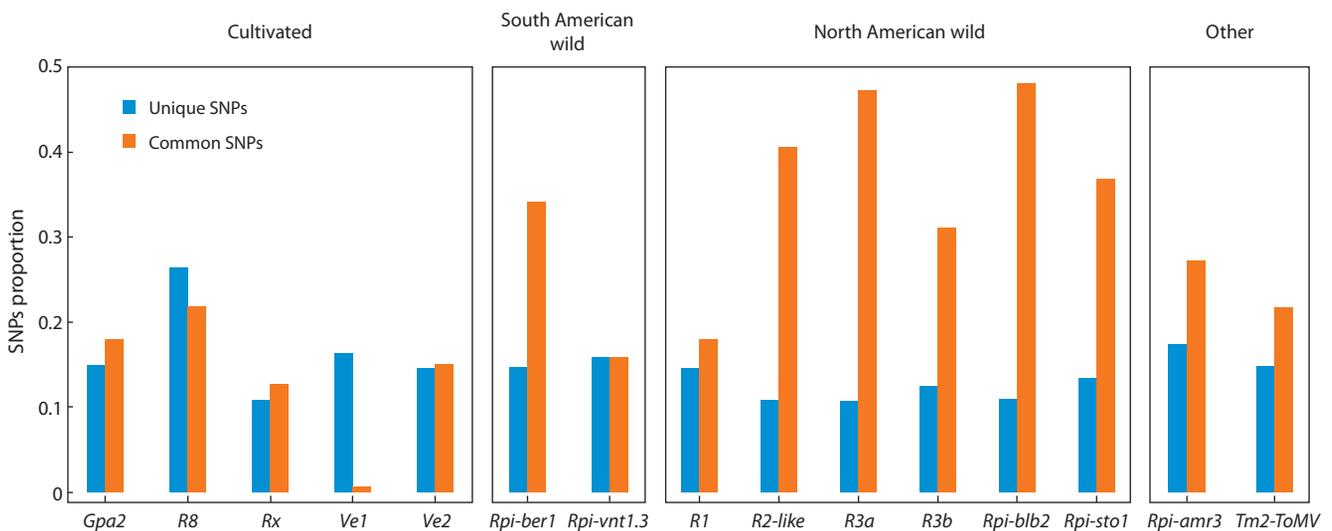


Fig. 1. Occurrence of SNPs in *R* gene homologs among samples of PCPS (according to SRA data).

For the 15 selected reference *R* genes, nucleotide and amino acid sequences were subsequently analyzed for the presence of polymorphism in PCPS; they were found to contain more than two thousand polymorphism sites in the coding sequences of resistance genes, most of which are single nucleotide polymorphisms (SNPs). Deletions or insertions that affect the predicted amino acid sequence were found in at least one copy of all genes. Nevertheless, variants capable of producing a protein similar to the reference one were found for all genes in the genomes of individual species.

The distribution of SNP occurrence corresponds to the concept of PCPS as a group of closely related species. Most commonly distributed are the SNPs common for all samples, compared to the unique SNPs found in only one sample. Many common SNPs (572, which is more than 25 % of all detected) distinguished the studied samples from the reference *R* gene sequences (Fig. 1). Almost 15 % of the detected SNPs were unique sites.

The distribution of SNPs occurrence in samples of PCPS depends on the origin of the source species of the reference

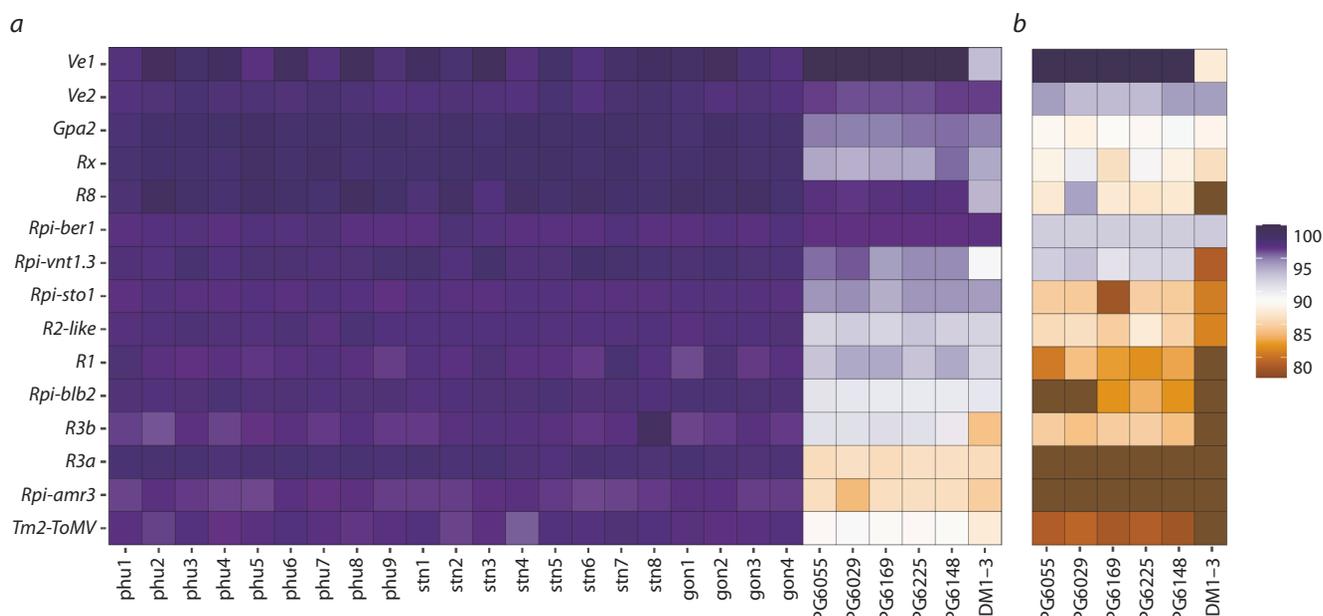


Fig. 2. Heat map of similarity of cds nucleotide sequences (a) and amino acid sequences (b) of *R* gene homologs in PCPS with the reference sequences of *R* genes.

The color scale represents the level of similarity with the reference gene.

R gene. We found significant differences (Kruskal–Wallis test value of 7.4044, p -value = 0.02467) between SNPs common for all studied samples of PCPS in genes from North American wild species and from cultivated potato species (see Fig. 1). We noted similar indicators in the analyzed genome assemblies. The number of variable sites common for PCPS was minimal for the reference genes *Rpi-R8*, *Rx*, *Ve1*, and *Ve2*, the sources of which are samples of cultivated potato. Meanwhile, more than 40 % of the SNPs detected in the genes *Rpi-R3a*, *Rpi-blb2*, and *Rpi-R2-like*, were the same for all samples of PCPS, the sources of which are samples of North American wild potato species phylogenetically distant from PCPS (see Fig. 1). The differences in the frequencies of unique SNPs are not so pronounced; the largest proportion is characteristic of the *Rpi-R8*, *Ve1*, and *Gpa2* genes, the sources of which are cultivated potato species.

Of the genes originating from North American wild potato species, only the *Rpi-R1* gene breaks the trend of a high proportion of common SNPs characteristic of other genes (in this gene, the proportion of SNPs common for PCPS is below 20 %) (see Fig. 1). This gene was introgressed into cultivated potato from a distant wild species *S. demissum*. However, when examining different copies of this gene present in genomes of PCPS, we found several indels in each of these copies, including deletions up to 15 nucleotides long. We observed a similar decrease in the proportion of common SNPs and the presence of indels in each copy when studying *Rpi-amr3*, another late blight resistance gene found in *S. americanum*, one of the nightshade species, also very distant from cultivated potato.

An analysis of genome assemblies and SRA data revealed differences in the degree of similarity between sequences found in PCPS and reference sequences of *R* genes (Fig. 2, a).

The differences are most pronounced in multicopy genes. The reason is that the analysis of SRA data actually takes into account the consensus sequence for all copies, while an increase in the number of copies reduces the likelihood of taking into account the SNPs, since the variant that occurs in the majority of copies is taken into account in the consensus sequence. The strongest differences between the estimates of the assembled genomes and raw reads are observed when analyzing the *Rpi-R3a* and *Rpi-amr3* genes (see Fig. 2, a). Most likely, they result from the chimeric alignment of reads from other genes to the reference ones. Further analysis of these sequences did not take into account the results of SRA data processing.

Samples of PCPS showed differences in terms of the degree of similarity of the found homologs (without taking into account the DM1-3 assembly) with the reference sequences of *R* genes. The median level of differences is 0.5 %. The copies of the *Rpi-ber1* and *Ve1* genes that are most similar to the reference gene differ in single SNPs (0.01 %) between different genomes of PCPS. The maximum level of differences between the nucleotide sequences of *R* gene homologs in PCPS was observed for the *Rpi-amr3* gene and amounted to 2.3 %. A level of polymorphism close to this value was noted when analyzing *Rx*, *R1*, *Rpi-vnt1.3*, and *Rpi-sto1* homologs. It was noted for the *Rpi-R8*, *Rpi-blb2*, and *Ve2* genes that one or several copies are quite stable and have a low level of polymorphism, while other copies differ significantly from each other. This corresponds to the notions about the evolutionary process of *R* genes, for which the copy number can reduce the influence of selection.

In the assemblies of PCPS (excluding the DM1-3 assembly), homologs of all *R* genes were found to contain copies potentially capable of producing amino acid sequences (no

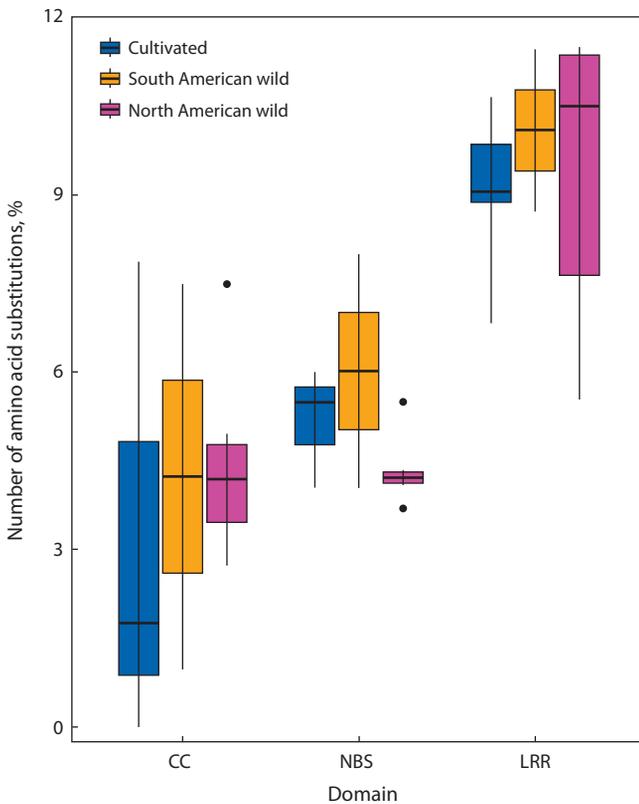


Fig. 3. Comparison of the proportion of amino acid substitutions in different domains of *R* genes.

Significant differences between LRR and other domains were assessed according to the Kruskal–Wallis test value of 17.343, p -value = 0.0001714.

reading frame shifts or premature stop codons were found in them). The degree of similarity to the reference amino acid sequence varies between 72–100 % (see Fig. 2, *b*).

The groups of genes, the origin of which is directly related to cultivated potatoes, are clearly divided into two groups based on the similarity of amino acid sequences to the reference ones: the genes of the RLP-RLK family, *Ve1* and *Ve2*, have the highest similarity to the reference gene: 100 and 92–94 %, respectively; the similarity of genes of the CC-NBS-LRR family (*Gpa2*, *Rx*, *Rpi-R8*) is significantly lower 85–90 % (see Fig. 2, *b*). The genes from the CC-NBS-LRR family, the origin of which is associated with South American wild species, have a higher amino acid sequence similarity to the reference ones than the *Rpi-R2-like*, *Rpi-sto1*, *Rpi-R1* and *Rpi-R3b* genes from North American wild species: about 95 and 80–85 %, respectively. Changes in the nucleotide sequence in the *Rpi-R3a*, *Rpi-amr3* and *Tm2-ToMV* genes lead to low similarity (75–80 %) with the reference amino acid sequence (see Fig. 2, *b*).

The functional analysis of amino acid sequences performed using InterPro showed that the homologs of the NBS-LRR genes in PCPS have no changes in domain organization, even in the sequences, the similarity of which to the reference sequence is below 80 %.

The distribution of amino acid substitutions across individual domains in genes of the CC-NBS-LRR family is shown

in Fig. 3. Only sequences forming complete proteins (free from premature stop codons) were taken into account. The *Ve1* and *Ve2* genes, which, according to InterPro and NCBI data, have a different domain organization compared to CC-NBS-LRR, are not included.

The percentage of substituted amino acids varies significantly depending on the domain (see Fig. 3). The LRR domain in all found *R* gene homologs demonstrated the largest number (on average more than 9 %) of amino acid substitutions compared to the reference. In the *Rpi-vnt1.3*, *Rpi-sto1*, *Rpi-R1*, and *Rpi-R3b* genes, the proportion of amino acids substituted in the LRR domain exceeds 10 %. The proportion of substitutions in the CC domain varies significantly between genes. This domain was found to contain no amino acid substitutions in homologs of the *R8*, *Rpi-vnt1.3*, and *Rx* genes, while there were 7 % amino acid substitutions in *Rpi-blb2* and *Gpa2*, and 11 % in the *Tm2-ToMV* gene. The NBS domain contains about 5 % substituted amino acids on the average.

The distribution of substituted amino acids in different domains among PCPS is also uneven. In the CC domain of almost all genes, most substitutions occur only in individual samples or small groups. Substituted amino acids common for all samples of PCPS are represented by 0–3 sites. An exception is the *Gpa2* gene, in which five substitutions in the CC domain are common, and several of them are located either directly in the RanGAP2 interaction site (according to NCBI data), or in close proximity to it (neighboring amino acids). In the NBS domain, substitutions common for PCPS are prevalent, and *Rx* was the only gene where they are absent. In the LRR domain, amino acid substitutes that are common or characteristic only of some samples are distributed relatively evenly.

Discussion

Potato is a heterozygous tetraploid, the genetic analysis of which is a challenge. Genomic studies of potatoes and related tuber-forming species, the search for and annotation of genes that determine agronomically important characteristics make a significant contribution to genetics proper of the crop and the improvement of breeding technologies (Bradshaw, 2021). PCPS belong to a systematic group of interest from the evolutionary point of view and as a source of important potato traits of value for breeding, such as disease resistance.

We conducted a bioinformatics analysis of whole-genome sequencing data from 27 samples of *S. phureja*, *S. stenotomum* and *S. goniocalyx* (21 from SRA data and 6 genome assemblies) for 26 genes of resistance to pathogens and pests. All samples contained sequences homologous to the coding regions of all analyzed *R* genes, and their significant similarity (> 85 %) with the corresponding regions of reference genes was revealed. The noncoding sequences (UTRs and/or introns) in most putative homologs differed greatly from those in the reference genes, and the degree of their similarity, as a rule, did not exceed 60 %.

Many of the reference sequences were obtained from species distant from PCPS. For instance, the *Rpi-amr3* gene sequence was isolated from the nightshade *S. americanum*, and *Rpi-blb2*, from *S. bulbocastanum*, a representative of the North American wild potato species. The latter does not directly cross with cultivated species, and the discovery of

R gene homologs with a fairly high level of similarity in PCPS is of interest from the point of view of peculiarities of these genes' evolution. For the first time, homologs of the *Rpi-R3b*, *Rpi-amr3*, and *Rpi-ber1* genes were discovered in *S. phureja* and *S. stenotomum*.

Cultivated potatoes grown in South America were found to contain homologs of genes providing reliable protection against late blight, which had been found in the South American wild species *S. venturii* and the Mexican species *S. bulbocastanum* and *S. stoloniferum*, as well as homologs of race-specific late blight resistance genes from the North American species *S. demissum*. Previously, molecular genetic screening of accessions from the VIR potato collection revealed the presence of SCAR markers of the *Rpi* genes from wild potato species of the North and South American series, and in samples of PCPS (Muratova et al., 2020; Gurina et al., 2022; Rogozina et al., 2023). The similarity of wild and cultivated potatoes in genes of resistance to the highly specialized pathogen *Phytophthora infestans* (Mont.) de Bary confirms the conclusion about the inheritance of part of the genetic material from wild species by diploid cultivated potatoes (Hardigan et al., 2017).

The discovery of not only genes protecting potatoes from late blight, viruses, and pale nematodes, but also homologs of *R* genes from tomato and nightshade in representatives of cultivated potatoes, is in good agreement with the data on a significantly larger repertoire of disease resistance genes in potatoes compared to the closely related Solanaceous crops (Tang et al., 2022).

The previously conducted bioinformatics analysis of the potato reference genome DM1-3 (Jupe et al., 2012; Lozano et al., 2012) was performed without taking into account the *R3b*, *Rpi-amr3*, *Ve1*, *Ve2*, and *Rpi-ber1* genes, since their sequences were not yet known. In a later work devoted to the search for NBS-LRR genes in the cultivated potato *S. stenotomum* subsp. *goniocalyx*, the *R8*, *Ve1*, *Ve2*, *Rpi-ber1*, and *Tm2-ToMV* gene sequences were not used (Liu, 2020). For the majority of the *R* gene homologs that we identified, the degree of their similarity to the reference sequences corresponds to the published data (Lozano, 2012; Liu, 2020). However, in the studied samples of *S. goniocalyx*, sufficiently reliable homologous sequences for the *Rx*, *R3b*, and *Gpa2* genes were not found (Liu, 2020). Most likely, this is due to the extremely high degree of polymorphism in the non-coding parts of these genes, which, moreover, occupy rather extensive areas, which prevents their full-length analysis.

The diploid species *S. phureja* and *S. stenotomum* are the closest to the early domesticated forms of tuber-forming species of the genus *Solanum* L. Representatives of these species were found to contain homologs of *R* genes with different evolutionary histories. The ancient origin of the extracellular receptor encoded by the *Ve1* gene is indicated by the wide distribution of functional and nonfunctional *Ve1* homologs in plants of the *Solanacea* family and other phylogenetically distant species (Song et al., 2017). In PCPS, copies of the *Ve1* gene were found in all assemblies, the amino acid sequences of which do not differ from the reference, thus indicating prospects of this group as a potential donor of resistance to Verticillium wilt.

The evolution of *Rpi-blb1* and *Rpi-blb3*, the genes of the Mexican species *S. bulbocastanum*, proceeded at different rates and in different ways, and only sequences identical to the reference gene were found in *S. bulbocastanum* samples for the *Rpi-blb2* gene (Lokossou et al., 2010). Therefore, the discovery of *Rpi-blb2* and *Rpi-sto1* homologs (*Rpi-blb1* ortholog) in cultivated potatoes growing on another continent and non-crossable with *S. bulbocastanum* is of particular interest.

As a rule, *R* genes, especially representatives of the NBS-LRR family, are presented in the genome not singly, but in clusters, which are numerous copies of homologous genes (Prakash, 2020). In genomes of PCPS, most of the studied *R* genes (except for *R3a*, *Rpi-ber1*, and *Ve2*) are also represented by two or more copies. The *R2-like* gene, which belongs to a large cluster of *R* genes located on the fourth chromosome, especially stands out. Its copy number among PCPS was the highest, and the number of identified copies varied from 24 to 40. According to the literature, the presence/absence of a gene and its copy number is a common type of polymorphism among resistance genes, since their organization contributes to unequal recombination, which results in differences in the number of copies of one gene. For some genes, we did observe differences in copy number, but since the assemblies are incomplete, and the method only showed homologs with similarities above the threshold, we cannot say for sure that the revealed differences really occur instead of being a consequence of methodological aspects of the work.

Many authors associate the process of domestication with the loss of diversity in many genes, primarily the economically valuable ones, which include *R* genes (Li Y. et al., 2018; Tang et al., 2022). Y. Li et al. (2018) noted lower variability in the genetic material of cultivated species compared to wild ones. This agrees well with the high proportion of common SNPs in PCPS, as well as with the potato taxonomy developed by D. Spooner, who considers the species of this plant as a group within one species *S. tuberosum* (Spooner et al., 2014).

The *R* genes under study have a complex structure, and the unequal rate of amino acid substitutions in different domains has been previously shown (Prakash et al., 2020). For instance, the NBS domain is more conservative, since it is associated with the activation of the plant protection mechanism inside the cell, while the LRR domain responsible for pathogen recognition is more variable, and the rate of substitutions in it is higher. The uneven distribution of amino acid substitutions across the domains of NBS-LRR proteins discovered in our research corresponds to literature data on the rate of these proteins evolution.

In terms of the prevalence of substituted amino acids among PCPS samples, the *R* gene domains are also not uniform. The most widespread are changes in the NBS domain, which are common for this group, and are likely associated with changes in proteins subsequently activated by *R* genes. On the contrary, the CC domain has the smallest number of amino acid substitutions common for PCPS. Little is understood about the function of this domain; in some proteins, it is known to trigger the mechanism of cell death (Huang J. et al., 2021), but for other genes, its interaction with the pathogen recognition system has been shown (Rairdan et al., 2008). It is quite difficult to guess what causes such differences in PCPS and

what they can lead to, but when studying different copies, we found that indels were located exactly in the CC domain in the majority of clearly non-functional sequences (containing a premature stop codon). This may reflect the evolutionary nature of diversification of paralogs in the genome while a relatively stable variant is maintained.

The LRR domain was found to contain substituted amino acids, both common ones and those characteristic of individual samples or large groups; most likely, these substitutions are part of the mechanism of PCPS ancestral forms adaptation to the races of pathogens common in those times. The mechanism of the LRR domain operation is poorly understood. It has been shown that even a small number of changes in specific amino acids can result in a change in pathogen recognition from virus to nematode (Slootweg et al., 2017). At the same time, large groups of orthologs that recognize the same pathogen are known to be only 85–90 % similar to each other (Park et al., 2005). Apparently, only changes in specific patterns in the sequence lead to changes in the effector, which is recognized by the LRR domain. A detailed study of this phenomenon, however, requires a significantly larger number of well-phenotyped samples, as well as the use of methods that allow the assessment of the interaction of products of different genes.

Conclusion

Primitive cultivated potato species were found to contain sequences homologous to *R* genes that provide protection of potatoes from late blight, viruses, potato cyst nematode, of potatoes and tomatoes from *Verticillium* wilt, and of tomatoes from mosaic viruses. For the first time, a search was carried out and showed the presence of homologs of the *R3b*, *Rpi-amr3*, *Ve1*, *Ve2*, and *Rpi-ber1* genes in PCPS. The similarity of the coding sequences found in PCPS to the reference *R* genes is over 85 % for all genes. It was shown for the first time that homologs of all *R* genes in all the studied genome assemblies have copies, among which all genomes except DM1-3 have at least one sequence that does not contain stop codons or frameshifts.

The studied group of samples showed differences in the nature of *R* gene polymorphism, depending on the source of the reference gene. The changes characterizing the PCPS group as a whole are significantly more represented for homologs of genes from North American wild species, compared to genes from cultivated potato species. It was shown for the first time that one of the copies of the *Ve1* gene in PCPS does not contain amino acid substitutions relative to the reference gene, which indicates the potential resistance of PCPS to *Verticillium dahliae* Kleb.

References

Armstrong M.R., Vossen J., Lim T.Y., Hutten R.C.B., Xu J., Strachan S.M., Harrower B., Champouret N., Gilroy E.M., Hein I. Tracking disease resistance deployment in potato breeding by enrichment sequencing. *Plant Biotechnol. J.* 2019;17(2):540-549. DOI 10.1111/pbi.12997

Ballvora A., Ercolano M.R., Weiß J., Meksem K., Bormann C.A., Oberhagemann P., Salamini F., Gebhardt C. The *R1* gene for potato resistance to late blight (*Phytophthora infestans*) belongs to the leucine zipper/NBS/LRR class of plant resistance genes. *Plant J.* 2002;30(3):361-371. DOI 10.1046/J.1365-313X.2001.01292.X

Bendahmane A., Kanyuka K., Baulcombe D.C. The *Rx* gene from potato controls separate virus resistance and cell death responses. *Plant Cell.* 1999;11(5):781-791. DOI 10.1105/tpc.11.5.781

Bolger A.M., Lohse M., Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014;30(15):2114-2120. DOI 10.1093/bioinformatics/btu170

Bradeen J.M., Kole C. (Eds.). Genetics, Genomics and Breeding of Potato. CRC Press, 2011. DOI 10.1201/b10881

Bradshaw J.E. Potato Breeding: Theory and Practice. Cham (Switzerland): Springer, 2021. DOI 10.1007/978-3-030-64414-7

Deng Y., Ning Y., Yang D.-L., Zhai K., Wang G.-L., He Z. Molecular basis of disease resistance and perspectives on breeding strategies for resistance improvement in crops. *Mol. Plant.* 2020;13(10):1402-1419. DOI 10.1016/j.molp.2020.09.018

Fang Y., Jiang J., Hou X., Guo J., Li X., Zhao D., Xie X. Plant protein-coding gene families: Their origin and evolution. *Front. Plant Sci.* 2022;13:995746. DOI 10.3389/fpls.2022.995746

FAO. The Second Report on the World's Plant Genetic Resources for Food and Agriculture. Rome, 2010

FAO. Land Use Statistics and Indicators. Global, Regional and Country Trends, 2000–2020. No. 48. Rome: FAOSTAT analytical Brief, 2020

Foster S.J., Park T.H., Pel M., Brigneti G., Sliwka J., Jagger L., van der Vossen E., Jones J.D.G. *Rpi-vnt1.1*, a *Tm-2²* homolog from *Solanum venturii*, confers resistance to potato late blight. *Mol. Plant Microbe Interact.* 2009;22(5):589-600. DOI 10.1094/MPMI-22-5-0589

Gurina A.A., Alpatieva N.V., Chalaya N.A., Mironenko N.V., Khiuti A.V., Rogozina E.V. Homologs of late blight resistance genes in representatives of tuber-bearing species of the genus *Solanum* L. *Russ. J. Genet.* 2022;58(12):1473-1484. DOI 10.1134/S1022795422120043

Hardigan M.A., Laimbeer F.P.E., Newton L., Crisovan E., Hamilton J.P., Vaillancourt B., Wiegert-Rininger K., Wood J.C., Douches D.S., Farré E.M., Veilleux R.E., Buell C.R. Genome diversity of tuber-bearing *Solanum* uncovers complex evolutionary history and targets of domestication in the cultivated potato. *Proc. Natl. Acad. Sci. USA.* 2017;114(46):E9999-E10008. DOI 10.1073/pnas.1714380114

Hawkes J.G. The Potato: Evolution, Biodiversity and Genetic Resources. Belhaven Press, 1990

Huang S., van der Vossen E.A.G., Kuang H., Vleeshouwers V.G.A.A., Zhang N., Borm T.J.A., van Eck H.J., Baker B., Jacobsen E., Visser R.G.F. Comparative genomics enabled the isolation of the *R3a* late blight resistance gene in potato. *Plant J.* 2005;42(2):251-261. DOI 10.1111/J.1365-313X.2005.02365.X

Huang J., Wu X., Sun K., Gao Z. Structure and function analysis of a CC-NBS-LRR protein AT1G12290. *Biochem. Biophys. Res. Commun.* 2021;1(534):206-211. DOI 10.1016/j.bbrc.2020.11.111

Jones J., Foster S.J., Chu Z., Park T.H., Pel M.A. Late Blight Resistance Genes and Methods. US Patent WO 2009013468 A3. 2013

Jupe F., Pritchard L., Etherington G.J., MacKenzie K., Cock P.J.A., Wright F., Sharma S.K., Bolser D., Bryan G.J., Jones J.D.G., Hein I. Identification and localisation of the NB-LRR gene family within the potato genome. *BMC Genomics.* 2012;13:75. DOI 10.1186/1471-2164-13-75

Koboldt D.C., Chen K., Wylie T., Larson D.E., McLellan M.D., Mardis E.R., Weinstock G.M., Wilson R.K., Ding L. VarScan: variant detection in massively parallel sequencing of individual and pooled samples. *Bioinformatics.* 2009;25(17):2283-2285. DOI 10.1093/Bioinformatics/BTP373

Kourelis J., van der Hoorn R.A.L. Defended to the nines: 25 years of resistance gene cloning identifies nine mechanisms for R protein function. *Plant Cell.* 2018;30(2):285-299. DOI 10.1105/tpc.17.00579

Kumar S., Stecher G., Li M., Knyaz C., Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 2018;35(6):1547-1549. DOI 10.1093/MOLBEV/MSY096

- Ladunga I. Finding similar nucleotide sequences using Network BLAST searches. *Curr. Protoc. Bioinformatics*. 2017;58(1):3.3.1-3.3.25. DOI 10.1002/CPBI.29
- Langmead B., Salzberg S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods*. 2012;9(4):357-359. DOI 10.1038/nmeth.1923
- Li G., Huang S., Guo X., Li Y., Yang Y., Guo Z., Kuang H., Rietman H., Bergervoet M., Vleeshouwers V.G.G.A., van der Vossen E.A.G., Qu D., Visser R.G.F., Jacobsen E., Vossen J.H. Cloning and characterization of *R3b*; members of the *R3* superfamily of late blight resistance genes show sequence and functional divergence. 2011;24(10):1132-1142. DOI 10.1094/MPMI-11-10-0276
- Li Y., Colleoni C., Zhang J., Liang Q., Hu Y., Ruess H., Simon R., Liu Y., Liu H., Yu G., Schmitt E., Ponitzki C., Liu G., Huang H., Zhan F., Chen L., Huang Y., Spooner D., Huang B. Genomic analyses yield markers for identifying agronomically important genes in potato. *Mol. Plant*. 2018;11(3):473-484. DOI 10.1016/j.molp.2018.01.009
- Liu Z. Diversity of NB-LRR genes in the genome of *Solanum stenotomum* subsp. *goniocalyx*. Thesis. Montreal: McGill University, 2020
- Lokossou A.A., Park T.H., van Arkel G., Arens M., Ruyter-Spira C., Morales J., Whisson S.C., Birch P.R.J., Visser R.G.F., Jacobsen E., van der Vossen E.A.G. Exploiting knowledge of *R/Avr* genes to rapidly clone a new LZ-NBS-LRR family of late blight resistance genes from potato linkage group IV. *Mol. Plant Microbe Interact*. 2009;22(6):630-641. DOI 10.1094/MPMI-22-6-0630
- Lokossou A.A., Rietman H., Wang M.Q., Krenek P., Schoot H., Henken B., Hoekstra R., Vleeshouwers V.G.A.A., van der Vossen E.A.G., Visser R.G.F., Jacobsen E., Vosman B. Diversity, distribution, and evolution of *Solanum bulbocastanum* late blight resistance genes. *Mol. Plant Microbe Interact*. 2010;23(9):1206-1216. DOI 10.1094/MPMI-23-9-1206
- Lozano R., Ponce O., Ramirez M., Mostajo N., Orjeda G. Genome-wide identification and mapping of NBS-encoding resistance genes in *Solanum tuberosum* group phureja. *PLoS One*. 2012;7(4):e34775. DOI 10.1371/journal.pone.0034775
- Monino-Lopez D., Nijenhuis M., Kodde L., Kamoun S., Salehian H., Schentsnyi K., Stam R., Lokossou A., Abd-El-Halim A., Visser R.G.F., Vossen J.H. Allelic variants of the NLR protein Rpi-chc1 differentially recognize members of the *Phytophthora infestans* PexRD12/31 effector superfamily through the leucine-rich repeat domain. *Plant J*. 2021;107(1):182-197. DOI 10.1111/tpj.15284
- Muratova (Fadina) O.A., Beketova M.P., Kuznetsova M.A., Rogozina E.V., Khavkin E.E. South American species *Solanum alandiae* Card. and *S. okadae* Hawkes et Hjerting as potential sources of genes for potato late blight resistance. *Proc. Appl. Bot. Genet. Breed*. 2020;181(1):73-83. DOI 10.30901/2227-8834-2020-1-73-83
- Oosumi T., Rockhold D.R., Maccree M.M., Deahl K.L., McCue K.F., Belknap W.R. Gene *Rpi-bt1* from *Solanum bulbocastanum* confers resistance to late blight in transgenic potatoes. *Am. J. Pot. Res*. 2009;86:456-465. DOI 10.1007/s12230-009-9100-4
- Paal J., Henselewski H., Muth J., Meksem K., Menéndez C.M., Salamini F., Ballvora A., Gebhardt C. Molecular cloning of the potato *Gro1-4* gene conferring resistance to pathotype Ro1 of the root cyst nematode *Globodera rostochiensis*, based on a candidate gene approach. *Plant J*. 2004;38(2):285-97. DOI 10.1111/j.1365-313X.2004.02047.x
- Park T.H., Gros J., Sikkema A., Vleeshouwers V.G., Muskens M., Allefs S., Jacobsen E., Visser R.G., van der Vossen E.A. The late blight resistance locus *Rpi-bib3* from *Solanum bulbocastanum* belongs to a major late blight *R* gene cluster on chromosome 4 of potato. *Mol. Plant Microbe Interact*. 2005;18(7):722-729. DOI 10.1094/MPMI-18-0722
- Paysan-Lafosse T., Blum M., Chuguransky S., Grego T., Pinto B.L., Salazar G.A., Bileschi M.L., ... Sillitoe I., Thanki N., Thomas P.D., Tosatto S.C.E., Wu C.H., Bateman A. InterPro in 2022. *Nucleic Acids Res*. 2023;51(D1):D418-D427. DOI 10.1093/NAR/GKAC993
- Potato Genome Sequencing Consortium; Xu X., Pan S., Cheng S., Zhang B., Mu D., Ni P., Zhang G., Yang S., ... Kloosterman B., van Eck H., Datema E., Hekkert Bt., Goverse A., van Ham R.C., Visser R.G. Genome sequence and analysis of the tuber crop potato. *Nature*. 2011;475(7355):189-195. DOI 10.1038/nature10158
- Prakash C., Trognitz F.C., Venhuizen P., von Haeseler A., Trognitz B. A compendium of genome-wide sequence reads from NBS (nucleotide binding site) domains of resistance genes in the common potato. *Sci. Rep*. 2020;10(1):11392. DOI 10.1038/s41598-020-67848-z
- Quinlan A.R., Hall I.M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*. 2010;26(6):841-842. DOI 10.1093/bioinformatics/btq033
- Rairdan G.J., Collier S.M., Sacco M.A., Baldwin T.T., Boettrich T., Moffett P. The coiled-coil and nucleotide binding domains of the Potato Rx disease resistance protein function in pathogen recognition and signaling. *Plant Cell*. 2008;20(3):739-751. DOI 10.1105/tpc.107.056036
- Rogozina E.V., Khavkin E.E. Interspecific potato hybrids as donors of durable resistance to pathogens. *Vavilovskii Zhurnal Genetiki i Selekcii = Vavilov Journal of Genetics and Breeding*. 2017;21(1):30-41. DOI 10.18699/VJ17.221 (in Russian)
- Rogozina E.V., Gurina A.A., Chalaya N.A., Zoteyeva N.M., Kuznetsova M.A., Beketova M.P., Muratova O.A., Sokolova E.A., Drobyazina P.E., Khavkin E.E. Diversity of late blight resistance genes in the VIR potato collection. *Plants*. 2023;12(2):273. DOI 10.3390/plants12020273
- Sliwka J., Jakuczun H., Chmielarz M., Hara-Skrzypiec A., Tomczyńska I., Kilian A., Zimnoch-Guzowska E. A resistance gene against potato late blight originating from *Solanum × michoacanum* maps to potato chromosome VII. *Theor. Appl. Genet*. 2012;124(2):397-406. DOI 10.1007/s00122-011-1715-4
- Slootweg E., Koropacka K., Roosien J., Dees R., Overmars H., Lankhorst R.K., van Schaik C., Pomp R., Bouwman L., Helder J., Schots A., Bakker J., Smant G., Goverse A. Sequence exchange between homologous NB-LRR genes converts virus resistance into nematode resistance, and vice versa. *Plant Physiol*. 2017;175(1):498-510. DOI 10.1104/pp.17.00485
- Song Y., Zhang Z., Seidl M.F., Majer A., Jakse J., Javornik B., Thoma B.P.H.J. Broad taxonomic characterization of *Verticillium* wilt resistance genes reveals an ancient origin of the tomato Ve1 immune receptor. *Mol. Plant Pathol*. 2017;18(2):195-209. DOI 10.1111/MPP.12390
- Spooner D.M., Ghislain M., Simon R., Jansky S.H., Gavrilenko T. Systematics, diversity, genetics, and evolution of wild and cultivated potatoes. *Bot. Rev*. 2014;80:283-383. DOI 10.1007/s12229-014-9146-y
- 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 Y., Zhou Y., 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
- Usadel B. *Solanaceae* pangenomes are coming of graphical age to bring heritability back. *aBIOTECH*. 2022;3(4):233-236. DOI 10.1007/s42994-022-00087-0
- van der Vossen E.A., van der Voort J.N., Kanyuka K., Bendahmane A., Sandbrink H., Baulcombe D.C., Bakker J., Stiekema W.J., Klein-Lankhorst R.M. Homologues of a single resistance-gene cluster in potato confer resistance to distinct pathogens: a virus and a nematode. *Plant J*. 2000;23(5):567-576. DOI 10.1046/j.1365-313x.2000.00814.x. PMID: 10972883
- van der Vossen E., Sikkema A., Hekkert Bt., Gros J., Stevens P., Muskens M., Wouters D., Pereira A., Stiekema W., Allefs S. An ancient R gene from the wild potato species *Solanum bulbocastanum* confers broad-spectrum resistance to *Phytophthora infestans* in cultivated potato and tomato. *Plant J*. 2003;36(6):867-882. DOI 10.1046/j.1365-313x.2003.01934.x

- van der Vossen E.A.G., Gros J., Sikkema A., Muskens M., Wouters D., Wolters P., Pereira A., Allefs S. The *Rpi-blb2* gene from *Solanum bulbocastanum* is an *Mi-1* gene homolog conferring broad-spectrum late blight resistance in potato. *Plant J.* 2005;44(2):208-222. DOI 10.1111/j.1365-313X.2005.02527.x
- Vleeshouwers V.G.A.A., Rietman H., Krenek P., Champouret N., Young C., Oh S.K., Wang M., Bouwmeester K., Vosman B., Visser R.G.F., Jacobsen E., Govers F., Kamoun S., van der Vossen E.A.G. Effector genomics accelerates discovery and functional profiling of potato disease resistance and phytophthora infestans avirulence genes. *PLoS One.* 2008;3(8):2875. DOI 10.1371/journal.pone.0002875
- Vossen J.H., van Arkel G., Bergervoet M., Jo K.R., Jacobsen E., Visser R.G.F. The *Solanum demissum* *R8* late blight resistance gene is an *Sw-5* homologue that has been deployed worldwide in late blight resistant varieties. *Theor. Appl. Genet.* 2016;129(9):1785-1796. DOI 10.1007/S00122-016-2740-0
- Wingett S.W., Andrews S. FastQ Screen: A tool for multi-genome mapping and quality control. *F1000Res.* 2018;7:1338. DOI 10.12688/f1000research.15931.2
- Witek K., Lin X., Karki H.S., Jupe F., Witek A.I., Steuernagel B., Stam R., van Oosterhout C., Fairhead S., Heal R., Cocker J.M., Bhanvadia S., Barrett W., Wu C.H., Adachi H., Song T., Kamoun S., Vleeshouwers V.G.A.A., Tomlinson L., Wulff B.B.H., Jones J.D.G. A complex resistance locus in *Solanum americanum* recognizes a conserved *Phytophthora* effector. *Nat. Plants.* 2021;7(2):198-208. DOI 10.1038/S41477-021-00854-9
- Zhang R., Zheng F., Wei S., Zhang S., Li G., Cao P., Zhao S. Evolution of disease defense genes and their regulators in plants. *Int. J. Mol. Sci.* 2019;20(2):335. DOI 10.3390/ijms20020335

Funding. The work was supported by the Russian Science Foundation Project No. 22-26-00111 "Genes of potato resistance to late blight in the context of the evolution of cultivated and wild tuber-bearing species of *Solanum* L."

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

Received July 16, 2023. Revised August 29, 2023. Accepted August 30, 2023.

DOI 10.18699/vjgb-24-22

Fertility differences between two wild-type *Drosophila melanogaster* lines correlate with differences in the expression of the *Jheh1* gene, which codes for an enzyme degrading juvenile hormone

O.V. Andreenkova , N.V. Adonyeva, V.M. Efimov , N.E. Gruntenko  

Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
 nataly@bionet.nsc.ru

Abstract. Juvenile hormone plays a “status quo” role in *Drosophila melanogaster* larvae, preventing the untimely metamorphosis, and performs a gonadotropic function in imagoes, ensuring the ovaries’ preparedness for vitellogenesis. The decreased level of juvenile hormone results in reproductive disorders in *D. melanogaster* females including a delay in the oviposition onset and a fertility decrease. Another factor that can affect the insect reproduction is an infection with the maternally inherited symbiotic α -proteobacterium *Wolbachia*. The present study is devoted to the analysis of the expression of two juvenile hormone metabolism genes encoding enzymes of its synthesis and degradation, juvenile hormone acid O-methyltransferase (*jhamt*) and juvenile hormone epoxide hydrase (*Jheh1*), respectively, in four wild-type *D. melanogaster* lines, two of them being infected with *Wolbachia*. Lines *w153* and *Bi90* were both derived from an individual wild-caught females infected with *Wolbachia*, while lines *w153^T* and *Bi90^T* were derived from them by tetracycline treatment and are free of infection. Line *Bi90* is known to be infected with the *Wolbachia* strain *wMel*, and line *w153*, with the *Wolbachia* strain *wMelPlus* belonging to the *wMelCS* genotype. It was found that infection with either *Wolbachia* strain does not affect the expression of the studied genes. At the same time, it was shown that the *w153* and *w153^T* lines differ from the *Bi90* and *Bi90^T* lines by an increased level of the *Jheh1* gene expression and do not differ in the *jhamt* gene expression level. Analysis of the fertility of these four lines showed that it does not depend on *Wolbachia* infection either, but differs between lines with different nuclear genotypes: in *w153* and *w153^T*, it is significantly lower than in lines *Bi90* and *Bi90^T*. The data obtained allow us to reasonably propose that the inter-line *D. melanogaster* polymorphism in the metabolism of the juvenile hormone is determined by its degradation (not by its synthesis) and correlates with the fertility level.

Key words: *Drosophila melanogaster*; *Wolbachia*; *jhamt*; *Jheh1*; gene expression; fertility; juvenile hormone metabolism.

For citation: Andreenkova O.V., Adonyeva N.V., Efimov V.M., Gruntenko N.E. Fertility differences between two wild-type *Drosophila melanogaster* lines correlate with differences in the expression of the *Jheh1* gene, which codes for an enzyme degrading juvenile hormone. *Vavilovskii Zhurnal Genetiki i Seleksii* = *Vavilov Journal of Genetics and Breeding*. 2024;28(2): 185-189. DOI 10.18699/vjgb-24-22

Различия в плодовитости между двумя линиями *Drosophila melanogaster* дикого типа коррелируют с различиями в экспрессии гена *Jheh1*, кодирующего фермент деградации ювенильного гормона

О.В. Андрееenkova , Н.В. Адоньева, В.М. Ефимов , Н.Е. Грунтенко  

Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
 nataly@bionet.nsc.ru

Аннотация. Ювенильный гормон играет у личинок *Drosophila melanogaster* роль “status quo” гормона, препятствуя преждевременному наступлению метаморфоза, а у имаго выполняет гонадотропную функцию, обеспечивая подготовку яичников к вителлогенезу. При снижении уровня ювенильного гормона у самок *D. melanogaster* наблюдаются нарушения репродукции, выражающиеся в задержке начала откладки яиц и снижении плодовитости. Еще одним фактором, способным повлиять на репродуктивную функцию насекомых, является инфицирование матерински наследуемой симбиотической α -протеобактерией *Wolbachia*. Настоящее исследование посвящено анализу экспрессии двух генов метаболизма ювенильного гормона, кодирующих ферменты его синтеза и деградации – кислую О-метилтрансферазу ювенильного гормона (*jhamt*) и эпоксидгидразу ювенильного гормона

(*Jheh1*) соответственно, у четырех линий *D. melanogaster* дикого типа, две из которых инфицированы *Wolbachia*. Линии *w153* и *Bi90* происходят от отдельных самок, отловленных в дикой природе, и инфицированы *Wolbachia*, а линии *w153^T* и *Bi90^T* получены на их основе посредством тетрациклиновой обработки и характеризуются отсутствием инфекции. Известно, что линия *Bi90* инфицирована штаммом *Wolbachia* генотипа *wMel*, а линия *w153* – штаммом *Wolbachia wMelPlus*, относящимся к генотипу *wMelCS*. Обнаружено, что инфицирование как одним, так и другим штаммом *Wolbachia* не влияет на экспрессию исследованных генов. В то же время показано, что обе линии с ядерным генотипом *w153* отличаются от линий с генотипом *Bi90* повышенным уровнем экспрессии гена *Jheh1* и не различаются по уровню экспрессии гена *jhamt*. Уровень плодовитости не зависел от инфицирования *Wolbachia*, но у линий с ядерным генотипом *w153* он оказался существенно ниже, чем у линий с ядерным генотипом *Bi90*. Полученные данные позволяют сделать обоснованное предположение о том, что межлинейный полиморфизм *D. melanogaster* по метаболизму ювенильного гормона коррелирует с уровнем плодовитости и определяется деградацией гормона, а не его синтезом.

Ключевые слова: *Drosophila melanogaster*; *Wolbachia*; *jhamt*; *Jheh1*; экспрессия генов; плодовитость; метаболизм ювенильного гормона.

Introduction

According to the current understanding of the genetic control of Diptera reproduction, a key role is played by 20-hydroxyecdysone (20E), while juvenile hormone (JH) only prepares ovaries for vitellogenesis, unlike in most other insect orders, where JH has the function that in Diptera is performed by 20E (Roy et al., 2018; Wu et al., 2021). The balance between these two hormones determines many events in the life of holometabolic insects from the larval period, where 20E initiates the start of moulting, and the JH level determines whether it will be larval moulting (if it is high), or the onset of metamorphosis (if it is low) (Truman, Riddiford, 2007), the neurohormonal stress response, which involves both hormones, and the regulation of changes in ovaries under heat stress or starvation (Gruntenko et al., 2003a; Terashima et al., 2005; Gruntenko, Rauschenbach, 2008).

Despite the secondary role of JH in oogenesis regulation and reproduction control in *Drosophila*, there are data indicating that in flies with a decreased level of JH, the reproduction process is disrupted, which is expressed as a delay in oviposition onset and a fertility decrease (Altartz et al., 1991; Gruntenko et al., 2003b; Yamamoto et al., 2013; Meiselman et al., 2017), and endogenic JH treatment of females speeds up egg maturation (Richard et al., 2001). Thus, we can assume that through controlling vitellogenins uptake by oocytes (Berger, Dubrovsky, 2005), JH takes part in the determination of the fertility level in *Drosophila*.

The intracellular signaling of JH is well described in the literature (Jindra et al., 2015; Roy et al., 2018), including the JH receptor complex Methoprene-tolerant (Met) – Taiman – Germ cell-expressed (Gce), the heat shock protein HSP83 and nucleoporin Nup358, which interact with Met and ensure JH transfer into the nucleus and the activation of the transcriptional factor Kr-h1 by it. At the same time, the mechanisms of the JH level regulation are still underresearched.

To add to the knowledge regarding this subject, we have estimated the level of fertility and expression of the genes responsible for JH synthesis and degradation, *jhamt* and *Jheh1*, in four *Drosophila melanogaster* lines, two of which were earlier demonstrated to differ in the fertility level (Adonyeva et al., 2021). *jhamt* codes for juvenile hormone acid O-methyltransferase (JHAMT), transforming JH acid or inactive JH precursors into the active form of the hormone at the

final stage of JH biosynthesis in insects (Niwa et al., 2008). *Jheh1* codes for one of the forms of JH epoxide hydrolase that inactivates the hormone via hydrolysis of the epoxide functional group producing JH diol (Flatt et al., 2005).

Notably, another factor capable of affecting fly fertility as well as JH metabolism is the infection with the maternally-inherited symbiotic α -proteobacterium *Wolbachia pipientis* (Werren et al., 2008; Burdina, Gruntenko, 2022). *Wolbachium* is a widely spread intracellular insect symbiont infecting more than 40 % of the studied species and greatly affecting host physiology (Werren et al., 2008; Burdina, Gruntenko, 2022). As lines *w153^T* and *Bi90^T*, the differences in the fertility of which were shown earlier, were derived from lines *w153* and *Bi90*, which, in turn, were derived from single females caught in nature and were initially infected with *Wolbachia*, we decided to use for analysis lines *w153* and *Bi90*, carrying the infection, and lines *w153^T* and *Bi90^T*, having undergone antibacterial therapy, to search for possible effects of *Wolbachia* on the fertility level and the expression of the JH metabolism genes.

Materials and methods

***Drosophila* lines.** In the work, we used four *D. melanogaster* lines: the *w153* and *Bi90* lines, derived from single females and carrying *Wolbachia* strains of the *wMelCS* and *wMel* genotypes, respectively (Ilinsky, 2013), and their derivatives, *w153^T* and *Bi90^T*, which underwent antibacterial therapy prior to the start of experiments. The lines were received from the collection of the Institute of Cytology and Genetics SB RAS. Notably, the *Wolbachia wMelPlus* strain, infecting the *w153* line, differs from other published strains of *wMelCS* by a large chromosomal inversion (Korenskaia et al., 2022).

Flies were kept on a standard medium (agar-agar, 7 g/l; corn flour 50 g/l; dry yeast 18 g/l; sugar 40 g/l) in an incubator (Sanyo, Japan) at a temperature of 25 °C, relative humidity of 50%, and 12:12 h light cycle. For the experiments, flies were synchronized at eclosion (they were collected 3–4 h afterwards). To analyze fertility and gene expression levels, 10-days-old females were taken.

Total RNA isolation and real-time RT-PCR. To assess the number of mRNA of the *jhamt* and *Jheh1* genes, 15 females per biological replicate per line were frozen in liquid nitrogen in 1.5 ml Eppendorf tubes. In total, three biological

replicates of all four *Drosophila* lines were performed. After removing the tubes from liquid nitrogen, 150 µl of TRI reagent No. BCBT8883 (Sigma, USA) was added to each tube and flies were homogenized. To remove large tissue fragments from the homogenate, the tubes were centrifuged for 5 min at 10,000 rpm in an Eppendorf centrifuge at a temperature of 7 °C and then the homogenate was transferred to clean 0.5 µl tubes. 30 µl of cold chloroform was added, and after shaking the tubes were left for 15 min at room temperature. Afterwards, the homogenate was centrifuged for 15 min at 12,000 rpm and a temperature of 7 °C. 75 µl of cold isopropanol was added to the supernatant, and after shaking the tubes were left for 10 min at room temperature. After centrifugation (12,000 rpm; 10 min), pellets were washed with 150 µl of 75° ethanol twice with centrifugation in-between, dried and dissolved in 100 µl of deionized water. RNA concentration was measured with the use of Nanodrop OneC (Thermo Scientific, USA) and adjusted to 200 ng/µl by the addition of deionized water. cDNA synthesis was performed with the use of ABScript III RT Master Mix for qPCR with gDNA Remover No. RK20429 (ABclonal Technology, China) in accordance with the manufacturer’s protocol.

jhamt and *Jheh1* expression were analyzed with CFX96 Touch amplificador (Bio-Rad, USA) using real-time RT-PCR with the M-427 set with SYBR-Green I (Syntol, Russia). Data were normalized on *Act5C*. For every sample, three technical replicates were performed. Primer sequences used in the study are presented in Table 1.

Fertility. To assess fertility, three male-female pairs aged 0–5 h were placed in cultivation vials (10 vials per experimental group), where they were left to lay eggs under standard conditions; flies were transferred to new vials every 24 h for 10 days. Fertility was calculated as the number of offspring (imagoes) eclosed from the eggs laid by experimental flies during the 10th day per one parent female.

Statistical analysis. Statistical significance of the differences in fertility (number of eggs per day per female) in the experimental groups was assessed using Student’s *t*-test. Pairwise comparisons were performed using the Benjamini–Hochberg correction. In all cases, $p < 0.05$ was considered statistically significant. Histogram data are presented as average means ± SEM.

Data on gene expression were analyzed by $2^{-\Delta\Delta CT}$ method (Livak, Schmittgen, 2001) using three biological replicates, each of which was obtained from three technical ones. Since Real-Time CFX96 Touch amplificador (Bio-Rad) provides

only the average mean of three technical replicates and standard mean error, it is impossible to check normality, use non-parametric criteria or bootstrap. However, these data are sufficient to calculate sum of squares of the three replicates.

The general formula for the squared error of the mean:

$$SEM^2(\bar{x}) = (\sum x_i^2 - N\bar{x}^2) / (N \times (N - 1)).$$

For each biological replicate, $N = 3$, and

$$\sum x_i^2 = 6SEM^2(\bar{x}) + 3\bar{x}^2.$$

This is sufficient to calculate both the average mean and its error:

$$\bar{x} = (\sum \bar{x}_k) / 3;$$

$$SEM^2(\bar{x}) = (6\sum SEM^2(\bar{x}_k) + 3\sum \bar{x}_k^2 - 9\bar{x}^2) / 72.$$

The total sum of technical values for each average mean equals 9. When calculating Student’s criterion of the significance of the difference between two average means, $2 \times 9 - 2 = 16$ degrees of freedom are obtained. It is known that significance criteria like Student’s criterion are resistant to deviations from normality (Kendall, Stewart, 1961) due to the distribution of means approaching normality with increasing sample size.

As we made six comparisons in total, the Benjamini–Hochberg correction was additionally calculated for the *p*-value to compare with three standard significance levels (Narkevich, Vinogradov, 2020).

Results and discussion

Real-time RT-PCR did not reveal significant differences in the expression level of the gene of juvenile hormone acid O-methyltransferase, *jhamt*, both between *D. melanogaster* lines *Bi90/Bi90^T* and *w153/w153^T* (the same genetic background but infected/uninfected with *Wolbachia*), and between lines *Bi90/w153* and *Bi90^T/w153^T* (the same infection status but different genetic background) (Fig. 1, a). At the same time, the expression of the JH epoxide hydrolase gene, *Jheh1*, was significantly decreased ($p < 0.001$) in lines *Bi90* and *Bi90^T* compared to lines *w153* and *w153^T* (see Fig. 1, b). This means that *Wolbachia* infection has no influence on the expression level of *jhamt* and *Jheh1*, which allows us to assume that synthesis and degradation of JH do not change under the bacteria effect. On the other hand, the existing differences in the gene expression level of *Jheh1*, which encodes an enzyme degrading JH, between *D. melanogaster* lines *Bi90/Bi90^T* and *w153/w153^T* allow us to assume the existence of differences in the enzyme activity level and, as a result, in JH content,

Primer sequences used in RT-PCR

Gene name		5'→3'	Length	Tm, °C	Reference
<i>jhamt</i>	F	GTCAGTTGTTGAACGATGTGGGT	23	58	Qu et al., 2017
	R	CCTCATTATTTTCACCTTGCTGC	24	55	
<i>Jheh1</i>	F	AGTTTACCAGGTTATGGCTGGTC	23	57	Guio et al., 2014
	R	CAGTCTCCACCCTGGATAAGAA	23	56	
<i>Act5C</i>	F	GCGCCCTTACTCTTTCACCA	20	58	Guio et al., 2014
	R	ATGTCACGGACGATTTCACG	20	55	

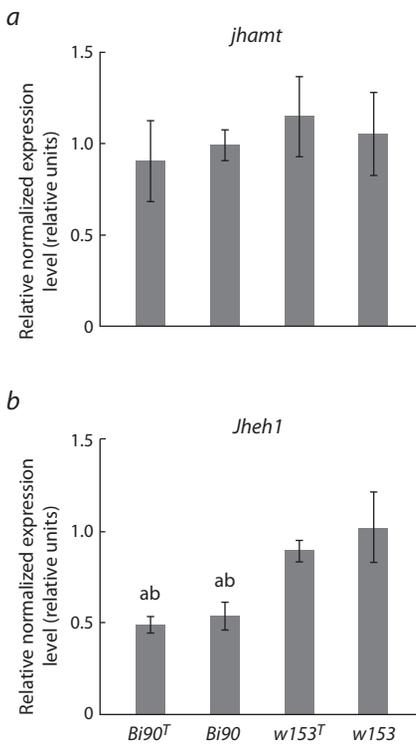


Fig. 1. Relative expression level of the *jhamt* (a) and *Jheh1* (b) genes in females of *D. melanogaster* lines *Bi90* (infected with the *wMel* strain of *Wolbachia*), *w153* (infected with the *wMelPlus* strain of *Wolbachia*), *Bi90^T* (uninfected), *w153^T* (uninfected).

Each value is the average mean of three biological replicates \pm SEM. a – the significance of the differences from females of the *w153* line ($p < 0.001$); b – the significance of the differences from females of the *w153^T* line ($p < 0.001$).

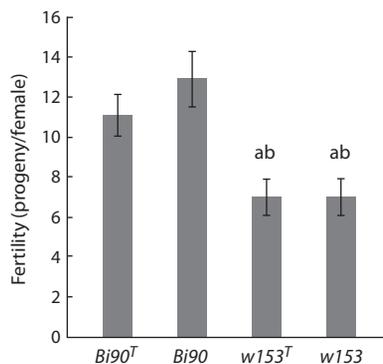


Fig. 2. Fertility level of *D. melanogaster* lines *Bi90* (infected with the *wMel* strain of *Wolbachia*), *w153* (infected with the *wMelPlus* strain of *Wolbachia*), *Bi90^T* (uninfected), *w153^T* (uninfected).

Each value is the average mean of 10 replicates (three females per replicate) \pm SEM. a – the significance of the differences from females of the *Bi90* line ($p < 0.001$); b – the significance of the differences from females of the *Bi90^T* line ($p < 0.001$).

the level of which should be elevated in the *Bi90* and *Bi90^T* lines with a decreased expression level of the gene coding for an enzyme degrading the hormone.

Fertility analysis of the *Bi90*, *Bi90^T*, *w153* and *w153^T* lines revealed that the line with a supposedly lower JH level (*w153* and *w153^T*) are characterized by significantly lower fertility ($p < 0.001$) compared to both the *Bi90* line and the *Bi90^T* line (Fig. 2). This agrees well with earlier data on the correlation of low fertility level with low JH level (Altaratz et al., 1991; Gruntenko et al., 2003b; Yamamoto et al., 2013; Meiselman et al., 2017) or with the *Met²⁷* mutation in the gene of the JH receptor (Gruntenko et al., 2000). Data regarding a decrease of the number of germline stem cells in the ovaries of *D. melanogaster*, carrying mutations in the *jhamt* and *Met* genes or with a knockdown of the latter, also indicate the important role JH plays in fertility regulation (Luo et al., 2020).

However, it is worth noting that most researchers attribute the demonstrated disruptions in reproduction and the decrease in fertility to disturbances in JH synthesis or the functioning of its receptor (Altaratz et al., 1991; Yamamoto et al., 2013; Meiselman et al., 2017; Luo et al., 2020), whereas our results indicate a correlation of interline differences in fertility in *D. melanogaster* with differences in the expression of gene encoding the enzyme that does not synthesize but degrades JH.

The lack of difference in fertility between the lines with the same genetic background infected and the uninfected with *Wolbachia* (*Bi90/Bi90^T* and *w153/w153^T*) correlates with the lack of difference in the expression of genes responsible for JH synthesis and degradation and allows us to assume that *Wolbachia* does not influence this trait. There is a slight contradiction with our earlier data obtained on the *Bi90^{wMelPlus}* line, where fertility and JH degradation differed from those of the *Bi90* line (Gruntenko et al., 2019).

However, it is necessary to note that the *Bi90^{wMelPlus}* line was received by transferring cytoplasm carrying a *Wolbachia* strain from the *w153* line to the nuclear background of the *Bi90* line (via 20 generations of backcrossing females carrying the corresponding *Wolbachia* strain with males of the *Bi90^T* line), and there is a non-zero probability that some aspects of *Wolbachia* influence on the physiology of the host can be attributed to the recent bacteria transfer and not merely to its presence in the cytoplasm. This hypothesis is indirectly confirmed by the lack of *Wolbachia* effect on fertility and JH degradation in the *Bi90* line, which was discovered in the same study (Gruntenko et al., 2019).

Additionally, transcriptomic analysis data obtained on infected *D. melanogaster* females revealed no changes in the differential expression of genes of the JH signaling pathway and the enzymes of its metabolism compared to uninfected females, which correlates with the lack of *Wolbachia* effect on the expression of the *jhamt* and *Jheh1* genes in *D. melanogaster* females in our work (Detcharoen et al., 2021; Lindsey et al., 2021).

Conclusion

To sum up, three reasonable hypotheses could be made based on our data: 1) JH does play a certain part in the regulation *D. melanogaster* reproduction; 2) JH catabolism has no less, or perhaps more, of a role in providing interline polymorphism by the JH level; 3) *Wolbachia* does not affect the JH level and fertility in *D. melanogaster* given the long history of symbiosis between a certain bacterium strain and host line.

References

- Adonyeva N.V., Menshanov P.N., Gruntenko N. A link between atmospheric pressure and fertility of *Drosophila* laboratory strains. *Insects*. 2021;12(10):947. DOI 10.3390/insects12100947
- Altaratz M., Applebaum Sh.W., Richard D.S., Gilbert L.I., Segal D. Regulation of juvenile hormone synthetis in wild-type and apterous mutant *Drosophila*. *Mol. Cell. Endocrinol.* 1991; 81(1-3):205-216. DOI 10.1016/0303-7207(91)90219-i
- Berger E.M., Dubrovsky E.B. Juvenile hormone molecular actions and interactions during development of *Drosophila melanogaster*. *Vitam. Horm.* 2005;73:175-215. DOI 10.1016/S0083-6729(05)73006-5
- Burdina E.V., Gruntenko N.E. Physiological aspects of *Wolbachia pipientis*–*Drosophila melanogaster* relationship. *J. Evol. Biochem. Phys.* 2022;58(2):303-317. DOI 10.1134/S0022093022020016

- Detcharoen M., Schilling M.P., Arthofer W., Schlick-Steiner B.C., Steiner F.M. Differential gene expression in *Drosophila melanogaster* and *D. nigrosarsa* infected with the same *Wolbachia* strain. *Sci. Rep.* 2021;11(1):11336. DOI 10.1038/s41598-021-90857-5
- Flatt T., Tu M.P., Tatar M. Hormonal pleiotropy and the juvenile hormone regulation of *Drosophila* development and life history. *BioEssays.* 2005;27(10):999-1010. DOI 10.1002/bies. 20290
- Gruntenko N.E., Rauschenbach I.Y. Interplay of JH, 20E and biogenic amines under normal and stress conditions and its effects on reproduction. *J. Insect Physiol.* 2008;54(6):902-908. DOI 10.1016/j.jinsphys.2008.04.004
- Gruntenko N.E., Khlebodarova T.M., Vasenkova I.A., Sukhanova M.J., Wilson T.G., Rauschenbach I.Y. Stress-reactivity of a *Drosophila melanogaster* strain with impaired juvenile hormone action. *J. Insect Physiol.* 2000;46(4):451-456. DOI 10.1016/S0022-1910(99)00131-6
- Gruntenko N.E., Bownes M., Terashima J., Sukhanova M.Zh., Rauschenbach I.Y. Heat stress affects oogenesis differently in wild-type *Drosophila virilis* and a mutant with altered juvenile hormone and 20-hydroxyecdysone levels. *Insect. Mol. Biol.* 2003a;12(4):393-404. DOI 10.1046/j.1365-2583.2003.00424.x
- Gruntenko N.E., Chentsova N.A., Andreenkova E.V., Bownes M., Segal D., Adonyeva N.V., Rauschenbach I.Y. Stress response in a juvenile hormone-deficient *Drosophila melanogaster* mutant *apterous^{56f}*. *Insect Mol. Biol.* 2003b;12(4):353-363. DOI 10.1046/j.1365-2583.2003.00419.x
- Gruntenko N.E., Karpova E.K., Adonyeva N.V., Andreenkova O.V., Burdina E.V., Ilinsky Y.Y., Bykov R.A., Mentshanov P.N., Rauschenbach I.Y. *Drosophila* female fertility and juvenile hormone metabolism depends on the type of *Wolbachia* infection. *J. Exp. Biol.* 2019; 222(Pt. 4):jeb195347. DOI 10.1242/jeb.195347
- Guio L., Barron M.G., Gonzalez J. The transposable element Bari-Jheh mediates oxidative stress response in *Drosophila*. *Mol. Ecol.* 2014; 23(8):2020-2030. DOI 10.1111/mec.12711
- Ilinsky Y.Y. Coevolution of *Drosophila melanogaster* mtDNA and *Wolbachia* genotypes. *PLoS One.* 2013;8(1):e54373. DOI 10.1371/journal.pone.0054373
- Jindra M., Bellés X., Shinoda T. Molecular basis of juvenile hormone signaling. *Curr. Opin. Insect Sci.* 2015;11:39-46. DOI 10.1016/j.cois.2015.08.004
- Kendall M.G., Stuart A. The Advanced Theory of Statistics. Vol. 2. Inference and Relationship. London: Charles Griffin, 1961
- Korenskaia A.E., Shishkina O.D., Klimentko A.I., Andreenkova O.V., Bobrovskikh M.A., Shatskaya N.V., Vasiliev G.V., Gruntenko N.E. New *Wolbachia pipiensis* genotype increasing heat stress resistance of *Drosophila melanogaster* host is characterized by a large chromosomal inversion. *Int. J. Mol. Sci.* 2022;23(24):16212. DOI 10.3390/ijms232416212
- Lindsey A.R.I., Bhattacharya T., Hardy R.W., Newton I.L.G. *Wolbachia* and virus alter the host transcriptome at the interface of nucleotide metabolism pathways. *mBio.* 2021;12(1):e03472-20. DOI 10.1128/mBio.03472-20
- Livak K.J., Schmittgen T.D. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods.* 2001;25(4):402-408. DOI 10.1006/meth.2001.1262
- Luo W., Veeran S., Wang J., Li S., Li K., Liu S.N. Dual roles of juvenile hormone signaling during early oogenesis in *Drosophila*. *Insect Sci.* 2020;27(4):665-674. DOI 10.1111/1744-7917.12698
- Meiselman M., Lee S.S., Tran R.T., Dai H., Ding Y., Rivera-Perez C., Wijesekera T.P., Dauwalder B., Noriega F.G., Adams M.E. Endocrine network essential for reproductive success in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA.* 2017;114(19):E3849-E3858. DOI 10.1073/pnas.1620760114
- Narkevich A.N., Vinogradov K.A., Grjibovski A.M. Multiple comparisons in biomedical research: the problem and its solutions. *Ekologiya Cheloveka = Human Ecology.* 2020;27(10):55-64. DOI 10.33396/1728-0869-2020-10-55-64 (in Russian)
- Niwa R., Niimi T., Honda N., Yoshiyama M., Itoyama K., Kataoka H., Shinoda T. Juvenile hormone acid O-methyltransferase in *Drosophila melanogaster*. *Insect Biochem. Mol. Biol.* 2008;38(7):714-720. DOI 10.1016/j.ibmb.2008.04.003
- Qu Z., Bendena W.G., Nong W., Siggins K.W., Noriega F.G., Kai Z.P., Zang Y.Y., Koon A.C., Chan H.Y.E., Chan T.F., Chu K.H., Lam H.M., Akam M., Tobe S.S., Lam Hui J.H. MicroRNAs regulate the sesquiterpenoid hormonal pathway in *Drosophila* and other arthropods. *Proc. Biol. Sci.* 2017;284(1869):20171827. DOI 10.1098/rspb.2017.1827
- Richard D.S., Jones J.M., Barbarito M.R., Cerula S., Detweiler J.P., Fisher S.J., Brannigan D.M., Scheswohl D.M. Vitellogenesis in diapausing and mutant *Drosophila melanogaster*: further evidence for the relative roles of ecdysteroids and juvenile hormones. *J. Insect Physiol.* 2001;47(8):905-913. DOI 10.1016/S0022-1910(01)00063-4
- Roy S., Saha T.T., Zou Z., Raikhel A.S. Regulatory pathways controlling female insect reproduction. *Annu. Rev. Entomol.* 2018;63:489-511. DOI 10.1146/annurev-ento-020117-043258
- Terashima J., Takaki K., Sakurai S., Bownes M. Nutritional status affects 20-hydroxyecdysone concentration and progression of oogenesis in *Drosophila melanogaster*. *J. Endocrinol.* 2005;187(1):69-79. DOI 10.1677/joe.1.06220
- Truman J.W., Riddiford L.M. The morphostatic actions of juvenile hormone. *Insect Biochem. Mol. Biol.* 2007;37(8):761-770. DOI 10.1016/j.ibmb.2007.05.011
- Werren J.H., Baldo L., Clark M.E. *Wolbachia*: master manipulators of invertebrate biology. *Nat. Rev. Microbiol.* 2008;6(10):741-751. DOI 10.1038/nrmicro1969
- Wu Z., Yang L., He Q., Zhou S. Regulatory mechanisms of vitellogenesis in insects. *Front. Cell Dev. Biol.* 2021;8:593613. DOI 10.3389/fcell.2020.593613
- Yamamoto R., Bai H., Dolezal A.G., Amdam G., Tatar M. Juvenile hormone regulation of *Drosophila* aging. *BMC Biol.* 2013;11:85. DOI 10.1186/1741-7007-11-85

Funding. The work was done with the support of the Russian Science Foundation grant No. 21-14-00090. The maintenance of experimental *D. melanogaster* lines was carried out in the *Drosophila* collection of the Institute of Cytology and Genetics SB RAS and was supported by BP #FWNR-2022-0019 of the Ministry of Science and Higher Education of the Russian Federation.

Acknowledgements. The authors thank Darya Kochetova for the translation of the article.

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

Received December 1, 2023. Revised January 17, 2024. Accepted January 18, 2024.

DOI 10.18699/vjgb-24-23

Influence of breed and environment on leukocyte telomere length in cattle

N.S. Yudin ^{1#}, A.V. Igoshin ^{1#}, G.A. Romashov¹, A.A. Martynov², D.M. Larkin ³ 

¹ Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia

² Arctic State Agrotechnological University, Yakutsk, Republic of Sakha (Yakutia), Russia

³ Royal Veterinary College, University of London, London United Kingdom

 dlarkin@rvc.ac.uk

Abstract. High milk yield is associated with reduced longevity in high-producing dairy cattle breeds. Pre-term culling leads to high replacement heifer demand and economic losses for the dairy industry. Selection for this trait is limited because of low heritability and difficulties in phenotype measurement. Telomeres are elements found at the ends of chromosomes, consisting of repetitive DNA sequences, several thousand base pairs in length, coupled with nucleoprotein complexes. Eventually, in humans and most other animals, telomere length reduces with age. When telomeric DNA is truncated to a critical length, cell ageing, cell cycle arrest, and apoptosis are induced. As a result, telomere length can be considered as a predictor of health risks and an individual's lifespan. The leukocyte telomere length may be used as a proxy phenotype of productive lifespan to improve cattle selection. Our objectives were to assess the effects of breed and breed group (dairy vs. beef) on the leukocyte telomere length and to estimate the effect of cold climate on this trait in Kalmyk cattle populations from the South (Rostov Oblast) and Far North (Republic of Sakha) regions of Russia. The leukocyte telomere lengths were estimated computationally from whole-genome resequencing data. We leveraged data on leukocyte telomere length, sex, and age of 239 animals from 17 cattle breeds. The breed factor had a significant effect on leukocyte telomere length across our sample. There was no difference in leukocyte telomere length between dairy and beef groups. The population factor had a significant effect on leukocyte telomere length in Kalmyk animals. In conclusion, we found that breed, but not breed group (dairy vs. beef), was significantly associated with leukocyte telomere length in cattle. Residence in colder climates was associated with longer leukocyte telomere length in Kalmyk breed cattle.

Key words: longevity; selection; cattle; breed; dairy; beef; environment; cold climate; leukocyte telomere length.

For citation: Yudin N.S., Igoshin A.V., Romashov G.A., Martynov A.A., Larkin D.M. Influence of breed and environment on leukocyte telomere length in cattle. *Vavilovskii Zhurnal Genetiki i Seleksii* = *Vavilov Journal of Genetics and Breeding*. 2024;28(2):190-197. DOI 10.18699/vjgb-24-23

Влияние породы и среды на длину теломер лейкоцитов у крупного рогатого скота

Н.С. Юдин ^{1#}, А.В. Игошин ^{1#}, Г.А. Ромашов¹, А.А. Мартынов², Д.М. Ларкин ³ 

¹ Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия

² Арктический государственный агротехнологический университет, Якутск, Россия

³ Королевский ветеринарный колледж, Лондон, Великобритания

 dlarkin@rvc.ac.uk

Аннотация. Высокие удои молока сопряжены с сокращением продолжительности жизни у высокопродуктивных молочных пород скота. Преждевременная выбраковка приводит к значительным экономическим потерям в молочном животноводстве и увеличению потребности в ремонтных телках. Отбор по этому признаку затруднен из-за низкой наследуемости и сложности измерения данного фенотипа. Теломеры – это структуры, находящиеся на концах хромосом, состоящие из повторяющихся последовательностей ДНК длиной в несколько тысяч пар оснований, связанных с нуклеопротеиновыми комплексами. У людей и большинства других животных длина теломер уменьшается с возрастом. Когда теломерная ДНК сокращается до критической длины, индуцируются процессы старения клеток, остановки клеточного цикла и апоптоза. В результате длину теломер можно рассматривать как предиктор рисков для здоровья и продолжительности жизни индивида. Длина теломер лейкоцитов может быть использована в качестве суррогатного фенотипа для признака продуктивного долголетия для улучшения селекции крупного рогатого скота. Целью нашей работы было – оценить влияние породы и направления продуктивности (молочное или мясное) на длину теломер лейкоцитов, а также проанализировать влияние холодного климата на этот признак в популяциях крупного рогатого скота калмыцкой породы на Юге (Ростовская область) и Крайнем Севере (Республика Саха) России. Измерение длины теломер лейкоцитов осуществлено с помощью компьютерных методов на основе данных полногеномного

ресеквенирования. Мы использовали данные о длине теломер лейкоцитов, половой принадлежности и возрасте 239 животных, относящихся к 17 породам крупного рогатого скота. Фактор породы оказывает существенное влияние на длину теломер лейкоцитов в нашей выборке. Достоверных различий в длине теломер лейкоцитов между молочными и мясными группами нами не выявлено. Значительное влияние на длину теломер лейкоцитов у животных калмыцкой породы оказывает фактор популяции. Таким образом, мы обнаружили, что именно порода, но не направление продуктивности (молочное или мясное), достоверно влияла на длину теломер лейкоцитов у крупного рогатого скота. Разведение в более холодном климате было ассоциировано с большей длиной теломер лейкоцитов у крупного рогатого скота калмыцкой породы.

Ключевые слова: долголетие; селекция; крупный рогатый скот; молочный; мясной; холодный климат; длина теломер лейкоцитов.

Introduction

High milk production correlates with poor longevity in high-producing cattle breeds, primarily Holstein Friesian (Hu et al., 2021). Pre-term culling leads to high replacement heifer demand and economic losses for the dairy industry (Grandl et al., 2019). There is a need to improve the longevity traits of dairy cattle. However, selection for these traits is limited because of low heritability and difficulties in phenotype measurement (Zhang H. et al., 2021).

Telomeres are elements found at the ends of chromosomes, consisting of repetitive DNA sequences, several thousand base pairs in length, coupled with nucleoprotein complexes (Jenner et al., 2022). They protect the chromosomes from degradation and inhibit aberrant rearrangements during cell division (Monaghan, Ozanne, 2018). Telomeres become shorter with every cell division due to the end replication problem (Chakravarti et al., 2021) but can be maintained through telomerase activity (Schrumpfová, Fajkus, 2020) and shelterin complex (de Lange, 2018). Eventually, in humans and most other animals, telomere length reduces with age (Blackburn et al., 2015; Whitemore et al., 2019). When telomeric DNA is truncated to a critical length, cell ageing, cell cycle arrest, and apoptosis are induced (Chakravarti et al., 2021; López-Otín et al., 2023). As a result, telomere length can be considered as a predictor of health risks and an individual's lifespan. Telomere length is correlated with many age-related conditions in humans (Armanios, 2022; Rossiello et al., 2022) and reduced life expectancy in humans and other species (Wilbourn et al., 2018; Liu et al., 2019; Crocco et al., 2021).

Telomeres in cattle shorten with age, similar to most other animals (Miyashita et al., 2002). Adult Holstein dams with short telomeres are more likely to be culled than dams with long telomeres (Brown et al., 2012). Productive lifespan in Holsteins was correlated with telomere length at birth (Ilska-Warner et al., 2019), at the age of one year (Seeker et al., 2018a), as well as telomere attrition rate (Seeker et al., 2021). The telomere length in the Agerolese breed, having a long lifespan, was significantly higher than that in the Holstein breed of the same age (Iannuzzi et al., 2022). Therefore, telomere length may be used as a proxy trait of lifespan and health to improve cattle selection.

Telomere length in cattle is a complex trait controlled by both genetics and environment. However, it is still unclear to what extent these factors influence this trait. A recent meta-analysis of the heritability of telomere length showed a moderate mean heritability of this trait (0.45) in 18 vertebrate species (Chik et al., 2022). Estimates of the heritability of telomere length, even in a single species (human), may range from 0.36

(Andrew et al., 2006) to 0.70 (Broer et al., 2013). This variability in estimates appears to be due to different research methodologies. For example, in most studies, parents and offspring are of different ages. In statistical analyses, age is counted as a covariate, but this implies a linear relationship between telomere length and age, which is not always the case (Dugdale, Richardson, 2018).

The influence of environmental factors on telomere length has been well studied in human epidemiological studies. For example, a negative correlation was found between telomere length and emotional stress (Law et al., 2016), Western pattern diet (Rafie et al., 2017), cigarette smoking (Astuti et al., 2017), and environmental chemicals (Zhang X. et al., 2013). In birds, short telomeres or a high rate of telomere shortening have been associated with malaria infection (Asghar et al., 2016), increased brood size (Reichert et al., 2014), early postnatal stress (Herborn et al., 2014) and sibling competition (Mizutani et al., 2016).

Heritability estimates of leukocyte telomere length in Holstein cattle ranged from 0.32 to 0.47 (Seeker et al., 2018a, b). Fourteen candidate genes at birth and nine at first lactation were associated with this trait in this breed using a genome-wide association study (Ilska-Warner et al., 2019). Our genome-wide association study of seventeen cattle breeds revealed several SNPs associated with bovine telomere length. We also confirmed the effects of loci reported by previous studies (Igochin et al., 2023). Mastitis (Ilska-Warner et al., 2019), bovine leukaemia virus infection (Szczołka et al., 2019), oxidative stress (Ribas-Maynou et al., 2022), parturition and raising the first calf (O'Daniel et al., 2023), lameness (Ilska-Warner et al., 2019), and lactation (Laubenthal et al., 2016) have been found to be associated with cattle telomere length. The management of the farm and genetics are herd-related factors that can significantly affect telomere length (Brown et al., 2012). However, the question remains unanswered: to what extent is telomere length in cattle determined by breed and influenced by environmental stressors, such as weather conditions?

There are two studies on the association of the animal breed and telomere length in cattle (Tilesi et al., 2010; Iannuzzi et al., 2022). Both studies found differences in telomere lengths between the two cattle breeds in the same tissues. However, it is unclear how widespread this phenomenon is across multiple cattle breeds with different phylogenetic origins and ecogeographic breeding conditions. P. Kordowitzki et al. hypothesized that severely disturbed energy balance in high-producing dairy cows eventually leads to decreased regenerative capacity and premature senescence, which can be

The climatic conditions for the sampling locations in Rostov Oblast and the Republic of Sakha (according to <https://climatecharts.net> (Zepner et al., 2021), accessed on 30 August 2023)

Location	Coordinates	Mean January temperature, °C	Mean July temperature, °C	Mean annual temperature, °C	Precipitation sum, mm
Mechetny settlement (Rostov Oblast)	48°03'54"N, 40°38'23"E	-4.4	23.7	9.8	508.7
Kyuyorelyakh settlement (Republic of Sakha)	62°34'17"N, 126°50'26"E	-37.2	18.0	-9.1	312.7

assessed by telomere length (Kordowitzki et al., 2021). The fraction of short telomeres in PBMCs of the high-producing Holstein-Friesian breed was higher than in the dual-purpose Polish Red breed, but this observation was not supported by a statistical test significance. Therefore, it remains unclear whether different breeds of cattle (dairy vs. beef) exhibit variation in telomere length.

Seeker et al. suggested that heat may be an environmental stressor capable of causing telomere attrition (Seeker et al., 2021). They found a strong correlation between maximum summer temperature and telomere attrition in Holstein-Friesian cattle. Heat stress during gestation also affected the telomere length in newborn Holstein calves. A higher median temperature-humidity index during gestation resulted in calves born with shorter telomere lengths (Meesters et al., 2023). There are, however, no studies that investigate the influence of cold weather on telomere length in cattle. A single human study showed that prenatal temperature exposure below 5 °C was associated with longer telomere length in newborn babies (Martens et al., 2019).

There are only two native beef cattle breeds in Russia: Kalmyk and its derivative Kazakh Whiteheaded breeds. It is believed that the Kalmyk cattle originated in Northwest China (Dzungaria) and was brought to Russia, to the Volga area, by migrating nomadic tribes in the seventeenth century (Dmitriev, Ernst, 1989). The Kalmyk breed was created under harsh conditions: the icy wind in winter or the hot sun in summer, frequent epizootics, etc. The specific traits of the Kalmyk breed include high viability, adaptation to the harsh climate, resistance to infections, long lifespan, thickening of the epidermis at the expense of the dermis in winter, abundance of sebaceous and sweat glands in the skin compared to other breeds (Dmitriev, Ernst, 1989). In Russia, the Kalmyk beef herd is mainly found in two regions: the Republic of Kalmykia and Rostov Oblast (Kayumov et al., 2014). This breed has also been reared in the Republic of Sakha (Yakutia) since 2013 when about 200 Kalmyk cattle animals were imported from the Republic of Kalmykia (Sleptsov, Machakhtyrova, 2019). Yakutia has an extreme and severe climate, with the average winter temperature below -35 °C.

The objectives of our study were (1) to assess the effects of breed and breed group (dairy vs. beef) on the leukocyte telomere length in the sample of 239 animals from 17 cattle breeds and (2) to estimate the effect of cold climate on this trait in Kalmyk cattle populations from the South and Far North regions of Russia. We hypothesized that high milk yield or extreme cold weather may be stress factors that may have led to a change in telomere length in cattle.

Materials and methods

Samples. In this work, we leveraged data on leukocyte telomere length, sex, and age of 239 animals from cattle breeds used in our previous study (Igoshin et al., 2023). The leukocyte telomere lengths have been estimated computationally from whole-genome resequencing data using TelSeq software (Ding et al., 2014), which is a frequently used program for this purpose and which has been confirmed by multiple experimental techniques (Ding et al., 2014; Cook et al., 2016; Pinese et al., 2020; Taub et al., 2022; Zhang D. et al., 2022). The details of the estimation procedure can be found in our previous work (Igoshin et al., 2023).

The breeds investigated are dairy (Russian Black Pied, Holstein, Kholmogory, Red Steppe, Yaroslavl), beef (Charolais, Hereford, Kalmyk, Kazakh Whiteheaded, Wagyu), and dual-purpose (Alatau, Bestuzhev, Buryat, Kostroma, Tagil, Ukrainian Grey, Yakut) (Dunin, Dankvert, 2013; Lhasaranov, 2020). Among 30 individuals of the Kalmyk cattle breed (Supplementary Material 1)¹, one group of animals ($n = 10$) was reared in Rostov Oblast (Mechetny settlement), while the other ($n = 20$) was from the Republic of Sakha (Kyuyorelyakh settlement). The climatic conditions in these locations differ substantially (see the Table). All the Kalmyk animals from both locations were raised in a stall-pasture system.

Population structure. Even in the absence of selection, a founder effect or, more broadly, genetic drift could lead to genetic differentiation between two isolated populations of common origin. To ensure that two populations of the Kalmyk breed are genetically indistinguishable, we performed the principal component analysis using PLINK v.1.9 (Purcell et al., 2007) (--pca option) and the analysis of population structure using fastSTRUCTURE v1.0 (Raj et al., 2014). The fastSTRUCTURE program was run with K ranging from $K = 2$ to $K = 8$. The resulting cluster memberships were visualized with PONG v.1.5 software (Behr et al., 2016). For both methods, we used an LD-pruned (PLINK: --indep-pairwise 5000 100 0.1) dataset containing genotypes of 20,184 SNPs in 116 animals (Kalmyk animals and individuals having SRA ID from Supplementary Material 1).

Statistical analysis. Like in many other studies, the distribution of LTL in our work was skewed. If not corrected, this violates the assumptions of parametric tests, thus affecting statistical power (Lantz, 2013). Therefore, associations with LTL were tested by using log-transformed LTL values (e.g. Leung et al., 2014; Lynch et al., 2016). To find out whether a breed factor contributes to LTL variation in cattle, we used

¹ Supplementary Materials 1–6 are available at <https://vavilovj-icg.ru/download/pict-2024-28/appx10.pdf>

ANOVA (“*aov*” R function) with log-transformed LTLs ($\log_{10}(\text{LTLs})$) as a response variable and breed as a factor variable, accounting for age and sex: *aov(logLTL ~ Age + Sex + Breed)*. To find out which breeds significantly differ from each other, we additionally performed a standard ANOVA *post hoc* test – Tukey’s HSD test utilising the “*glht*” function from the “*multcomp*” R package (Hothorn et al., 2008). Also, we combined beef and dairy breeds into two groups and tested for a difference between them: *aov(logLTL ~ Age + Sex + Group)*. As all the Kalmyk individuals used were dams, the test for differences between this breed’s populations was conducted by accounting only for age: *aov(logLTL ~ Age + Population)*.

For statistically significant variables we estimated the variance explained (η^2) using the “*eta_squared*” function from the “*effectsize*” R package (Ben-Shachar et al., 2020) with the “*partial = TRUE*” option.

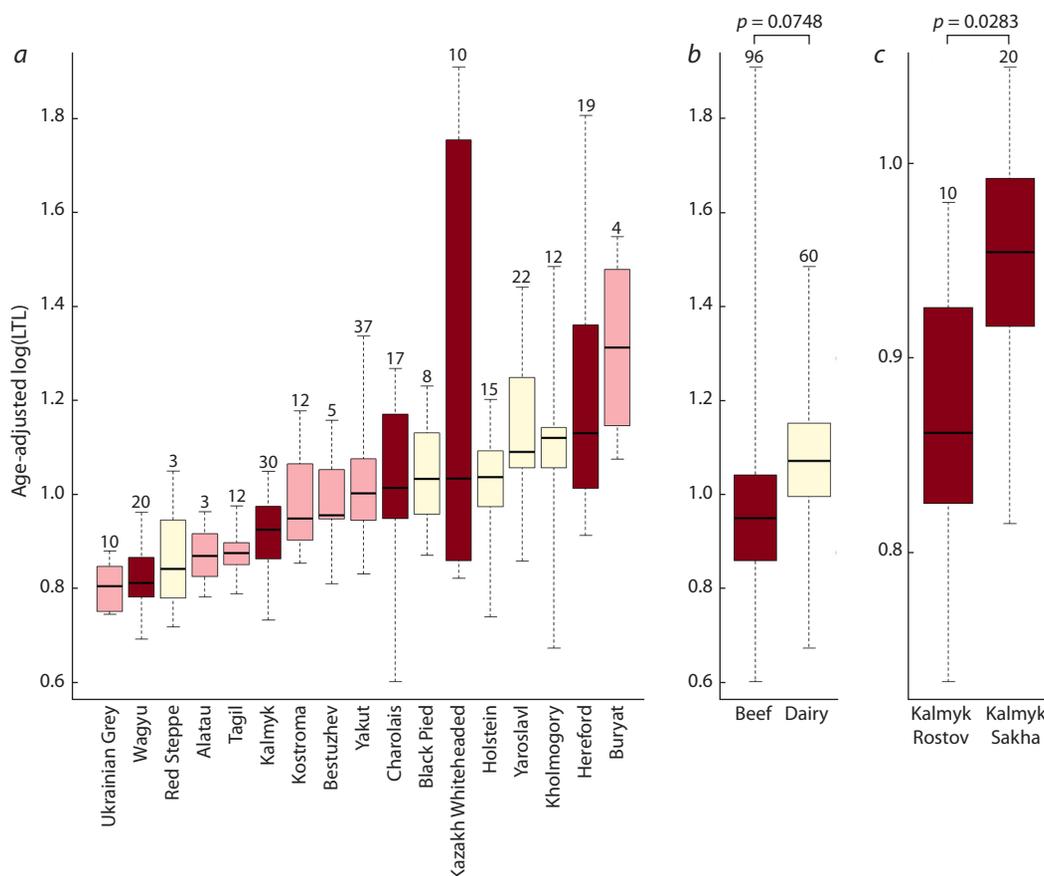
Preparing data for visualization and descriptive statistics. As confirmed in our study, telomere length typically decreases with age (Spearman’s $\rho = -0.305$, $p = 1.58 \times 10^{-6}$ for raw and log-transformed LTLs) (Supplementary Material 2). Therefore, for boxplot visualization and descriptive statistics, we calculated the $\log_{10}(\text{LTL})$ values expected given the constant age. For this purpose, we fitted a regression model (“*lm*” R function) with logLTL as the response variable, and

age, sex (coded by 0/1) and breed (16 covariates coded by 0/1) variables as predictors. As a result, we obtained regression parameters (intercept and slopes for each predictor) and residuals. For each animal, we summed up: the intercept, the animal’s residual, and products of each predictor value and its respective slope. For the age predictor, however, actual values were substituted by the average value of 4.5 years. The resulting values represent the expectation for logLTL given the constant age of 4.5 years. These age-adjusted $\log_{10}(\text{LTLs})$ and their corresponding values in kilobases (hereafter “age-adjusted LTLs”) are shown in Supplementary Material 1. The descriptive statistics for breeds can be found in Supplementary Material 3.

Results

The statistical analysis shows that breed factor has a significant ($p = 6.12 \times 10^{-15}$, $\eta^2 = 0.37$) effect on leukocyte telomere length across our sample of 239 individuals (see Figure, a). Tukey’s HSD test showed significant differences for 28 breed pairs (Supplementary Material 4). At the same time, there is no significant difference in LTL between dairy and beef groups ($p = 0.0748$) (see Figure, b).

The statistical testing performed for the Kalmyk breed shows that the population factor has a significant ($p = 0.0283$,



Boxplots illustrating the log-transformed and adjusted for age (expectation at 4.5 years) leukocyte telomere lengths in a) different breeds; b) beef and dairy breed groups, and c) in two populations of the Kalmyk breed.

The *p*-values designate the statistical significance for differences between the abovementioned categories. The brown, pink and light-yellow colours correspond to beef, dual-purpose, and dairy breeds. The numbers at the top of boxplots indicate the number of animals.

$\eta^2 = 0.17$) effect on LTL in studied Kalmyk animals, with the individuals from Rostov Oblast having shorter telomeres (see Figure, c). The results of the principal component analysis show that the Rostov and Sakha populations form two highly overlapping clusters (Supplementary Material 5). Also, fastSTRUCTURE results suggest that the two Kalmyk populations are homogeneous and possible genetic differences between them do not exceed the level of variation within other breeds (Supplementary Material 6). Therefore, the LTL differences between the two groups are most likely explained by environmental conditions.

It should also be mentioned that the age variable significantly affects LTL in all tests: $p = 1.46 \times 10^{-6}$, $\eta^2 = 0.10$ (testing for the effect of breed); $p = 0.0248$, $\eta^2 = 0.07$ (dairy vs. beef group); and $p = 0.0038$, $\eta^2 = 0.27$ (Kalmyk Sakha vs. Kalmyk Rostov). However, the sex factor has no significant effect on LTL either in the test for breed factor ($p = 0.205$) or in the test for the factor of breed group ($p = 0.8752$).

Discussion

The primary aim of this study was to investigate the correlation between the breed type and leukocyte telomere length (LTL) in cattle. Our additional goal was to check the effect of the environment (e. g., colder climate) on the LTL within populations of the same breed grown in different climates. Based on the resequencing data of 17 cattle breeds, our findings indicate that the breed factor has a significant impact on LTL. These results align with earlier studies that reported LTL differences in pairwise comparisons between cattle breeds (Tilesi et al., 2010; Iannuzzi et al., 2022). We also found evidence for an association between LTL and differences in climatic conditions for a single cattle breed reared in different regions of Russia.

It was shown in our analysis that the breed factor contributes more to total LTL variation compared to age. The possible practical implication for this could be the use of breeds characterised by long telomeres in crossbreeding programs aimed to improve telomere-length-associated phenotypes in cattle.

Apart from cattle, studies are reporting LTL differences between *Caenorhabditis elegans* strains (Cook et al., 2016), outbred populations and inbred strains of mice (*Mus musculus* and *Peromyscus leucopus*) (Manning et al., 2002), and dog breeds (Fick et al., 2012). These reports suggest the existence of a genetic basis for such variability. Based on heritability estimates for LTL in Holsteins (0.32–0.47) (Seeker et al., 2018a, b), we propose that genetic factors may largely explain inter-breed differences observed in our study.

Herein we compared the LTL in dairy and beef breeds. The results However did not reveal any statistically significant difference between these two groups. This finding is consistent with a previous study that compared LTLs between dairy and dual-purpose cattle breeds, which also showed no difference (Kordowitzki et al., 2021). Our study focused on distinct groups of cattle breeds, specifically dairy and beef, covering a wider range of genetics than the previous study. Also, each production breed type was represented by five breeds. Therefore, the results reported herein could provide stronger support for the lack of LTL differences between the cattle breed types. Our results are also consistent with similar studies done in other domestic species, e. g., dogs, where no difference in LTL was reported for breed groups (working,

herding, hunting) (Fick et al., 2012). It appears that complementary contributions of many factors affecting a particular breed (e. g. genetic makeup, management practices, veterinary care, climate conditions, etc.) have a greater influence on the LTL than physiological features associated with different production types. This result suggests that the selection for telomere-length-associated traits will probably not lead to substantial changes in milk or meat yields.

To investigate the possible effects of environments on telomere lengths of the same cattle breed, we compared the LTLs between two populations of the Kalmyk cattle reared in different climatic conditions. We observed a significant difference in agreement with a previous human study showing an association between prenatal cold exposure and longer blood telomere length in newborns (Martens et al., 2019). Indeed, longer telomere lengths detected in animals from the Sakha Republic with colder climates compared to the control population from Rostov Oblast imply that there could be a mechanism of telomere maintenance in colder climates in cattle. In ectotherms, however, there are reports that at cooler conditions telomere shortening happens during development (Friesen et al., 2022; Burraco et al., 2023), but other authors did not confirm this observation (McLennan et al., 2018).

In bat species, *Myotis myotis*, average and minimum temperatures, rainfall and wind speed during the spring when bats emerge from hibernation, give birth and rear young were associated with higher telomere attrition (Foley et al., 2020). The authors, however, did not report which variable is the driver for telomere length change. The comparison of telomere length between two species of rodents hibernating at either 3 or 14 °C revealed that individuals hibernating at the warmer temperature had longer telomeres than individuals hibernating at the colder temperature (Nowack et al., 2019). The authors hypothesized that the observed effect was not related to cold climate, but rather was associated with restoration of telomere length during frequent arousals when the body temperature returns to normal values.

The mechanisms by which cold climate impacts leukocyte telomere length in cattle remain unclear. On the one hand, cold exposure may inhibit telomere shortening, since low temperature reduces the rate of cell proliferation in mammalian cells (Kanagawa et al., 2006; Fulbert et al., 2019). On the other hand, cold exposure may induce telomere elongation by influencing the components of the telomerase complex. It was shown that cold-inducible RNA-binding protein (CIRP) was essential for telomere maintenance at hypothermia conditions *in vitro* by regulating both reverse transcriptase TERT and the RNA subunit TERC in the telomerase core complex (Zhang Y. et al., 2016). The transcription of telomeric repeat-containing RNAs (TERRAs), which are associated with telomere stability, was induced in mice exposed to cold (Galigniana et al., 2020).

One could ask if there is an optimal ambient temperature at which the LTL in cattle would be the longest. The few studies on the effect of ambient temperature on the LTL of endothermic mammals do not allow us to answer this question unambiguously. Extremely low or high ambient temperatures lead to hypo- and hyperthermia when the body temperature deviates substantially below and above the narrow limits of the regulated range, i. e., cause stress. The influence of a large

number of stressors, including extreme environmental factors, is known to be associated with shorter telomeres or an increased rate of telomere shortening (Chatelain et al., 2020; Lin, Epel, 2022).

Indirect information on the effect of moderate cold on LTL, when body temperature remains within the normal range, could be obtained from studies of the effect of body temperature on the ageing process. On the one hand, in endotherms, a small decrease in body temperature is associated with an increase in life expectancy (Conti et al., 2006; Carrillo, Flouris, 2011). On the other hand, in some cases, there is an inverse relationship (Zhao et al., 2022). For example, human females tend to have a longer life expectancy than males, but their body temperature is higher (Waaen, Buxbaum, 2011). The mechanisms that control the relationship between body temperature and life expectancy involve not only a decrease in metabolic rate when the temperature declines, but also neuroendocrine processes that indirectly affect a variety of physiological responses when temperature changes. Therefore, the optimal limits of ambient temperature at which the LTL in cattle will be longest could exist, but this question requires further study.

Conclusion

In conclusion, we found that breed, but not breed group (dairy vs. beef), significantly influenced leukocyte telomere length in cattle. Residence in colder climates was associated with longer leukocyte telomere length in the Kalmyk cattle breed. Our results add to the evidence regarding the influence of breed origin and cold climate on this trait in farm animals.

References

- Andrew T., Aviv A., Falchi M., Surdulescu G.L., Gardner J.P., Lu X., Kimura M., Kato B.S., Valdes A.M., Spector T.D. Mapping genetic loci that determine leukocyte telomere length in a large sample of unselected female sibling pairs. *Am. J. Hum. Genet.* 2006;78(3): 480-486. DOI 10.1086/500052
- Armanios M. The role of telomeres in human disease. *Annu. Rev. Genomics Hum. Genet.* 2022;23:363-381. DOI 10.1146/annurev-genom-010422-091101
- Asghar M., Palinauskas V., Zaghdoudi-Allan N., Valkiūnas G., Mukhin A., Platonova E., Färmert A., Bensch S., Hasselquist D. Parallel telomere shortening in multiple body tissues owing to malaria infection. *Proc. Biol. Sci.* 2016;283(1836):20161184. DOI 10.1098/rspb.2016.1184
- Astuti Y., Wardhana A., Watkins J., Wulaningsih W.; PILAR Research Network. Cigarette smoking and telomere length: A systematic review of 84 studies and meta-analysis. *Environ. Res.* 2017;158:480-489. DOI 10.1016/j.envres.2017.06.038
- Behr A.A., Liu K.Z., Liu-Fang G., Nakka P., Ramachandran S. pong: fast analysis and visualization of latent clusters in population genetic data. *Bioinformatics.* 2016;32(18):2817-2823. DOI 10.1093/bioinformatics/btw327
- Ben-Shachar M.S., Lüdtke D., Makowski D. effectsize: Estimation of effect size indices and standardized parameters. *J. Open Source Softw.* 2020;5(56):2815. DOI 10.21105/joss.02815
- Blackburn E.H., Epel E.S., Lin J. Human telomere biology: A contributory and interactive factor in aging, disease risks, and protection. *Science.* 2015;350(6265):1193-1198. DOI 10.1126/science.aab3389
- Broer L., Codd V., Nyholt D.R., Deelen J., Mangino M., Willemsen G., Albrecht E., ... Vink J.M., Spector T.D., Slagboom P.E., Martin N.G., Samani N.J., van Duijn C.M., Boomsma D.I. Meta-analysis of telomere length in 19,713 subjects reveals high heritability, stronger maternal inheritance and a paternal age effect. *Eur. J. Hum. Genet.* 2013;21(10):1163-1168. DOI 10.1038/ejhg.2012.303
- Brown D.E., Dechow C.D., Liu W.S., Harvatine K.J., Ott T.L. Hot topic: association of telomere length with age, herd, and culling in lactating Holsteins. *J. Dairy Sci.* 2012;95(11):6384-6387. DOI 10.3168/jds.2012-5593
- Burraco P., Hernandez-Gonzalez M., Metcalfe N.B., Monaghan P. Ageing across the great divide: tissue transformation, organismal growth and temperature shape telomere dynamics through the metamorphic transition. *Proc. Biol. Sci.* 2023;290(1992):20222448. DOI 10.1098/rspb.2022.2448
- Carrillo A.E., Flouris A.D. Caloric restriction and longevity: effects of reduced body temperature. *Ageing Res. Rev.* 2011;10(1):153-162. DOI 10.1016/j.arr.2010.10.001
- Chakravarti D., LaBella K.A., DePinho R.A. Telomeres: history, health, and hallmarks of aging. *Cell.* 2021;184(2):306-322. DOI 10.1016/j.cell.2020.12.028
- Chatelain M., Drobniak S.M., Szulkin M. The association between stressors and telomeres in non-human vertebrates: a meta-analysis. *Ecol. Lett.* 2020;23(2):381-398. DOI 10.1111/ele.13426
- Chik H.Y.J., Sparks A.M., Schroeder J., Dugdale H.L. A meta-analysis on the heritability of vertebrate telomere length. *J. Evol. Biol.* 2022;35(10):1283-1295. DOI 10.1111/jeb.14071
- Conti B., Sanchez-Alavez M., Winsky-Sommerer R., Morale M.C., Lucero J., Brownell S., Fabre V., Huitron-Resendiz S., Henriksen S., Zorrilla E.P., de Lecea L., Bartfai T. Transgenic mice with a reduced core body temperature have an increased life span. *Science.* 2006; 314(5800):825-828. DOI 10.1126/science.1132191
- Cook D.E., Zdraljjevic S., Tanny R.E., Seo B., Riccardi D.D., Noble L.M., Rockman M.V., Alkema M.J., Braendle C., Kammenga J.E., Wang J., Brownell S., Félix M.A., Lee J., Andersen E.C. The genetic basis of natural variation in *Caenorhabditis elegans* telomere length. *Genetics.* 2016;204(1):371-383. DOI 10.1534/genetics.116.191148
- Crocco P., De Rango F., Dato S., Rose G., Passarino G. Telomere length as a function of age at population level parallels human survival curves. *Ageing (Albany NY).* 2021;13(1):204-218. DOI 10.18632/aging.202498
- de Lange T. Shelterin-mediated telomere protection. *Annu. Rev. Genet.* 2018;52:223-247. DOI 10.1146/annurev-genet-032918-021921
- Ding Z., Mangino M., Aviv A., Spector T., Durbin R. Estimating telomere length from whole genome sequence data. *Nucleic Acids Res.* 2014;42(9):e75. DOI 10.1093/nar/gku181
- Dmitriev N.G., Ernst L.K. (Eds.). Animal Genetics Resources of the USSR. Rome: Food and Agriculture Organization of the United Nations, 1989
- Dugdale H.L., Richardson D.S. Heritability of telomere variation: it is all about the environment! *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 2018;373(1741):20160450. DOI 10.1098/rstb.2016.0450
- Dunin I.M., Dankvert A.G. (Eds.). Breeds and Types of Farm Animals in the Russian Federation. Moscow: All-Russia Research Institute of Animal Breeding, 2013 (in Russian)
- Fick L.J., Fick G.H., Li Z., Cao E., Bao B., Heffelfinger D., Parker H.G., Ostrander E.A., Riabowol K. Telomere length correlates with life span of dog breeds. *Cell Rep.* 2012;2(6):1530-1536. DOI 10.1016/j.celrep.2012.11.021
- Foley N.M., Petit E.J., Brazier T., Finarelli J.A., Hughes G.M., Touzalin F., Puechmaile 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
- Friesen C.R., Wapstra E., Olsson M. Of telomeres and temperature: Measuring thermal effects on telomeres in ectothermic animals. *Mol. Ecol.* 2022;31(23):6069-6086. DOI 10.1111/mec.16154
- Fulbert C., Gaude C., Sulpice E., Chabardès S., Ratel D. Moderate hypothermia inhibits both proliferation and migration of human glioblastoma cells. *J. Neurooncol.* 2019;144(3):489-499. DOI 10.1007/s11060-019-03263-3
- Galigniana N.M., Charó N.L., Uranga R., Cabanillas A.M., Pivien-Pi-

- lipuk G. Oxidative stress induces transcription of telomeric repeat-containing RNA (TERRA) by engaging PKA signaling and cytoskeleton dynamics. *Biochim. Biophys. Acta Mol. Cell. Res.* 2020; 1867(4):118643. DOI 10.1016/j.bbamer.2020.118643
- Grandl F., Furger M., Kreuzer M., Zehetmeier M. Impact of longevity on greenhouse gas emissions and profitability of individual dairy cows analysed with different system boundaries. *Animal*. 2019; 13(1):198-208. DOI 10.1017/S175173111800112X
- Herborn K.A., Heidinger B.J., Boner W., Noguera J.C., Adam A., Daunt F., Monaghan P. Stress exposure in early post-natal life reduces telomere length: an experimental demonstration in a long-lived seabird. *Proc. Biol. Sci.* 2014;281(1782):20133151. DOI 10.1098/rspb.2013.3151
- Hothorn T., Bretz F., Westfall P. Simultaneous inference in general parametric models. *Biom. J.* 2008;50(3):346-363. DOI 10.1002/bimj.20081042
- Hu H., Mu T., Ma Y., Wang X., Ma Y. Analysis of longevity traits in Holstein cattle: A review. *Front. Genet.* 2021;12:695543. DOI 10.3389/fgene.2021.695543
- Iannuzzi A., Albarella S., Parma P., Galdiero G., D'Anza E., Pistucci R., Peretti V., Ciotola F. Characterization of telomere length in Agerolese cattle breed, correlating blood and milk samples. *Anim. Genet.* 2022;53(5):676-679. DOI 10.1111/age.13227
- Igoshin A.V., Yudin N.S., Romashov G.A., Larkin D.M. A multi-breed genome-wide association study for cattle leukocyte telomere length. *Genes (Basel)*. 2023;14(8):1596. DOI 10.3390/genes14081596
- Ilska-Warner J.J., Psifidi A., Seeker L.A., Wilbourn R.V., Underwood S.L., Fairlie J., Whitelaw B., Nussey D.H., Coffey M.P., Banos G. The genetic architecture of bovine telomere length in early life and association with animal fitness. *Front. Genet.* 2019;10:1048. DOI 10.3389/fgene.2019.01048
- Jenner L.P., Peska V., Fulnečková J., Sýkorová E. Telomeres and their neighbors. *Genes (Basel)*. 2022;13(9):1663. DOI 10.3390/genes13091663
- Kanagawa T., Fukuda H., Tsubouchi H., Komoto Y., Hayashi S., Fukui O., Shimoya K., Murata Y. A decrease of cell proliferation by hypothermia in the hippocampus of the neonatal rat. *Brain Res.* 2006;1111(1):36-40. DOI 10.1016/j.brainres.2006.06.112
- Kayumov F.G., Chernomyrdin V.N., Mayevskaya L.A., Surundaeva L.G., Polskikh S.S. The use of Kalmyk cattle on animal breeding farms in Russia. *Izv. Orenbg. State Agrar. Univ.* 2014;5(49):116-119 (in Russian)
- Kordowitzki P., Merle R., Hass P.-K., Plendl J., Rieger J., Kaesmeyer S. Influence of age and breed on bovine ovarian capillary blood supply, ovarian mitochondria and telomere length. *Cells*. 2021;10(10):2661. DOI 10.3390/cells10102661
- Lantz B. The impact of sample non-normality on ANOVA and alternative methods. *Br. J. Math. Stat. Psychol.* 2013;66(2):224-244. DOI 10.1111/j.2044-8317.2012.02047.x
- Laubenthal L., Hoelker M., Frahm J., Dänicke S., Gerlach K., Südekum K.-H., Sauerwein H., Häussler S. Short communication: Telomere lengths in different tissues of dairy cows during early and late lactation. *J. Dairy Sci.* 2016;99(6):4881-4885. DOI 10.3168/jds.2015-10095
- Law E., Girgis A., Lambert S., Sylvie L., Levesque J., Pickett H. Telomeres and stress: promising avenues for research in psycho-oncology. *Asia-Pacific J. Oncol. Nurs.* 2016;3(2):137-147. DOI 10.4103/2347-5625.182931
- Leung C.W., Lاراia B.A., Needham B.L., Rehkopf D.H., Adler N.E., Lin J., Blackburn E.H., Epel E.S. Soda and cell aging: associations between sugar-sweetened beverage consumption and leukocyte telomere length in healthy adults from the National Health and Nutrition Examination Surveys. *Am. J. Public Health*. 2014;104(12):2425-2431. DOI 10.2105/AJPH.2014.302151
- Lhasaranov B. Pasture animal husbandry in Eastern Siberia. *Biomed. J. Sci. Tech. Res.* 2020;31(3):24160-24163. DOI 10.26717/BJSTR.2020.31.005094
- 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
- Liu J., Wang L., Wang Z., Liu J.-P. Roles of telomere biology in cell senescence, replicative and chronological ageing. *Cells*. 2019;8(1):54. DOI 10.3390/cells8010054
- López-Otín C., Blasco M.A., Partridge L., Serrano M., Kroemer G. Hallmarks of aging: An expanding universe. *Cell*. 2023;186(2):243-278. DOI 10.1016/j.cell.2022.11.001
- Lynch S.M., Peek M.K., Mitra N., Ravichandran K., Branas C., Spangler E., Zhou W., Paskett E.D., Gehlert S., DeGraffinreid C., Rebbeck T.R., Riethman H. Race, ethnicity, psychosocial factors, and telomere length in a multicenter setting. *PLoS One*. 2016;11(1):e0146723. DOI 10.1371/journal.pone.0146723
- Manning E.L., Crossland J., Dewey M.J., Van Zant G. Influences of inbreeding and genetics on telomere length in mice. *Mamm. Genome*. 2002;13(5):234-238. DOI 10.1007/s003350020027
- Martens D.S., Plusquin M., Cox B., Nawrot T.S. Early biological aging and fetal exposure to high and low ambient temperature: A birth cohort study. *Environ. Health Perspect.* 2019;127(11):117001. DOI 10.1289/EHP5153
- McLennan D., Armstrong J.D., Stewart D.C., Mckelvey S., Boner W., Monaghan P., Metcalfe N.B. Telomere elongation during early development is independent of environmental temperatures in Atlantic salmon. *J. Exp. Biol.* 2018;221(Pt. 11):jeb178616. DOI 10.1242/jeb.178616
- Meesters M., Van Eetvelde M., Martens D.S., Nawrot T.S., Dewulf M., Govaere J., Opsomer G. Prenatal environment impacts telomere length in newborn dairy heifers. *Sci. Rep.* 2023;13(1):4672. DOI 10.1038/s41598-023-31943-8
- Miyashita N., Shiga K., Yonai M., Kaneyama K., Kobayashi S., Kojima T., Goto Y., Kishi M., Aso H., Suzuki T., Sakaguchi M., Nagai T. Remarkable differences in telomere lengths among cloned cattle derived from different cell types. *Biol. Reprod.* 2002;66(6):1649-1655. DOI 10.1095/biolreprod66.6.1649
- Mizutani Y., Niizuma Y., Yoda K. How do growth and sibling competition affect telomere dynamics in the first month of life of long-lived seabird? *PLoS One*. 2016;11(11):e0167261. DOI 10.1371/journal.pone.0167261
- Monaghan P., Ozanne S.E. Somatic growth and telomere dynamics in vertebrates: relationships, mechanisms and consequences. *Philos. Trans. R. Soc. London Ser. B. Biol. Sci.* 2018;373(1741):20160446. DOI 10.1098/rstb.2016.0446
- Nowack J., Tarmann I., Hoelzl F., Smith S., Giroud S., Ruf T. Always a price to pay: hibernation at low temperatures comes with a trade-off between energy savings and telomere damage. *Biol. Lett.* 2019;15(10):20190466. DOI 10.1098/rsbl.2019.0466
- O'Daniel S.E., Kochan K.J., Long C.R., Riley D.G., Randel R.D., Welsh T.H.J. Comparison of telomere length in age-matched primiparous and multiparous Brahman cows. *Animals (Basel)*. 2023; 13(14):2325. DOI 10.3390/ani13142325
- Pinese M., Lacaze P., Rath E.M., Stone A., Brion M.-J., Ameur A., Nagpal S., ... Kaplan W., Gibson G., Gyllensten U., Cairns M.J., McNamara M., Dinger M.E., Thomas D.M. The Medical Genome Reference Bank contains whole genome and phenotype data of 2570 healthy elderly. *Nat. Commun.* 2020;11(1):435. DOI 10.1038/s41467-019-14079-0
- Purcell S., Neale B., Todd-Brown K., Thomas L., Ferreira M.A.R., Bender D., Maller J., Sklar P., de Bakker P.I.W., Daly M.J., Sham P.C. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 2007;81(3):559-575. DOI 10.1086/519795
- Rafie N., Golpour Hamedani S., Barak F., Safavi S.M., Miraghajani M. Dietary patterns, food groups and telomere length: a systematic review of current studies. *Eur. J. Clin. Nutr.* 2017;71(2):151-158. DOI 10.1038/ejcn.2016.149
- Raj A., Stephens M., Pritchard J.K. fastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics*. 2014;197(2):573-589. DOI 10.1534/genetics.114.164350

- Reichert S., Stier A., Zahn S., Arrivé M., Bize P., Massemin S., Criscuolo F. Increased brood size leads to persistent eroded telomeres. *Front. Ecol. Evol.* 2014;2:9. DOI 10.3389/fevo.2014.00009
- Ribas-Maynou J., Llavanera M., Mateo-Otero Y., Ruiz N., Muiño R., Bonet S., Yeste M. Telomere length in bovine sperm is related to the production of reactive oxygen species, but not to reproductive performance. *Theriogenology.* 2022;189:290-300. DOI 10.1016/j.theriogenology.2022.06.025
- Rossiello F., Jurk D., Passos J.F., d'Adda di Fagnagna F. Telomere dysfunction in ageing and age-related diseases. *Nat. Cell Biol.* 2022; 24(2):135-147. DOI 10.1038/s41556-022-00842-x
- Schrumpfová P.P., Fajkus J. Composition and function of telomerase-A polymerase associated with the origin of eukaryotes. *Biomolecules.* 2020;10(10):1425. DOI 10.3390/biom10101425
- Seeker L.A., Ilkska J.J., Psifidi A., Wilbourn R.V., Underwood S.L., Fairlie J., Holland R., Froy H., Salvo-Chirside E., Bagnall A., Whitelaw B., Coffey M.P., Nussey D.H., Banos G. Bovine telomere dynamics and the association between telomere length and productive lifespan. *Sci. Rep.* 2018a;8(1):12748. DOI 10.1038/s41598-018-31185-z
- Seeker L.A., Ilkska J.J., Psifidi A., Wilbourn R.V., Underwood S.L., Fairlie J., Holland R., Froy H., Bagnall A., Whitelaw B., Coffey M., Nussey D.H., Banos G. Longitudinal changes in telomere length and associated genetic parameters in dairy cattle analysed using random regression models. *PLoS One.* 2018b;13(2):e0192864. DOI 10.1371/journal.pone.0192864
- Seeker L.A., Underwood S.L., Wilbourn R.V., Dorrens J., Froy H., Holland R., Ilkska J.J., Psifidi A., Bagnall A., Whitelaw B., Coffey M., Banos G., Nussey D.H. Telomere attrition rates are associated with weather conditions and predict productive lifespan in dairy cattle. *Sci. Rep.* 2021;11(1):5589. DOI 10.1038/s41598-021-84984-2
- Sleptsov I.I., Machakhtyrova V.A., Ivanova N.P. Clinical and physiological indicators of the Kalmyk cattle breed in Yakutia conditions. *Bull. Kurgan State Agric. Acad.* 2019;4(32):44-46 (in Russian)
- Szczotka M., Kocki J., Iwan E., Pluta A. Determination of telomere length and telomerase activity in cattle infected with bovine leukaemia virus (BLV). *Pol. J. Vet. Sci.* 2019;22(2):391-403. DOI 10.24425/pjvs.2019.129299
- Taub M.A., Conomos M.P., Keener R., Iyer K.R., Weinstock J.S., Yanek L.R., Lane J., ... de Andrade M., Correa A., Chen Y.I., Boerwinkle E., Barnes K.C., Ashley-Koch A.E., Arnett D.K.; NHLBI Trans-Omics for Precision Medicine (TOPMed) Consortium; TOPMed Hematology and Hemostasis Working Group; TOPMed Structural Variation Working Group; Laurie C.C., Abecasis G., Nickerson D.A., Wilson J.G., Rich S.S., Levy D., Ruczinski I., Aviv A., Blackwell T.W., Thornton T., O'Connell J., Cox N.J., Perry J.A., Armanios M., Battle A., Pankratz N., Reiner A.P., Mathias R.A. Genetic determinants of telomere length from 109,122 ancestrally diverse whole-genome sequences in TOPMed. *Cell Genom.* 2022;2(1):100084. DOI 10.1016/j.xgen.2021.100084
- Tilesi F., Di Domenico E.G., Pariset L., Bosco L., Willems D., Valentini A., Ascenzioni F. Telomere length diversity in cattle breeds. *Diversity.* 2010;2(9):1118-1129. DOI 10.3390/d2091118
- Waalén J., Buxbaum J.N. Is older colder or colder older? The association of age with body temperature in 18,630 individuals. *J. Gerontol. A Biol. Sci. Med. Sci.* 2011;66(5):487-492. DOI 10.1093/geronol/glr001
- Whittemore K., Vera E., Martínez-Nevaldo E., Sanpera C., Blasco M.A. Telomere shortening rate predicts species life span. *Proc. Natl. Acad. Sci. USA.* 2019;116(30):15122-15127. DOI 10.1073/pnas.1902452116
- Wilbourn R.V., Moatt J.P., Froy H., Walling C.A., Nussey D.H., Boonekamp J.J. The relationship between telomere length and mortality risk in non-model vertebrate systems: a meta-analysis. *Philos. Trans. R. Soc. London Ser. B. Biol. Sci.* 2018;373(1741):20160447. DOI 10.1098/rstb.2016.0447
- Zepner L., Karrasch P., Wiemann F., Bernard L. ClimateCharts.net – an interactive climate analysis web platform. *Int. J. Digit. Earth.* 2021; 14(3):338-356. DOI 10.1080/17538947.2020.1829112
- Zhang D., Newton C.A., Wang B., Povysil G., Noth I., Martinez F.J., Raghu G., Goldstein D., Garcia C.K. Utility of whole genome sequencing in assessing risk and clinically relevant outcomes for pulmonary fibrosis. *Eur. Respir. J.* 2022;60(6):2200577. DOI 10.1183/13993003.00577-2022
- Zhang H., Liu A., Wang Y., Luo H., Yan X., Guo X., Li X., Liu L., Su G. Genetic parameters and genome-wide association studies of eight longevity traits representing either full or partial lifespan in Chinese Holsteins. *Front. Genet.* 2021;12:634986. DOI 10.3389/fgene.2021.634986
- Zhang X., Lin S., Funk W.E., Hou L. Environmental and occupational exposure to chemicals and telomere length in human studies. *Occup. Environ. Med.* 2013;70(10):743-749. DOI 10.1136/oemed-2012-101350
- Zhang Y., Wu Y., Mao P., Li F., Han X., Zhang Y., Jiang S., Chen Y., Huang J., Liu D., Zhao Y., Ma W., Songyang Z. Cold-inducible RNA-binding protein CIRP/hnRNP A18 regulates telomerase activity in a temperature-dependent manner. *Nucleic Acids Res.* 2016;44(2): 761-775. DOI 10.1093/nar/gkv1465
- Zhao Z., Cao J., Niu C., Bao M., Xu J., Huo D., Liao S., Liu W., Speakman J.R. Body temperature is a more important modulator of lifespan than metabolic rate in two small mammals. *Nat. Metab.* 2022; 4(3):320-326. DOI 10.1038/s42255-022-00545-5

Funding. The work was supported by the Russian Scientific Foundation (RSF) grant No. 22-26-00143 (<https://rscf.ru/project/22-26-00143/>).

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

Received December 3, 2023. Revised December 25, 2023. Accepted December 25, 2023.

DOI 10.18699/vjgb-24-24

Aberrant methylation of placental development genes in chorionic villi of spontaneous abortions with trisomy 16

O.Yu. Vasilyeva , E.N. Tolmacheva , A.E. Dmitriev , Ya.A. Darkova , E.A. Sazhenova ,
T.V. Nikitina , I.N. Lebedev , S.A. Vasilyev 

¹ Research Institute of Medical Genetics of the Tomsk National Research Medical Center of the Russian Academy of Sciences, Tomsk, Russia

² National Research Tomsk State University, Tomsk, Russia

 oksana.vasilyeva@medgenetics.ru

Abstract. In humans, aneuploidy is incompatible with the birth of healthy children and mainly leads to the death of embryos in the early stages of development in the first trimester of pregnancy. Trisomy 16 is the most common aneuploidy among spontaneous abortions of the first trimester of pregnancy. However, the mechanisms leading to the death of embryos with trisomy 16 remain insufficiently investigated. One of these potential mechanisms is abnormal placental development, including aberrant remodeling of spiral arteries. Spiral artery remodeling involves the migration of trophoblast cells into the maternal spiral arteries, replacing their endothelium and remodeling to ensure a stable embryonic nutrition and oxygen supply. This is a complex process which depends on many factors from both the embryo and the mother. We analyzed the methylation level of seven genes (*ADORA2B*, *NPR3*, *PRDM1*, *PSG2*, *PHTLH*, *SV2C*, and *TICAM2*) involved in placental development in the chorionic villi of spontaneous abortions with trisomy 16 ($n = 14$), compared with spontaneous abortions with a normal karyotype ($n = 31$) and the control group of induced abortions ($n = 10$). To obtain sequencing libraries, targeted amplification of individual gene regions using designed oligonucleotide primers for bisulfite-converted DNA was used. The analysis was carried out using targeted bisulfite massive parallel sequencing. In the group of spontaneous abortions with trisomy 16, the level of methylation of the *PRDM1* and *PSG2* genes was significantly increased compared to induced abortions ($p = 0.0004$ and $p = 0.0015$, respectively). In the group of spontaneous abortions, there was no increase in the level of methylation of the *PRDM1* and *PSG2* genes, but the level of methylation of the *ADORA2B* gene was significantly increased compared to the induced abortions ($p = 0.032$). The results obtained indicate the potential mechanisms of the pathogenetic effect of trisomy 16 on the placental development with the participation of the studied genes.

Key words: aneuploidy; trisomy 16; DNA methylation; chorionic villi; miscarriage; bisulfite sequencing; spontaneous abortions.

For citation: Vasilyeva O.Yu., Tolmacheva E.N., Dmitriev A.E., Darkova Ya.A., Sazhenova E.A., Nikitina T.V., Lebedev I.N., Vasilyev S.A. Aberrant methylation of placental development genes in chorionic villi of spontaneous abortions with trisomy 16. *Vavilovskii Zhurnal Genetiki i Seleksii* = *Vavilov Journal of Genetics and Breeding*. 2024;28(2):198-203. DOI 10.18699/vjgb-24-24

Аберрантное метилирование генов развития плаценты в ворсинах хориона спонтанных абортусов с трисомией 16

О.Ю. Васильева , Е.Н. Толмачева , А.Э. Дмитриев , Я.А. Даркова , Е.А. Саженова ,
Т.В. Никитина , И.Н. Лебедев , С.А. Васильев 

¹ Научно-исследовательский институт медицинской генетики Томского национального исследовательского медицинского центра Российской академии наук, Томск, Россия

² Национальный исследовательский Томский государственный университет, Томск, Россия

 oksana.vasilyeva@medgenetics.ru

Аннотация. У человека анеуплоидия не совместима с рождением здоровых детей и в основном приводит к гибели эмбрионов на ранних стадиях развития в первом триместре беременности. Наиболее частая анеуплоидия среди спонтанных абортусов первого триместра беременности – трисомия 16. Однако механизмы, приводящие к гибели эмбрионов с трисомией 16, остаются недостаточно исследованными. Одним из таких потенциальных механизмов является нарушение развития плаценты, в том числе ремоделирования спиральных артерий. Ремоделирование спиральных артерий заключается в миграции клеток трофобласта в спиральные артерии матери и замещении их эндотелия для обеспечения стабильного питания эмбриона и снабжения кислородом. Это комплексный процесс, зависящий от множества факторов как со стороны эм-

бриона, так и со стороны матери. Нами проведен анализ уровня метилирования семи генов (*ADORA2B*, *NPR3*, *PRDM1*, *PSG2*, *PTHLH*, *SV2C* и *TICAM2*), участвующих в развитии плаценты, в ворсинах хориона спонтанных абортусов с трисомией 16 ($n = 14$), по сравнению со спонтанными абортусами с нормальными кариотипом ($n = 31$) и контрольной группой медицинских абортусов ($n = 10$). Для получения библиотек для секвенирования использована таргетная амплификация отдельных регионов генов с помощью разработанных олигонуклеотидных праймеров на бисульфит-конвертированной ДНК. Анализ проводили с применением таргетного бисульфитного массового параллельного секвенирования. В группе спонтанных абортусов с трисомией 16 был значимо повышен уровень метилирования генов *PRDM1* и *PSG2* по сравнению с медицинскими абортусами ($p = 0.0004$ и $p = 0.0015$ соответственно). В группе спонтанных абортусов с нормальным кариотипом не обнаружено повышения уровня метилирования генов *PRDM1* и *PSG2*, но был значимо повышен уровень метилирования гена *ADORA2B* по сравнению с медицинскими абортусами ($p = 0.032$). Полученные результаты указывают на потенциальные механизмы патогенетического действия трисомии 16 на развитие плаценты с участием изученных генов.

Ключевые слова: анеуплоидия; трисомия 16; метилирование ДНК; ворсины хориона; невынашивание беременности; бисульфитное секвенирование; спонтанные абортусы.

Introduction

Spontaneous abortion (miscarriage) is the spontaneous death of an embryo or fetus before 20 weeks of gestation. In the vast majority of cases, pregnancy is terminated when the embryo has life-incompatible genetic abnormalities. Slightly more than 50 % of all clinically diagnosed miscarriages are caused by aneuploidy, which mainly occurs during spermatogenesis or oogenesis, or in the early stages of embryonic development, and the most common aneuploidy among spontaneous abortions of the first trimester is trisomy 16 (Nikitina et al., 2016). Most of the aneuploid embryos die at the implantation stage, and the next peak of embryonic mortality is observed around 8–9 weeks of pregnancy. However, the mechanisms leading to the death of embryos with aneuploidy remain poorly understood.

In the first trimester of pregnancy, the most important process occurs: remodeling of spiral arteries, which consists in the migration of trophoblast cells into the spiral arteries of the mother, replacement of their endothelium and remodeling to ensure stable nutrition of the embryo and oxygen supply (Red-Horse et al., 2004; Jauniaux et al., 2006). This is a complex process that depends on many factors from both the embryo and the mother. Our preliminary results show that spontaneous first trimester abortions with an aneuploid karyotype have large-scale methylation disorders of repetitive sequences (Vasilyev et al., 2021) and genes playing an important role in placental development (Tolmacheva et al., 2022).

In this work, we conducted a more detailed study of part of the genes, the methylation disorders of which were previously detected in spontaneous abortions with trisomy 16 (*PRDM1*, *PTHLH*) (Tolmacheva et al., 2022) and a normal karyotype (*ADORA2B*, *NPR3*, *PSG2*, *SV2C*, and *TICAM2*) (unpublished data).

The *ADORA2B* gene is associated with remodeling of spiral arteries, and its hypermethylation is associated with impaired placental development, fetal growth retardation and the development of preeclampsia (Jia et al., 2012; Yeung et al., 2016). The *NPR3* gene is a receptor for natriuretic peptide A, which plays an important role in the remodeling of the spiral arteries of the uterine wall (Zhang et al., 2021). Deficiency of natriuretic peptide A impairs trophoblast invasion and remodeling of spiral arteries, which leads to a phenotype similar to preeclampsia. The *PRDM1* gene is a key regulator of

terminal differentiation of giant trophoblast cells that replace the endothelium of the spiral arteries of the mother (Maioli et al., 2004). The *PSG2* gene encodes pregnancy-specific beta-1 glycoprotein 2, the expression of which is increased in the trophoblast, and its increased level is observed in circulating trophoblast cells with true placenta accreta (Grunblatt et al., 2004). The *PTHLH* gene encodes osteostatin (parathyroid hormone-related protein), which is a precursor to a signaling peptide that plays a role in the differentiation of giant mouse trophoblast cells (Sandor et al., 2017). For the *SV2C* and *TICAM2* genes, expression disorders are known in other pregnancy pathologies potentially associated with abnormal placentation (McMaster et al., 2004). The expression of the *SV2C* gene increases in exosomes in the blood of the mother with gestational diabetes compared with the group with normal pregnancy (Fang et al., 2021). Hypomethylation and high expression of the *TICAM2* gene are also associated with preeclampsia and premature birth (Mason et al., 2011; Lim et al., 2020).

The aim of this study was to analyze the aberrant methylation of the genes of placental development *ADORA2B*, *NPR3*, *PRDM1*, *PSG2*, *PTHLH*, *SV2C*, and *TICAM2* among spontaneous abortions of the first trimester of pregnancy with trisomy 16.

Materials and methods

The analysis was performed on chorionic villi of spontaneous abortions with trisomy 16 ($n = 14$, gestational age 8.7 ± 1.6 weeks), spontaneous abortions with a normal karyotype ($n = 31$, gestational age 10.0 ± 2.2 weeks) and induced abortions ($n = 10$, gestational age 8.3 ± 1.2 weeks). Samples from the bio-collection “Biobank of the population of North Eurasia” of the Research Institute of Medical Genetics of the Tomsk National Research Medical Center of the Russian Academy of Sciences were used. Tissue samples were stored at a temperature of -80 °C. Informed parental consent was obtained for all samples to use the biomaterial for biobanking and research. The study was approved by the Committee on Biomedical Ethics of the Research Institute of Medical Genetics of the Tomsk NRMC (09.11.2020/No. 7).

The karyotype was determined using conventional cytogenetic analysis on direct preparations of chorionic villi and in fibroblast cultures of extraembryonic mesoderm (Lebedev et

Sequences of oligonucleotide primers used to obtain libraries for targeted bisulfite massive parallel sequencing

Name		Sequence	Product, bp	Coordinates (hg38), gene region
ADORA2B_m1	F	5'-GATAGATATTTGGTTATTTGTGTTT-3'	519	chr17:15945552-15946070, exon 1-intron 1
	R	5'-TACCTTACCCTTAATAAAAACCTCC-3'		
ADORA2B_m2	F	5'-GTTATGTTGTTGGAGATATAGGA-3'	331	chr17:15945246-15945576, exon 1
	R	5'-AAACACAAATAACCAATATCTATC-3'		
NPR3	F	5'-TAGGGAGGAGTTTTGATGTAAGAAT-3'	465	chr5:32711142-32711606, exon 1
	R	5'-TCCTTCTACAATATCACTAATT-3'		
PRDM1	F	5'-AGTTTGTGTATTTGAAATTGTATAAG-3'	372	chr6:106107017-106107388, exon 7
	R	5'-ACTCATAAAAACCTACACCTAAAAA-3'		
PSG2	F	5'-GTTGTTGTGTGTAGAGGAGGAATAG-3'	344	chr19:43083681-43084024, promoter
	R	5'-ATCCCAAACCAACCTAAAACCTAAC-3'		
PTHLH	F	5'-GGGTTTTGATAGTAATTATTATTTT-3'	476	chr12:27969063-27969538, exon 4-intron 4
	R	5'-ATCCCATCCCTTCTATTACAATC-3'		
SV2C	F	5'-AGAGTAAGAGAGGGTTGGAGAGTAG-3'	393	chr5:76084117-76084509, intron 1
	R	5'-ATACTAATAACCCCAAACCAATTC-3'		
TICAM2	F	5'-ATTTGTTTATGGTTTTGAGGTGTTT-3'	671	chr5:115601810-115602480, exon 1-intron 1
	R	5'-CACATTAACCCCTAACTACAACCTAA-3'		

al., 2004). The presence of trisomy 16 was verified by fluorescent *in situ* hybridization (FISH) using subtelomeric DNA probes (16q and 16p) according to the described technique (Vasilyev et al., 2010).

Tissue separation was carried out morphologically, then chorionic villi cells were incubated overnight at 37 °C with proteinase K. The standard phenol-chloroform method was used to isolate DNA. Bisulfite DNA modification was performed using the EZ DNA Methylation-Direct kit (Zymo Research, USA) according to the manufacturer's protocol. During bisulfite conversion, unmethylated cytosine is modified into uracil, which is replaced by thymine during further PCR, and methylated cytosine is not modified.

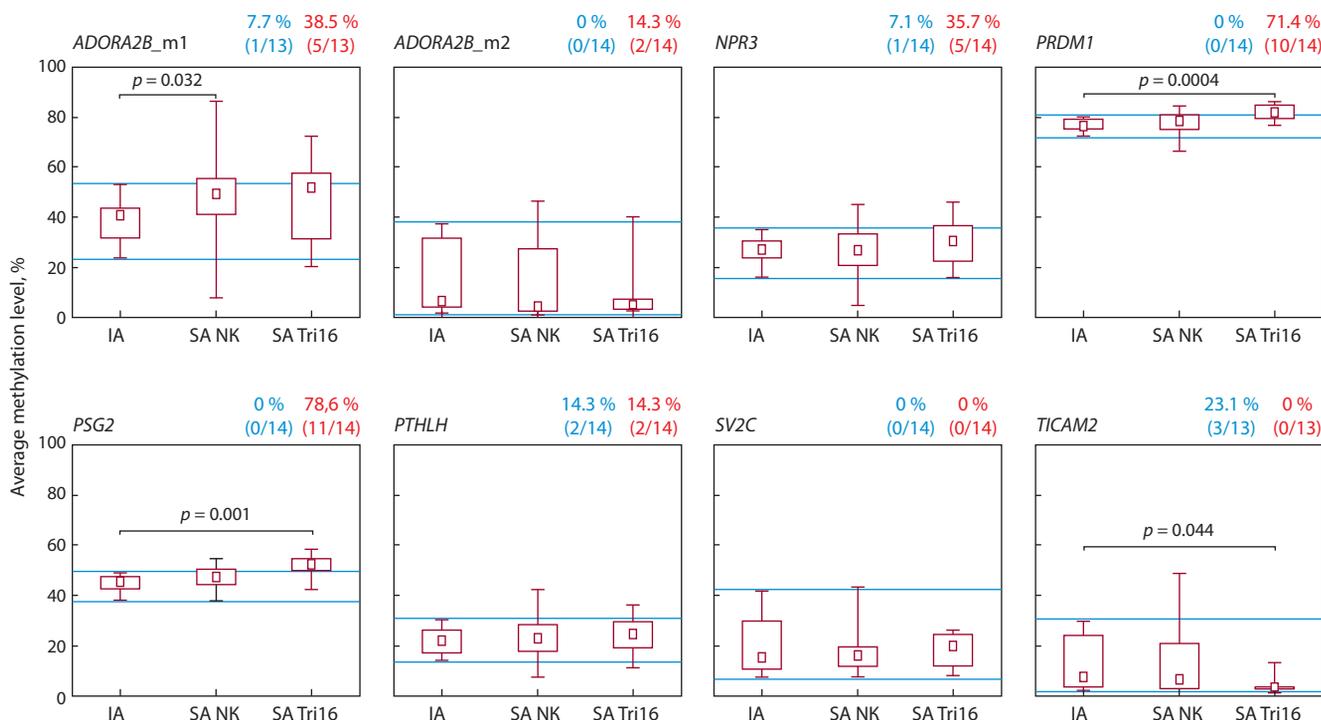
The methylation profile was analyzed using targeted bisulfite massive parallel sequencing. To obtain the libraries, the designed oligonucleotide primers were used to amplify the target regions of the *ADORA2B*, *NPR3*, *PRDM1*, *PSG2*, *PTHLH*, *SV2C*, and *TICAM2* genes from bisulfite-converted DNA (see the Table). The choice of genes and target sites in them was determined by the differentially methylated CpG sites in chorionic villi of spontaneous abortions according to our preliminary results obtained using a large-scale methylation analysis (Tolmacheva et al., 2022), and their participation in the development of the placenta.

Amplification of target fragments was carried out using a set of HS-Taq PCR mastermix (2×) (Biolabmix, Russia) according to the manufacturer's protocol with the following PCR

conditions: 95 °C for 5 min; 36 cycles: 95 °C for 30 s, 60 °C for 45 s, 72 °C for 45 s. The concentration of the target fragments was determined using a Qubit 4.0 fluorimeter (Thermo Fisher Scientific, USA). The reaction products were purified using Sephadex G50 solution (Sigma, USA).

Targeted bisulfite massive parallel sequencing was performed on a MiSeq device (Illumina, USA) using a Micro kit (2x150). The quality of the reads was evaluated using FastQC v0.11.8, after which the remaining adapter sequences and low-quality reads were trimmed using Trim-Galore. The reads were then mapped to bisulfite-converted target sequences using the bwa-meth tool (v0.2.2) with default parameters. Methylation data in the context of CpG were extracted from the resulting BAM files using the MethylDackel tool. The results were presented as the methylation level, which is the ratio of the number of cytosines to the total number of cytosines and thymines in a individual CpG site. In addition, the average methylation level was calculated along all target sites. The statistical analysis was performed using the Statistica 10.0 software package (StatSoft, USA). The Mann–Whitney rank test was used to compare the methylation level between groups of samples. The differences were considered statistically significant at $p < 0.05$.

The study was conducted using the equipment of the center for collective use “Medical Genomics” of the Tomsk National Research Medical Center of the Russian Academy of Sciences.



The average level of methylation of CpG sites in the target regions of the *ADORA2B*, *NPR3*, *PRDM1*, *PSG2*, *PTHLH*, *SV2C*, and *TICAM2* genes in chorionic villi of spontaneous abortions with trisomy 16 compared with spontaneous abortions with a normal karyotype and induced abortions.

IA – induced abortions; SA NK – spontaneous abortions with a normal karyotype; SA Tri16 – spontaneous abortions with trisomy 16. The boxes represent the 25th and 75th percentiles, and the whiskers mark the minimum and maximum values. The square in the center of the box indicates the median, with blue lines marking the minimum and maximum values in the IA group. Above the figures, the proportion and number of spontaneous abortions with trisomy 16 with methylation levels of target genes beyond the limits of variation in the group of induced abortions are indicated (lowered – blue, elevated – red).

Results

Significant differences in spontaneous abortions with trisomy 16 compared with induced abortions were observed in the following genes: *PRDM1* ($81.9 \pm 2.8\%$ vs. $76.5 \pm 2.6\%$, $p = 0.0004$), *PSG2* ($51.6 \pm 4.4\%$ vs. $44.6 \pm 3.6\%$, $p = 0.001$), and *TICAM2* ($4.5 \pm 3.6\%$ vs. $12.5 \pm 11.0\%$, $p = 0.044$) (see the Figure). At the same time, the level of methylation of the *PRDM1* and *PSG2* genes in the group of spontaneous abortions with trisomy 16 was higher, and that of the *TICAM2* gene was lower compared with induced abortions. In the group of spontaneous abortions with a normal karyotype, the level of methylation of the *ADORA2B* gene (m1 region) was significantly higher compared with the group of induced abortions ($48.8 \pm 15.3\%$ vs. $38.7 \pm 10.2\%$, $p = 0.032$) (see the Figure).

Some spontaneous abortions had methylation levels beyond the normal variability in the control group of induced abortions (see the Figure). The maximum number of spontaneous abortions with trisomy 16 with an increased level of methylation was found for the *PSG2* gene (11 samples out of 14, which is 78.6% of the total number of the studied samples) and for the *PRDM1* gene (10 samples out of 14, 71.4%) (see the Figure). Also, an increased level of methylation was observed in some spontaneous abortions with trisomy 16 for the following genes: *ADORA2B_m1* (38.5%), *ADORA2B_m2* (14.3%), *NPR3* (35.7%), and *PTHLH* (13.3%). Lowered methylation levels in some spontaneous abortions were recorded only for the *ADORA2B_m1* (7.7%); *NPR3* (7.1%); *PTHLH* (14.3%),

and *TICAM2* (23.1%) genes. No spontaneous abortions with impaired methylation levels were detected for the *SV2C* gene (see the Figure).

Discussion

Previously, our group and others showed methylation disorders in chorionic villi of spontaneous abortions with trisomy 16: increased methylation of retrotransposon LINE-1 (Vasilyev et al., 2021) and large-scale methylation disorders throughout the genome (Blair et al., 2014; Tolmacheva et al., 2022). In addition, methylation disorders in trisomy 16 were found to overlap with those in early-onset preeclampsia (Blair et al., 2014). Considering that one of the mechanisms of the development of preeclampsia is considered to be impaired placentation and remodeling of spiral arteries (McMaster et al., 2004), a possible mechanism leading to the death of embryos with trisomy 16 is abnormal methylation of placental development genes.

In this work, it was shown that in the studied group of spontaneous abortions with trisomy 16, the level of methylation of the *PRDM1* and *PSG2* genes was significantly increased compared with induced abortions. At the same time, the level of methylation of these genes did not significantly increase in the group of spontaneous abortions with a normal karyotype. Since the *PRDM1* gene is a transcription factor, its effect is observed in many processes in the body. Recently, it was found that hypomethylation of the *PRDM1* gene and a cor-

responding increase in gene expression in chorionic villi is associated with recurrent miscarriage (Du et al., 2020). Potentially, the negative effect of hypomethylation of *PRDM1* may be associated with abnormal trophoblast migration and an increased level of trophoblast cell apoptosis (Du et al., 2020), as well as with the role of *PRDM1* in regulating the transcription factor *GATA2*, which is key for trophoblast development (Paul et al., 2017).

It is possible that the increased methylation of the *PRDM1* and *PSG2* genes in the group of spontaneous abortions with trisomy 16 is associated with the effect of a supernumerary chromosome on the DNA methylation profile (Tolmacheva et al., 2013). This process is probably triggered by specific genes located on chromosome 16, which may be involved in the regulation of DNA methylation (Tolmacheva et al., 2022). In the context of this work, it is interesting that *CTCF*, one of the key regulators of chromatin conformation located on chromosome 16, can regulate the transcription of human *PSG* genes (*PSG1–PSG9*, *PSG11*) in trophoblast cells. Suppression of *CTGF* expression increased or suppressed the expression of several *PSG* genes, and this effect was accompanied by epigenetic changes (Jeong et al., 2021).

An interesting result requiring further study is a significant increase in the level of methylation of the *ADORA2B* gene in spontaneous abortions with a normal karyotype. It is likely that in individual embryos with a normal karyotype, methylation disorders of some genes involved in the development of the placenta may be caused by other causes unrelated to the influence of aneuploidy.

Conclusion

The results indicate that the aberrant level of methylation of placental development genes may be an important factor associated with the death of embryos with trisomy 16 in the first trimester of pregnancy.

References

- Blair J.D., Langlois S., Mcfadden D.E., Robinson W.P. Overlapping DNA methylation profile between placentas with trisomy 16 and early-onset preeclampsia. *Placenta*. 2014;35(3):216-222. DOI 10.1016/j.placenta.2014.01.001
- Du G., Yu M., Xu Q., Huang Z., Huang X., Han L., Fan Y., Zhang Y., Wang R., Xu S., Han X., Fu G., Lv S., Qin Y., Wang X., Lu C., Xia Y. Hypomethylation of *PRDM1* is associated with recurrent pregnancy loss. *J. Cell. Mol. Med.* 2020;24(12):7072-7077. DOI 10.1111/jcmm.15335
- Fang Y., Wan C., Wen Y., Wu Z., Pan J., Zhong M., Zhong N. Autism-associated synaptic vesicle transcripts are differentially expressed in maternal plasma exosomes of physiopathologic pregnancies. *J. Transl. Med.* 2021;19(1):154. DOI 10.1186/s12967-021-02821-6
- Grunblatt E., Mandel S., Jacob-Hirsch J., Zeligson S., Amariglio N., Rechavi G., Li J., Ravid R., Roggendorf W., Riederer P., Youdim M.B. Gene expression profiling of parkinsonian substantia nigra pars compacta; alterations in ubiquitin-proteasome, heat shock protein, iron and oxidative stress regulated proteins, cell adhesion/cellular matrix and vesicle trafficking genes. *J. Neural Transm. (Vienna)*. 2004;111(12):1543-1573. DOI 10.1007/s00702-004-0212-1
- Jauniaux E., Poston L., Burton G.J. Placental-related diseases of pregnancy: Involvement of oxidative stress and implications in human evolution. *Hum. Reprod. Update*. 2006;12(6):747-755. DOI 10.1093/humupd/dml016
- Jeong D.S., Kim M.H., Lee J.Y. Depletion of *CTCF* disrupts *PSG* gene expression in the human trophoblast cell line Swan 71. *FEBS Open Bio*. 2021;11(3):804-812. DOI 10.1002/2211-5463.13087
- Jia R.Z., Zhang X., Hu P., Liu X.M., Hua X.D., Wang X., Ding H.J. Screening for differential methylation status in human placenta in preeclampsia using a CpG island plus promoter microarray. *Int. J. Mol. Med.* 2012;30(1):133-141. DOI 10.3892/ijmm.2012.983
- Lebedev I.N., Ostroverkhova N.V., Nikitina T.V., Sukhanova N.N., Nazarenko S.A. Features of chromosomal abnormalities in spontaneous abortion cell culture failures detected by interphase FISH analysis. *Eur. J. Hum. Genet.* 2004;12(7):513-520. DOI 10.1038/sj.ejhg.5201178
- Lim J.H., Kang Y.J., Bak H.J., Kim M.S., Lee H.J., Kwak D.W., Han Y.J., Kim M.Y., Boo H., Kim S.Y., Ryu H.M. Epigenome-wide DNA methylation profiling of preeclamptic placenta according to severe features. *Clin. Epigenetics*. 2020;12(1):128. DOI 10.1186/s13148-020-00918-1
- Maioli E., Fortino V., Pacini A. Parathyroid hormone-related protein in preeclampsia: a linkage between maternal and fetal failures. *Biol. Reprod.* 2004;71(6):1779-1784. DOI 10.1095/biolreprod.104.030932
- Mason C.W., Buhimschi I.A., Buhimschi C.S., Dong Y., Weiner C.P., Swaan P.W. ATP-binding cassette transporter expression in human placenta as a function of pregnancy condition. *Drug Metab. Dispos.* 2011;39(6):1000-1007. DOI 10.1124/dmd.111.038166
- McMaster M.T., Zhou Y., Fisher S.J. Abnormal placentation and the syndrome of preeclampsia. *Semin. Nephrol.* 2004;24(6):540-547. DOI 10.1016/s0270-9295(04)00124-x
- Nikitina T.V., Sazhenova E.A., Tolmacheva E.N., Sukhanova N.N., Kashevarova A.A., Skryabin N.A., Vasilyev S.A., Nemtseva T.N., Yuriev S.Y., Lebedev I.N. Comparative cytogenetic analysis of spontaneous abortions in recurrent and sporadic pregnancy losses. *Biomed. Hub*. 2016;1(1):1-11. DOI 10.1159/000446099
- Paul S., Home P., Bhattacharya B., Ray S. *GATA* factors: Master regulators of gene expression in trophoblast progenitors. *Placenta*. 2017;60(Suppl. 1):S61-S66. DOI 10.1016/j.placenta.2017.05.005
- Red-Horse K., Zhou Y., Genbacev O., Prakobphol A., Foulk R., McMaster M., Fisher S.J. Trophoblast differentiation during embryo implantation and formation of the maternal-fetal interface. *J. Clin. Invest.* 2004;114(6):744-754. DOI 10.1172/JCI22991
- Sandor C., Robertson P., Lang C., Heger A., Booth H., Vowles J., Witty L., Bowden R., Hu M., Cowley S.A., Wade-Martins R., Webber C. Transcriptomic profiling of purified patient-derived dopamine neurons identifies convergent perturbations and therapeutics for Parkinson's disease. *Hum. Mol. Genet.* 2017;26(3):552-566. DOI 10.1093/hmg/ddw412
- Tolmacheva E.N., Kashevarova A.A., Skryabin N.A., Lebedev I.N. Epigenetic effects of trisomy 16 in human placenta. *Mol. Biol.* 2013;47(3):373-381. DOI 10.1134/s0026893313030175
- Tolmacheva E.N., Vasilyev S.A., Nikitina T.V., Lytkina E.S., Sazhenova E.A., Zhigalina D.I., Vasilyeva O.Y., Markov A.V., Demeneva V.V., Tashireva L.A., Kashevarova A.A., Lebedev I.N. Identification of differentially methylated genes in first-trimester placentas with trisomy 16. *Sci. Rep.* 2022;12(1):1166. DOI 10.1038/s41598-021-04107-9
- Vasilyev S.A., Timoshevsky V.A., Lebedev I.N. Cytogenetic mechanisms of aneuploidy in somatic cells of chemonuclear indus-

- try professionals with incorporated plutonium-239. *Russ. J. Genet.* 2010;46(11):1381-1385. DOI 10.1134/s1022795410110141
- Vasilyev S.A., Tolmacheva E.N., Vasilyeva O.Y., Markov A.V., Zhigalina D.I., Zatula L.A., Lee V.A., Serdyukova E.S., Sazhenova E.A., Nikitina T.V., Kashevarova A.A., Lebedev I.N. LINE-1 retrotransposon methylation in chorionic villi of first trimester miscarriages with aneuploidy. *J. Assist. Reprod. Genet.* 2021;38(1):139-149. DOI 10.1007/s10815-020-02003-1
- Yeung K.R., Chiu C.L., Pidsley R., Makris A., Hennessy A., Lind J.M. DNA methylation profiles in preeclampsia and healthy control placentas. *Am. J. Physiol. Heart Circ. Physiol.* 2016; 310(10):H1295-H1303. DOI 10.1152/ajpheart.00958.2015
- Zhang W., Li S., Lou J., Li H., Liu M., Dong N., Wu Q. Atrial natriuretic peptide promotes uterine decidualization and a TRAIL-dependent mechanism in spiral artery remodeling. *J. Clin. Invest.* 2021; 131(20):e151053. DOI 10.1172/JCI151053

Funding. The study was supported by the grant of the Russian Science Foundation No. 23-15-00341.

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

Received November 17, 2023. Revised January 18, 2024. Accepted January 18, 2024.

DOI 10.18699/vjgb-24-25

Composition of the sputum bacterial microbiome of patients with different pathomorphological forms of non-small-cell lung cancer

V.G. Druzhinin ^{1,2} , E.D. Baranova ¹, P.S. Demenkov ³, L.V. Matskova ⁴, A.V. Larionov ¹

¹ Kemerovo State University, Kemerovo, Russia

² Kemerovo State Medical University, Kemerovo, Russia

³ Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia

⁴ Karolinska Institute, Stockholm, Sweden

 druzhinin_vladim@mail.ru

Abstract. Recent studies have shown that the bacterial microbiome of the respiratory tract influences the development of lung cancer. Changes in the composition of the microbiome are observed in patients with chronic inflammatory processes. Such microbiome changes may include the occurrence of bacteria that cause oxidative stress and that are capable of causing genome damage in the cells of the host organism directly and indirectly. To date, the composition of the respiratory microbiome in patients with various histological variants of lung cancer has not been studied. In the present study, we determined the taxonomic composition of the sputum microbiome of 52 patients with squamous cell carcinoma of the lung, 52 patients with lung adenocarcinoma and 52 healthy control donors, using next-generation sequencing (NGS) on the V3-V4 region of the bacterial gene encoding 16S rRNA. The sputum microbiomes of patients with different histological types of lung cancer and controls did not show significant differences in terms of the species richness index (Shannon); however, the patients differed from the controls in terms of evenness index (Pielou). The structures of bacterial communities (beta diversity) in the adenocarcinoma and squamous cell carcinoma groups were also similar; however, when analyzed according to the matrix constructed by the Bray–Curtis method, there were differences between patients with squamous cell carcinoma and healthy subjects, but not between those with adenocarcinoma and controls. Using the LEFSe method it was possible to identify an increase in the content of *Bacillota* (*Streptococcus* and *Bacillus*) and *Actinomycetota* (*Rothia*) in the sputum of patients with squamous cell carcinoma when compared with samples from patients with adenocarcinoma. There were no differences in the content of bacteria between the samples of patients with adenocarcinoma and the control ones. The content of representatives of the genera *Streptococcus*, *Bacillus*, *Peptostreptococcus* (phylum *Bacillota*), *Prevotella*, *Mucellibacteroides* (phylum *Bacteroidota*), *Rothia* (phylum *Actinomycetota*) and *Actinobacillus* (phylum *Pseudomonadota*) was increased in the microbiome of sputum samples from patients with squamous cell carcinoma, compared with the control. Thus, the sputum bacterial microbiome of patients with different histological types of non-small-cell lung cancer has significant differences. Further research should be devoted to the search for microbiome biomarkers of lung cancer at the level of bacterial species using whole-genome sequencing.

Key words: non-small cell lung cancer; squamous cell lung cancer; lung adenocarcinoma; bacterial microbiome; sputum; taxonomic composition; 16S rRNA; NGS sequencing.

For citation: Druzhinin V.G., Baranova E.D., Demenkov P.S., Matskova L.V., Larionov A.V. Composition of the sputum bacterial microbiome of patients with different pathomorphological forms of non-small-cell lung cancer. *Vavilovskii Zhurnal Genetiki i Seleksii* = *Vavilov Journal of Genetics and Breeding*. 2024;28(2):204-214. DOI 10.18699/vjgb-24-25

Состав бактериального микробиома мокроты пациентов с разными патоморфологическими формами немелкоклеточного рака легкого

В.Г. Дружинин ^{1,2} , Е.Д. Баранова ¹, П.С. Деменков ³, Л.В. Мацкова ⁴, А.В. Ларионов ¹

¹ Кемеровский государственный университет, Кемерово, Россия

² Кемеровский государственный медицинский университет, Кемерово, Россия

³ Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия

⁴ Каролинский институт, Стокгольм, Швеция

 druzhinin_vladim@mail.ru

Аннотация. Исследования последних лет показали, что бактериальный микробиом респираторного тракта влияет на развитие рака легкого. Изменение состава микробиома у пациентов связывают с хроническими воспалительными процессами, так как многие бактерии вызывают окислительный стресс, а также способны прямо

или опосредованно повреждать геном в клетках организма хозяина. До настоящего времени состав респираторного микробиома у больных с различными гистологическими вариантами рака легкого не изучен. В настоящем исследовании для анализа таксономического состава микробиома мокроты 52 пациентов с плоскоклеточным раком легкого, 52 пациентов с аденокарциномой легкого и 52 здоровых доноров контрольной группы использовали технологию массового параллельного секвенирования региона V3-V4 16S рPHK. Микробиомы мокроты больных с разными гистологическими типами рака легкого и контроля не имели значимых различий по индексу видового богатства (Шеннона), однако у пациентов они отличались от контроля по индексу выравненности (Пиелу). Структуры бактериальных сообществ (бета-разнообразие) между аденокарциномой и плоскоклеточным раком также были близкими. Тем не менее матрица, построенная по Брэю–Кёртису, позволила выявить различия между пациентами с плоскоклеточным раком и здоровыми субъектами, но не между аденокарциномой и контролем. Метод LEFse позволил идентифицировать в мокроте больных плоскоклеточным раком увеличение содержания *Bacillota* (*Streptococcus* и *Bacillus*) и *Actinomycetota* (*Rothia*) при сопоставлении с образцами пациентов с аденокарциномой. Не найдено различий в содержании бактерий между образцами больных аденокарциномой и контроля. В микробиоме образцов мокроты пациентов с плоскоклеточным раком по сравнению с контролем было повышено содержание представителей родов *Streptococcus*, *Bacillus*, *Peptostreptococcus* (филум *Bacillota*), *Prevotella*, *Macellibacteroides* (филум *Bacteroidota*), *Rothia* (филум *Actinomycetota*) и *Actinobacillus* (филум *Pseudomonadota*). Таким образом, бактериальный микробиом мокроты пациентов с разными гистологическими типами немелкоклеточного рака легкого имеет существенные различия. Дальнейшие исследования должны быть посвящены поиску микробиомных биомаркеров рака легкого на уровне бактериальных видов с использованием полногеномного секвенирования.

Ключевые слова: немелкоклеточный рак легкого; плоскоклеточный рак легкого; аденокарцинома легкого; бактериальный микробиом; мокрота; таксономический состав; 16S рPHK; NGS секвенирование.

Introduction

Recent studies show that many bacteria living in the human body are related to the development of malignant tumors. Microbial ecosystems capable of initiating oncogenic transformation, inducing metabolic changes in the tumor microenvironment, or modulating responses to cancer immunotherapy have already been described (Xavier et al., 2020; Chen et al., 2022). Integrated metagenomic approaches are expected to accurately identify tumor-associated microbiome profiles and uncover mechanisms of bacterial influence on cancer initiation and progression (Chiu, Miller, 2019). Moreover, recent studies have identified microbial profiles specific to certain cancer types that may serve as biomarkers for diagnosing tumor risk (Wu et al., 2021).

Lung cancer (LC) originates in the lung parenchyma or bronchi and is diagnosed in approximately 1.2 million people worldwide each year (Cheng T.Y. et al., 2016). Mortality from LC remains high, in part due to the lack of early detection of diagnostic biomarkers, including metagenomic markers. Therefore, the search for bacteria associated with the risk of developing LC has intensified dramatically in recent years, especially with the application of massively parallel DNA sequencing technology (Mao et al., 2018; Maddi et al., 2019). Previous studies have shown that there are features of microbiota composition in saliva, bronchoalveolar lavage, and lung tissue samples that may be associated with LC, but the results of these studies regarding the significance of specific bacteria are largely contradictory (Hasegawa et al., 2014; Lee et al., 2016; Liu H.X. et al., 2018; Tsay et al., 2018; Peters et al., 2019; Wang et al., 2019; Zhang et al., 2019; Cheng C. et al., 2020; Zhuo et al., 2020).

An important source of information on the composition of the respiratory tract microbiota is sputum, which has so far been little studied in LC patients (Hosgood et al., 2014, 2019; Cameron et al., 2017; Druzhinin et al., 2020; Ran et al., 2020). Although sputum does not reflect the microbiome of any specific part of the respiratory tract, it may be useful for

searching for metagenomic biomarkers of LC because its collection is relatively simple and non-invasive.

Despite the fact that all forms of LC originate from epithelial cells of the airway mucosa, the current classification includes several different histologic types of this disease (Tsao, Yoon, 2018). LC is commonly divided into small cell lung cancer and non-small cell lung cancer (NSCLC), which accounts for 85 % of all LC cases (Molina et al., 2008). NSCLC is in turn subdivided into large cell lung cancer, adenocarcinoma of the lung (AD), and squamous cell lung cancer (LUSC). Different histological types of LC are characterized by distinctive biological patterns, different molecular markers and specific treatment strategies (Herbst et al., 2008). Based on this, it can be hypothesized that the composition of the respiratory tract microbiome may also differ between AD and LUSC patients. To date, this question remains open, given the very few published studies comparing the respiratory microbiome with individual histologic types of LC.

Here, we present the results of a comparative study of the taxonomic composition of the bacterial microbiome of the sputum of AD, LUSC patients and healthy donors, residents of the Kuzbass region of Western Siberia, for the first time.

Material and methods

Microbiota composition was studied in sputum samples from 52 patients with AD (37 men, 15 women; mean age 62.5 years); 52 patients with LUSC (49 men, 3 women; mean age 59.9 years) and 52 healthy donors (39 men, 13 women; mean age 62.5 years). The cohort of patients with NSCLC was formed from individuals who were first admitted for examination to the Kemerovo Regional Oncology Center (Kemerovo, Russian Federation). The material for the study was collected from March 2018 to March 2022. A questionnaire was filled out for each participant with information on place and date of birth, living environment, occupation, exposure to occupational hazards, health status, medication intake, radiologic procedures, smoking and alcohol consumption. For patients

with NSCLC, the results of clinical and histological analyses, primary tumor localization, and disease stage according to the TNM classification were additionally taken into account (Goldstraw, 2013). Demographic and clinical data on patients and control donors are presented in Table 1.

Inclusion criteria were male and female age ≥ 40 years, sputum donation, and signing written informed consent. Exclusion criteria were any acute or chronic condition that would limit the patient's ability to participate in the study, use of antibiotics within 4 weeks prior to collection, inability to obtain a sputum sample, or refusal to give informed consent. All participants were informed about the aims, possible risks of the study and signed informed consent. The study was approved by the Biomedical Ethics Commission of Kemerovo State University (protocol # 17/2021 dated 05.04.2021). When patients and control donors were included in the study, ethical principles required by the World Medical Association Declaration of Helsinki (World Medical Association Declaration of Helsinki, 1964, 2000) were followed.

To analyze the taxonomic composition of the respiratory microbiome, sputum samples (2–3 ml) from patients with

NSCLC and control group donors were obtained noninvasively through productive coughing. The obtained samples were immediately placed in sterile plastic vials and frozen ($-20\text{ }^{\circ}\text{C}$). Frozen samples were transported to the laboratory and stored at $-80\text{ }^{\circ}\text{C}$ until bacterial DNA extraction.

DNA extraction, amplification, and sequencing of 16S rRNA on a MiSeq instrument (Illumina, USA) were performed according to the manufacturer's recommendations. A detailed description of the procedures is given in a previous publication (Druzhinin et al., 2021).

Microbiome sequencing data were processed using the QIIME2 software package (Bolyen et al., 2019). Quality assurance was performed and a sequence library was created. Sequences were combined into operational taxonomic units (OTUs) based on a 99 % nucleotide similarity threshold using the Greengenes (version 13-8) and SILVA (version 138) reference sequence libraries, followed by removal of singletons (OTUs containing only one sequence). The correspondence of bacterial phylum names to current international nomenclature was determined using the LPSN resource (Parte et al., 2020).

Table 1. Characteristics of the study cohorts

Variables	NSCLC, <i>n</i> = 52	AD, <i>n</i> = 52	Control, <i>n</i> = 52
Age, years (mean)	59.9	62.5	62.5
Gender (<i>n</i> /%):			
Men	49/94.0	37/71.0	39/75.0
Women	3/6.0	15/29.0	13/25.0
Place of residence (<i>n</i> /%):			
City	35/67.0	40/77.0	46/88.0
Village	17/33.0	12/23.0	6/12.0
Occupational hazards (<i>n</i> /%):			
Yes	19/37.0	23/44.0	12/23.0
No	33/63.0	29/56.0	40/77.0
Smoking status (<i>n</i> /%):			
Yes	38/73.0	25/48.0	20/38.5
No	14/27.0	27/52.0	32/61.5
Alcohol consumption (<i>n</i> /%):			
Yes	34/65.0	35/67.0	39/75.0
No	18/35.0	17/33.0	13/25.0
Chronic diseases (<i>n</i> /%):			
Cardiovascular	29/56.0	40/77.0	20/38.5
Bronchitis	16/31.0	12/23.0	4/8.0
Chronic obstructive pulmonary disease	24/45.0	6/12.0	0
Gastrointestinal	7/15.0	7/15.0	11/21.0
Diabetes	1/2.0	3/6.0	4/8.0
Asthma	3/6.0	1/2.0	1/2.0
Obesity	4/8.0	15.0	1/2.0
TNM [#] (<i>n</i> /%):			
I, II	28/54.0	32/61.5	–
III, IV	24/46.0	20/38.5	
Tumor localization (<i>n</i> /%):			
Central	27/52.0	3/6.0	–
Peripheral	22/42.0	47/90.0	
Not established	3/6.0	2/4.0	

[#] TNM – tumor-node-metastasis.

The total diversity (alpha-diversity) of sputum prokaryotic communities was estimated by the number of isolated OTUs (analogous to species richness) and Shannon indices ($H = \sum p_i \ln p_i$, where p_i is the proportion of the i -th species in the community). The evenness of species distribution in terms of their abundance in the community was assessed by the Pielou index. The difference in the structure of bacterial communities of different samples (beta diversity) was analyzed using UniFrac (Lozupone, Knight, 2005), a method common in microbial ecology that assesses the difference between communities based on the phylogenetic relatedness of the represented taxa. Normalization of samples by 1070 sequences (minimum number of sequences obtained per sample) was used to calculate diversity indices. The significance of differences between groups of samples was assessed by the PERMANOVA method (Adonis). The construction of the principal coordinate analysis (PCOA) graph was performed using the QIIME2 package. A linear discriminant analysis (LEFse) effect size measure (Segata et al., 2011) was used to compare the relative percentages of individual bacterial taxonomic units in the microbiomes of the matched groups.

Statistical processing of the study results was performed using the STATISTICA.10 program package (Statsoft, USA). Quantitative parameters were evaluated by calculating mean values (M). The Mann–Whitney rank U-test was used to assess the reliability of differences in the relative percentages of individual bacterial taxa in the samples. Differences were considered reliable at $p < 0.05$. To eliminate the effect of multiple comparisons, the False Discovery Rate (FDR) correction was used to assess the significance of differences. Multiple regression analysis was used to assess the relationships between the content of individual bacteria in the sputum of patients with the presence of comorbidities, smoking, alcohol consumption, place of residence, and occupational harmful factors.

Results

Sequencing of the V3-V4 region of the 16S rRNA gene in sputum identified a total of nine bacterial types with a relative frequency above 0.1 %. The predominant bacterial types in the microbiomes of LUSC patients, AD patients and controls were *Bacillota* and *Bacteroidota*, which together accounted for about 70 % of the total microbiota. Overall, the relative percentages as well as the ratio of dominant bacterial types in sputum appeared close to the parameters previously described for the sputum microbiome in LC patients (Hosgood et al., 2014; Huang et al., 2019).

The Shannon index was used to assess alpha diversity. The results of the analysis showed that there were no differences between the matched samples of patients and healthy donors (Fig. 1). However, a significant reduction in alpha diversity according to the Pielou index (evenness) was found in the sputum of patients with AD and with LUSC compared to controls (Kruskal–Wallis test; $p = 0.0001$). There were no significant differences in uniformity (Pielou index) between the different histologic types of LC (Fig. 2).

Differences in bacterial community structure (beta diversity) in sputum samples from AD patients, LUSC patients and healthy individuals were evaluated with the PERMANOVA (Adonis) test using a Bray–Curtis difference matrix (Fig. 3). The analysis showed that there were differences in beta

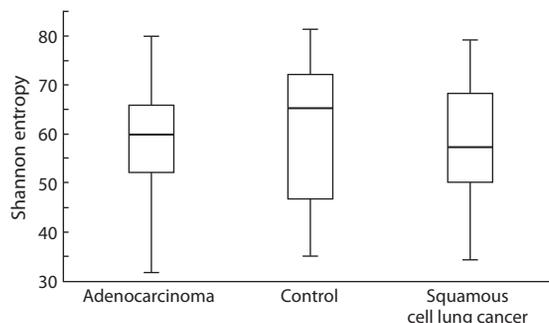


Fig. 1. Shannon diversity index of microbiomes of patients with adenocarcinoma, squamous cell lung cancer and control donors.

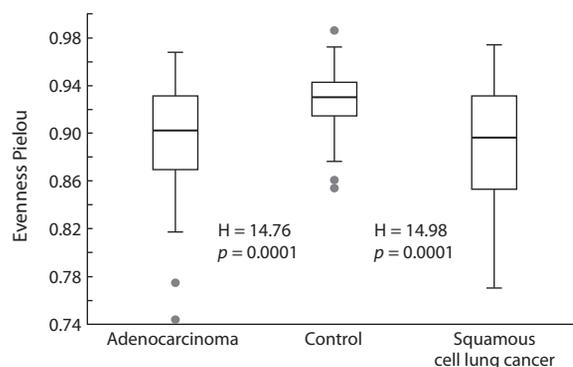


Fig. 2. Pielou index of microbiomes of patients with adenocarcinoma, squamous cell lung cancer and control donors.

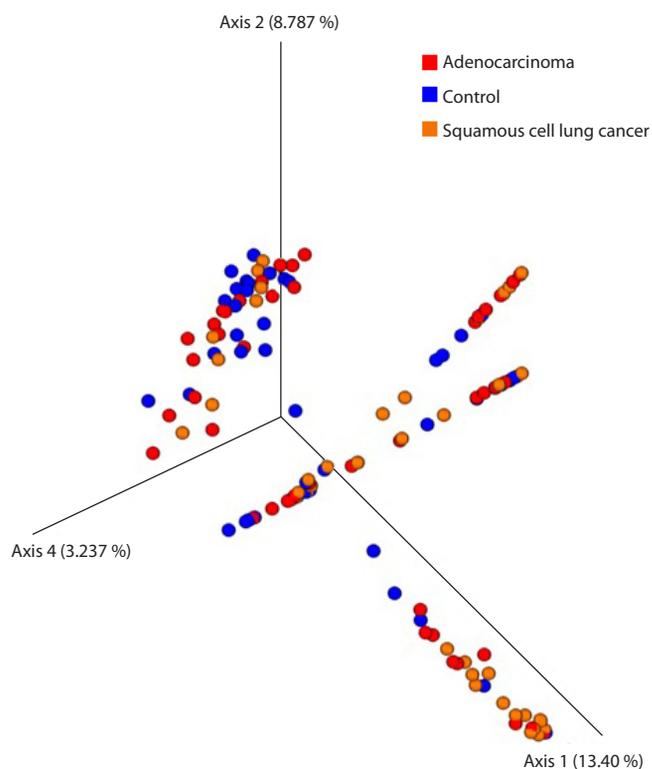


Fig. 3. Three-dimensional diagram constructed by principal component analysis showing the phylogenetic diversity of prokaryotic communities in the sputum of patients with adenocarcinoma, squamous cell lung cancer and control donors.

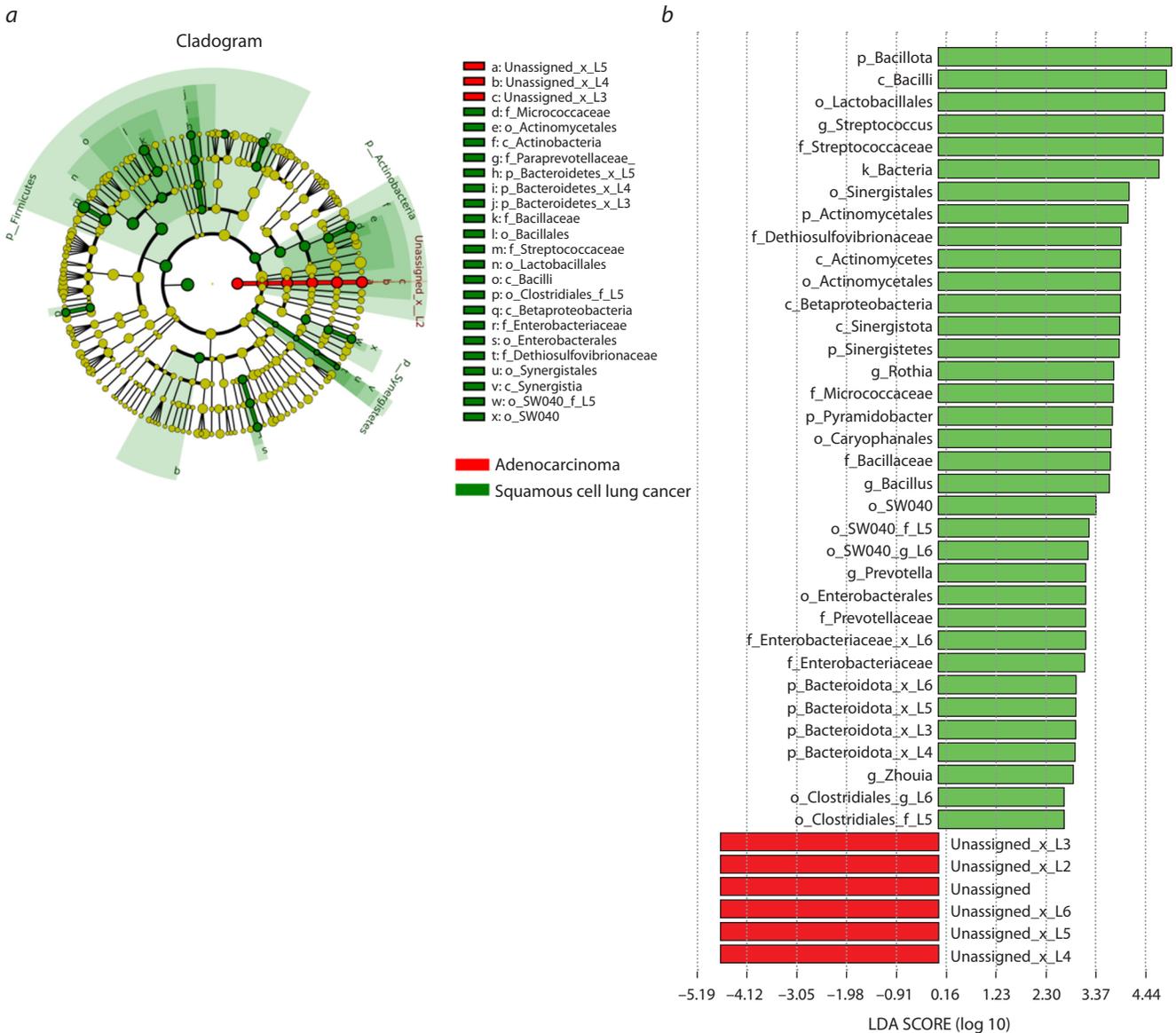


Fig. 4. Different representation of bacterial taxa in sputum samples of patients with squamous cell carcinoma and adenocarcinoma of the lung. Here and in Fig. 5, 6: *a* – cladogram giving an idea of the proximity of the differing taxonomic groups; *b* – graph representing the results of LEFse analysis. LDA – linear discriminant analysis.

diversity only between LUSC and control communities (pseudo-F = 3.89; *p* = 0.007).

Differences in bacterial taxonomic composition between the study samples were examined using linear discriminant analysis (LEFse), which estimates the effect size of the representation of different bacteria. The LEFse method revealed a significant increase in the representation of selected bacterial taxa in the sputum of patients with LUSC compared with AD. This applies in particular to the type *Bacillota*, the class *Bacilli* and the genus *Streptococcus* (Fig. 4). Comparison of the taxonomic composition of LUSC patients and healthy donors showed an increase in the content of representatives of *Bacillota* and *Pseudomonadota* types, *Bacilli* class, genera *Streptococcus*, *Rothia*, *Bacillus*, *Macellibacteroides*, etc. in the sputum of patients (Fig. 5). There were significantly fewer bacterial taxa for which LEFse analysis revealed dif-

ferences between healthy donors and AD patients (Fig. 6). Specifically, the sputum of healthy individuals showed an increase in representatives of the order *Clostridiales*, the class *Clostridia* and the genus *Moryella*, whereas in patients with AD the representation of the order *Flavobacteriales* and the class *Flavobacteriia* was increased.

Multiple regression analysis (MRA) was used to assess possible relationships between the content of individual bacteria in the sputum of LUSC patients with a range of other factors potentially affecting the composition of the microbiota. In addition to the bacterial genera (*Streptococcus*, *Rothia*, *Bacillus*, *Macellibacteroides*) significant for LUSC, MRA models included sex, patient age, comorbidities, smoking, alcohol consumption, place of residence, and presence of occupational hazards (see Table 1). As a result, it was found that among the confounders studied, only the presence of

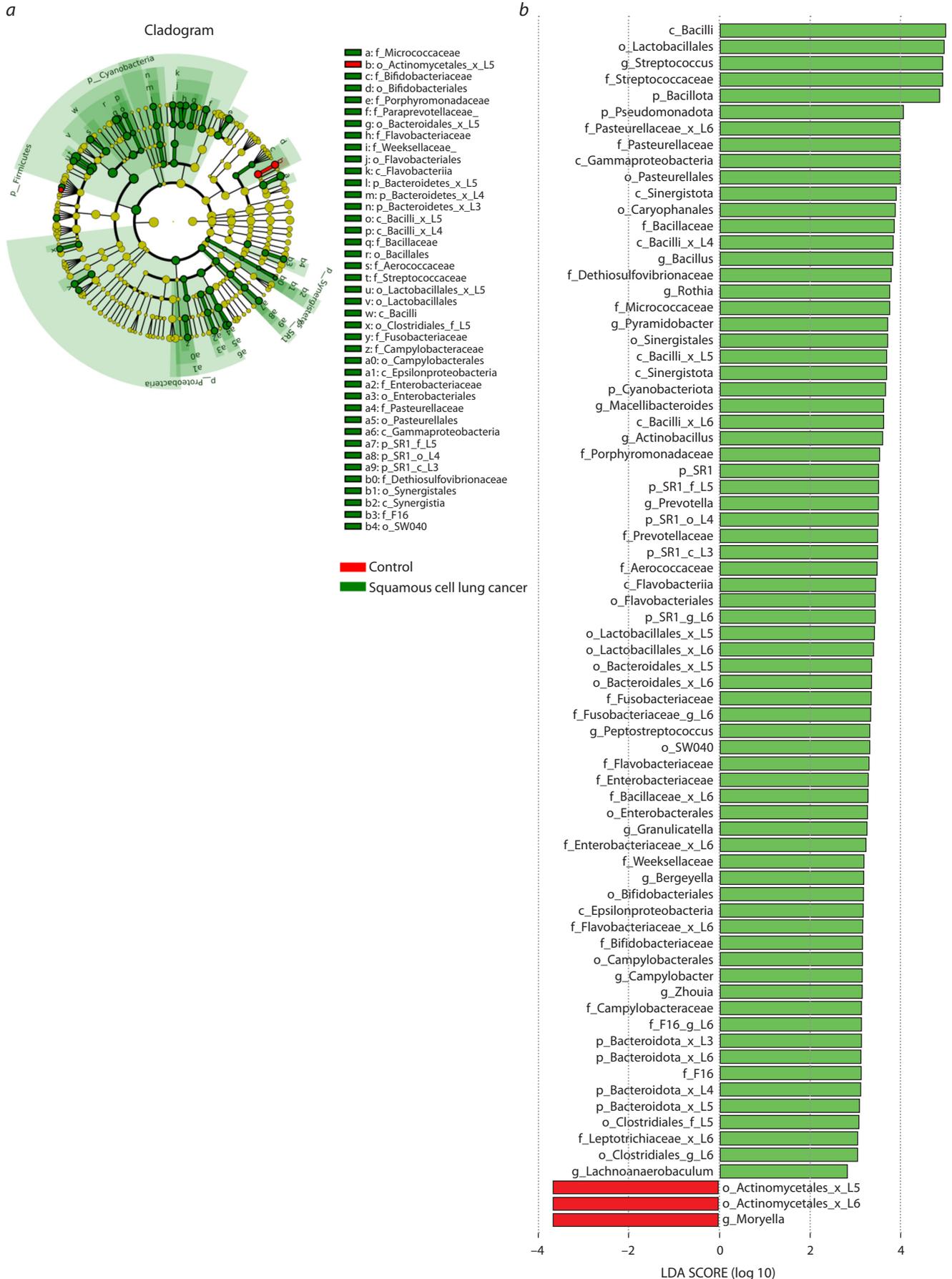


Fig. 5. Different representation of bacterial taxa in sputum samples of patients with squamous cell cancer and healthy donors.

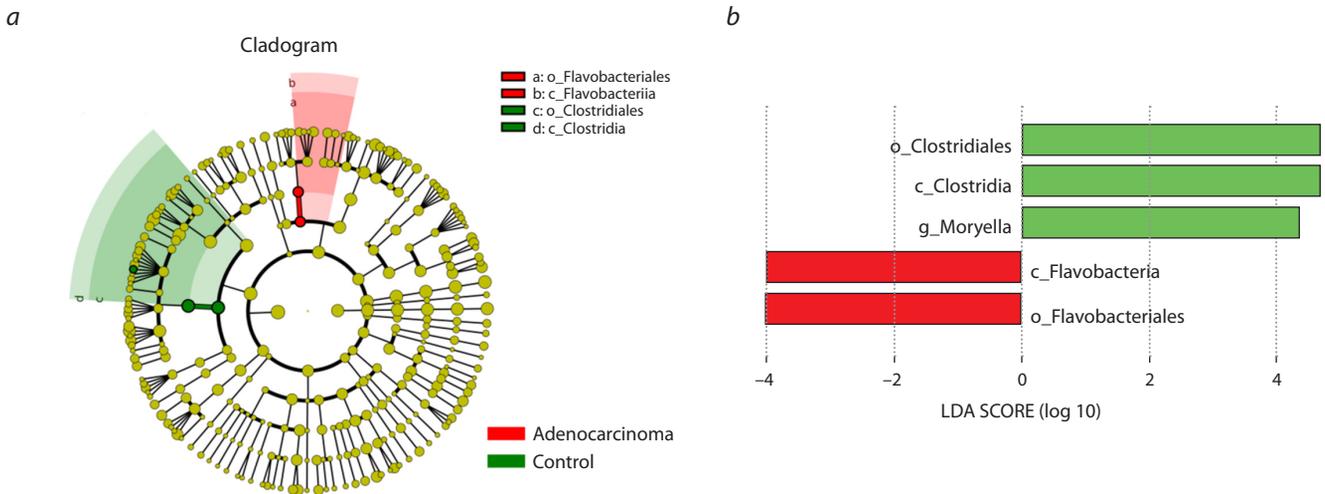


Fig. 6. Different representation of bacterial taxa in sputum samples of patients with lung adenocarcinoma and healthy donors.

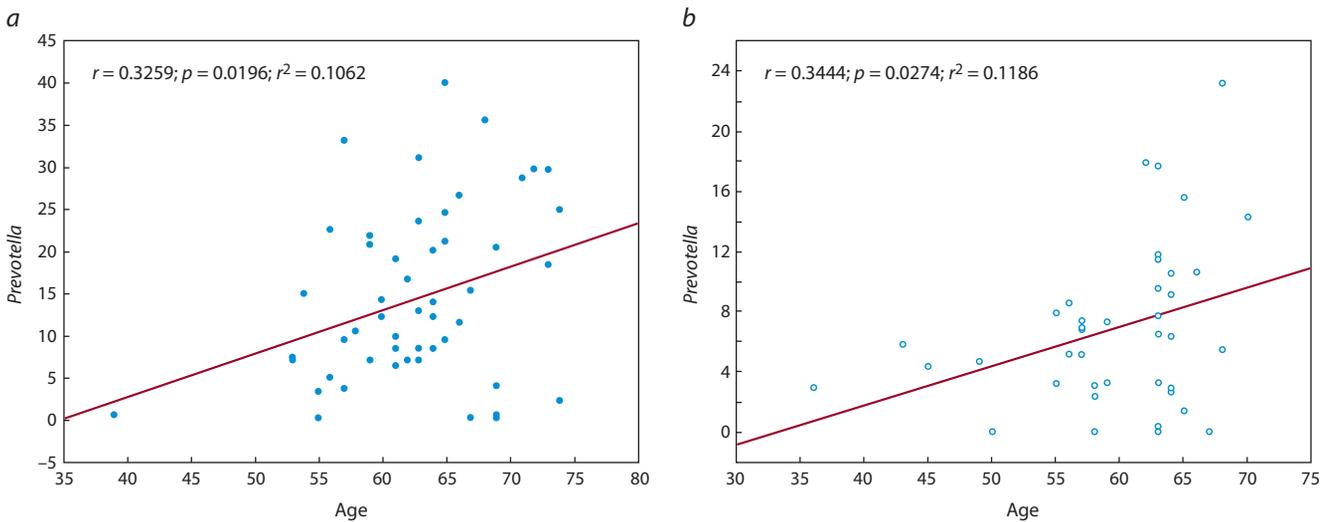


Fig. 7. Content of *Prevotella* representatives in the sputum of patients with lung adenocarcinoma (a) and squamous cell lung cancer (b) as a function of age.

cardiovascular disease (ischemia, hypertension, etc.), chronic bronchitis, and/or chronic obstructive pulmonary disease was associated with LUSC.

The effect of age and smoking status on microbiota composition in patients and controls was studied separately. Correlation analysis (Spearman) revealed a significant increase with age in the content of *Prevotella* species in the sputum of patients with AD ($p = 0.0196$) and in patients with PRL ($p = 0.0274$) (Fig. 7). At the same time, a positive correlation of age with the content of representatives of the genera *Atopobium* ($p = 0.03$) and *Leptotrichia* ($p = 0.03$) was observed in patients with AD. In the control samples, an increase with age in the content of bacteria from the genera *Porphyromonas* ($p = 0.01$) and *Veillonella* ($p = 0.045$) and, at the same time, a decrease in the content of representatives of the genera *Lachnoanaerobaculum* ($p = 0.02$), *Stomatobaculum* ($p = 0.006$) and *Oribacterium* ($p = 0.02$) were found.

Smoking status had no effect on sputum microbiome composition in patients with AD and LUSC. For the control sample, there was an increase in *Streptococcus* in the sputum of smoking donors compared to nonsmoking donors (20.87 vs. 15.16 %; $p = 0.0007$), and a significant decrease in *Neisseria* in the sputum of smokers (2.75 vs. 5.68 %; $p = 0.001$).

A question of separate interest is the possible influence of the stage of the tumor process on the composition of bacteria in sputum. The results summarized in Table 2 show that the percentage of bacterial taxa differs significantly between patients with NSCL in stages I–II as compared to stages III–IV of the disease. From the analysis of this data, it can be concluded that there is an increase in bacteria belonging to four genera in the sputum of patients in advanced stages of tumorigenesis.

Primary tumor localization in LC may be another factor potentially influencing the composition of the bacterial microbiota of the respiratory tract. Therefore, we compared the

Table 2. Average percentage of bacterial taxa in the sputum microbiome of patients with non-small cell lung cancer at different stages of the disease

Genus	I–II (n = 60), M, %	III–IV (n = 44), M, %	p*
<i>Porphyromonas</i>	3.09	3.99	0.004*
<i>Alloprevotella</i>	1.75	3.86	0.002*
<i>Selenomonas</i>	1.1	1.4	0.007*
<i>Megasphaera</i>	0.88	1.45	0.03
<i>Oribacterium</i>	0.48	0.72	0.01*
<i>Filifactor</i>	0.05	0.09	0.03

* Here and in Table 3: the p-value is less than the FDR-corrected p-value.

mean percentage of bacterial genera in the sputum of patients with central NSCLC and peripheral NSCLC (Table 3).

As follows from these data, the central localization of the tumor is accompanied by an increase in representatives of the genus *Bacteroides*. At the same time, an increase in bacteria of the genus *Veillonella* was observed in patients with peripheral NSCLC as compared to a central tumor localization (12.79 vs. 7.99%; $p = 0.01$).

Discussion

Differences in the taxonomic composition of the bacterial microbiome of the human respiratory tract have already been recognized as an important pathogenetic factor in lung cancer (Maddi et al., 2019; Yagi et al., 2021), but to date, the importance of the microflora in patients with different histological types of NSCLC remains an open question. Here, we compared the taxonomic composition of the microbiome in sputum samples from patients with the two most common forms of NSCLC: adenocarcinoma and squamous cell lung cancer.

According to previous studies, the respiratory microbiota of LC patients tends to have lower alpha diversity compared to healthy individuals, while beta diversity is not significantly different (Lee et al., 2016; Liu N.N. et al., 2020). The same trend has been observed for the microbiomes of cancer-affected and non-cancerous lung tissues (Kim et al., 2022).

Evidence of similarities or differences between airway and lung tissue community diversity parameters of patients with different histologic types of LC to date is scarce and these results are inconsistent. For example, alpha diversity of the microbiome was found to be higher in the sputum of AD patients compared to LUSC and a significant difference in beta diversity was also found between these groups, but the samples compared were too small (6 and 7 cases, respectively) (Ran et al., 2020). Another study showed that the bronchoalveolar lavage (BAL) microbiota was more diverse in LUSC than in AD (Gomes et al., 2019). No differences in alpha diversity as well as beta diversity of microbiomes from sputum and BAL samples were found between samples of patients with AD and LUSC (Huang et al., 2019). The microbiome of tumor tissues of patients with AD did not differ in alpha diversity from LUSC, although a significant increase in the content of

Table 3. Average percentage of bacterial taxa in the sputum microbiome of patients with non-small cell lung cancer with different tumor localization

Genus	Central NSCLC (n = 31), M, %	Peripheral NSCLC (n = 68), M, %	p
<i>Veillonella</i>	8.76	12.79	0.01*
<i>Bacillus</i>	3.52	1.87	0.03
<i>Granulicatella</i>	1.63	0.94	0.04
<i>Bacteroides</i>	1.62	0.52	0.002*
<i>Oribacterium</i>	0.43	0.18	0.03

Gram-positive bacteria was recorded in the adenocarcinoma group (Kovaleva et al., 2020).

Our study showed that the values of the Shannon index, which reflects the species richness of the microbiota, are close in the matched cohorts of patients and control donors. A significant decrease in both patient cohorts compared to controls was observed for the evenness index, which is based on measuring the relative abundance of different species in a community and is one of the metrics characterizing alpha diversity. Bacterial community structures (beta diversity) between AD and LUSC were also similar, but according to the Bray–Curtis matrix, differences were present between the bacterial communities of LUSC patients and healthy subjects, but not between AD and controls (see Fig. 3). Thus, our study showed that the α -diversity and β -diversity of bacterial communities of sputum from patients with different histologic types are similar. Nevertheless, it is worth noting that the microbiome of LUSC patients differs significantly from that of healthy individuals.

To answer the question of differences between sputum microbiome compositions in cohorts of patients with different histologic types of LC, we used the LEFse method, which is the most commonly used method in microbiome studies. LEFse analysis allowed identification of differences between the compared patient samples (see Fig. 4). The sputum of LUSC patients had a significant enrichment of *Bacillota* (genus *Streptococcus* and *Bacillus*) and *Actinomycetota* (genus *Rothia*) when compared with samples from AD patients. Comparison of respiratory microbiome composition in groups of patients with LUSC and healthy subjects also revealed a number of significant differences. According to the results of LEFse analysis (see Fig. 5), the content of representatives of the phylum *Bacillota* and *Pseudomonadota*; genera *Streptococcus*, *Bacillus*, *Rothia*, *Macellibacteroides*, *Prevotella*, *Actinobacillus* and *Peptostreptococcus* was increased in the sputum of patients compared to controls. In healthy study participants, an increase in the representation of the *Actinomycetales* and the genus *Moryella* was observed.

Several previous studies have shown that the composition of the bacterial microbiota in the respiratory tract of LC patients may be histologically dependent. For example, Q. Leng and

colleagues (Leng et al., 2021) used digital droplet PCR to analyze 25 genera of bacteria commonly associated with NSCLC in the sputum of 17 NSCLC patients and 10 healthy subjects. A significant increase in the content of representatives of the genera *Acidovorax*, *Streptococcus*, *H. pylori* and *Veillonella* was detected in the sputum of LUSC patients, whereas an increased abundance of *Capnocytophaga* was found in the sputum of AD patients. These same sputum bacterial biomarkers were then confirmed in another cohort consisting of 69 NSCLC cases and 79 control donors. In another study, the relationship between saliva microflora and lung cancer was examined. DNA samples from 20 LC patients (10 LUSC and 10 AD) and control subjects ($n = 10$) were sequenced (Yan et al., 2015). At the level of bacterial genera, *Capnocytophaga*, *Selenomonas*, and *Veillonella* were elevated in both AD and squamous cell cancer, and *Neisseria* was reduced in both AD and LUSC.

In our study, patients with LUSC had a significant increase in members of the genera *Streptococcus*, *Bacillus* and *Rothia* compared to AD. There was an increase in *Capnocytophaga* (1.46 vs. 1.08 %) in the sputum of AD patients compared to LUSC, as in a previous study (Leng et al., 2021), but these differences were not significant. Thus, it can be stated on the one hand that the two main histologic forms of LC have distinct respiratory microbiomes, however, there is no uniform set of bacterial taxa marking these differences. Perhaps, this fact reflects the initially different composition of bacteria inhabiting the respiratory tract of patients with NSCLC living in different regions of the world, i.e. it is a consequence of environmental factors (Costello et al., 2012).

An important finding of this study is the significant difference in the content of bacterial taxa in the sputum microbiome of patients with different histologic forms of LC compared to healthy subjects. While for LUSC there is a significant enrichment of *Streptococcus*, *Bacillus*, *Rothia*, *Macellibacteroides*, *Prevotella*, *Actinobacillus* and *Peptostreptococcus* genera in sputum (see Fig. 5), no significant differences in bacterial composition were found in the sample of patients with adenocarcinoma compared to controls (see Fig. 6). This fact means that the search for metagenomic biomarkers associated with LC can be correct only after separate analysis of microbiota composition depending on the histological classification of the tumor.

The sample size used in our study allowed us to examine, in addition to the histological type of tumor, other individual factors (age, smoking status, stage of malignant process, tumor localization) potentially capable of influencing the composition of the microbiota in NSCLC. Of interest is the age-correlated increase in the content of representatives of the genus *Prevotella*, which was registered in both samples of patients (see Fig. 7). This is in disagreement with the results of a study of BAL samples from NSCLC patients, where a subgroup of patients older than 60 years recorded a decrease in *Prevotella* (*P. oryzae*) compared to younger patients (Zheng et al., 2021).

Comparison of the composition of the sputum microbiome in smoking and nonsmoking patients with PRL and ACL showed no differences in bacterial composition. However, the control group showed an increase in *Streptococcus* as well as a marked decrease in *Neisseria* in the sputum of smokers

compared to nonsmokers, which is consistent with previously published results (Huang, Shi, 2019; Ying et al., 2022). Noteworthy is the fact that smokers also show increased *Streptococcus* representation and decreased *Neisseria* representation in the upper gastrointestinal tract compared to non-smoking donors (Shanahan et al., 2018). According to recent findings (Haldar et al., 2020), the effect of smoking on sputum microbiota remains unclear and requires further investigation.

The evaluation of the possible influence of NSCLC stage on the structure of sputum microbiome has shown that in the sputum of patients at advanced stages of tumor progression there is an increase in the content of bacteria belonging to the genera *Porphyromonas*, *Alloprevotella*, *Selenomonas*, *Megasphaera*, *Oribacterium* and *Filifactor*. NSCLC patients with central lung cancer had increased sputum levels of bacteria from the genus *Bacteroides*. At the same time, an increase in *Veillonella* content was noted in patients with peripheral lung cancer compared to central tumor localization. These results should be considered as preliminary, as the analysis was performed for the total sample of patients without taking into account the histologic type of NSCLC.

Conclusion

In this study, a comparative analysis of the taxonomic composition of the bacterial microbiome of sputum from patients with two major histologic types of NSCLC and healthy sputum donors was performed based on the sequence of the 16S rRNA coding gene region identified using massively parallel sequencing technology. Significant differences in the content of representatives of a number of bacterial genera in the sputum of patients with LUSC and AD were revealed. In particular, the presence of *Streptococcus*, *Bacillus*, *Rothia* and other genera was elevated in the sputum of LUSC patients compared to healthy subjects.

The present findings require confirmation in independent large-scale studies to further understand the role of the sputum microbiota in the development of NSCLC. In addition, the search for bacterial “signatures” associated with lung cancer risk requires whole-genome sequencing to obtain an accurate assessment of taxonomic composition at the species level.

References

- Bolyen E., Rideout J.R., Dillon M.R., Bokulich N.A., Abnet C.C., Al-Ghalith G.A., Alexander H., ... Willis A.D., Xu Z.Z., Zaneveld J.R., Zhang Y., Zhu Q., Knight R., Caporaso J.G. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* 2019;37(8):852-857. DOI 10.1038/s41587-019-0209-9
- Cameron S.J.S., Lewis K.E., Huws S.A., Hegarty M.J., Lewis P.D., Pachebat J.A., Mur L.A.J. A pilot study using metagenomic sequencing of the sputum microbiome suggests potential bacterial biomarkers for lung cancer. *PLoS One.* 2017;12(5):e0177062. DOI 10.1371/journal.pone.0177062
- Chen Y., Wu F.H., Wu P.Q., Xing H.Y., Ma T. The role of the tumor microbiome in tumor development and its treatment. *Front. Immunol.* 2022;13:935846. DOI 10.3389/fimmu.2022.935846
- Cheng C., Wang Z., Wang J., Ding C., Sun C., Liu P., Xu X., Liu Y., Chen B., Gu B. Characterization of the lung microbiome and exploration of potential bacterial biomarkers for lung cancer. *Transl. Lung. Cancer Res.* 2020;9(3):693-704. DOI 10.21037/tlcr-19-590

- Cheng T.Y., Cramb S.M., Baade P.D., Youlten D.R., Nwogu C., Reid M.E. The international epidemiology of lung cancer: latest trends, disparities, and tumor characteristics. *J. Thorac. Oncol.* 2016; 11(10):1653-1671. DOI 10.1016/j.jtho.2016.05.021
- Chiu C.Y., Miller S.A. Clinical metagenomics. *Nat. Rev. Genet.* 2019; 20(6):341-355. DOI 10.1038/s41576-019-0113-7
- Costello E.K., Stagaman K., Dethlefsen L., Bohannan B.J., Relman D.A. The application of ecological theory toward an understanding of the human microbiome. *Science.* 2012;336(6086):1255-1262. DOI 10.1126/science.1224203
- Druzhinin V.G., Matskova L.V., Demenkov P.S., Baranova E.D., Volo-baev V.P., Minina V.I., Apalko S.V., Churina M.A., Romanyuk S.A., Shcherbak S.G., Ivanov V.I., Lariонов A.V. Taxonomic diversity of sputum microbiome in lung cancer patients and its relationship with chromosomal aberrations in blood lymphocytes. *Sci. Rep.* 2020; 10(1):9681. DOI 10.1038/s41598-020-66654-x
- Druzhinin V.G., Matskova L.V., Demenkov P.S., Baranova E.D., Volo-baev V.P., Minina V.I., Lariонов A.V., Titov V.A., Fucic A. Genetic damage in lymphocytes of lung cancer patients is correlated to the composition of the respiratory tract microbiome. *Mutagenesis.* 2021;36(2):143-153. DOI 10.1093/tract/geab004
- Goldstraw P. New staging system: How does it affect our practice? *J. Clin. Oncol.* 2013;31(8):984-991. DOI 10.1200/JCO.2012.42.7922
- Gomes S., Cavadas B., Ferreira J.C., Marques P.I., Monteiro C., Suce-na M., Sousa C., Vaz Rodrigues L., Teixeira G., Pinto P., Tavares de Abreu T., Bárbara C., Semedo J., Mota L., Carvalho A.S., Matthei-sen R., Pereira L., Seixas S. Profiling of lung microbiota discloses differences in adenocarcinoma and squamous cell carcinoma. *Sci. Rep.* 2019;9(1):12838. DOI 10.1038/s41598-019-49195-w
- Haldar K., George L., Wang Z., Mistry V., Ramsheh M.Y., Free R.C., John C., Reeve N.F., Miller B.E., Tal-Singer R., Webb A.J., Brookes A.J., Tobin M.D., Singh D., Donaldson G.C., Wedzi-cha J.A., Brown J.R., Barer M.R., Brightling C.E. The sputum microbiome is distinct between COPD and health, independent of smoking history. *Respir. Res.* 2020;21(1):183. DOI 10.1186/s12931-020-01448-3
- Hasegawa A., Sato T., Hoshikawa Y., Ishida N., Tanda N., Kawamu-ura Y., Kondo T., Takahashi N. Detection and identification of oral anaerobes in intraoperative bronchial fluids of patients with pul-monary carcinoma. *Microbiol. Immunol.* 2014;58(7):375-381. DOI 10.1111/1348-0421.12157
- Herbst R.S., Heymach J.V., Lippman S.M. Lung cancer. *N. Engl. J. Med.* 2008;359(13):1367-1380. DOI 10.1056/NEJMra0802714
- Hosgood H.D. 3rd, Sapkota A.R., Rothman N., Rohan T., Hu W., Xu J., Vermeulen R., He X., White J.R., Wu G., Wei F., Mongodin E.F., Lan Q. The potential role of lung microbiota in lung cancer attributed to household coal burning exposures. *Environ. Mol. Mutagen.* 2014;55(8):643-651. DOI 10.1002/em.21878
- Hosgood H.D. 3rd, Mongodin E.F., Wan Y., Hua X., Rothman N., Hu W., Vermeulen R., Seow W.J., Rohan T., Xu J., Li J., He J., Huang Y., Yang K., Wu G., Wei F., Shi J., Sapkota A.R., Lan Q. The respiratory tract microbiome and its relationship to lung cancer and environmental exposures found in rural China. *Environ. Mol. Muta-gen.* 2019;60(7):617-623. DOI 10.1002/em.22291
- Huang C., Shi G. Smoking and microbiome in oral, airway, gut and some systemic diseases. *J. Transl. Med.* 2019;17(1):225. DOI 10.1186/s12967-019-1971-7
- Huang D., Su X., Yuan M., Zhang S., He J., Deng Q., Qiu W., Dong H., Cai S. The characterization of lung microbiome in lung cancer pa-tients with different clinicopathology. *Am. J. Cancer Res.* 2019;9(9): 2047-2063
- Kim O.H., Choi B.Y., Kim D.K., Kim N.H., Rho J.K., Sul W.J., Lee S.W. The microbiome of lung cancer tissue and its associa-tion with pathological and clinical parameters. *Am. J. Cancer Res.* 2022;12(5):2350-2362
- Kovaleva O., Podlesnaya P., Rashidova M., Samoiloa D., Petrenko A., Zborovskaya I., Mochalnikova V., Kataev V., Khlopko Y., Plotni-kov A., Gratchev A. Lung microbiome differentially impacts sur-vival of patients with non-small cell lung cancer depending on tu-mor stroma phenotype. *Biomedicines.* 2020;8(9):349. DOI 10.3390/biomedicines8090349
- Lee S.H., Sung J.Y., Yong D., Chun J., Kim S.Y., Song J.H., Chung K.S., Kim E.Y., Jung J.Y., Kang Y.A., Kim Y.S., Kim S.K., Chang J., Park M.S. Characterization of microbiome in bronchoalveolar la-vage fluid of patients with lung cancer comparing with benign mass like lesions. *Lung Cancer.* 2016;102:89-95. DOI 10.1016/j.lungcan. 2016.10.016
- Leng Q., Holden V.K., Deepak J., Todd N.W., Jiang F. Microbiota bio-markers for lung cancer. *Diagnostics (Basel).* 2021;11(3):407. DOI 10.3390/diagnostics11030407
- Liu H.X., Tao L.L., Zhang J., Zhu Y.G., Zheng Y., Liu D., Zhou M., Ke H., Shi M.M., Qu J.M. Difference of lower airway microbiome in bilateral protected specimen brush between lung cancer patients with unilateral lobar masses and control subjects. *Int. J. Cancer.* 2018;142(4):769-778. DOI 10.1002/ijc.31098
- Liu N.N., Ma Q., Ge Y., Yi C.X., Wei L.Q., Tan J.C., Chu Q., Li J.Q., Zhang P., Wang H. Microbiome dysbiosis in lung cancer: from com-position to therapy. *NPJ Precis. Oncol.* 2020;4(1):33. DOI 10.1038/s41698-020-00138-z
- Lozupone C., Knight R. UniFrac: a new phylogenetic method for com-paring microbial communities. *Appl. Environ. Microbiol.* 2005; 71(12):8228-8235. DOI 10.1128/AEM.71.12.8228-8235.2005
- Maddi A., Sabharwal A., Violante T., Manuballa S., Genco R., Pat-naik S., Yendamuri S. The microbiome and lung cancer. *J. Thorac. Dis.* 2019;11(1):280-291. DOI 10.21037/jtd.2018.12.88
- Mao Q., Jiang F., Yin R., Wang J., Xia W., Dong G., Ma W., Yang Y., Xu L., Hu J. Interplay between the lung microbiome and lung cancer. *Cancer Lett.* 2018;415:40-48. DOI 10.1016/j.canlet.2017.11.036
- Molina J.R., Yang P., Cassivi S.D., Schild S.E., Adjei A.A. Non-small cell lung cancer: epidemiology, risk factors, treatment, and survivor-ship. *Mayo Clin. Proc.* 2008;83(5):584-594. DOI 10.4065/83.5.584
- Parte A.C., Sardà Carbasse J., Meier-Kolthoff J.P., Reimer L.C., Göker M. List of Prokaryotic names with Standing in Nomenclature (LPSN) moves to the DSMZ. *Int. J. Syst. Evol. Microbiol.* 2020; 70(11):5607-5612. DOI 10.1099/ijsem.0.004332
- Peters B.A., Hayes R.B., Goparaju C., Reid C., Pass H.I., Ahn J. The microbiome in lung cancer tissue and recurrence-free survival. *Cancer Epidemiol. Biomark. Prev.* 2019;28(4):731-740. DOI 10.1158/1055-9965.EPI-18-0966
- Ran Z., Liu J., Wang F., Xin C., Shen X., Zeng S., Song Z., Xiong B. Analysis of pulmonary microbial diversity in patients with advanced lung cancer based on high-throughput sequencing technology. *Zhongguo Fei Ai Za Zhi.* 2020;23(12):1031-1038. DOI 10.3779/j.issn.1009-3419.2020.103.16 (in Chinese)
- Segata N., Izard J., Waldron L., Gevers D., Miropolsky L., Garrett W.S., Huttenhower C. Metagenomic biomarker discovery and explana-tion. *Genome Biol.* 2011;12(6):R60. DOI 10.1186/gb-2011-12-6-r60
- Shanahan E.R., Shah A., Koloski N., Walker M.M., Talley N.J., Morris-son M., Holtmann G.J. Influence of cigarette smoking on the human duodenal mucosa-associated microbiota. *Microbiome.* 2018;6(1): 150. DOI 10.1186/s40168-018-0531-3
- Tsao M.S., Yoon J.Y. The eighth TNM classification for lung can-cer – What is next? *Lung Cancer.* 2018;121:97-98. DOI 10.1016/j.lungcan.2018.04.018
- Tsay J.J., Wu B.G., Badri M.H., Clemente J.C., Shen N., Meyn P., Li Y., Yie T.A., Lhakhang T., Olsen E., Murthy V., Michaud G., Sulai-man I., Tsirigos A., Heguy A., Pass H., Weiden M.D., Rom W.N., Sterman D.H., Bonneau R., Blaser M.J., Segal L.N. Airway mro-biota is associated with upregulation of the PI3K pathway in lung cancer. *Am. J. Respir. Crit. Care Med.* 2018;198:1188-1198. DOI 10.1164/rccm.201710-2118OC
- Wang K., Huang Y., Zhang Z., Liao J., Ding Y., Fang X., Liu L., Luo J., Kong J. A preliminary study of microbiota diversity in saliva and bronchoalveolar lavage fluid from patients with primary bron-chogenic carcinoma. *Med. Sci. Monit.* 2019;25:2819-2834. DOI 10.12659/MSM.915332

- Wu Y., Jiao N., Zhu R., Zhang Y., Wu D., Wang A.J., Fang S., Tao L., Li Y., Cheng S., He X., Lan P., Tian C., Liu N.N., Zhu L. Identification of microbial markers across populations in early detection of colorectal cancer. *Nat. Commun.* 2021;12(1):3063. DOI 10.1038/s41467-021-23265-y
- Xavier J.B., Young V.B., Skufca J., Ginty F., Testerman T., Pearson A.T., Macklin P., ... Johnson W.E., Jobin C., Ridlon J.M., Koh A.Y., Yu M., Kelly L., Wargo J.A. The cancer microbiome: distinguishing direct and indirect effects requires a systemic view. *Trends Cancer.* 2020;6(3):192-204. DOI 10.1016/j.trecan.2020.01.004
- Yagi K., Huffnagle G.B., Lukacs N.W., Asai N. The lung microbiome during health and disease. *Int. J. Mol. Sci.* 2021;22(19):10872. DOI 10.3390/ijms221910872
- Yan X., Yang M., Liu J., Gao R., Hu J., Li J., Zhang L., Shi Y., Guo H., Cheng J., Razi M., Pang S., Yu X., Hu S. Discovery and validation of potential bacterial biomarkers for lung cancer. *Am. J. Cancer Res.* 2015;5(10):3111-3122
- Ying K.L., Brasky T.M., Freudenheim J.L., McElroy J.P., Nickerson Q.A., Song M.A., Weng D.Y., Wewers M.D., Whiteman N.B., Mathe E.A., Shields P.G. Saliva and lung microbiome associations with electronic cigarette use and smoking. *Cancer Prev. Res. (Phila).* 2022;15(7):435-446. DOI 10.1158/1940-6207.CAPR-21-0601
- Zhang W., Luo J., Dong X., Zhao S., Hao Y., Peng C., Shi H., Zhou Y., Shan L., Sun Q., Li Y., Zhao X. Salivary microbial dysbiosis is associated with systemic inflammatory markers and predicted oral metabolites in non-small cell lung cancer patients. *J. Cancer.* 2019;10(7):1651-1662. DOI 10.7150/jca.28077
- Zheng L., Sun R., Zhu Y., Li Z., She X., Jian X., Yu F., Deng X., Sai B., Wang L., Zhou W., Wu M., Li G., Tang J., Jia W., Xiang J. Lung microbiome alterations in NSCLC patients. *Sci. Rep.* 2021;11(1):11736. DOI 10.1038/s41598-021-91195-2
- Zhuo M., An T., Zhang C., Wang Z. Characterization of microbiota in cancerous lung and the contralateral non-cancerous lung within lung cancer patients. *Front. Oncol.* 2020;10:1584. DOI 10.3389/fonc.2020.01584

Funding. This work was financially supported by the Russian Science Foundation (grant No. 18-14-00022R).

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

Received April 14, 2023. Revised July 21, 2023. Accepted July 23, 2023.

DOI 10.18699/vjgb-24-26

Role of sirtuins in epigenetic regulation and aging control

E.M. Samoilova ^{1, 2#}, S.E. Romanov ^{1, 3#}, D.A. Chudakova ⁴, P.P. Laktionov ^{1, 3} ¹ Novosibirsk State University, Novosibirsk, Russia² Engelhardt Institute of Molecular Biology of the Russian Academy of Sciences, Moscow, Russia³ Institute of Molecular and Cellular Biology of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia⁴ Federal Center of Brain Research and Neurotechnologies of the Federal Medical Biological Agency of Russia, Moscow, Russia laktionov@mcb.nsc.ru

Abstract. Advances in modern healthcare in developed countries make it possible to extend the human lifespan, which is why maintaining active longevity is becoming increasingly important. After the sirtuin (SIRT) protein family was discovered, it started to be considered as a significant regulator of the physiological processes associated with aging. SIRT has deacetylase, deacylase, and ADP-ribosyltransferase activity and modifies a variety of protein substrates, including chromatin components and regulatory proteins. This multifactorial regulatory system affects many processes: cellular metabolism, mitochondrial functions, epigenetic regulation, DNA repair and more. As is expected, the activity of sirtuin proteins affects the manifestation of classic signs of aging in the body, such as cellular senescence, metabolic disorders, mitochondrial dysfunction, genomic instability, and the disruption of epigenetic regulation. Changes in the SIRT activity in human cells can also be considered a marker of aging and are involved in the genesis of various age-dependent disorders. Additionally, experimental data obtained in animal models, as well as data from population genomic studies, suggest a SIRT effect on life expectancy. At the same time, the diversity of sirtuin functions and biochemical substrates makes it extremely complicated to identify cause-and-effect relationships and the direct role of SIRT in controlling the functional state of the body. However, the SIRT influence on the epigenetic regulation of gene expression during the aging process and the development of disorders is one of the most important aspects of maintaining the homeostasis of organs and tissues. The presented review centers on the diversity of SIRT in humans and model animals. In addition to a brief description of the main SIRT enzymatic and biological activity, the review discusses its role in the epigenetic regulation of chromatin structure, including the context of the development of genome instability associated with aging. Studies on the functional connection between SIRT and longevity, as well as its effect on pathological processes associated with aging, such as chronic inflammation, fibrosis, and neuroinflammation, have been critically analyzed.

Key words: sirtuins; aging; protein deacetylation; epigenetic regulation.

For citation: Samoilova E.M., Romanov S.E., Chudakova D.A., Laktionov P.P. Role of sirtuins in epigenetic regulation and aging control. *Vavilovskii Zhurnal Genetiki i Seleksii = Vavilov Journal of Genetics and Breeding*. 2024;28(2):215-227. DOI 10.18699/vjgb-24-26

Роль сиртуинов в эпигенетической регуляции и контроле старения

E.M. Самойлова ^{1, 2#}, С.Е. Романов ^{1, 3#}, Д.А. Чудакова ⁴, П.П. Лактионов ^{1, 3} ¹ Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия² Институт молекулярной биологии им. В.А. Энгельгардта Российской академии наук, Москва, Россия³ Институт молекулярной и клеточной биологии Сибирского отделения Российской академии наук, Новосибирск, Россия⁴ Федеральный центр мозга и нейротехнологий Федерального медико-биологического агентства России, Москва, Россия laktionov@mcb.nsc.ru

Аннотация. Достижения современного здравоохранения в развитых странах позволили увеличить продолжительность жизни, из-за чего все более актуальным становится сохранение активного долголетия. С момента открытия белки семейства сиртуинов рассматривались в качестве значимых регуляторов физиологических процессов, ассоциированных со старением. Сиртуины проявляют деацетилазную, деацетилазную, АДФ-рибозилтрансферазную активность и модифицируют множество белковых субстратов, включая компоненты хроматина и регуляторные белки. Столь многофакторная система регуляции затрагивает ряд процессов, таких как клеточный метаболизм, функции митохондрий, эпигенетическую регуляцию, репарацию ДНК и прочие. Неудивительно, что активность белков-сиртуинов затрагивает проявление классических признаков старения организма: клеточное старение, нарушения метаболизма, митохондриальную дисфункцию, геномную нестабильность и нарушение эпигенетической регуляции. Непосредственное изменение активности сиртуи-

нов в клетках человека также рассматривается в качестве маркера старения и вовлечено в генез различных возраст-зависимых патологических состояний. Кроме того, экспериментальные данные, полученные на модельных животных, а также результаты популяционных геномных исследований позволяют предположить влияние сиртуинов на продолжительность жизни. Вместе с тем многообразие функций сиртуинов и биохимических субстратов делает крайне нетривиальным выявление причинно-следственных связей и непосредственной роли сиртуинов в контроле функционального состояния организма. Однако влияние сиртуинов на эпигенетическую регуляцию экспрессии генов в ходе старения и при патологиях – один из наиболее важных аспектов поддержания гомеостаза органов и тканей. Представленный обзор посвящен разнообразию белков-сиртуинов у человека и модельных животных. Помимо краткого описания основных ферментативных и биологических активностей сиртуинов, рассматривается роль сиртуинов в эпигенетической регуляции структуры хроматина, в том числе в контексте развития нестабильности генома, ассоциированной со старением. Проведен критический анализ работ по исследованию функциональной связи сиртуинов и долголетия, а также влияния сиртуинов на ассоциированные со старением патологические процессы, такие как хроническое воспаление, фиброз и нейровоспаление.

Ключевые слова: сиртуины; старение; деацетилирование белков; эпигенетическая регуляция.

Introduction

The first representative of sirtuin proteins, Sir2 (silent information regulator 2), was discovered in the budding yeast *Saccharomyces cerevisiae*. It was initially described as a key transcription repressor at HM loci responsible for the mating-type switching of yeast (Ivy et al., 1986). Subsequently, it was confirmed that Sir2 is needed to suppress the expression of transgenes near telomeres and the silencing of retrotransposons that have been integrated into tandem repeats of ribosomal DNA (Gottschling et al., 1990; Bryk et al., 1997). Its main function is the NAD⁺-dependent histone deacetylation (Imai et al., 2000; Smith et al., 2000). Deletion of the *Sir2* gene was found to approximately halve the lifespan of *S. cerevisiae*. Conversely, the overexpression of *Sir2* increases it by a third (Sinclair, Guarente, 1997; Kaerberlein et al., 1999). Sir2 homologues, united by the name “sirtuins” (silent information regulator two proteins), were found in all domains of living organisms from bacteria and archaea to humans (Frye, 2000). Therefore, a direct relationship between Sir2 activity and yeast lifespan initiated research into the role of the NAD-dependent Sir2 family deacetylases in the regulation of the aging processes. Seven sirtuins (SIRT1–SIRT7) were found in mammals, where SIRT1 has the greatest homology to yeast Sir2 (Frye, 2000).

Sirtuins of higher eukaryotes were grouped into class III HDAC due to their specific function and structural features (Gray, Ekström, 2001). All human sirtuins are characterized by a common conserved core region of 250–270 amino acids in length (Fig. 1). This protein fragment consists of a Rossmann fold domain, which is characteristic of many NAD⁺-dependent proteins, and a small domain, which consists of a Zn-binding and a helical modules (Moniot et al., 2013).

The main differences between homologues are found in the N- and C-terminal domains, which can contain signals for nuclear or nucleolar localization (NLS and NoLS, respectively), nuclear export (NES) or mitochondrial transport (MTS) in different proteins (see Fig. 1). It is worth noting that the seven human sirtuin genes encode at least 23 protein isoforms (Rack et al., 2014; Zhang X. et al., 2021). Features of minor isoforms can be either an absence of transport signal sequences or an altered core structure, due to which they can have original functions (Rack et al., 2014; Du Y. et al., 2018).

Sirtuins are involved in the regulation of different intracellular processes: cellular metabolism, mitochondrial functions, chromatin remodeling, and response to oxidative stress (Wu et al., 2022). At the body level, sirtuins affect metabolism, aging, and carcinogenesis. Changes in the activity of sirtuins in human cells and model organisms are considered to be markers of aging (Kumar et al., 2014; Zhang J. et al., 2016), as well as factors affecting overall life expectancy (Roichman

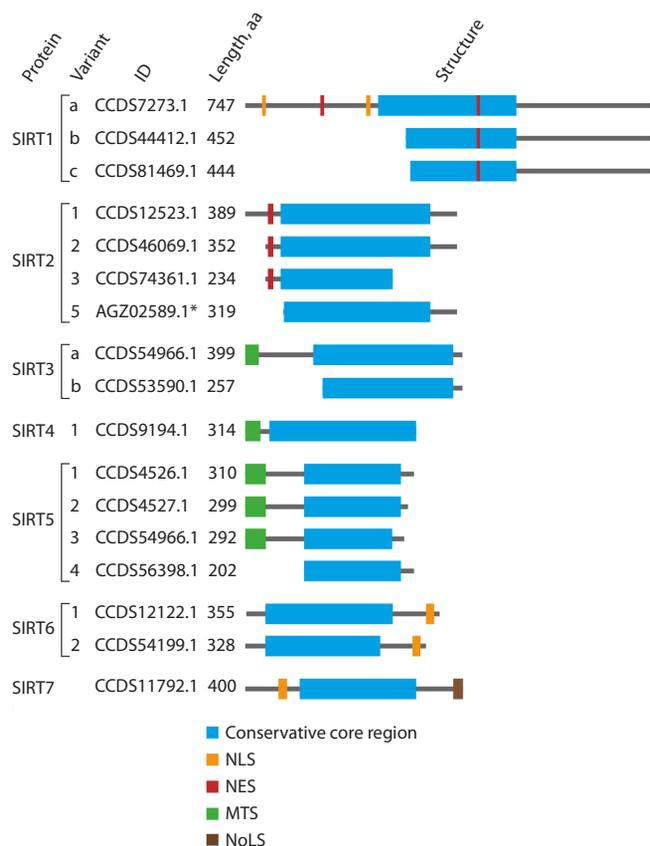


Fig. 1. Human sirtuins diversity. Information on protein isoforms annotated in the CCDS database is provided (Pruitt et al., 2009).

Functional protein regions are marked with colored rectangles. * Amino acid sequence identifier from the NCBI Proteins Database (Sayers et al., 2022).

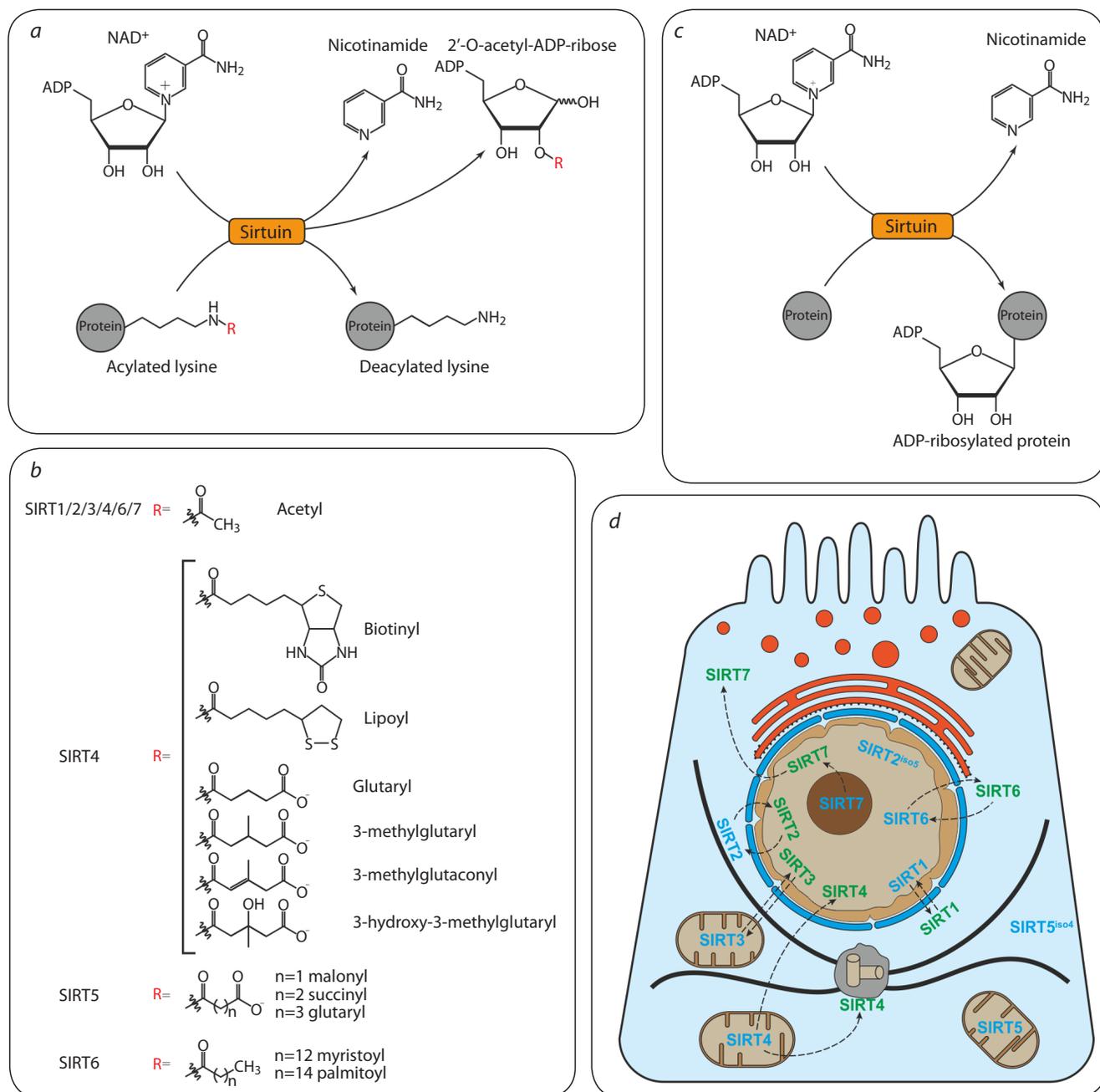


Fig. 2. Human sirtuins molecular function.

a – deacetylase activity of sirtuins; *b* – some functional groups that can be removed by different sirtuins; *c* – ADP-ribosyltransferase reaction involving sirtuins; *d* – localization of human sirtuins in cells. The protein is marked with blue text if it is placed in the area of constitutive localization in the diagram, or green otherwise.

et al., 2021). Within the frames of the presented review, the regulatory activities of sirtuins, their participation in epigenetic regulation and involvement in the genesis of age-related diseases are analyzed.

Sirtuins biochemical activity

First of all, sirtuins are known as enzymes-deacetylases of histones and non-histone proteins (Fig. 2, *a*) (Sauve et al., 2006). In addition, sirtuins are able to remove different acyl residues (see Fig. 2, *a, b*). For example, SIRT5 predominantly

deacetylates lysine residues which are modified by succinyl, malonyl or glutaryl groups (Du J. et al., 2011; Tan et al., 2014). Some sirtuins combine deacetylase and deacylase activity. For instance, SIRT6 can remove residues of myristic and palmitic fatty acids (Jiang et al., 2013; Zhang X. et al., 2017), and SIRT4 can remove residues of lipoic acid, biotin, glutarate and its derivatives (Laurent et al., 2013; Mathias et al., 2014).

Sirtuins are also able to perform mono-ADP-ribosylation of proteins (Frye, 1999). In such cases, sirtuins transfer ADP-

ribose from NAD⁺ directly to the amino acid residues of arginine, yielding nicotinamide (see Fig. 2, c) (Fahie et al., 2009). ADP-ribosyltransferase activity in mammals has now been described for SIRT4, SIRT6, and SIRT7. However, the total number of their targets is small. For example, SIRT4 ADP-ribosylates glutamate dehydrogenase (GDH) in pancreatic beta cells and inhibits its activity by limiting glutamate and glutamine metabolism (Haigis et al., 2006). Under oxidative stress, SIRT6 is recruited to DNA breakpoints and induces damage repair by ADP-ribosylation of poly-ADP-ribosyltransferase PARP1, one of the most important regulators of DNA repair (Mao et al., 2011). Moreover, SIRT6 is recruited in fibroblasts to the 5' untranslated region of LINE1 retrotransposon and suppresses its expression by ADP-ribosylation of transcriptional repressor KAP1 (Van Meter et al., 2014). Auto-ADP-ribosylation of SIRT7 in humans regulates its binding to genes enriched with mH2A1.1 histone modification, which is important for glucose homeostasis (Simonet et al., 2020).

Sirtuin function and modulation in cells

Sirtuins are divided into predominantly nuclear (SIRT1/6/7), cytoplasmic (SIRT2), and mitochondrial (SIRT3/4/5) (Michishita et al., 2005). However, their intracellular localization can change both during the cell cycle and under various stimuli (see Fig. 2, d).

SIRT1 is most often found in the nucleus, where it regulates the structure of chromatin and the activity of many regulatory proteins. It has also been found in the cytoplasm (Bai, Zhang, 2016). SIRT1 is assumed to be present in the nucleus, but under the influence of unknown stimuli it can be transported to the cytoplasm due to a nuclear export signal (NES) (Sun, Fang, 2016). The transfer of SIRT1 into cytoplasm is also observed during cellular aging, accompanied by autophagocytosis of SIRT1 in lysosomes (Xu et al., 2020; Wang L. et al., 2021).

SIRT6 is more commonly found in the nucleus, but it can be observed in the cytoplasm of liver cells in response to increased levels of saturated fatty acids. In this case, it deacetylates and activates acyl-CoA synthetase 5 (ACSL5), one of the fatty acid oxidation enzymes (Hou et al., 2022). In mouse macrophages, the SIRT6 fraction is constantly present in the cytoplasm and stimulates the secretion of TNF- α protein by removing its meristyl modification (Bresque et al., 2022).

SIRT7 is the only sirtuin that is enriched in the nucleoli, where it is involved in the transcription of ribosomal genes (Ford et al., 2006; Kiran et al., 2013). Under stress, when rRNA production is disrupted, SIRT7 moves to the nucleoplasm or cytoplasm (Chen et al., 2013; Zhang P.-Y. et al., 2016), where it interacts with a variety of proteins (Tsai et al., 2012; Lee et al., 2014). Although the exact function of cytoplasmic SIRT7 remains unknown, it is presumably associated with the regulation of replicative senescence (Kiran et al., 2013).

SIRT2 is more often found in the cytoplasm. It is involved in the regulation of cell cycle, fatty acids and carbohydrates metabolism, oxidative stress response and many other processes. In the interphase, it localizes to microtubules and deacetylates α -tubulin (North et al., 2003). During the G2/M

transition, SIRT2 temporarily migrates to the nuclei and deacetylates histone H4 by lysine 16, thereby modulating chromatin condensation (Vaquero et al., 2006). SIRT2 passes from the nucleus to the cytoplasm due to the nuclear export signal (NES) at the N-terminus (North, Verdin, 2007). Its nuclear export depends on posttranslational modifications and can be disrupted by various stimuli (North, Verdin, 2007). For example, infection of HeLa cells with *Listeria monocytogenes* bacterium leads to dephosphorylation of SIRT2 by S25 residue, which leads to an increase in the nuclear concentration of the protein. Here, it mediates the repression of immune response genes by deacetylation of H3K18ac (Eskandarian et al., 2013; Pereira et al., 2018).

The transfer of SIRT2 to the nucleus is also observed in glioblastoma cells and other types of tumors, and patients with higher SIRT2 in the nuclei of tumor cells have a worse prognosis for glioma (Imaoka et al., 2012; Eldridge et al., 2020). Reduction in nuclear exports can also be achieved through alternative splicing. For example, the recently discovered SIRT2^{iso5} isoform with unknown enzymatic activity does not have a nuclear export signal and is constitutively present in the nucleus, where it interacts with histone methyltransferases, and also suppresses transcription and replication of the hepatitis B virus (HBV) (Rack et al., 2014; Piracha et al., 2020).

SIRT3, SIRT4, and SIRT5 are predominantly found in the mitochondrial matrix and play a key role in cellular processes such as oxidative stress response, dissimilation, and apoptosis (Michishita et al., 2005). At the same time, in mice, a disruption of the SIRT3 function leads to a significant increase in acetylation of mitochondrial proteins. Knockout of SIRT4 and SIRT5 has a significantly weaker effect (Lombard et al., 2007; Finkel et al., 2009). However, mitochondrial proteins are also found in other cellular compartments. SIRT3 is detected in the cell nucleus, where it is involved in the regulation of heterochromatin structure and NHEJ-dependent DNA repair (Sengupta, Haldar, 2018; Diao et al., 2021). SIRT4 interacts with the centrosome at the end of the G2 phase, and under mitochondrial stress moves to the nucleus, although the function of this remains unclear (Ramadani-Muja et al., 2019; Bergmann et al., 2020). In addition, only three out of four studied SIRT5 isoforms contain a signal of mitochondrial localization (see Fig. 1). At the same time, the shortest isoform, SIRT5^{iso4}, does not have it, and is found in the cytoplasm of cells (Du Y. et al., 2018).

Sirtuins in chromatin and epigenetic regulation

The key parts of epigenetic mechanisms of gene regulation are specific marks – histone modifications and DNA methylation – as well as effector proteins capable of establishing and recognizing such tags. The coordinated work of such a system determines the properties of chromatin, which, in its turn, determines the activity of genes and the stability of the genome. Sirtuins are involved both in the direct control of histone modifications and in the activity and stability of regulatory factors.

SIRT1 is involved in the formation of heterochromatin by removing H4K16Ac, H3K9Ac, and H1K26Ac marks, as

well as interacting with Suv39h1 histone methyltransferase, which is responsible for the installation of a key histone modification of the H3K9me3 constitutive heterochromatin (Vaquero et al., 2004, 2007; Bosch-Presegué et al., 2011). Thus, SIRT1 participates in the establishment of H3K9me3 by removing H3K9Ac, as well as by direct interaction with Suv39h1, which increases the specific activity of the latter as a result of conformational changes, deacetylation of K266 in the catalytic SET domain, and also increases resistance to proteasomal degradation by suppressing ubiquitination at the K87 site in Suv39h1 chromodomain (Vaquero et al., 2007; Bosch-Presegué et al., 2011). Impaired SIRT1 function leads to a significant loss of HP1 and H3K9me3 in the composition of pericentromeric heterochromatin (Vaquero et al., 2007; Wang R.-H. et al., 2008; El Ramy et al., 2009).

It is important to note that such a significant effect on chromatin structure cannot but affect the expression of other regulatory factors. For example, the activation of SIRT1 has been found to increase proliferation, invasion, and accelerate epithelial-mesenchymal transition in pancreatic tumor cells, which is associated at least in part with suppression of gene expression of the FOXO3 and GRHL3 transcription factors (Leng et al., 2021).

SIRT1 is also involved in the regulation of DNA methylation, both at the level of transcription regulation and directly modulating the activity of DNA methyltransferases. Thus, in a mouse embryonic stem cell model, it was shown that SIRT1 deacetylates H1 and H4 histones in the promoter region of the *Dnmt3l* gene, suppressing its expression (Heo et al., 2017). A deficiency of SIRT1 leads to increased methylation of genomic DNA, as well as to deregulation of imprinted genes (Heo et al., 2017). It is interesting to note that the same study showed that SIRT1 is able to deacetylate the Dnmt3l protein, which reduces its stability (Heo et al., 2017). The effect of SIRT1 on human DNMT1 DNA methyltransferase has also been demonstrated. Moreover, deacetylation of lysine residues in the catalytic domain led to an increase in methyltransferase activity, and deacetylation in GK linker region led to its decrease (Peng et al., 2011). During differentiation of human macrophages, SIRT1 and SIRT2 physically interact with the DNMT3B DNA methyltransferase to prevent aberrant activation of pro-inflammatory genes (Li T. et al., 2020).

In addition to DNA methyltransferases, SIRT1 also affects the activity of many other non-histone targets. Thus, deacetylation of p53 protein under the action of SIRT1 led to the repression of apoptosis in H1299 cells (Luo et al., 2001; Vaziri et al., 2001). In addition, deacetylation of p53 leads to suppression of its regulatory activity as an oncosuppressor (Ong, Ramasamy, 2018).

Deacetylation of Ku70 protein, one of the key components of NHEJ DNA repair, by SIRT1 activates DNA repair (Jeong et al., 2007). These examples show that SIRT1 stimulates cell survival in case of DNA damage. However, SIRT1 is also able to deacetylate the p65 subunit of NF- κ B, which on the contrary leads to activation of TNF- α -induced apoptosis in non-small cell lung cancer cells (Yeung et al., 2004).

Deacetylation of FOXO transcription factors (FOXO1, FOXO3, FOXO4) under the action of SIRT1 can lead to both

activation of apoptosis (FOXO1) and cell cycle arrest and suppression of apoptosis (FOXO3) (Brunet et al., 2004; Yang et al., 2005). In turn, deacetylation of FOXO4 increases its protective effect under oxidative stress (van der Horst et al., 2004). In addition to the above, SIRT1 is involved in regulating the activity of transcription factors that control the response to hypoxia, metabolism, cell invasion, and proliferation.

SIRT3 is largely described as a regulator of mitochondrial functions. However, in nuclei, its role in NHEJ DNA repair due to the removal of H3K56ac histone modification has been shown (Sengupta, Haldar, 2018). SIRT3 also plays a role in deacetylation of H3 histone at lysine residue 27, which is associated with repression of transcription of the FOS transcription factor gene and prevention of TNF- α -induced inflammatory and profibrotic responses in rat cardiomyocytes (Palomer et al., 2020).

At the chromatin level in HEK293T cells, SIRT3 has been shown to be able to directly interact with nuclear lamina components LaminB1 and LBR, as well as the HP1 α , HP1 γ , and KAP1 heterochromatin proteins (Diao et al., 2021). Moreover, deletion of SIRT3 in human mesenchymal stem cells (MSCs) resulted in dissociation of lamina-associated domains and a decrease in the abundance of the H3K9me3, HP1 α , and KAP1 heterochromatin protein markers, as well as the LaminB1 nuclear membrane component (Diao et al., 2021). Restoring SIRT3 had the opposite effect (Diao et al., 2021). At the same time, it is important to note that deletion of SIRT3 led to accelerated cellular senescence, and all the detected effects are (among other things) its classical manifestations. In this regard, despite the obviousness of the phenotype, as well as the participation of SIRT3 in the regulation of other processes associated with cellular aging, the cause-and-effect relationship may be more complex.

SIRT6 is also involved in epigenetic regulation. One of its first discovered activities at the level of chromatin regulation was the ability to deacetylate the H3 histone by K9 and K56 (Kawahara et al., 2009). SIRT6 acts as a co-repressor of such transcription factors as NF- κ B and HIF-1 α by regulating the H3K9 acetylation (Kawahara et al., 2009; Zhong et al., 2010). Deacetylation of H3K9 under the action of SIRT6 is involved in the regulation of telomeres and gene expression (Michishita et al., 2008; Zhong et al., 2010). For example, the control of mouse embryonic stem cell differentiation depends on the removal of the H3K56ac and H3K9ac marks in the promoter regions of the *Oct4*, *Sox2*, and *Nanog* genes (Etchegaray et al., 2015). In addition, SIRT6 is able to directly deacetylate DNMT1 DNA methyltransferase, which reduces its stability (Jia et al., 2021; Subramani et al., 2023).

Increased expression of SIRT6 leads to destabilization of DNMT1 and a decrease in methylation of the *NOTCH1* and *NOTCH2* gene promoters, which leads to a predominant osteogenic differentiation of the MSCs from adipose tissue (Jia et al., 2021). In non-small cell lung cancer cell cultures, decreased SIRT6 expression leads to stabilization of DNMT1 and methylation of the promoter of the *NOTCH1* gene, which is involved in oncogenesis and metastasis (Subramani et al., 2023). As mentioned earlier, SIRT6, through ADP-ribosylation of the KAP1 transcription repressor, is involved in suppress-

ing LINE1 expression and maintaining genome stability (Van Meter et al., 2014).

SIRT7 is the only sirtuin localized in the nucleolus, where it plays a key role in the formation of transcriptionally inactive heterochromatin by recruiting DNMT1, SIRT1, and SMARCA5 to ribosomal DNA repeats (Ianni et al., 2017; Paredes et al., 2018). Compact state of chromatin is needed to prevent homologous recombination between repetitive rDNA sequences. Therefore, a disruption of the SIRT7 function leads to the formation of active chromatin, instability of the rDNA region and genome, and accelerated cellular senescence (Ianni et al., 2017; Paredes et al., 2018). In addition, SIRT7 is involved in the regulation of R-loop – RNA-DNA complexes, which are also a potential factor in genome instability (Aguilera, García-Muse, 2012; Song et al., 2017).

It is interesting to note that in the *Drosophila* model, it was shown that the area of distribution of R-loops increases with age, while their number remains unchanged (Hall, 2023). Defects in the processing of R-loops can lead to accumulation of DNA/RNA hybrids, single-stranded DNA fragments in the cytoplasm, which stimulates the immune response, chronic inflammation, apoptosis and senescence (Chatzidoukaki et al., 2021; Crossley et al., 2023). SIRT7 deacetylates and activates the DDX21 helicase involved in R-loop resolving (Song et al., 2017).

In addition, the role of SIRT7 in the deacetylation of the H3K18ac histone modification, which is needed to activate the repair of double-strand breaks (DSBs) in DNA, has been shown (Barber et al., 2012; Lin et al., 2016b; Vazquez et al., 2016). Direct interaction with the KAP1, HP1 α , and HP1 γ heterochromatin proteins, as well as components of the LBR and LaminB1 nuclear lamina in the HEK293T cells was also demonstrated for SIRT7 (Bi et al., 2020). Reduced SIRT7 in human MSCs led to accelerated senescence, heterochromatin destabilization, awakening of repeated sequences, and the cGAS-STING proinflammatory signaling pathway activation (Bi et al., 2020).

Sirtuins and longevity

For the first time, the possible influence of sirtuins on longevity was discovered in trials with an overexpression of Sir2, which led to an increase in the number of budding cycles in the *S. cerevisiae* yeast (Kaerberlein et al., 1999). In studies of the Sir2 homologues, with SIR-2.1 protein in the *Caenorhabditis elegans* nematode and dSirt1 in *Drosophila melanogaster*, an increased lifespan was observed with their overexpression (Tissenbaum, Guarente, 2001; Rogina, Helfand, 2004). At the same time, an overexpression of Sir2 in yeast increases the replicative potential, but does not regulate the lifespan of quiescent cells, which is a key parameter in the chronological aging model of *S. cerevisiae* (Fabrizio et al., 2005). In nematodes, the positive effect of sirtuins depends on the genetic background and is not detected in some laboratories (Burnett et al., 2011; Viswanathan, Guarente, 2011; Schmeisser et al., 2013; Zhao et al., 2019).

In *Drosophila*, the effect of dSirt1 overexpression on longevity is dose-dependent. At the same time, exceeding a certain expression threshold may shorten the lifespan due to

its toxicity to certain organs (Griswold et al., 2008; Burnett et al., 2011; Whitaker et al., 2013). In particular, it has been demonstrated that induced tissue-specific overexpression of dSirt1 increases the median lifespan of flies only when the transgene is activated in adipose tissue, but not in muscles (Banerjee et al., 2012). Similarly, a 9–16 % increase in mouse lifespan was achieved by triggering transgenic SIRT1 specifically in hypothalamic cells (Satoh et al., 2013), whereas in an earlier study, an overexpression of SIRT1 in mice did not affect longevity, although it reduced the likelihood of cancer (Herranz et al., 2010).

The effect of overexpression of sirtuins on longevity has also been shown for dSirt4 and dSirt6, the *Drosophila* sirtuins (Wood et al., 2018; Taylor et al., 2022). In a recent study, slowed aging and increased maximum lifespan under the influence of SIRT6 were shown in mice (Roichman et al., 2021). The effect of SIRT6 on longevity is associated with its participation in DNA repair (Tian et al., 2019). Moreover, the activity of species-specific SIRT6 variants in the context of repair correlates with the maximum lifespan of different rodent species (Tian et al., 2019).

Conflicting data have been obtained for SIRT7. In particular, it has recently been demonstrated that male mice with SIRT7 knockout have an increased median lifespan and exhibit a slower decrease in physiological parameters (Mizumoto et al., 2022). This result contrasts with observations from earlier studies, in which the SIRT7 knockout significantly shortened the lifespan. It is important to note that none of the experiments evaluating the effect of sirtuins on organismal longevity showed an extreme increase in the lifespan, and along with the difficulty in selecting correct experimental controls – as genetically close as possible – this leads to ambiguous conclusions about the role of sirtuins as autonomous factors of longevity (Brenner, 2022).

Data on the possible relationship between sirtuins and human life expectancy are, to some extent, confirmed by population genetics data. For example, the research of Dutch centenarians showed that carriers of the rs12778366 single nucleotide polymorphism have better glucose tolerance and a reduced risk of death (Figarska et al., 2013). The research on the Americans of European descent and populations of Georgia and Louisiana demonstrated the association of the rs7896005 polymorphism of the *SIRT1* gene with longevity and telomere length in lymphocytes (Kim et al., 2012). In addition, *SIRT1* rs3758391 and rs4746720 polymorphisms were associated with healthy aging in the Han Chinese (Zhang W.-G. et al., 2010). However, in a similar research of another population of Chinese centenarians, the relationship of these loci with life expectancy was not revealed (Lin et al., 2016a). A number of other studies have not revealed the association of genetic variants of the *SIRT1* gene with longevity (Flachsbart et al., 2006; Willcox et al., 2008; Soerensen et al., 2013).

In the fifth intron of the *SIRT3* gene, variability in the number of tandem repeats (VNTR) was identified, some variants of which acquire the properties of an allele-specific enhancer. The allele without this enhancer activity was practically not found among men over 90 years old, while no such correlation was

observed in women (Bellizzi et al., 2005). Polymorphisms of the *SIRT3* gene, rs11555236 and rs4980329, were associated with the life expectancy of women in an Italian population (Albani et al., 2014).

The rs107251 nucleotide polymorphism of the *SIRT6* gene was associated with a more than five-year increase in life expectancy in the elderly in the United States, and the rs117385980 polymorphism was associated with longevity in Finns (TenNapel et al., 2014; Hirvonen et al., 2017). The *SIRT6* allele with N308K/A313S substitutions, which has strong ADP-ribosyltransferase activity, was enriched in a group of centenarians among Ashkenazi Jews (Simon et al., 2022).

Sirtuins in chronic inflammation

In addition to some evidence of a possible functional relationship with longevity, sirtuins have been shown to play a significant role in the development of age-associated diseases, in particular those arising from chronic inflammation. The *SIRT1*, *SIRT2*, and *SIRT6* proteins counteract the inflammatory response by suppressing the NF- κ B signaling pathway (Vazquez et al., 2021). The key element of this pathway – the transcription factor NF- κ B, which controls the expression of immune response genes, consists of five subunits: p50, p52, p65 (RelA), RelB, and c-Rel (Vazquez et al., 2021).

There are several mechanisms that sirtuins can use to suppress the activity of the NF- κ B signaling pathway. *SIRT1* and *SIRT2* are able to deacetylate the p65 subunit at lysine 310, directly inhibiting the activity of NF- κ B. They can also prevent methylation of neighboring lysine residues (K314 and K315), which contribute to ubiquitination and degradation of p65 (Rothgiesser et al., 2010). *SIRT6* interacts with p65 and deacetylates H3K9 in promoters of NF- κ B target genes, thereby reducing inflammation (Kawahara et al., 2009). *SIRT1* is also able to deacetylate and inhibit the activity of NF- κ B transcriptional coactivators, such as PARP-1 and p300 histone acetyl transferase (Rajamohan et al., 2009).

Sirtuins also influence the TGF- β signaling pathway – which plays a key role in tubulointerstitial renal fibrosis – by stimulating the production of connective tissue growth factor (CTGF) (Isaka, 2018). Overexpression of *SIRT1* suppresses TGF- β 1-induced cellular apoptosis and fibrosis, and reduces CTGF expression by stimulating TGF- β 1 in the kidneys of mice with unilateral ureteral obstruction (Ren et al., 2015). *SIRT1* is also able to weaken TGF- β -dependent signaling by deacetylating SMAD3 and SMAD4 molecules, which inhibits the production of collagen, fibronectin, and MMP7 metalloprotease (Zhang Y. et al., 2017).

The effect of sirtuins on the Smad transcription factors is also important in cardiac fibrosis. This is because systematic knockout of the mouse *SIRT6* gene disrupts inhibition of the TGF- β /Smad3 signaling pathway, the cause of cardiac fibrosis (Maity et al., 2020). In addition, *SIRT1* produces a cardioprotective effect by deacetylating SMAD2/3 and reducing the activity of the TGF- β signaling pathway in mouse cardiac fibroblasts (Bugyei-Twum et al., 2018).

The *SIRT3* protein has antifibrotic properties that weaken TGF- β -dependent signaling, and the suppression of *SIRT3*

activity can lead to the transformation of mouse and human cardiac fibroblasts into myofibroblasts – cells capable of producing extracellular matrix (Sundaresan et al., 2016). As expected, activators of sirtuins counteract fibrosis. Thus, honokiol – a *SIRT3* activator – counteracts kidney fibrosis in mice with unilateral ureteral obstruction (Quan et al., 2020). Similarly, activation of both *SIRT1* and *SIRT3* by resveratrol attenuates cardiac fibrosis in mice by inhibiting the TGF- β /Smad3 pathway (Liu et al., 2019).

The physiological effect of sirtuins in inflammation is also directly related to the effect on immune cells. For example, *SIRT1* is involved in the transmission of inflammatory signals in mouse dendritic cells by modulating the balance of type 1 pro-inflammatory T helper cells and Foxp3(+) anti-inflammatory regulatory T cells. A deficiency of *SIRT6* in macrophages leads to inflammation with increased acetylation and greater stability of FoxO1 (Woo et al., 2016, 2018). *SIRT4* also has an anti-inflammatory effect, since its deficiency can increase inflammation and promote macrophage infiltration and the development of cellular hepatocarcinoma in humans (Li Z. et al., 2019). In mouse liver cells, *SIRT3* inhibits the production of pro-inflammatory chemokines and some profibrotic factors (LoBianco et al., 2020).

As they do in the case of chronic inflammation, in neuroinflammation, sirtuins have mainly an anti-inflammatory effect. However, the literature describes exceptions. For example, inhibition of *SIRT2* in mice with accelerated cellular senescence reduced neuroinflammation, as evidenced by reduced glial fibrillar acid protein, IL-1 β , IL-6, and TNF- α and increased glutamate receptor subunits GluN2A, GluN2B, and GluA1. However, inhibition of *SIRT2* could not reverse cognitive decline or neuroinflammation (Diaz-Perdigon et al., 2020). In this case, *SIRT2* demonstrated a temporary pro-inflammatory effect.

Neurodegenerative diseases correlate with aging, as do changes in sirtuin expression (Julien et al., 2009; Jiao, Gong, 2020). It is interesting to note that age-related changes in serum sirtuin can be used as a diagnostic tool (Kumar et al., 2014). For instance, the expression of *SIRT1* and *SIRT6* is reduced against the background of neurodegenerative diseases (Jiao, Gong, 2020; Pukhalskaia et al., 2020). A high content of *SIRT2* is found in Alzheimer's and Parkinson's diseases, suggesting that it may contribute to neurodegeneration (Cabelos et al., 2019).

Conclusion

Since the discovery of yeast Sir2, studies of sirtuins have focused on their functions in regulating processes associated with aging (Pukhalskaia et al., 2022). Recent research confirms the key role of sirtuins in the pathogenesis of age-related diseases. This makes sirtuins promising targets for research in the field of age-related disease therapy. Indeed, currently underway are many clinical trials aimed at pharmacological modulation of sirtuin activity for the treatment of metabolic, immune, and neurological disorders, as well as cardiovascular and oncological diseases (Curry et al., 2021). Unfortunately, these clinical trials do not always show positive results, which, in all likelihood, may be due to the highly multifaceted functions of

sirtuins. However, detailed information about their functions is dynamically accumulated, which hopefully will allow for the implementation of such progressive methods of therapy for age-dependent diseases as soon as reasonably possible.

References

- Aguilera A., García-Muse T. R loops: from transcription byproducts to threats to genome stability. *Mol. Cell.* 2012;46(2):115-124. DOI 10.1016/j.molcel.2012.04.009
- Albani D., Ateri E., Mazzuco S., Ghilardi A., Rodilossi S., Biella G., Ongaro F., Antuono P., Boldrini P., Di Giorgi E., Frigato A., Durante E., Caberlotto L., Zanardo A., Siculi M., Gallucci M., Forloni G. Modulation of human longevity by *SIRT3* single nucleotide polymorphisms in the prospective study “Treviso Longeva (TRELONG).” *Age (Dordr.)*. 2014;36(1):469-478. DOI 10.1007/s11357-013-9559-2
- Bai W., Zhang X. Nucleus or cytoplasm? The mysterious case of SIRT1’s subcellular localization. *Cell Cycle*. 2016;15(24):3337-3338. DOI 10.1080/15384101.2016.1237170
- Banerjee K.Kr., Ayyub C., Ali S.Z., Mandot V., Prasad N.G., Kolthur-Seetharam U. dSir2 in the adult fat body, but not in muscles, regulates life span in a diet-dependent manner. *Cell Rep.* 2012;2(6):1485-1491. DOI 10.1016/j.celrep.2012.11.013
- Barber M.F., Michishita-Kioi E., Xi Y., Tasselli L., Kioi M., Moqtaderi Z., Tennen R.I., Paredes S., Young N.L., Chen K., Struhl K., Garcia B.A., Gozani O., Li W., Chua K.F. SIRT7 links H3K18 deacetylation to maintenance of oncogenic transformation. *Nature*. 2012;487(7405):114-118. DOI 10.1038/nature11043
- Bellizzi D., Rose G., Cavalcante P., Covello G., Dato S., De Rango F., Greco V., Maggiolini M., Feraco E., Mari V., Franceschi C., Passarino G., De Benedictis G. A novel VNTR enhancer within the *SIRT3* gene, a human homologue of *SIR2*, is associated with survival at oldest ages. *Genomics*. 2005;85(2):258-263. DOI 10.1016/j.ygeno.2004.11.003
- Bergmann L., Lang A., Bross C., Altinolu-Hambüchen S., Fey I., Overbeck N., Stefanski A., Wiek C., Kefalas A., Verhülsdonk P., Mielke C., Sohn D., Stühler K., Hanenberg H., Jänicke R.U., Scheller J., Reichert A.S., Ahmadian M.R., Piekorz R.P. Subcellular localization and mitotic interactome analyses identify SIRT4 as a centrosomally localized and microtubule associated protein. *Cells*. 2020;9(9):1950. DOI 10.3390/cells9091950
- Bi S., Liu Z., Wu Z., Wang Z., Liu X., Wang S., Ren J., Yao Y., Zhang W., Song M., Liu G.-H., Qu J. SIRT7 antagonizes human stem cell aging as a heterochromatin stabilizer. *Protein Cell*. 2020;11(7):483-504. DOI 10.1007/s13238-020-00728-4
- Bosch-Presegué L., Raurell-Vila H., Marazuela-Duque A., Kane-Goldsmith N., Valle A., Oliver J., Serrano L., Vaquero A. Stabilization of Suv39H1 by SirT1 is part of oxidative stress response and ensures genome protection. *Mol. Cell*. 2011;42(2):210-223. DOI 10.1016/j.molcel.2011.02.034
- Brenner C. Sirtuins are not conserved longevity genes. *Life Metab.* 2022;1(2):122-133. DOI 10.1093/lifemeta/loac025
- Bresque M., Cal K., Pérez-Torrado V., Colman L., Rodríguez-Duarte J., Vilaseca C., Santos L., Garat M.P., Ruiz S., Evans F., Dapuerto R., Contreras P., Calliari A., Escande C. SIRT6 stabilization and cytoplasmic localization in macrophages regulates acute and chronic inflammation in mice. *J. Biol. Chem.* 2022;298(3):101711. DOI 10.1016/j.jbc.2022.101711
- Brunet A., Sweeney L.B., Sturgill J.F., Chua K.F., Greer P.L., Lin Y., Tran H., Ross S.E., Mostoslavsky R., Cohen H.Y., Hu L.S., Cheng H.-L., Jedrychowski M.P., Gygi S.P., Sinclair D.A., Alt F.W., Greenberg M.E. Stress-dependent regulation of FOXO transcription factors by the SIRT1 deacetylase. *Science*. 2004;303(5666):2011-2015. DOI 10.1126/science.1094637
- Bryk M., Banerjee M., Murphy M., Knudsen K.E., Garfinkel D.J., Curcio M.J. Transcriptional silencing of Ty1 elements in the *RDNI* locus of yeast. *Genes Dev.* 1997;11(2):255-269. DOI 10.1101/gad.11.2.255
- Bugyei-Twum A., Ford C., Civitarese R., Seegobin J., Advani S.L., Desjardins J.-F., Kabir G., Zhang Y., Mitchell M., Switzer J., Thai K., Shen V., Abadeh A., Singh K.K., Billia F., Advani A., Gilbert R.E., Connelly K.A. Sirtuin 1 activation attenuates cardiac fibrosis in a rodent pressure overload model by modifying Smad2/3 transactivation. *Cardiovasc. Res.* 2018;114(12):1629-1641. DOI 10.1093/cvr/cvy131
- Burnett C., Valentini S., Cabreiro F., Goss M., Somogyvári M., Piper M.D., Hodginott M., Sutphin G.L., Leko V., McElwee J.J., Vazquez-Manrique R.P., Orfila A.-M., Ackerman D., Au C., Vinti G., Riesen M., Howard K., Neri C., Bedalov A., Kaeberlein M., Söti C., Partridge L., Gems D. Absence of effects of Sir2 overexpression on lifespan in *C. elegans* and *Drosophila*. *Nature*. 2011;477(7365):482-485. DOI 10.1038/nature10296
- Cacabelos R., Carril J., Cacabelos N., Kazantsev A., Vostrov A., Corzo L., Cacabelos P., Goldgaber D. Sirtuins in Alzheimer’s disease: SIRT2-related genotypes and implications for pharmacogenetics. *Int. J. Mol. Sci.* 2019;20(5):1249. DOI 10.3390/ijms20051249
- Chatzidoukaki O., Stratigi K., Goulielmaki E., Niotis G., Akalestou-Clocher A., Gkirtzimanaki K., Zafeiropoulos A., Altmüller J., Topalis P., Garinis G.A. R-loops trigger the release of cytoplasmic ssDNAs leading to chronic inflammation upon DNA damage. *Sci. Adv.* 2021;7(47):eabj5769. DOI 10.1126/sciadv.abj5769
- Chen S., Seiler J., Santiago-Reichert M., Felbel K., Grummt I., Voit R. Repression of RNA polymerase I upon stress is caused by inhibition of RNA-dependent deacetylation of PAF53 by SIRT7. *Mol. Cell*. 2013;52(3):303-313. DOI 10.1016/j.molcel.2013.10.010
- Crossley M.P., Song C., Bocek M.J., Choi J.-H., Kousorous J., Sathirachinda A., Lin C., Brickner J.R., Bai G., Lans H., Vermeulen W., Abu-Remaileh M., Cimprich K.A. R-loop-derived cytoplasmic RNA-DNA hybrids activate an immune response. *Nature*. 2023;613(7942):187-194. DOI 10.1038/s41586-022-05545-9
- Curry A.M., White D.S., Donu D., Cen Y. Human sirtuin regulators: The “success” stories. *Front. Physiol.* 2021;12:752117. DOI 10.3389/fphys.2021.752117
- Diao Z., Ji Q., Wu Z., Zhang W., Cai Y., Wang Z., Hu J., Liu Z., Wang Q., Bi S., Huang D., Ji Z., Liu G.-H., Wang S., Song M., Qu J. SIRT3 consolidates heterochromatin and counteracts senescence. *Nucleic Acids Res.* 2021;49(8):4203-4219. DOI 10.1093/nar/gkab161
- Diaz-Perdigon T., Belloch F.B., Ricobaraza A., Elboray E.E., Suzuki T., Tordera R.M., Puerta E. Early sirtuin 2 inhibition prevents age-related cognitive decline in a senescence-accelerated mouse model. *Neuropharmacology*. 2020;45(2):347-357. DOI 10.1038/s41386-019-0503-8
- Du J., Zhou Y., Su X., Yu J.J., Khan S., Jiang H., Kim J., Woo J., Kim J.H., Choi B.H., He B., Chen W., Zhang S., Cerione R.A., Auwerx J., Hao Q., Lin H. Sirt5 is a NAD-dependent protein lysine demalonylase and desuccinylase. *Science*. 2011;334(6057):806-809. DOI 10.1126/science.1207861
- Du Y., Hu H., Hua C., Du K., Wei T. Tissue distribution, subcellular localization, and enzymatic activity analysis of human SIRT5 isoforms. *Biochem. Biophys. Res. Commun.* 2018;503(2):763-769. DOI 10.1016/j.bbrc.2018.06.073
- El Ramy R., Magroun N., Messadecq N., Gauthier L.R., Boussin F.D., Kolthur-Seetharam U., Schreiber V., McBurney M.W., Sassone-Corsi P., Dantzer F. Functional interplay between Parp-1 and SirT1 in genome integrity and chromatin-based processes. *Cell. Mol. Life Sci.* 2009;66(19):3219-3234. DOI 10.1007/s00018-009-0105-4
- Eldridge M.J.G., Pereira J.M., Impens F., Hamon M.A. Active nuclear import of the deacetylase Sirtuin-2 is controlled by its C-terminus

- and importins. *Sci. Rep.* 2020;10(1):2034. DOI 10.1038/s41598-020-58397-6
- Eskandarian H.A., Impens F., Nahori M.-A., Soubigou G., Coppe J.-Y., Cossart P., Hamon M.A. A role for SIRT2-dependent histone H3K18 deacetylation in bacterial infection. *Science.* 2013; 341(6145):1238858. DOI 10.1126/science.1238858
- Etcheagaray J.-P., Chavez L., Huang Y., Ross K.N., Choi J., Martinez-Pastor B., Walsh R.M., Sommer C.A., Lienhard M., Gladden A., Kugel S., Silberman D.M., Ramaswamy S., Mostoslavsky G., Hochedlinger K., Goren A., Rao A., Mostoslavsky R. The histone deacetylase SIRT6 controls embryonic stem cell fate via TET-mediated production of 5-hydroxymethylcytosine. *Nat. Cell Biol.* 2015; 17(5): 545-557. DOI 10.1038/ncb3147
- Fabrizio P., Gattazzo C., Battistella L., Wei M., Cheng C., McGrew K., Longo V.D. Sir2 blocks extreme life-span extension. *Cell.* 2005; 123(4):655-667. DOI 10.1016/j.cell.2005.08.042
- Fahie K., Hu P., Swatkoski S., Cotter R.J., Zhang Y., Wolberger C. Side chain specificity of ADP-ribosylation by a sirtuin. *FEBS J.* 2009;276(23):7159-7176. DOI 10.1111/j.1742-4658.2009.07427.x
- Figarska S.M., Vonk J.M., Boezen H.M. *SIRT1* polymorphism, long-term survival and glucose tolerance in the general population. *PLoS One.* 2013;8(3):e58636. DOI 10.1371/journal.pone.0058636
- Finkel T., Deng C.-X., Mostoslavsky R. Recent progress in the biology and physiology of sirtuins. *Nature.* 2009;460(7255):587-591. DOI 10.1038/nature08197
- Flachsbart F., Croucher P.J.P., Nikolaus S., Hampe J., Cordes C., Schreiber S., Nebel A. Sirtuin 1 (*SIRT1*) sequence variation is not associated with exceptional human longevity. *Exp. Gerontol.* 2006; 41(1):98-102. DOI 10.1016/j.exger.2005.09.008
- Ford E., Voit R., Liszt G., Magin C., Grummt I., Guarente L. Mammalian Sir2 homolog SIRT7 is an activator of RNA polymerase I transcription. *Genes Dev.* 2006;20(9):1075-1080. DOI 10.1101/gad.1399706
- Frye R.A. Characterization of five human cDNAs with homology to the yeast SIR2 gene: Sir2-like proteins (sirtuins) metabolize NAD and may have protein ADP-ribosyltransferase activity. *Biochem. Biophys. Res. Commun.* 1999;260(1):273-279. DOI 10.1006/bbrc.1999.0897
- Frye R.A. Phylogenetic classification of prokaryotic and eukaryotic Sir2-like proteins. *Biochem. Biophys. Res. Commun.* 2000;273(2): 793-798. DOI 10.1006/bbrc.2000.3000
- Gottschling D.E., Aparicio O.M., Billington B.L., Zakian V.A. Position effect at *S. cerevisiae* telomeres: Reversible repression of Pol II transcription. *Cell.* 1990;63(4):751-762. DOI 10.1016/0092-8674(90) 90141-Z
- Gray S.G., Ekström T.J. The human histone deacetylase family. *Exp. Cell Res.* 2001;262(2):75-83. DOI 10.1006/excr.2000.5080
- Griswold A.J., Chang K.T., Runko A.P., Knight M.A., Min K.-T. Sir2 mediates apoptosis through JNK-dependent pathways in *Drosophila*. *Proc. Natl. Acad. Sci. USA.* 2008;105(25):8673-8678. DOI 10.1073/pnas.0803837105
- Haigis M.C., Mostoslavsky R., Haigis K.M., Fahie K., Christodoulou D.C., Murphy A.J., Valenzuela D.M., Yancopoulos G.D., Karow M., Blander G., Wolberger C., Prolla T.A., Weindrich R., Alt F.W., Guarente L. SIRT4 inhibits glutamate dehydrogenase and opposes the effects of calorie restriction in pancreatic β cells. *Cell.* 2006;126(5):941-954. DOI 10.1016/j.cell.2006.06.057
- Hall H. R-loops in neuronal aging. *Aging.* 2023;15(17):8535-8536. DOI 10.18632/aging.205070
- Heo J., Lim J., Lee S., Jeong J., Kang H., Kim Y., Kang J.W., Yu H.Y., Jeong E.M., Kim K., Kucia M., Waigel S.J., Zacharias W., Chen Y., Kim I.-G., Ratajczak M.Z., Shin D.-M. *Sirt1* regulates DNA methylation and differentiation potential of embryonic stem cells by antagonizing *Dnmt3l*. *Cell Rep.* 2017;18(8):1930-1945. DOI 10.1016/j.celrep.2017.01.074
- Herranz D., Muñoz-Martin M., Cañamero M., Mulero F., Martinez-Pastor B., Fernandez-Capetillo O., Serrano M. Sirt1 improves healthy ageing and protects from metabolic syndrome-associated cancer. *Nat. Commun.* 2010;1(1):3. DOI 10.1038/ncomms1001
- Hirvonen K., Laivuori H., Lahti J., Strandberg T., Eriksson J.G., Hackman P. SIRT6 polymorphism rs117385980 is associated with longevity and healthy aging in Finnish men. *BMC Med. Genet.* 2017; 18(1):41. DOI 10.1186/s12881-017-0401-z
- Hou T., Tian Y., Cao Z., Zhang J., Feng T., Tao W., Sun H., Wen H., Lu Xiaopeng, Zhu Q., Li M., Lu X., Liu B., Zhao Y., Yang Y., Zhu W.-G. Cytoplasmic SIRT6-mediated ACSL5 deacetylation impedes nonalcoholic fatty liver disease by facilitating hepatic fatty acid oxidation. *Mol. Cell.* 2022;82(21):4099-4115.e9. DOI 10.1016/j.molcel.2022.09.018
- Ianni A., Hoelper S., Krueger M., Braun T., Bober E. Sirt7 stabilizes rDNA heterochromatin through recruitment of DNMT1 and Sirt1. *Biochem. Biophys. Res. Commun.* 2017;492(3):434-440. DOI 10.1016/j.bbrc.2017.08.081
- Imai S., Armstrong C.M., Kaerberlein M., Guarente L. Transcriptional silencing and longevity protein Sir2 is an NAD-dependent histone deacetylase. *Nature.* 2000;403(6771):795-800. DOI 10.1038/35001622
- Imaoka N., Hiratsuka M., Osaki M., Kamitani H., Kambe A., Fukuoaka J., Kurimoto M., Nagai S., Okada F., Watanabe T., Ohama E., Kato S., Oshimura M. Prognostic significance of sirtuin 2 protein nuclear localization in glioma: an immunohistochemical study. *Oncol. Rep.* 2012;28(3):923-230. DOI 10.3892/or.2012.1872
- Isaka Y. Targeting TGF- β signaling in kidney fibrosis. *Int. J. Mol. Sci.* 2018;19(9):2532. DOI 10.3390/ijms19092532
- Ivy J.M., Klar A.J., Hicks J.B. Cloning and characterization of four *SIR* genes of *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 1986;6(2): 688-702. DOI 10.1128/MCB.6.2.688
- Jeong J., Juhn K., Lee H., Kim S.-H., Min B.-H., Lee K.-M., Cho M.-H., Park G.-H., Lee K.-H. SIRT1 promotes DNA repair activity and deacetylation of Ku70. *Exp. Mol. Med.* 2007;39(1):8-13. DOI 10.1038/emm.2007.2
- Jia B., Chen J., Wang Q., Sun X., Han J., Guastaldi F., Xiang S., Ye Q., He Y. SIRT6 promotes osteogenic differentiation of adipose-derived mesenchymal stem cells through antagonizing DNMT1. *Front. Cell Dev. Biol.* 2021;9:648627. DOI 10.3389/fcell.2021.648627
- Jiang H., Khan S., Wang Y., Charron G., He B., Sebastian C., Du J., Kim R., Ge E., Mostoslavsky R., Hang H.C., Hao Q., Lin H. SIRT6 regulates TNF- α secretion through hydrolysis of long-chain fatty acyl lysine. *Nature.* 2013;496(7443):110-113. DOI 10.1038/nature 12038
- Jiao F., Gong Z. The beneficial roles of SIRT1 in neuroinflammation-related diseases. *Oxid. Med. Cell. Longev.* 2020;2020:6782872. DOI 10.1155/2020/6782872
- Julien C., Tremblay C., Émond V., Lebbadi M., Salem N., Bennett D.A., Calon F. Sirtuin 1 reduction parallels the accumulation of tau in Alzheimer disease. *J. Neuropathol. Exp. Neurol.* 2009;68(1):48-58. DOI 10.1097/NEN.0b013e3181922348
- Kaerberlein M., McVey M., Guarente L. The *SIR2/3/4* complex and *SIR2* alone promote longevity in *Saccharomyces cerevisiae* by two different mechanisms. *Genes Dev.* 1999;13(19):2570-2580. DOI 10.1101/gad.13.19.2570
- Kawahara T.L.A., Michishita E., Adler A.S., Damian M., Berber E., Lin M., McCord R.A., Ongaiqui K.C.L., Boxer L.D., Chang H.Y., Chua K.F. SIRT6 links histone H3 lysine 9 deacetylation to NF- κ B-dependent gene expression and organismal life span. *Cell.* 2009; 136(1):62-74. DOI 10.1016/j.cell.2008.10.052

- Kim S., Bi X., Czarny-Ratajczak M., Dai J., Welsh D.A., Myers L., Welsch M.A., Cherry K.E., Arnold J., Poon L.W., Jazwinski S.M. Telomere maintenance genes *SIRT1* and *XRCC6* impact age-related decline in telomere length but only *SIRT1* is associated with human longevity. *Biogerontology*. 2012;13(2):119-131. DOI 10.1007/s10522-011-9360-5
- Kiran S., Chatterjee N., Singh S., Kaul S.C., Wadhwa R., Ramakrishna G. Intracellular distribution of human SIRT7 and mapping of the nuclear/nucleolar localization signal. *FEBS J*. 2013;280(14):3451-3466. DOI 10.1111/febs.12346
- Kumar R., Mohan N., Upadhyay A.D., Singh A.P., Sahu V., Dwivedi S., Dey A.B., Dey S. Identification of serum sirtuins as novel noninvasive protein markers for frailty. *Aging Cell*. 2014;13(6):975-980. DOI 10.1111/acel.12260
- Laurent G., German N.J., Saha A.K., de Boer V.C.J., Davies M., Koves T.R., Dephore N., Fischer F., Boanca G., Vaitheesvaran B., Lovitch S.B., Sharpe A.H., Kurland I.J., Steegborn C., Gygi S.P., Muoio D.M., Ruderman N.B., Haigis M.C. SIRT4 coordinates the balance between lipid synthesis and catabolism by repressing malonyl CoA decarboxylase. *Mol. Cell*. 2013;50(5):686-698. DOI 10.1016/j.molcel.2013.05.012
- Lee N., Kim D.-K., Kim E.-S., Park S.J., Kwon J.-H., Shin J., Park S.-M., Moon Y.H., Wang H.J., Gho Y.S., Choi K.Y. Comparative interactomes of SIRT6 and SIRT7: Implication of functional links to aging. *Proteomics*. 2014;14(13-14):1610-1622. DOI 10.1002/pmic.201400001
- Leng S., Huang W., Chen Y., Yang Ya., Feng D., Liu W., Gao T., Ren Y., Huo M., Zhang J., Yang Yu., Wang Y. SIRT1 coordinates with the CRL4B complex to regulate pancreatic cancer stem cells to promote tumorigenesis. *Cell Death Differ*. 2021;28(12):3329-3343. DOI 10.1038/s41418-021-00821-z
- Li T., Garcia-Gomez A., Morante-Palacios O., Ciudad L., Özkaramemet S., Van Dijk E., Rodríguez-Ubrea J., Vaquero A., Ballestar E. SIRT1/2 orchestrate acquisition of DNA methylation and loss of histone H3 activating marks to prevent premature activation of inflammatory genes in macrophages. *Nucleic Acids Res*. 2020;48(2):665-681. DOI 10.1093/nar/gkz1127
- Li Z., Li H., Zhao Z.-B., Zhu W., Feng P.-P., Zhu X.-W., Gong J.-P. SIRT4 silencing in tumor-associated macrophages promotes HCC development via PPAR δ signalling-mediated alternative activation of macrophages. *J. Exp. Clin. Cancer Res*. 2019;38(1):469. DOI 10.1186/s13046-019-1456-9
- Lin R., Yan D., Zhang Y., Liao X., Gong G., Hu J., Fu Y., Cai W. Common variants in *SIRT1* and human longevity in a Chinese population. *BMC Med. Genet*. 2016a;17(1):31. DOI 10.1186/s12881-016-0293-3
- Lin R., Zhang Y., Yan D., Liao X., Gong G., Hu J., Fu Y., Cai W. Lack of association between polymorphisms in the *SIRT6* gene and longevity in a Chinese population. *Mol. Cell. Probes*. 2016b;30(2):79-82. DOI 10.1016/j.mcp.2016.01.005
- Liu Z.-H., Zhang Ya., Wang X., Fan X.-F., Zhang Yu., Li X., Gong Y.-Sh., Han L.-P. SIRT1 activation attenuates cardiac fibrosis by endothelial-to-mesenchymal transition. *Biomed. Pharmacother*. 2019;118:109227. DOI 10.1016/j.biopha.2019.109227
- LoBianco F.V., Krager K.J., Carter G.S., Alam S., Yuan Y., Lavoie E.G., Dranoff J.A., Aikin-Burns N. The role of Sirtuin 3 in radiation-induced long-term persistent liver injury. *Antioxidants*. 2020;9(5):409. DOI 10.3390/antiox9050409
- Lombard D.B., Alt F.W., Cheng H.-L., Bunkenborg J., Streeper R.S., Mostoslavsky R., Kim J., Yancopoulos G., Valenzuela D., Murphy A., Yang Y., Chen Y., Hirschey M.D., Bronson R.T., Haigis M., Guarente L.P., Fares R.V., Weissman S., Verdin E., Schwer B. Mammalian Sir2 homolog SIRT3 regulates global mitochondrial lysine acetylation. *Mol. Cell. Biol*. 2007;27(24):8807-8814. DOI 10.1128/MCB.01636-07
- Luo J., Nikolaev A.Y., Imai S., Chen D., Su F., Shiloh A., Guarente L., Gu W. Negative control of p53 by Sir2a promotes cell survival under stress. *Cell*. 2001;107(2):137-148. DOI 10.1016/S0092-8674(01)00524-4
- Maity S., Muhamed J., Sarikhani M., Kumar S., Ahamed F., Spurthi K.M., Ravi V., Jain A., Khan D., Arathi B.P., Desingu P.A., Sundaresan N.R. Sirtuin 6 deficiency transcriptionally up-regulates TGF- β signaling and induces fibrosis in mice. *J. Biol. Chem*. 2020;295(2):415-434. DOI 10.1074/jbc.RA118.007212
- Mao Z., Hine C., Tian X., Van Meter M., Au M., Vaidya A., Seluanov A., Gorbunova V. SIRT6 promotes DNA repair under stress by activating PARP1. *Science*. 2011;332(6036):1443-1446. DOI 10.1126/science.1202723
- Mathias R.A., Greco T.M., Oberstein A., Budayeva H.G., Chakrabarti R., Rowland E.A., Kang Y., Shenk T., Cristea I.M. Sirtuin 4 is a lipoamidase regulating pyruvate dehydrogenase complex activity. *Cell*. 2014;159(7):1615-1625. DOI 10.1016/j.cell.2014.11.046
- Michishita E., Park J.Y., Burneskis J.M., Barrett J.C., Horikawa I. Evolutionarily conserved and nonconserved cellular localizations and functions of human SIRT proteins. *Mol. Biol. Cell*. 2005;16(10):4623-4635. DOI 10.1091/mbc.e05-01-0033
- Michishita E., McCord R.A., Berber E., Kioi M., Padilla-Nash H., Damian M., Cheung P., Kusumoto R., Kawahara T.L.A., Barrett J.C., Chang H.Y., Bohr V.A., Ried T., Gozani O., Chua K.F. SIRT6 is a histone H3 lysine 9 deacetylase that modulates telomeric chromatin. *Nature*. 2008;452(7186):492-496. DOI 10.1038/nature06736
- Mizumoto T., Yoshizawa T., Sato Y., Ito T., Tsuyama T., Satoh A., Araki S., Tsujita K., Tamura M., Oike Y., Yamagata K. SIRT7 deficiency protects against aging-associated glucose intolerance and extends lifespan in male mice. *Cells*. 2022;11(22):3609. DOI 10.3390/cells11223609
- Moniot S., Schutkowski M., Steegborn C. Crystal structure analysis of human Sirt2 and its ADP-ribose complex. *J. Struct. Biol*. 2013;182(2):136-143. DOI 10.1016/j.jsb.2013.02.012
- North B.J., Verdin E. Interphase nucleo-cytoplasmic shuttling and localization of SIRT2 during mitosis. *PLoS One*. 2007;2(8):e784. DOI 10.1371/journal.pone.0000784
- North B.J., Marshall B.L., Borra M.T., Denu J.M., Verdin E. The human Sir2 ortholog, SIRT2, is an NAD⁺-dependent tubulin deacetylase. *Mol. Cell*. 2003;11(2):437-444. DOI 10.1016/S1097-2765(03)00038-8
- Ong A.L.C., Ramasamy T.S. Role of Sirtuin1-p53 regulatory axis in aging, cancer and cellular reprogramming. *Ageing Res. Rev*. 2018;43:64-80. DOI 10.1016/j.arr.2018.02.004
- Palomer X., Román-Azcona M.S., Pizarro-Delgado J., Planavila A., Villarroya F., Valenzuela-Alcaraz B., Crispi F., Sepúlveda-Martínez Á., Miguel-Escalada I., Ferrer J., Nistal J.F., García R., Davidson M.M., Barroso E., Vázquez-Carrera M. SIRT3-mediated inhibition of FOS through histone H3 deacetylation prevents cardiac fibrosis and inflammation. *Signal Transduct. Target. Ther*. 2020;5(1):14. DOI 10.1038/s41392-020-0114-1
- Paredes S., Angulo-Ibanez M., Tasselli L., Carlson S.M., Zheng W., Li T.-M., Chua K.F. The epigenetic regulator SIRT7 guards against mammalian cellular senescence induced by ribosomal DNA instability. *J. Biol. Chem*. 2018;293(28):11242-11250. DOI 10.1074/jbc.AC118.003325
- Peng L., Yuan Z., Ling H., Fukasawa K., Robertson K., Olashaw N., Koomen J., Chen J., Lane W.S., Seto E. SIRT1 deacetylates the DNA methyltransferase 1 (DNMT1) protein and alters its activities. *Mol. Cell. Biol*. 2011;31(23):4720-4734. DOI 10.1128/MCB.06147-11
- Pereira J.M., Chevalier C., Chaze T., Gianetto Q., Impens F., Matondo M., Cossart P., Hamon M.A. Infection reveals a modification of SIRT2 critical for chromatin association. *Cell Rep*. 2018;23(4):1124-1137. DOI 10.1016/j.celrep.2018.03.116

- Piracha Z.Z., Saeed U., Kim J., Kwon H., Chwae Y.-J., Lee H.W., Lim J.H., Park S., Shin H.-J., Kim K. An alternatively spliced Sirtuin 2 isoform 5 inhibits Hepatitis B virus replication from cccDNA by repressing epigenetic modifications made by histone lysine methyltransferases. *J. Virol.* 2020;94(16):e00926-20. DOI 10.1128/JVI.00926-20
- Pruitt K.D., Harrow J., Harte R.A., Wallin C., Diekhans M., Maglott D.R., Searle S., ... Wu W., Birney E., Haussler D., Hubbard T., Ostell J., Durbin R., Lipman D. The consensus coding sequence (CCDS) project: Identifying a common protein-coding gene set for the human and mouse genomes. *Genome Res.* 2009;19(7):1316-1323. DOI 10.1101/gr.080531.108
- Pukhalskaia A.E., Dyatlova A.S., Linkova N.S., Kozlov K.L., Kvetnaia T.V., Koroleva M.V., Kvetnoy I.M. Sirtuins as possible predictors of aging and Alzheimer's disease development: verification in the hippocampus and saliva. *Bull. Exp. Biol. Med.* 2020;169(6):821-824. DOI 10.1007/s10517-020-04986-4
- Pukhalskaia A.E., Kvetnoy I.M., Linkova N.S., Diatlova A.S., Gutop E.O., Kozlov K.L., Paltsev M.A. Sirtuins and aging. *Uspekhi Fiziologicheskikh Nauk = Progress in Physiological Science.* 2022; 53(1):16-27. DOI 10.31857/S0301179821040056 (in Russian)
- Quan Y., Park W., Jin J., Kim W., Park S.K., Kang K.P. Sirtuin 3 activation by honokiol decreases unilateral ureteral obstruction-induced renal inflammation and fibrosis via regulation of mitochondrial dynamics and the renal NF- κ B-TGF- β 1/Smad signaling pathway. *Int. J. Mol. Sci.* 2020;21(2):402. DOI 10.3390/ijms21020402
- Rack J.G.M., VanLinden M.R., Lutter T., Aasland R., Ziegler M. Constitutive nuclear localization of an alternatively spliced Sirtuin-2 isoform. *J. Mol. Biol.* 2014;426(8):1677-1691. DOI 10.1016/j.jmb.2013.10.027
- Rajamohan S.B., Pillai V.B., Gupta M., Sundaresan N.R., Birukov K.G., Samant S., Hottiger M.O., Gupta M.P. SIRT1 promotes cell survival under stress by deacetylation-dependent deactivation of poly(ADP-ribose) polymerase 1. *Mol. Cell. Biol.* 2009;29(15):4116-4129. DOI 10.1128/MCB.00121-09
- Ramadani-Muja J., Gottschalk B., Pfeil K., Burgstaller S., Rauter T., Bischof H., Waldeck-Weiermair M., Bugger H., Graier W.F., Mallick R. Visualization of Sirtuin 4 distribution between mitochondria and the nucleus, based on bimolecular fluorescence self-complementation. *Cells.* 2019;8(12):1583. DOI 10.3390/cells8121583
- Ren Y., Du C., Yan L., Wei J., Wu H., Shi Y., Duan H. CTGF siRNA ameliorates tubular cell apoptosis and tubulointerstitial fibrosis in obstructed mouse kidneys in a Sirt1-independent manner. *Drug Des. Devel. Ther.* 2015;9:4155-4171. DOI 10.2147/DDDT.S86748
- Rogina B., Helfand S.L. Sir2 mediates longevity in the fly through a pathway related to calorie restriction. *Proc. Natl. Acad. Sci. USA.* 2004;101(45):15998-16003. DOI 10.1073/pnas.0404184101
- Roichman A., Elhanati S., Aon M.A., Abramovich I., Di Francesco A., Shahar Y., Avivi M.Y., Shurgi M., Rubinstein A., Wiesner Y., Shuchami A., Petrover Z., Leberthal-Loinger I., Yaron O., Lyashkov A., Ubaida-Mohien C., Kanfi Y., Lerrer B., Fernández-Marcos P.J., Serrano M., Gottlieb E., de Cabo R., Cohen H.Y. Restoration of energy homeostasis by SIRT6 extends healthy lifespan. *Nat. Commun.* 2021;12(1):3208. DOI 10.1038/s41467-021-23545-7
- Rothgiesser K.M., Erenner S., Waibel S., Lüscher B., Hottiger M.O. SIRT2 regulates NF- κ B-dependent gene expression through deacetylation of p65 Lys310. *J. Cell Sci.* 2010;123(24):4251-4258. DOI 10.1242/jcs.073783
- Satoh A., Brace C.S., Rensing N., Cliften P., Wozniak D.F., Herzog E.D., Yamada K.A., Imai S. Sirt1 extends life span and delays aging in mice through the regulation of Nk2 homeobox 1 in the DMH and LH. *Cell Metab.* 2013;18(3):416-430. DOI 10.1016/j.cmet.2013.07.013
- Sauve A.A., Wolberger C., Schramm V.L., Boeke J.D. The biochemistry of sirtuins. *Annu. Rev. Biochem.* 2006;75:435-465. DOI 10.1146/annurev.biochem.74.082803.133500
- Sayers E.W., Bolton E.E., Brister J.R., Canese K., Chan J., Coomeau D.C., Connor R., Funk K., Kelly C., Kim S., Madej T., Marchler-Bauer A., Lanczycki C., Lathrop S., Lu Z., Thibaud-Nissen F., Murphy T., Phan L., Skripchenko Y., Tse T., Wang J., Williams R., Trzwick B.W., Pruitt K.D., Sherry S.T. Database resources of the national center for biotechnology information. *Nucleic Acids Res.* 2022;50(D1):D20-D26. DOI 10.1093/nar/gkab1112
- Schmeisser K., Mansfeld J., Kuhlow D., Weimer S., Priebe S., Heiland I., Birringer M., Groth M., Segref A., Kanfi Y., Price N.L., Schmeisser S., Schuster S., Pfeiffer A.F.H., Guthke R., Platzer M., Hoppe T., Cohen H.Y., Zarse K., Sinclair D.A., Ristow M. Role of sirtuins in lifespan regulation is linked to methylation of nicotinamide. *Nat. Chem. Biol.* 2013;9(11):693-700. DOI 10.1038/nchembio.1352
- Sengupta A., Haldar D. Human sirtuin 3 (SIRT3) deacetylates histone H3 lysine 56 to promote nonhomologous end joining repair. *DNA Repair (Amst.).* 2018;61:1-16. DOI 10.1016/j.dnarep.2017.11.003
- Simon M., Yang J., Gigas J., Earley E.J., Hillpot E., Zhang L., Zagorulya M., Tomblin G., Gilbert M., Yuen S.L., Pope A., Van Meter M., Emmrich S., Firsanov D., Athreya A., Biashad S.A., Han J., Ryu S., Tare A., Zhu Y., Hudgins A., Atzmon G., Barzilai N., Wolfe A., Moody K., Garcia B.A., Thomas D.D., Robbins P.D., Vijg J., Seluanov A., Suh Y., Gorbunova V. A rare human centenarian variant of SIRT6 enhances genome stability and interaction with Lamin A. *EMBO J.* 2022;41(21):e110393. DOI 10.15252/embj.2021110393
- Simonet N.G., Thackray J.K., Vazquez B.N., Ianni A., Espinosa-Alcantud M., Morales-Sanfrutos J., Hurtado-Bagès S., Sabidó E., Buschbeck M., Tischfield J., De La Torre C., Esteller M., Braun T., Olivella M., Serrano L., Vaquero A. SirT7 auto-ADP-ribosylation regulates glucose starvation response through mH2A1. *Sci. Adv.* 2020;6(30):eaaz2590. DOI 10.1126/sciadv.aaz2590
- Sinclair D.A., Guarente L. Extrachromosomal rDNA circles – a cause of aging in yeast. *Cell.* 1997;91(7):1033-1042. DOI 10.1016/S0092-8674(00)80493-6
- Smith J.S., Brachmann C.B., Celic I., Kenna M.A., Muhammad S., Starai V.J., Avalos J.L., Escalante-Semerena J.C., Grubmeyer C., Wolberger C., Boeke J.D. A phylogenetically conserved NAD⁺-dependent protein deacetylase activity in the Sir2 protein family. *Proc. Natl. Acad. Sci. USA.* 2000;97(12):6658-6663. DOI 10.1073/pnas.97.12.6658
- Soerensen M., Dato S., Tan Q., Thinggaard M., Kleindorp R., Beekman M., Suchiman H.E.D., Jacobsen R., McGue M., Stevnsner T., Bohr V.A., de Craen A.J.M., Westendorp R.G.J., Schreiber S., Slagboom P.E., Nebel A., Vaupel J.W., Christensen K., Christiansen L. Evidence from case-control and longitudinal studies supports associations of genetic variation in *APOE*, *CETP*, and *IL6* with human longevity. *Age (Dordr.).* 2013;35(2):487-500. DOI 10.1007/s11357-011-9373-7
- Song C., Hotz-Wagenblatt A., Voit R., Grummt I. SIRT7 and the DEAD-box helicase DDX21 cooperate to resolve genomic R loops and safeguard genome stability. *Genes Dev.* 2017;31(13):1370-1381. DOI 10.1101/gad.300624.117
- Subramani P., Nagarajan N., Mariaraj S., Vilwanathan R. Knockdown of sirtuin6 positively regulates acetylation of DNMT1 to inhibit NOTCH signaling pathway in non-small cell lung cancer cell lines. *Cell. Signal.* 2023;105:110629. DOI 10.1016/j.cellsig.2023.110629
- Sun L., Fang J. Macromolecular crowding effect is critical for maintaining SIRT1's nuclear localization in cancer cells. *Cell Cycle.* 2016;15(19):2647-2655. DOI 10.1080/15384101.2016.1211214

- Sundaresan N.R., Bindu S., Pillai V.B., Samant S., Pan Y., Huang J.-Y., Gupta M., Nagalingam R.S., Wolfgeher D., Verdin E., Gupta M.P. SIRT3 blocks aging-associated tissue fibrosis in mice by deacetylating and activating glycogen synthase kinase 3 β . *Mol. Cell. Biol.* 2016;36(5):678-692. DOI 10.1128/MCB.00586-15
- Tan M., Peng C., Anderson K.A., Chhoy P., Xie Z., Dai L., Park J., Chen Y., Huang H., Zhang Y., Ro J., Wagner G.R., Green M.F., Madsen A.S., Schmiesing J., Peterson B.S., Xu G., Ilkayeva O.R., Muehlbauer M.J., Braluk T., Muhlhausen C., Backos D.S., Olsen C.A., McGuire P.J., Pletcher S.D., Lombard D.B., Hirschey M.D., Zhao Y. Lysine glutarylation is a protein posttranslational modification regulated by SIRT5. *Cell Metab.* 2014;19(4):605-617. DOI 10.1016/j.cmet.2014.03.014
- Taylor J.R., Wood J.G., Mizerak E., Hinthorn S., Liu J., Finn M., Gordon S., Zingas L., Chang C., Klein M.A., Denu J.M., Gorbunova V., Seluanov A., Boeke J.D., Sedivy J.M., Helfand S.L. Sirt6 regulates lifespan in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA.* 2022;119(5):e2111176119. DOI 10.1073/pnas.2111176119
- TenNapel M.J., Lynch C.F., Burns T.L., Wallace R., Smith B.J., Button A., Domann F.E. SIRT6 minor allele genotype is associated with >5-year decrease in lifespan in an aged cohort. *PLoS One.* 2014;9(12):e115616. DOI 10.1371/journal.pone.0115616
- Tian X., Firsanov D., Zhang Z., Cheng Y., Luo L., Tomblin G., Tan R., Simon M., Henderson S., Steffan J., Goldfarb A., Tam J., Zheng K., Cornwell A., Johnson A., Yang J.-N., Mao Z., Manta B., Dang W., Zhang Z., Vigg J., Wolfe A., Moody K., Kennedy B.K., Bohmann D., Gladyshev V.N., Seluanov A., Gorbunova V. SIRT6 is responsible for more efficient DNA double-strand break repair in long-lived species. *Cell.* 2019;177(3):622-638.e22. DOI 10.1016/j.cell.2019.03.043
- Tissenbaum H.A., Guarente L. Increased dosage of a *sir-2* gene extends lifespan in *Caenorhabditis elegans*. *Nature.* 2001;410(6825):227-230. DOI 10.1038/35065638
- Tsai Y.-C., Greco T.M., Boonmee A., Miteva Y., Cristea I.M. Functional proteomics establishes the interaction of SIRT7 with chromatin remodeling complexes and expands its role in regulation of RNA polymerase I transcription. *Mol. Cell. Proteomics.* 2012;11(5):60-76. DOI 10.1074/mcp.A111.015156
- van der Horst A., Tertoolen L.G.J., de Vries-Smits L.M.M., Frye R.A., Medema R.H., Burgering B.M.T. FOXO4 is acetylated upon peroxide stress and deacetylated by the longevity protein hSir2^{SIRT1}. *J. Biol. Chem.* 2004;279(28):28873-28879. DOI 10.1074/jbc.M401138200
- Van Meter M., Kashyap M., Rezazadeh S., Geneva A.J., Morello T.D., Seluanov A., Gorbunova V. SIRT6 represses LINE1 retrotransposons by ribosylating KAP1 but this repression fails with stress and age. *Nat. Commun.* 2014;5(1):5011. DOI 10.1038/ncomms6011
- Vaquero A., Scher M., Lee D., Erdjument-Bromage H., Tempst P., Reinberg D. Human sirt1 interacts with histone H1 and promotes formation of facultative heterochromatin. *Mol. Cell.* 2004;16(1):93-105. DOI 10.1016/j.molcel.2004.08.031
- Vaquero A., Scher M.B., Lee D.H., Sutton A., Cheng H.-L., Alt F.W., Serrano L., Sternglanz R., Reinberg D. SirT2 is a histone deacetylase with preference for histone H4 Lys 16 during mitosis. *Genes Dev.* 2006;20(10):1256-1261. DOI 10.1101/gad.1412706
- Vaquero A., Scher M., Erdjument-Bromage H., Tempst P., Serrano L., Reinberg D. SIRT1 regulates the histone methyltransferase SUV39H1 during heterochromatin formation. *Nature.* 2007;450(7168):440-444. DOI 10.1038/nature06268
- Vaziri H., Dessain S.K., Eaton E.N., Imai S.-I., Frye R.A., Pandita T.K., Guarente L., Weinberg R.A. hSIR2/SIRT1 functions as an NAD-dependent p53 deacetylase. *Cell.* 2001;107(2):149-159. DOI 10.1016/S0092-8674(01)00527-X
- Vazquez-Bonondo P., Nguyen T., Bunting S., Vaquero A., Tischfield J.A., Serrano L. SIRT7 promotes genome integrity and modulates non-homologous end joining DNA repair. *EMBO J.* 2016;35(14):1488-1503. DOI 10.15252/embj.201593499
- Vazquez B.N., Fernández-Duran I., Vaquero A. Sirtuins in hematopoiesis and blood malignancies. Chapter 23. In: Maiese K. (Ed.). Sirtuin Biology in Medicine. Academic Press, 2021;373-391. DOI 10.1016/B978-0-12-8141118-2.00020-3
- Viswanathan M., Guarente L. Regulation of *Caenorhabditis elegans* lifespan by *sir-2.1* transgenes. *Nature.* 2011;477(7365):E1-E2. DOI 10.1038/nature10440
- Wang L., Xu C., Johansen T., Berger S.L., Dou Z. SIRT1 – a new mammalian substrate of nuclear autophagy. *Autophagy.* 2021;17(2):593-595. DOI 10.1080/15548627.2020.1860541
- Wang R.-H., Sengupta K., Li C., Kim H.-S., Cao L., Xiao C., Kim S., Xu X., Zheng Y., Chilton B., Jia R., Zheng Z.-M., Appella E., Wang X.W., Ried T., Deng C.-X. Impaired DNA damage response, genome instability, and tumorigenesis in SIRT1 mutant mice. *Cancer Cell.* 2008;14(4):312-323. DOI 10.1016/j.ccr.2008.09.001
- Whitaker R., Faulkner S., Miyokawa R., Burhenn L., Henriksen M., Wood J.G., Helfand S.L. Increased expression of *Drosophila Sir2* extends life span in a dose-dependent manner. *Aging.* 2013;5(9):682-691. DOI 10.18632/aging.100599
- Willcox B.J., Donlon T.A., He Q., Chen R., Grove J.S., Yano K., Masaki K.H., Willcox D.C., Rodriguez B., Curb J.D. FOXO3A genotype is strongly associated with human longevity. *Proc. Natl. Acad. Sci. USA.* 2008;105(37):13987-13992. DOI 10.1073/pnas.0801030105
- Woo S.J., Lee S.-M., Lim H.S., Hah Y.-S., Jung I.D., Park Y.-M., Kim H.-O., Cheon Y.-H., Jeon M.-G., Jang K.Y., Kim K.M., Park B.-H., Lee S.-I. Myeloid deletion of SIRT1 suppresses collagen-induced arthritis in mice by modulating dendritic cell maturation. *Exp. Mol. Med.* 2016;48(3):e221. DOI 10.1038/emmm.2015.124
- Woo S.J., Noh H.S., Lee N.Y., Cheon Y.-H., Yi S.M., Jeon H.M., Bae E.J., Lee S.-I., Park B.-H. Myeloid sirtuin 6 deficiency accelerates experimental rheumatoid arthritis by enhancing macrophage activation and infiltration into synovium. *EBioMedicine.* 2018;38:228-237. DOI 10.1016/j.ebiom.2018.11.005
- Wood J.G., Schwer B., Wickremesinghe P.C., Hartnett D.A., Burhenn L., Garcia M., Li M., Verdin E., Helfand S.L. Sirt4 is a mitochondrial regulator of metabolism and lifespan in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA.* 2018;115(7):1564-1569. DOI 10.1073/pnas.1720673115
- Wu Q.J., Zhang T.N., Chen H.H., Yu X.F., Lv J.L., Liu Y.Y., Liu Y.S., Zheng G., Zhao J.Q., Wei Y.F., Guo J.Y., Liu F.H., Chang Q., Zhang Y.X., Liu C.G., Zhao Y.H. The sirtuin family in health and disease. *Signal Transduct. Target. Ther.* 2022;7(1):402. DOI 10.1038/s41392-022-01257-8
- Xu C., Wang L., Fozouni P., Evjen G., Chandra V., Jiang J., Lu C., Nicastri M., Bretz C., Winkler J.D., Amaravadi R., Garcia B.A., Adams P.D., Ott M., Tong W., Johansen T., Dou Z., Berger S.L. SIRT1 is downregulated by autophagy in senescence and ageing. *Nat. Cell Biol.* 2020;22(10):1170-1179. DOI 10.1038/s41556-020-00579-5
- Yang Y., Hou H., Haller E.M., Nicosia S.V., Bai W. Suppression of FOXO1 activity by FHL2 through SIRT1-mediated deacetylation. *EMBO J.* 2005;24(5):1021-1032. DOI 10.1038/sj.emboj.7600570
- Yeung F., Hoberg J.E., Ramsey C.S., Keller M.D., Jones D.R., Frye R.A., Mayo M.W. Modulation of NF- κ B-dependent transcription and cell survival by the SIRT1 deacetylase. *EMBO J.* 2004;23(12):2369-2380. DOI 10.1038/sj.emboj.7600244
- Zhang J., Fang L., Lu Z., Xiong J., Wu M., Shi L., Luo A., Wang S. Are sirtuins markers of ovarian aging? *Gene.* 2016;575(2 Pt. 3):680-686. DOI 10.1016/j.gene.2015.09.043
- Zhang P.-Y., Li G., Deng Z.-J., Liu L.-Y., Chen L., Tang J.-Z., Wang Y.-Q., Cao S.-T., Fang Y.-X., Wen F., Xu Y., Chen X., Shi K.-Q., Li W.-F., Xie C., Tang K.-F. Dicer interacts with SIRT7 and regulates H3K18 deacetylation in response to DNA damaging

- agents. *Nucleic Acids Res.* 2016;44(8):3629-3642. DOI 10.1093/nar/gkv1504
- Zhang W.-G., Bai X.-J., Chen X.-M. *SIRT1* variants are associated with aging in a healthy Han Chinese population. *Clin. Chim. Acta.* 2010;411(21-22):1679-1683. DOI 10.1016/j.cca.2010.06.030
- Zhang X., Spiegelman N.A., Nelson O.D., Jing H., Lin H. SIRT6 regulates Ras-related protein R-Ras2 by lysine defatty-acylation. *eLife.* 2017;6:e25158. DOI 10.7554/eLife.25158
- Zhang X., Ameer F.S., Azhar G., Wei J.Y. Alternative splicing increases sirtuin gene family diversity and modulates their subcellular localization and function. *Int. J. Mol. Sci.* 2021;22(2):473. DOI 10.3390/ijms22020473
- Zhang Y., Connelly K.A., Thai K., Wu X., Kapus A., Kepecs D., Gilbert R.E. Sirtuin 1 activation reduces transforming growth factor- β 1-induced fibrogenesis and affords organ protection in a model of progressive, experimental kidney and associated cardiac disease. *Am. J. Pathol.* 2017;187(1):80-90. DOI 10.1016/j.ajpath.2016.09.016
- Zhao Y., Wang H., Poole R.J., Gems D. A *fln-2* mutation affects lethal pathology and lifespan in *C. elegans*. *Nat. Commun.* 2019;10(1):5087. DOI 10.1038/s41467-019-13062-z
- Zhong L., D'Urso A., Toiber D., Sebastian C., Henry R.E., Vady-sirisack D.D., Guimaraes A., Marinelli B., Wikstrom J.D., Nir T., Clish C.B., Vaitheesvaran B., Iliopoulos O., Kurland I., Dor Y., Weissleder R., Shirihai O.S., Ellisen L.W., Espinosa J.M., Mostoslavsky R. The histone deacetylase SIRT6 regulates glucose homeostasis via Hif1 α . *Cell.* 2010;140(2):280-293. DOI 10.1016/j.cell.2009.12.041

Funding. The study was financially supported by the Russian Science Foundation under research project No. 22-74-10123. D.A. Chudakova's work in the field of brain aging and neurodegeneration research was carried out with financial support from the Federal Medical-Biological Agency of Russian Federation.

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

Received November 12, 2023. Revised January 15, 2024. Accepted January 16, 2024.

DOI 10.18699/vjgb-24-27

Involvement of transposable elements in Alzheimer's disease pathogenesis

R.N. Mustafin ¹ , E.K. Khusnutdinova ^{1, 2}¹ Bashkir State Medical University, Ufa, Russia² Institute of Biochemistry and Genetics – Subdivision of the Ufa Federal Research Centre of the Russian Academy of Sciences, Ufa, Russia ruji79@mail.ru

Abstract. Alzheimer's disease affects an average of 5 % of the population with a significant increase in prevalence with age, suggesting that the same mechanisms that underlie aging may influence this pathology. Investigation of these mechanisms is promising for effective methods of treatment and prevention of the disease. Possible participants in these mechanisms are transposons, which serve as drivers of epigenetic regulation, since they form species-specific distributions of non-coding RNA genes in genomes in evolution. Study of miRNA involvement in Alzheimer's disease pathogenesis is relevant, since the associations of protein-coding genes (*APOE4*, *ABCA7*, *BIN1*, *CLU*, *CR1*, *PICALM*, *TREM2*) with the disease revealed as a result of GWAS make it difficult to explain its complex pathogenesis. Specific expression changes of many genes were found in different brain parts of Alzheimer's patients, which may be due to global regulatory changes under the influence of transposons. Experimental and clinical studies have shown pathological activation of retroelements in Alzheimer's disease. Our analysis of scientific literature in accordance with MDTE DB revealed 28 miRNAs derived from transposons (17 from LINE, 5 from SINE, 4 from HERV, 2 from DNA transposons), the expression of which specifically changes in this disease (decreases in 17 and increases in 11 microRNA). Expression of 13 out of 28 miRNAs (miR-151a, -192, -211, -28, -31, -320c, -335, -340, -378a, -511, -576, -708, -885) also changes with aging and cancer development, which indicates the presence of possible common pathogenetic mechanisms. Most of these miRNAs originated from LINE retroelements, the pathological activation of which is associated with aging, carcinogenesis, and Alzheimer's disease, which supports the hypothesis that these three processes are based on the primary dysregulation of transposons that serve as drivers of epigenetic regulation of gene expression in ontogeny.

Key words: Alzheimer's disease; carcinogenesis; miRNA; aging; transposons; retroelements.

For citation: Mustafin R.N., Khusnutdinova E.K. Involvement of transposable elements in Alzheimer's disease pathogenesis. *Vavilovskii Zhurnal Genetiki i Seleksii* = *Vavilov Journal of Genetics and Breeding*. 2024;28(2):228-238. DOI 10.18699/vjgb-24-27

Влияние транспозонов на развитие болезни Альцгеймера

R.N. Мустафин ¹ , Э.К. Хуснутдинова ^{1, 2}¹ Башкирский государственный медицинский университет, Уфа, Россия² Институт биохимии и генетики – обособленное структурное подразделение Уфимского федерального исследовательского центра Российской академии наук, Уфа, Россия ruji79@mail.ru

Аннотация. Болезнь Альцгеймера поражает в среднем 5 % населения со значительным увеличением распространенности с возрастом, что свидетельствует о возможном влиянии на данную патологию тех же механизмов, которые лежат в основе старения человека. Исследование этих механизмов перспективно для разработки эффективных методов лечения и профилактики заболевания. Возможными участниками этих механизмов являются транспозоны, которые служат драйверами эпигенетической регуляции, поскольку формируют в эволюции видоспецифические распределения генов некодирующих РНК в геноме человека. Изучение роли микроРНК в развитии болезни Альцгеймера актуально, поскольку по результатам проведенных GWAS ассоциаций белок-кодирующих генов (*APOE4*, *ABCA7*, *BIN1*, *CLU*, *CR1*, *PICALM*, *TREM2*) трудно объяснить сложный патогенез заболевания. Кроме того, в различных долях головного мозга при болезни Альцгеймера были обнаружены специфические изменения экспрессии множества генов, что может быть обусловлено глобальными регуляторными изменениями под влиянием транспозонов. Действительно, экспериментальные и клинические исследования показали патологическую активацию ретроэлементов при болезни Альцгеймера. Проведенный нами анализ научной литературы в соответствии с базой данных MDTE DB (microRNAs derived from transposable elements) позволил выявить 28 различных микроРНК, происходящих от мобильных элементов (17 – от LINE, 5 – от SINE, 4 – от HERV, 2 – от ДНК-транспозонов), экспрессия которых специфически изменяется при данном заболевании (снижается у 17 и повышается у 11 микроРНК). Экспрес-

сия 13 из 28 микроРНК (miR-151a, -192, -211, -28, -31, -320с, -335, -340, -378а, -511, -576, -708, -885) меняется также при старении и развитии злокачественных новообразований, что подтверждает возможное наличие общих патогенетических механизмов. Большинство из этих микроРНК произошли от LINE-ретроэлементов, патологическая активация которых ассоциирована со старением, канцерогенезом и болезнью Альцгеймера, что свидетельствует в пользу гипотезы о том, что в основе этих трех процессов лежит первичная дисрегуляция транспозонов, которые служат драйверами эпигенетической регуляции экспрессии генов в онтогенезе.
Ключевые слова: болезнь Альцгеймера; канцерогенез; микроРНК; старение; транспозоны; ретроэлементы.

Introduction

Alzheimer's disease (AD) is the most common neurodegenerative disease. Disease pathogenesis is caused by extracellular deposition of beta-amyloid plaques and intracellular accumulation of tau protein tangles with cell death in the brain (Barak et al., 2013). AD is detected in 62 % of patients with dementia (Swarbrick et al., 2019). In 2017, a meta-analysis revealed a 5 % prevalence of AD in Europe (3.31 % in men and 7.13 % in women) and an increase in these rates with age (7.66 % in 75–84-year-olds, 22.53 % in 85-year-olds and older). In Japan, AD occurs in 7 % of people over 65 years of age, in the USA – in 9.51 % of people over 70 years of age. In Chinese residents, the incidence of AD is 1.27 % in people 65–69 years old and 18.54 % in people 85–89 years old (Niu et al., 2017). Twin studies showed the heritability of AD to be 58 %, regardless of gender (Gatz et al., 2006).

In 2018, a genome-wide association study (GWAS) of DNA samples from 314,278 patients showed an association of the *ACE*, *ADAM10*, *BCKDK/KAT8*, *TOMM40*, *VKORC1* genes with AD (Marioni et al., 2018). In 2019, a meta-analysis of GWAS results (53,042 AD patients and 355,900 healthy controls) identified 37 specific loci associated with AD in the human genome. Among them, the *APHIB*, *BINI*, *CASS4*, *CCDC6*, *NCK2*, *PILRA*, *PTK2B*, *SPRED2*, *TSPAN14* genes showed the greatest significance. However, it is difficult to explain the role of allelic variants of these genes in the pathogenesis of AD of these genes.

Possible mechanisms of other AD-associated genes are shown for *LILRB2* (encodes a receptor that recognizes multiple HLA alleles and may be involved in the growth of beta-amyloid fibrils), *ABCA1* (involved in the transfer of phospholipids to apolipoproteins), *AGRN* (involved in the formation of synapses of mature hippocampal neurons) (Schwartzentruuber et al., 2021). In 2021, a meta-analysis showed an association of 23 different SNPs with AD, among which the highest significance was determined for rs3865444 (in the *CD33* transmembrane receptor gene), rs7561528 (in the nucleocytoplasmic adapter protein gene (*BINI*)) and rs1801133 (in the methylenetetrahydrofolate reductase gene (*MTHFR*)) (GNS et al., 2021).

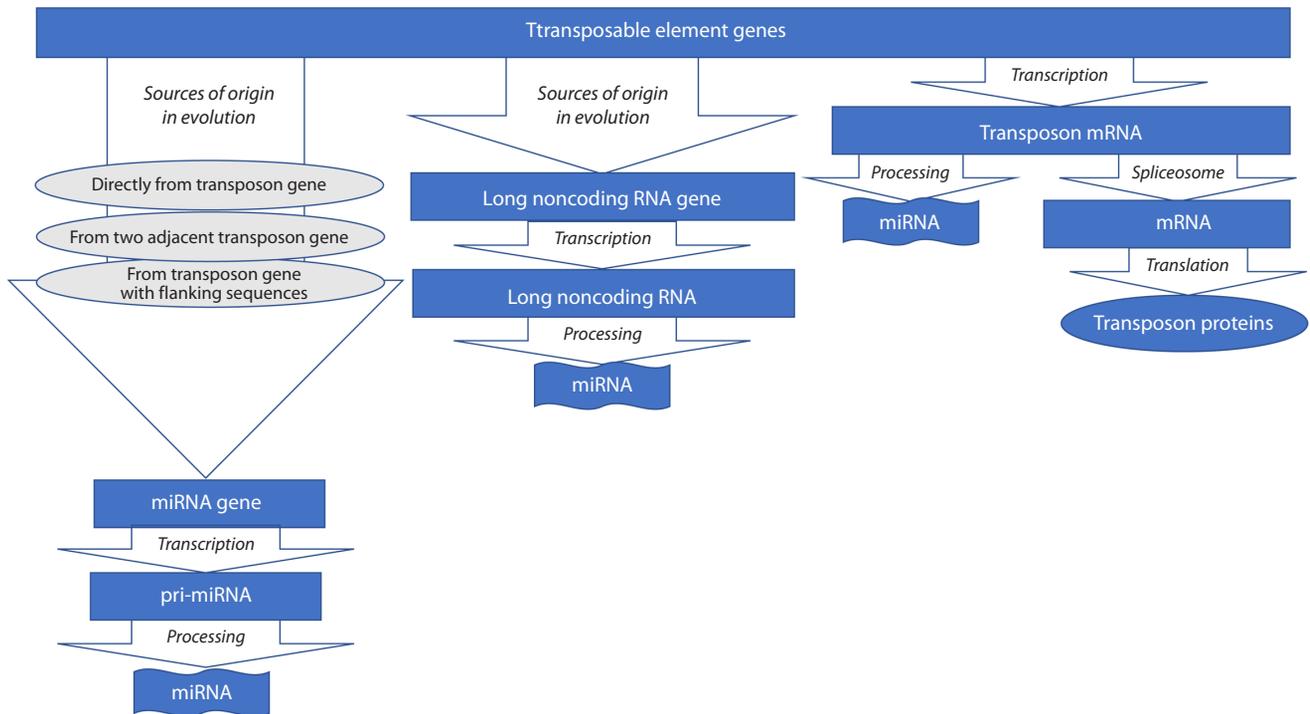
In GWASs, a significant association with AD was also shown for the *CLU* (*APOJ*, encodes apolipoprotein J), *CRI* (encodes complement component 3b/4b) (Lambert et al., 2009), *APOE* (encodes apolipoprotein E), *PICALM* (encodes protein phosphatidylinositol-binding clathrin assembly) (Harold et al., 2009; Ando et al., 2022), *BINI* (Ando et al., 2022) genes. Meta-analyses of GWAS results with Alzheimer's disease showed a significant association of allelic variants of the *TREM2* (encodes the trigger receptor expressed on protein 2 myeloid cells) (Guerreiro et al., 2013) and *ABCA7* (Ma et al., 2018) genes.

According to numerous genome-wide association meta-analyses and large-scale genome-wide association studies, the strongest genetic risk factor for sporadic AD is the *APOE* $\epsilon 4$ allele, while the most powerful protective genetic factor is the *APOE* $\epsilon 2$ allele. This is due to the effects of APOE on β -amyloid peptide aggregation and clearance, neurofibrillary tau degeneration, microglial and astrocyte responses, and the blood-brain barrier (Serrano-Pozo et al., 2021). The production and breakdown of amyloid are also directly influenced by the *ABPP*, *PSEN1* and *PSEN2* genes, the allelic variants of which contribute to increased toxic amyloid types aggregation (Robinson et al., 2017).

The key role of genetic factors in the development of AD is evidenced by the presence of monogenic hereditary forms of the disease with an autosomal dominant type of inheritance. These forms of the disease are caused by germline mutations in the *APP* (amyloid precursor protein) (Rogaev et al., 1994; Goate et al., 2006), *PSEN1* (presenilin-1) (Sherrington et al., 1995), *PSEN2* (presenilin-2) (Levy-Lahad et al., 1995) genes. Genomic instability is important in the pathogenesis of AD, as evidenced by the pronounced association of AD with age (which is characterized by genomic instability) (Hou et al., 2017). A genomic instability component in AD may be expression changes of long non-coding RNAs, such as XIST (X-inactive specific transcript), which is considered as a potential target for AD therapy (Chanda, Mukhopadhyay, 2020).

In addition to the association of allelic variants of specific genes from DNA samples of peripheral blood leukocytes of patients with AD, a number of studies have analyzed the expression of specific genes in the brain cells of patients. The data obtained could better explain the possible mechanisms of AD pathogenesis. In 2022, a meta-analysis identified 1915 differentially expressed genes in the entorhinal cortex (the hippocampus-related part of the temporal lobe) in AD patients compared to healthy controls (Fagone et al., 2022). Earlier, in 2019, a meta-analysis of the transcriptome in AD had showed differential expression of a large number of genes in different lobes of the brain: in the temporal lobe – 323, in the frontal – 435, parietal – 1023, cerebellar – 828 genes (Patel et al., 2019). This indicates a pronounced deregulation of gene expression in the brain in AD on a genome-wide scale, a possible cause of which is the pathological activation of transposable elements (TEs), which occupy 45 % of the human genome. TEs have a global regulatory effect on the expression of all genes (as drivers of epigenetic regulation (Mustafin, Khusnutdinova, 2017)) and binding sites for transcription factors (Mustafin, 2019).

The cause of genomic instability in neurons in AD may be somatic recombinations between TEs, such as Alus and LINE1s (Pascarella et al., 2022). This is evidenced by recent



Mechanisms of origin of microRNAs from transposable elements.

results obtained from fluorescent *in situ* hybridization (FISH) in individual neurons of AD patients' brain (Yurov et al., 2023), as well as experimental studies on mice with knockdown of one allele of the *BM11* gene (encodes a protein of the Polycomb group and regulates compaction of heterochromatin). In adults, *BM11* is normally expressed ubiquitously in brain neurons, but is reduced in AD. *Bmi1*^{+/-} mice are characterized by neurodegenerative changes similar to AD. In this case, the loss of heterochromatin is determined mainly in the regions with repeating sequences, which include TEs or originated from them in evolution (El Hajjar et al., 2019).

TEs carry out epigenetic regulation due to their interaction with microRNAs, which evolved from TEs by various mechanisms (see the Figure), as well as by processing their transcripts to form microRNAs (Wei et al., 2016). MicroRNAs have a post-transcriptional regulatory effect on gene expression (Barak et al., 2013) and are guides for binding to DNA methyltransferases (RdDM, RNA-directed DNA methylation) with specific genomic loci, regulating expression at the transcription level (Watcharanurak, Mutirangura, 2022). Therefore, it can be assumed that the observed hypermethylation of 236 specific CpG location loci in the cerebral cortex of AD patients occurs under the influence of microRNAs upon activation of TEs (Smith et al., 2021).

TEs are divided into classes of DNA transposons (moving by a cut-and-paste mechanism) and retroelements (REs). Transposition of REs occurs by "copy-and-paste" with an intermediate RNA, from which cDNA is formed by reverse transcription. Based on the presence of long terminal repeats (LTRs), REs are classified into LTR-containing REs and non-LTR-REs. The latter include autonomous LINES (long interspersed elements) and non-autonomous SINES (short

interspersed elements) and SVEs (SINE-VNTR-Alus). LTR-REs are endogenous retroviruses (ERVs), occupying 8 % of the human genome, while LINES (L1 and L2) occupy 21 % (Ravel-Godreuil et al., 2021).

Role of transposable elements in Alzheimer's disease development

The activity of TEs is under the control of epigenetic modifiers (DNA and histone methylation), as well as specific molecules, such as PRC2 (Polycomb repressive complex 2, which forms the H3K27me3 mark), DNMT1 (promotes the formation of H4K20me3), KAP1 protein (Kruppel-associated box associated protein 1, promotes the formation of H3K9me3 marks), sirtuin 6 (SIRT6, causes repression of L1 through ribosylation of KAP1, facilitating the interaction of KAP1 with its partners and the formation of heterochromatin in the L1 promoter region) (Ravel-Godreuil et al., 2021). According to recent data, TEs themselves are drivers of epigenetic regulation of genes through the formation of long non-coding RNAs and microRNAs from their transcripts (Mustafin, Khusnutdinova, 2017). That is, TEs are under the control of regulatory mechanisms that they drive, which indicates the presence of evolutionarily programmed self-control. Failure in this system is one of the factors of human aging (Wood, Helfand, 2013; Van Meter et al., 2014).

It is possible that the relationship between TEs and tau proteins reflects the system of mutual regulation of TEs and genes in the human genome. Indeed, inhibition of the *BM11* gene (component of the Polycomb repressive complex 1, which promotes chromatin compaction and gene silencing through E3-mono-ubiquitin ligation activity mediated by Ring1a/b on histone H2A at lysine 19 (H2A^{ub})) expression was found in

the brain of AD patients. *BMI1* gene knockout in postmitotic human neurons resulted in beta-amyloid deposition and accumulation of tau protein (because BMI1 suppresses tau protein transcription) (Flamier et al., 2018).

Modeling AD in mice by knocking out one allele of the *Bmi1* gene showed the development of neurodegeneration due to derepression of TEs (El Hajjar et al., 2019). Experiments in mice show enhanced processing of non-coding RNAs from SINE B2 transcripts in the hippocampus under the influence of amyloid deposition (Cheng et al., 2020). Transcriptomic analysis showed activation of TEs (mainly ERVs) induced by aging and tau in mouse brain. Transgenic mice expressing tau protein in the brain showed an increased number of TEs' DNA copies (Ramirez et al., 2022). G-quadruplex derived from evolutionarily conserved L1 suppresses gene expression in AD neurons (Hanna et al., 2021).

In 2018, an analysis of postmortem brain tissue samples showed that in tauopathies, decondensation of heterochromatin and decreased levels of piwi and piRNA cause deregulation of TEs. A significant increase in HERVs transcripts was also found in AD brains (Sun et al., 2018). In the same year, a study of postmortem brain tissue from AD patients (636 people) and a *Drosophila* model of the disease showed differential expression of several specific REs in association with the load of neurofibrillary tau tangles. In this case, global transcriptional activation of LINE1s and ERVs occurred. Tau protein-associated chromatin marks were detected at HERV-Fc1 location loci. Profiling of TEs in *Drosophila* throughout the brain showed heterogeneous response profiles, including those depending on age and genotype, activation of TEs under the influence of tau proteins (Guo C. et al., 2018).

Further studies of post-mortem brain tissue from patients with AD (60 individuals) confirmed the data on the activation of specific TEs (L1s and Alus) in AD compared with controls (Grundman et al., 2021). Analysis of blood samples from 25 late-onset AD patients revealed a significant increase in the expression of 1790 RE transcripts (LINE, LTR, SVA) before clinical phenocconversion (from normal cognitive indicators to the manifestation of AD), which the authors called a retrotransposon storm (Macciardi et al., 2022). It is possible that the data obtained by the researchers indicate the effect of a feedback relationship between TEs activated during aging and the influence of the resulting tau proteins on them, which triggers the cascade mechanism of the TEs > tau proteins > TEs relationship.

Activation of REs in AD depends on the transmission of redox signals (such as complex I of the mitochondrial respiratory chain) from mitochondria to the nucleus. It is believed that this phenomenon is a side effect of general signaling from mitochondria to the nucleus, aimed at facilitating the transcription of mitochondrial genes to restore mitochondrial function (Baeken et al., 2020). As a result, DNA hypomethylation and increased expression of REs, such as LINE1s, occur (Protasova et al., 2021). In this case, a vicious circle may develop when activated REs aggravate mitochondrial pathology due to insertions into genes involved in their functioning. Thus, frequent primate-specific retrotranspositions of Alu elements into the introns of the *TOMM40* gene, encoding the β -barrel protein necessary for mitochondrial transport of preproteins and associated with AD, were identified (Larsen et al., 2017).

Association of microRNAs derived from transposable elements with Alzheimer's disease

In 2016, G. Wei et al. created a database of microRNAs derived from TEs (MDTE DB: a database for microRNAs derived from Transposable element) (Wei et al., 2016). Due to the presence of data on the role of dysregulation of TEs in AD (Guo C. et al., 2018; Sun et al., 2018; Grudman et al., 2021; Macciardi et al., 2022), analysis of specific microRNAs presented in the MDTE DB may reveal one of the mechanisms of AD pathogenesis upon activation of TEs. In 2019, S. Swarbrick et al. conducted a systematic review of the accumulated data in the scientific literature on microRNAs associated with Alzheimer's disease. A significant role was identified for 44 microRNAs in blood plasma, 250 microRNAs in the brain, 153 microRNAs in cerebrospinal fluid (Swarbrick et al., 2019).

Our analysis of the scientific literature allowed us to determine the association of a number of microRNAs derived from TEs that are associated with AD. In 2014, a study of the brains of rabbits modeled for AD found decreased expression of miR-576-3p (Liu et al., 2014), which was derived from L1 (Wei et al., 2016). In 2022, a reduced level of miR-576-3p was detected in the serum of people with AD (Xu et al., 2022). In 2014, a GWAS of blood samples from 158 AD patients and 155 healthy controls showed a significant difference in the expression of miR-885-5p (derived from SINE/MIR (Wei et al., 2016)) in AD (Tan et al., 2014). Further studies showed that overexpression of miR-885-5p attenuates beta-amyloid-induced neuronal damage by suppressing KREMEN1 synthesis (Pan et al., 2022).

In 2015, a comparative analysis of microRNA levels in blood samples of 48 patients with AD and 22 controls showed an increase in the expression of miR-151a (Satoh et al., 2015), derived from L2 (Wei et al., 2016), miR-3200 (Satoh et al., 2015), derived from ERVL (Wei et al., 2016), and a decrease in the expression of miR-502 (derived from L2 (Wei et al., 2016)) (Satoh et al., 2015). In the same year, a study of 127 AD patients and 123 controls revealed a decrease in the level of miR-31 in AD (Dong H. et al., 2015), which originated from L2 (Wei et al., 2016). Experiments on AD mouse models showed significant improvement in neurological parameters with lentiviral-mediated expression of miR-31 due to a decrease of beta-amyloid in the hippocampus (Barros-Viegas et al., 2020).

In 2016, an experiment in mice modeled for AD demonstrated the role of miR-211 (derived from L2 (Wei et al., 2016)), affecting NUAK1, causing the accumulation of beta-amyloid and reducing neuronal survival (Fan et al., 2016). Elevated levels of miR-211 were found in another study in AD mouse models and beta-amyloid accumulation (Sierksma et al., 2018). A decrease in the expression of miR-511 (derived from L1 (Wei et al., 2016)) was found in AD, resulting in increased synthesis of the FKBP5 protein (Zheng et al., 2016). Treatment of AD mouse models with cauterization at acupuncture points of the control vessel contributed to the improvement of cognitive functions by increasing the expression of miR-511-3p (Jia et al., 2022).

In 2017, mouse models of AD showed increased levels of miR-28-3p (the miR-28 family is derived from L2 (Wei et al.,

2016)) in the cerebrospinal fluid (Hong et al., 2017). In the blood serum of people with asthma, an increased concentration of miR-28-3p was also determined, compared to healthy controls. The level of this microRNA decreased with effective donepezil therapy (Zhao et al., 2020).

ERV1-derived miR-1246 (Wei et al., 2016) has been proposed as a biomarker of AD to determine its level in the blood serum of patients (Guo R. et al., 2017). A decreased miR-545-3p level was determined in the blood plasma of patients with AD compared to controls (Cosin-Tomas et al., 2017). The miR-545 family originated from L2 (Wei et al., 2016). In AD, reduced expression of miR-325 (derived from L2 (Wei et al., 2016)) is determined, which has a post-transcriptional regulatory effect on tomosyn synthesis (impairs synaptic transmission in the brain) in the hippocampus (Barak et al., 2013). The pro-inflammatory microRNA miR-326, derived from the DNA transposon hAT-Tip100 (Wei et al., 2016), was characterized by increased expression in AD (Cai et al., 2017). Low levels of miR-342-5p (derived from SINE (Wei et al., 2016)) were detected in the worst course of AD (Dakterzada et al., 2021). SINE-derived miR-3646 was overexpressed in AD patients (Lu et al., 2021).

In 2018, increased expression of miR-320c (derived from L1 (Wei et al., 2016)) was determined in patients with AD compared to patients with amyotrophic lateral sclerosis (Raheja et al., 2018). Previously, significant association of miR-320 gene locus was determined in a genome-wide linkage analysis in patients with late-onset familial AD (Kunkle et al., 2016).

In mouse models of AD, increased expression of miR-320 in brain neurons was determined (Boese et al., 2016). Reduced level of miR-4487 (derived from L1 (Wei et al., 2016)) was detected in neurons of the brain of AD patients (Hu et al., 2018). In AD, miR-384 (derived from LINE/Dong-R4 (Wei et al., 2016)) is overexpressed. This microRNA interacts with mRNA of BACE1 protein (beta-secretase, which catalyzes the conversion of amyloid precursor to beta-amyloid (Samadian et al., 2021)). In the blood serum of AD patients, a reduced level of miR-4286 (Henriques et al., 2020), derived from ERV1 (Wei et al., 2016), and miR-4422-5p (derived from LTR/Gypsy (Wei et al., 2016)) was detected (Hajjari et al., 2021).

In 2019, in the search for potential biomarkers and therapeutic agents for AD, integration of transcriptomic data with protein-protein and transcriptional regulatory interactions revealed the role of miR-192-5p (derived from L2) and miR-335-5p (derived from SINE/MIR) (Wei et al., 2016) as key signaling and regulatory molecules associated with transcriptional changes in AD. Their levels decrease both in the blood of people with AD (Rahman et al., 2019) and in relation to miR-192-5p in the hippocampus of experimental mice. Further studies showed the potential protective efficacy of miR-192-5p in AD. The level of this microRNA decreased with exercise and contributed to a decrease in the expression of TNF- α , IL-6 and IL-1 β , which are involved in inflammation in AD (Qin et al., 2022). Similar results were obtained in experiments on cell cultures and AD mouse models regarding miR-335-5p, which can be used for targeted therapy of the disease (Wang et al., 2020).

In 2020, reduced expression of miR-340 (derived from TcMar-Mariner DNA TE (Wei et al., 2016)) was detected in mouse AD models (Tan et al., 2020). A low level of miR-

708-5p (derived from L2 (Wei et al., 2016)) was detected in blood samples of 28 patients with AD (Rahman et al., 2020). The data obtained were confirmed by studying the brain neurons of AD patients (Di Palo et al., 2022). Analysis of brain samples from patients who died from AD showed increased levels of miR-1202 (Henriques et al., 2020), derived from L1 (Wei et al., 2016).

In 2021, an analysis of DNA blood samples from 48 AD patients and 48 healthy controls showed a significant increase in the level of miR-378a (Dong Z. et al., 2021), which was derived from SINE/MIR (Wei et al., 2016). This microRNA has been proposed as a biomarker for AD. The brains of deceased AD patients showed reduced levels of miR-1271 (Majumder et al., 2021), which was derived from L2 (Wei et al., 2016). An increase in the expression of miR-4504 (derived from L1) in the brain of patients with AD was determined (Eysert et al., 2021). Our data on changes in the expression of specific microRNAs derived from TEs in AD are shown in Table 1.

Association of transposable element-derived microRNAs with aging, carcinogenesis and Alzheimer disease

Epidemiological studies indicate a significant increase in the risk of developing AD with age (Niu et al., 2017). In both aging and neurodegenerative diseases, genomic instability is observed in neurons, with activation of TEs by various mechanisms (Wood, Helfand, 2013; Guo C. et al., 2018), including loss of SIRT6 marks (Van Meter et al., 2014). Although AD and cancer are diseases associated with aging, an analysis of scientific literature has identified an inverse correlation between cancer and AD, which may be due to the influence of the proteins p53 and PIN1 (Peptidyl-prolyl cis-trans isomerase) (Lanni et al., 2021). At the same time, mortality from AD in people who survived cancer for 10 years or more was higher than in the general population (Abdel-Rahman, 2020), which may indicate the presence of common pathogenetic pathways of these diseases, possibly associated with TEs deregulation.

Changes in TEs activity during aging contribute to changes in the expression of microRNAs, which can contribute to the development of AD and suppress the growth of cancer (acting as tumor suppressors). To test this assumption, we analyzed an online resource created in 2018 by N.W. Wong et al. concerning changes in specific microRNAs in certain cancer types (Wong et al., 2018), as well as searched for scientific data on microRNAs associated with AD and aging. As a result, we identified 13 specific microRNAs derived from TEs associated with aging and simultaneously involved in the pathogenesis of AD and cancer (Table 2).

MiR-151a is associated with Alzheimer's disease (Satoh et al., 2015); expression changes of this microRNA are also characteristic of various cancers (Wong et al., 2018) and aging (Noren Hooten et al., 2013). LINE2-derived miR-192 (Wei et al., 2016), the level of which decreases in AD (Rahman et al., 2019; Qin et al., 2022), is associated with various cancers (Wong et al., 2018). miR-192 expression is significantly reduced in aging kidney tissue (Sataranatarajan et al., 2012). LINE/L2-derived miR-211 (Wei et al., 2016), the level of which increases in AD (Fan et al., 2016; Sierksma et al., 2018), is also associated with cancer (Wong et al., 2018). miR-211

Table 1. Association of transposon-derived microRNAs with Alzheimer’s disease

miRNA	Transposon, source of microRNA	Expression change	Reference
miR-1202	L1	Increased	Henriques et al., 2020
miR-1246	ERV1	Increased	Guo R. et al., 2017
miR-1271	L2	Decreased	Majumder et al., 2021
miR-151a	L2	Increased	Satoh et al., 2015
miR-192	L2	Decreased	Rahman et al., 2019; Qin et al., 2022
miR-211	L2	Increased	Fan et al., 2016; Sierksma et al., 2018
miR-28	L2	Increased	Hong et al., 2017; Zhao et al., 2020
miR-31	L2	Decreased	Dong H. et al., 2015; Barros-Viegas et al., 2020
miR-320	L1	Increased	Boese et al., 2016; Raheja et al., 2018
miR-3200	ERV1	Decreased	Satoh et al., 2015
miR-325	L2	Decreased	Barak et al., 2013
miR-326	DNA/hAT-Tip100	Increased	Cai et al., 2017
miR-335	SINE/MIR	Decreased	Rahman et al., 2019; Wang et al., 2020
miR-340	DNA/TcMar-Mariner	Decreased	Tan et al., 2020
miR-342	SINE	Decreased	Dakterzada et al., 2021
miR-3646	SINE/MIR	Increased	Lu et al., 2021
miR-378a	SINE/MIR	Increased	Dong Z. et al., 2021
miR-384	LINE/Dong-R4	Increased	Samadian et al., 2021
miR-4286	ERV1	Decreased	Henriques et al., 2020
miR-4422	LTR-Gypsy	Decreased	Hajjari et al., 2021
miR-4487	L1	Decreased	Hu et al., 2018
miR-4504	L1	Increased	Eysert et al., 2021
miR-502	L2	Decreased	Satoh et al., 2015
miR-511	L1	Decreased	Zheng et al., 2016; Jia et al., 2022
miR-545	L2	Decreased	Cosin-Tomas et al., 2017
miR-576	L1	Decreased	Liu et al., 2014; Xu et al., 2022
miR-708	L2	Decreased	Rahman et al., 2020; Di Palo et al., 2022
miR-885	SINE/MIR	Decreased	Tan et al., 2014; Pan et al., 2022

expression is increased in centenarians and may serve as a biomarker of aging (Smith-Vikos et al., 2016).

LINE2-derived miR-28 (Wei et al., 2016), the level of which increases in AD (Hong et al., 2017; Zhao et al., 2020), is also associated with specific cancers (Wong et al., 2018). Physiological aging is associated with decreased miR-28 production (Zhang T. et al., 2017). LINE2-derived miR-31 (Wei et al., 2016), the level of which decreases in AD (Dong H. et al., 2015; Barros-Viegas et al., 2020), is associated with cancer (Wong et al., 2018). Expression of this microRNA is increased during replicative aging (Dellago et al., 2013). MiR-320c, derived from LINE2 (Wei et al., 2016), the level of which is increased in AD (Boese et al., 2016; Raheja et al., 2018),

is also associated with specific cancers (Wong et al., 2018). MiR-320c levels decrease with aging (Ukai et al., 2012).

MiR-335, derived from SINE/MIR (Wei et al., 2016), which is reduced in AD (Rahman et al., 2019; Wang et al., 2020), is also associated with cancer (Wong et al., 2018) and aging (Raihan et al., 2018). MiR-340, derived from DNA-TE TcMar-Mariner (Wei et al., 2016), the expression of which is reduced in AD, is associated with cancer (Wong et al., 2018) and with aging (Zhang H. et al., 2015). SINE/MIR-derived miR-378a (Wei et al., 2016), the level of which is significantly increased in AD (Dong Z. et al., 2021), is associated with various cancers (Wong et al., 2018) and aging (Guo D. et al., 2017). The level of miR-511 (source: L1 (Wei et al., 2016)) decreases not only

Table 2. TE-derived miRNAs associated with aging, carcinogenesis and Alzheimer's disease

miRNA (TE-source)	Type of cancer (change in microRNA expression)	miRNA level change	
		in AD (reference)	in aging (reference)
miR-151a (LINE/L2)	BLCA, BRCA, CESC, COAD, ESCA, HNSC, KICH, KIRC, KIRP, LIHC, LUAD, LUSC, PRAD, READ, STAD, UCEC (increased)	Increased Satoh et al., 2015	Decreased Noren Hooten et al., 2013
miR-192 (LINE/L2)	BLCA, BRCA, COAD, KIRC, LUAD, LUSC, PRAD, READ, STAD, UCEC (increased); CHOL, KICH, KIRP, LIHC, THCA (decreased)	Decreased Rahman et al., 2019; Qin et al., 2022	Decreased Sataranatarajan et al., 2012
miR-211 (LINE/L2)	KIRC, KIRP, LIHC (increased); BRCA, HNSC, LUAD (decreased)	Increased Fan et al., 2016; Sierksma et al., 2018	Increased Smith-Vikos et al., 2016
miR-28 (LINE/L2)	HNSC, KIRC, LUAD, LUSC, PRAD (increased); BRCA, CHOL, COAD, ESCA, PCPG, READ, STAD, THCA (decreased)	Increased Hong et al., 2017; Zhao et al., 2020	Decreased Zhang T. et al., 2017
miR-31 (LINE/L2)	BLCA, CESC, HNSC, KIRP, LUAD, LUSC, STAD, THCA, UCEC (increased); KICH, KIRC, PRAD (decreased)	Decreased Dong H. et al., 2015; Barros-Viegas et al., 2020	Increased Cho et al., 2015; Dellago et al., 2013
miR-320c (LINE/L1, L2)	CHOL, KIRC, LUSC, STAD, UCEC (increased); COAD, READ (decreased)	Increased Boese et al., 2016; Raheja et al., 2018	Decreased Ukai et al., 2012
miR-335 (SINE/MIR)	BLCA, COAD, ESCA, HNSC, LUAD, LUSC, PRAD, STAD, THCA, UCEC (increased); BRCA, KICH, KIRC, LIHC (decreased)	Decreased Rahman et al., 2019; Wang et al., 2020	Increased Raihan et al., 2018
miR-340 (DNA/TcMar-Mariner)	BRCA, COAD, KICH, KIRC, KIRP, LUAD, LUSC, PRAD, UCEC (increased); CHOL, LIHC, PAAD (decreased)	Decreased Tan et al., 2020	Decreased Zhang H. et al., 2015
miR-378a (SINE/MIR)	PAAD (increased); BRCA, CHOL, COAD, HNSC, LIHC, LUAD, PAAD, PRAD, READ, STAD (decreased)	Increased Dong Z. et al., 2021	Increased Guo D. et al., 2017
miR-511 (LINE/L1)	HNSC, PRAD, READ, STAD (increased); BRCA, CHOL, KICH, KIRP, LIHC, LUSC, PCPG (decreased)	Decreased Zheng et al., 2016; Jia et al., 2022	Decreased Zheng et al., 2016
miR-576 (LINE/L1)	BLCA, BRCA, ESCA, HNSC, KICH, KIRC, KIRP, LUAD, LUSC, PRAD, READ, STAD, UCEC (increased); CHOL, LIHC, THCA (decreased)	Decreased Liu et al., 2014; Xu et al., 2022	Increased Ipson et al., 2018
miR-708 (LINE/L2)	BLCA, BRCA, CHOL, COAD, HNSC, KIRC, LUAD, LUSC, PRAD, READ, STAD (increased); KICH, THCA (decreased)	Decreased Rahman et al., 2020; Di Palo et al., 2022	Increased Lee et al., 2017
miR-885 (SINE/MIR)	KICH (increased); CHOL (decreased)	Decreased Tan et al., 2014; Pan et al., 2022	Increased Behbahanipour et al., 2019

Note. BLCA – bladder urothelial carcinoma; BRCA – breast invasive carcinoma; CESC – cervical squamous cell carcinoma and endocervical adenocarcinoma; CHOL – cholangiocarcinoma; COAD – colon adenocarcinoma; ESCA – esophageal carcinoma; HNSC – head and neck squamous cell carcinoma; KICH – kidney chromophobe carcinoma; KIRC – kidney renal clear cell carcinoma; KIRP – kidney renal papillary cell carcinoma; LIHC – liver hepatocellular carcinoma; LUAD – lung adenocarcinoma; LUSC – lung squamous cell carcinoma; PAAD – pancreatic adenocarcinoma; PRAD – prostate adenocarcinoma; PCPG – pheochromocytoma and paraganglioma; READ – rectal adenocarcinoma; STAD – stomach adenocarcinoma; THCA – thyroid carcinoma; UCEC – uterine corpus endometrial carcinoma.

in AD (Zheng et al., 2016; Jia et al., 2022), but also in aging. This microRNA is also associated with cancer (Wong et al., 2018). L1-derived miR-576 (Wei et al., 2016), the level of which is reduced in AD (Liu et al., 2014; Xu et al., 2022), is associated with cancer (Wong et al., 2018).

Increased expression of miR-576 is detected during aging (Ipson et al., 2018). MiR-708, derived from LINE2 (Wei et

al., 2016), a reduced level of which is observed in AD (Rahman et al., 2020; Di Palo et al., 2022), is associated with specific cancers (Wong et al., 2018) and with aging (Lee et al., 2017). The SINE/MIR-derived microRNA miR-885, which is downregulated in AD (Tan et al., 2014), is associated with chromophobe kidney cancer and cholangiocarcinoma (Wong et al., 2018) and aging (Behbahanipour et al., 2019).

Conclusion

The role of microRNAs in the development of AD indicates the possible potential of targeted therapy for the disease, as well as the search for the most optimal treatment regimens for AD. Examples are the decrease in miR-192-5p levels during exercise, which contributes to suppression of pro-inflammatory cytokines TNF- α , IL-6 and IL-1 β production (Qin et al., 2022); the decrease in miR-28-3p levels after donepezil therapy (which can be used as a diagnostic criterion for the effectiveness of treatment) (Zhao et al., 2020).

MicroRNAs can become not only therapeutic agents, but also highly accurate diagnostic markers, since changes in their levels are accompanied by regression in AD clinical picture, as was shown for miR-511 when exposed to acupuncture moxibustion in AD (Jia et al., 2022). The experimental effectiveness of miR-31 (Barros-Viegas et al., 2020) and miR-335-5p (Wang et al., 2020) in significantly reducing beta-amyloid accumulation in the hippocampus and its base indicates the potential of using this microRNA for targeting AD therapy (Barros-Viegas et al., 2020; Wang et al., 2020).

The association of TE-derived miRNAs with Alzheimer's disease indicates both the promise of their use in the treatment of AD and the need for a more detailed study of the mechanisms of action of these miRNAs, since the complementarity of their sequences with various TEs may become a likely basis for global changes in the expression of genes under TEs regulatory control.

References

- Abdel-Rahman O. Death from Alzheimer's disease among cancer survivors: a population-based study. *Curr. Med. Res. Opin.* 2020;36(5): 835-841. DOI 10.1080/03007995.2020.1734921
- Ando K., Nagaraj S., Kucukali F., de Fisenoy M.A., Kosa A.C., Doeraene E., Gutierrez L.L., Brion J.P., Leroy K. PICALM and Alzheimer's disease: An update and perspectives. *Nutrients.* 2022; 14(3):539. DOI 10.3390/nu14030539
- Baeken M.W., Moosmann B., Hajieva P. Retrotransposon activation by distressed mitochondria in neurons. *Biochem. Biophys. Res. Commun.* 2020;525(3):570-575. DOI 10.1016/j.bbrc.2020.02.106
- Barak B., Shvarts-Serebro I., Modai S., Gilam A., Okun E., Michaelson D.M., Mattson M.P., Shomron N., Ashery U. Opposing actions of environmental enrichment and Alzheimer's disease on the expression of hippocampal microRNA in mouse models. *Transl. Psychiatry.* 2013;3(9):e304. DOI 10.1038/tp.2013.77
- Barros-Viegas A.T., Carmona V., Ferreira E., Guedes J., Cardoso A.M., Cunha P., de Almeida L.P., de Oliveira C.R., de Magalhães J.P., Peça J., Cardoso A.L. miRNA-31 improves cognition and abolishes amyloid- β pathology by targeting APP and BACE1 in an animal model of Alzheimer's disease. *Mol. Ther. Nucleic Acids.* 2020;19: 1219-1236. DOI 10.1016/j.omtn.2020.01.010
- Behbahanipour M., Peymani M., Salari M., Hashemi M.S., Nasr-Esfahani M.H., Ghaedi K. Expression profiling of blood microRNAs 885, 361, and 17 in the patients with the Parkinson's disease: integrating interaction data to uncover the possible triggering age-related mechanisms. *Sci. Rep.* 2019;9:13759. DOI 10.1038/s41598-019-50256-3
- Boese A.S., Saba R., Campbell K., Majer A., Medina S., Burton L., Booth T.F., Chong P., Westmacott G., Dutta S.M., Saba J.A., Booth S.A. MicroRNA abundance is altered in synaptoneuroosomes during prion disease. *Mol. Cell. Neurosci.* 2016;71:13-24. DOI 10.1016/j.mcn.2015.12.001
- Cai Y., Sun Z., Jia H., Luo H., Ye X., Wu Q., Xiong Y., Zhang W., Wan J. *Rpph1* upregulates CDC42 expression and promotes hippocampal neuron dendritic spine formation by competing with miR-330-5p. *Front. Mol. Neurosci.* 2017;10:27. DOI 10.3389/fnmol.2017.00027
- Chanda K., Mukhopadhyay D. LncRNA Xist, X-chromosome instability and Alzheimer's disease. *Curr. Alzheimer Res.* 2020;17(6):499-507. DOI 10.2174/1567205017666200807185624
- Cheng Y., Saville L., Gollen B., Isaac C., Belay A., Mehla J., Patel K., Thakor N., Mohajerani M.H., Zovoilis A. Increased processing of SINE B2 ncRNAs unveils a novel type of transcriptome deregulation in amyloid beta neuropathology. *eLife.* 2020;9:e61265. DOI 10.7554/eLife.61265
- Cho J.H., Dimri M., Dimri G.P. MicroRNA-31 is a transcriptional target of histone deacetylase inhibitors and a regulator of cellular senescence. *J. Biol. Chem.* 2015;290(16):10555-10567. DOI 10.1074/jbc.M114.624361
- Cosin-Tomas M., Antonell A., Llado A., Alcolea D., Fortea J., Ezquerro M., Lleo A., Marti M.J., Pallas M., Sanchez-Valle R.S., Molinuevo J.L., Sanfeliu C., Kaliman P. Plasma miR-545-3p as early biomarkers of Alzheimer's disease: potential and limitations. *Mol. Neurobiol.* 2017;54(7):5550-5562. DOI 10.1007/s12035-016-0088-8
- Dakterzada F., Benítez I.D., Targa A., Lladó A., Torres G., Romero L., de Gonzalo-Calvo D., Moncusí-Moix A., Tort-Merino A., Huerto R., Sánchez-de-la-Torre M., Barbé F., Piñol-Ripoll G. Reduced levels of miR-342-5p in plasma are associated with worse cognitive evolution in patients with mild Alzheimer's disease. *Front. Aging Neurosci.* 2021;13:705989. DOI 10.3389/fnagi.2021.705989
- Dellago H., Preschitz-Kammerhofer B., Terlecki-Zaniewicz L., Schreiner C., Fortschegger K., Chang M.W., Hackl M., Monteforte R., Kuhnel H., Schosserer M., Gruber F., Tschachler E., Scheideler M., Grillari-Voglauer R., Grillari J., Wieser M. High levels of oncomiR-21 contribute to the senescence-induced growth arrest in normal human cells and its knock-down increases the replicative lifespan. *Aging Cell.* 2013;12(3):446-458. DOI 10.1111/accel.12069
- Di Palo A.D., Siniscalchi C., Crescente G., Leo I.D., Fiorentino A., Pacifico S., Russo A., Potenza N. Effect of cannabidiolic acid, *N-trans*-caffeoyltyramine and cannabisisin B from hemp seeds on microRNA expression in human neural cells. *Curr. Issues Mol. Biol.* 2022;44(10):5106-5116. DOI 10.3390/cimb44100347
- Dong H., Li J., Huang L., Chen X., Li D., Wang T., Hu C., Xu J., Zhang C., Zen K., Xiao S., Yan Q., Wang C., Zhang C.Y. Serum microRNA profiles serve as novel biomarkers for the diagnosis of Alzheimer's disease. *Dis. Markers.* 2015;2015:625659. DOI 10.1155/2015/625659
- Dong Z., Gu H., Guo Q., Liang S., Xue J., Yao F., Liu X., Li F., Liu H., Sun L., Zhao K. Profiling of serum exosome miRNA reveals the potential of a miRNA panel as diagnostic biomarker for Alzheimer's disease. *Mol. Neurobiol.* 2021;58(7):3084-3094. DOI 10.1007/s12035-021-02323-y
- El Hajjar J., Chatoo W., Hanna R., Nkanza P., Tétreault N., Tse Y.C., Wong T.P., Abdouh M., Bernier G. Heterochromatic genome instability and neurodegeneration sharing similarities with Alzheimer's disease in old *Bmi1*^{+/-} mice. *Sci. Rep.* 2019;9(1):594. DOI 10.1038/s41598-018-37444-3
- Eysert F., Coulon A., Boscher E., Vreulx A.C., Flaig A., Mendes T., Kilinc D., Lambert J., Chapuis J. Alzheimer's genetic risk factor FERMT2 (Kindlin-2) controls axonal growth and synaptic plasticity in an APP-dependent manner. *Mol. Psychiatry.* 2021;26(10):5592-5607. DOI 10.1038/s41380-020-00926-w
- Fagone P., Mangano K., Martino G., Quattropani M.C., Pennisi M., Bella R., Fiscaro F., Nicoletti F., Petralia M.C. Characterization of altered molecular pathways in the entorhinal cortex of Alzheimer's disease patients and *in silico* prediction of potential repurposable drugs. *Genes (Basel).* 2022;13(4):703. DOI 10.3390/genes13040703
- Fan C., Wu Q., Ye X., Luo H., Yan D., Xiong D., Xiong Y., Zhu H., Diaoy Y., Zhang W., Wan J. Role of miR-211 in neuronal differentiation and viability: implications to pathogenesis of Alzheimer's disease. *Front. Aging Neurosci.* 2016;8:166. DOI 10.3389/fnagi.2016.00166

- Flamier A., El Hajjar J., Adjaye J., Fernandes K.J., Abdouh M., Bernier G. Modeling late-onset sporadic Alzheimer's disease through BMI1 deficiency. *Cell Rep.* 2018;23(9):2653-2666. DOI 10.1016/j.celrep.2018.04.097
- Gatz M., Reynolds C.A., Fratiglioni L., Johansson B., Mortimer J.A., Berg S., Fiske A., Pedersen N.L. Role of genes and environments for explaining Alzheimer disease. *Arch. Gen. Psychiatry.* 2006;63(2): 168-174. DOI 10.1001/archpsyc.63.2.168
- GNS H.S., Marise V.L.P., Satish K.S., Yergolkar A.V., Krishnamurthy M., Rajalekshmi G.S., Radhika K., Burri R.R. Untangling huge literature to disinter genetic underpinnings of Alzheimer's disease: A systematic review and meta-analysis. *Ageing Res. Rev.* 2021;71: 101421. DOI 10.1016/j.arr.2021.101421
- Goate A. Segregation of a missense mutation in the amyloid beta-protein precursor gene with familial Alzheimer's disease. *J. Alzheimers Dis.* 2006;9(3 Suppl.):341-347. DOI 10.3233/jad-2006-9s338
- Grundman J., Spencer B., Sarsoza F., Rissman R.A. Transcriptome analyses reveal tau isoform-driven changes in transposable element and gene expression. *PLoS One.* 2021;16(9):e0251611. DOI 10.1371/journal.pone.0251611
- Guerreiro R., Wojtas A., Bras J., Carrasquillo M., Rogava E., Majounie E., Cruchaga C., Sassi C., Kauwe J.S., Younkin S., Hazrati L., Collinge J., Pocock J., Lashley T., Williams J., Lambert J.C., Amouyel P., Goate A., Rademakers R., Morgan K., Powell J., St. George-Hyslop P., Singleton A., Hardy J., Alzheimer Genetic Analysis Group. TREM2 variants in Alzheimer's disease. *N. Engl. J. Med.* 2013;368(2):117-127. DOI 10.1056/NEJMoa1211851
- Guo C., Jeong H.H., Hsieh Y.C., Klein H.U., Bennett D.A., De Jager P.L., Liu Z., Shulman J.M. Tau activates transposable elements in Alzheimer's disease. *Cell Rep.* 2018;23(10):2874-2880. DOI 10.1016/j.celrep.2018.05.004
- Guo D., Ye Y., Qi J., Tan X., Zhang Y., Ma Y., Li Y. Age and sex differences in microRNAs expression during the process of thymus aging. *Acta Biochim. Biophys. Sin. (Shanghai).* 2017;49(5):409-419. DOI 10.1093/abbs/gmx029
- Guo R., Fan G., Zhang J., Wu C., Du Y., Ye H., Li Z., Wang L., Zhang Z., Zhang L., Zhao Y., Lu Z. A 9-microRNA signature in serum serves as a noninvasive biomarker in early diagnosis of Alzheimer's disease. *J. Alzheimers Dis.* 2017;60(4):1365-1377. DOI 10.3233/JAD-170343
- Hajjari S.N., Sadigh-Eteghad S., Shanebandi D., Teimourian S., Shahbazi A., Mehdizadeh M. MicroRNA-4422-5p as a negative regulator of amyloidogenic secretases: A potential biomarker for Alzheimer's disease. *Neuroscience.* 2021;463:108-115. DOI 10.1016/j.neuroscience.2021.03.028
- Hanna R., Flamier A., Barabino A., Bernier G. G-quadruplexes originating from evolutionary conserved L1 elements interfere with neuronal gene expression in Alzheimer's disease. *Nat. Commun.* 2021; 12(1):1828. DOI 10.1038/s41467-021-22129-9
- Harold D., Abraham R., Hollingworth P., Sims R., Gerrish A., Hamshe M.L., Pahwa J.S., Moskva V., Dowzell K., Williams A., Jones N., Thomas C., Stretton A., Morgan A.R., Loveston S., Powell J., Proitsi P., Klopp N., Wichmann H.E., Carrasquillo M.M., Pankratz V.S., Younkin S.G., Holmans P.A., O'Donovan M., Owen M.J., Williams J. Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. *Nat. Genet.* 2009; 41(10):1088-1093. DOI 10.1038/ng.440
- Henriques A.D., Machado-Silva W., Leite R.E.P., Suemoto C.K., Leite K.R.M., Srougi M., Pereira A.C., Jacob-Filho W., Brazilian Aging Brain Study Group. Genome-wide profiling and predicted significance of post-mortem brain microRNA in Alzheimer's disease. *Mech. Ageing Dev.* 2020;191:111352. DOI 10.1016/j.mad.2020.111352
- Hong H., Li Y., Su B. Identification of circulating miR-125b as a potential biomarker of Alzheimer's disease in APP/PS1 transgenic mouse. *J. Alzheimers Dis.* 2017;59(4):1449-1458. DOI 10.3233/JAD-170156
- Hou Y., Song H., Croteau D.L., Akbari M., Bohr V.A. Genome instability in Alzheimer disease. *Mech. Ageing Dev.* 2017;161(Pt. A):83-94. DOI 10.1016/j.mad.2016.04.005
- Hu L., Zhang R., Yuan Q., Gao Y., Yang M.Q., Zhang C., Huang J., Sun Y., Yang W., Yang J.Y., Min Z., Cheng J., Deng Y., Hu X. The emerging role of microRNA-4487/6845-3p in Alzheimer's disease pathologies is induced by A β 25-35 triggered in SH-SY5Y cell. *BMC Syst. Biol.* 2018;12(Suppl. 7):119. DOI 10.1186/s12918-018-0633-3
- Ipson B.R., Fletcher M.B., Espinoza S.E., Fisher A.L. Identifying exosome-derived microRNAs as candidate biomarkers of frailty. *J. Frailty Aging.* 2018;7(2):100-103. DOI 10.14283/jfa.2017.45
- Jia Y.M., Zhu C.F., She Z.Y., Wu M.M., Wu Y.Y., Zhou B.Y., Zhang N. Effects on autophagy of moxibustion at governor vessel acupoints in APP/PS1 double-Transgenic Alzheimer's Disease Mice through the lncRNA Six3os1/miR-511-3p/AKT3 Molecular Axis. *Evid. Based Complement. Alternat. Med.* 2022;2022:3881962. DOI 10.1155/2022/3881962
- Kunkle B.W., Jaworski J., Barral S., Bardarajan B., Beecham G.W., Haines J.L., Pericak-Vance M. Genome-wide linkage analyses of non-Hispanic white families identify novel loci for familial late-onset Alzheimer's disease. *Alzheimer's Dement.* 2016;12(1):2-10. DOI 10.1016/j.jalz.2015.05.020
- Lambert J.C., Heath S., Even G., Campion D., Sleegers K., Hil-tunen M., Combarros O., Zelenika D., Bullido M.J., Tavernier B., Letenneur L., Bettens K., Berr C., Pasquier F., Fievet N., Barberger-Gateau P., Engelborghs S., Deyn P.D., Mateo I., Franck A., Helisalmi S., Tzourio C., Gut I., Van Broeckhoven C., Alperovitch A., Lathrop M., Amouyel P. Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease. *Nat. Genet.* 2009;41(10):1094-1099. DOI 10.1038/ng.439
- Lanni C., Masi M., Racchi M., Govoni S. Cancer and Alzheimer's disease inverse relationship: an age-associated diverging derailment of shared pathways. *Mol. Psychiatry.* 2021;26(1):280-295. DOI 10.1038/s41380-020-0760-2
- Larsen P.A., Lutz M.W., Hunnicutt K.E., Mihovilovic M., Saunders A.M., Yoder A.D., Roses A.D. The *Alu* neurodegeneration hypothesis: A primate-specific mechanism for neuronal transcription noise, mitochondrial dysfunction, and manifestation of neurodegenerative disease. *Alzheimer's Dement.* 2017;13(7):828-838. DOI 10.1016/j.jalz.2017.01.017
- Lee B.P., Buric I., George-Pandeth A., Flurkey K., Harrison D.E., Yuan R., Peters L.L., Kuchel G.A., Melzer D., Harries L.W. MicroRNAs miR-203-3p, miR-664-3p and miR-708-5p are associated with median strain lifespan in mice. *Sci. Rep.* 2017;7:44620. DOI 10.1038/srep44620
- Levy-Lahad E., Wasco W., Poorkaj P., Romano D.M., Oshima J., Pettingell W.H., Yu C.E., Jondro P.D., Schmidt S.D., Wang K. Candidate gene for the chromosome 1 familial Alzheimer's disease locus. *Science.* 1995;269(5226):973-977. DOI 10.1126/science.7638622
- Liu Q.Y., Chang M.N.V., Lei J.X., Koukiekolo R., Smith B., Zhang D., Ghribi O. Identification of microRNAs involved in Alzheimer's progression using a rabbit model of the disease. *Am. J. Neurodegener. Dis.* 2014;3(1):33-44
- Lu L., Dai W., Zhu X., Ma T. Analysis of serum miRNAs in Alzheimer's disease. *Am. J. Alzheimers Dis. Other Dement.* 2021;36: 15333175211021712. DOI 10.1177/15333175211021712
- Ma F.C., Wang H.F., Cao X.P., Tan C.C., Tan L., Yu J.T. Meta-analysis of the association between variants in ABCA7 and Alzheimer's disease. *J. Alzheimers Dis.* 2018;63(4):1261-1267. DOI 10.3233/JAD-180107
- Macciardi F., Bacalini M.G., Miramontes R., Boattini A., Taccioli C., Modenini G., Malhas R., Anderlucci L., Gusev Y., Gross T.J., Paddilla R.M., Fiandaca M.S., Head E., Guffanti G., Federoff H.J., Mapstone M. A retrotransposon storm marks clinical phenocopy conversion to late-onset Alzheimer's disease. *Geroscience.* 2022;44(3):1525-1550. DOI 10.1007/s11357-022-00580-w

- Majumder P., Chanda K., Das D., Singh B.K., Charkrabarti P., Jana N.R., Mukhopadhyay D. A nexus of miR-1271, PAX4 and ALK/Ryk influences the cytoskeletal architectures in Alzheimer's disease and type 2 diabetes. *Biochem. J.* 2021;478(17):3297-3317. DOI 10.1042/BCJ20210175
- Marioni R.E., Harris S.E., Zhang Q., McRae A.F., Hagenaars S.P., Hill W.D., Davies G., Ritchie C.W., Gale C.R., Starr J.M., Goate A.M., Porteous D.J., Yang J., Evans K.L., Deary I.J., Wray N.R., Visscher P.M. GWAS on family history of Alzheimer's disease. *Transl. Psychiatry.* 2018;8(1):99. DOI 10.1038/s41398-018-0150-6
- Mustafin R.N. The relationship between transposons and transcription factors in the evolution of eukaryotes. *Zhurnal Evolyutsionnoi Biokhimii i Fiziologii = Journal of Evolutionary Biochemistry and Physiology.* 2019;55(1):14-22. DOI 10.1134/S004445291901008X (in Russian)
- Mustafin R.N., Khusnutdinova E.K. Non-coding parts of genomes as the basis of epigenetic heredity. *Vavilovskii Zhurnal Genetiki i Selektzii = Vavilov Journal of Genetics and Breeding.* 2017;21(6):742-749. DOI 10.18699/VJ17.30-o (in Russian)
- Niu H., Alvarez-Alvarez I., Guillen-Grima F., Aguinaga-Ontoso I. Prevalence and incidence of Alzheimer's disease in Europe: A meta-analysis. *Neurologia.* 2017;32(8):523-532. DOI 10.1016/j.nrl.2016.02.016
- Noren Hooten N., Fitzpatrick M., Wood W.H. 3rd, De S., Ejiogu N., Zhang Y., Mattison J.A., Becker K.G., Zonderman A.B., Evans M.K. Age-related changes in microRNA levels in serum. *Aging (Albany N. Y.).* 2013;5(10):725-740. DOI 10.18632/aging.100603
- Pan W., Hu Y., Wang L., Li J. Circ_0003611 acts as a miR-885-5p sponge to aggravate the amyloid- β -induced neuronal injury in Alzheimer's disease. *Metab. Brain Dis.* 2022;37(4):961-971. DOI 10.1007/s11011-022-00912-x
- Pascarella G., Hon C.C., Hashimoto K., Busch A., Luginbuhl J., Parr C., Yip W.H., Abe K., Kratz A., Bonetti A., Agostini F., Severin J., Murayama S., Suzuki Y., Gustincich S., Frith M., Carninci P. Recombination of repeat elements generates somatic complexity in human genomes. *Cell.* 2022;185(16):3025-3040.e6. DOI 10.1016/j.cell.2022.06.032
- Patel H., Dobson R.J.B., Newhouse S.J. A meta-analysis of Alzheimer's disease brain transcriptomic data. *J. Alzheimers Dis.* 2019; 68(4):1635-1656. DOI 10.3233/JAD-181085
- Protasova M.S., Andreeva T.V., Rogaev E.I. Factors regulating the activity of LINE1 retrotransposons. *Genes (Basel).* 2021;12(10):1562. DOI 10.3390/genes12101562
- Qin Z., Han X., Ran J., Guo S., Lv L. Exercise-mediated alteration of miR-192-5p is associated with cognitive improvement in Alzheimer's disease. *Neuroimmunomodulation.* 2022;29(1):36-43. DOI 10.1159/000516928
- Raheja R., Regev K., Healy B.C., Mazzola M.A., Beynon V., Glehn F.V., Paul A., Diaz-Cruz C., Gholipour T., Glanz B.I., Kivisakk P., Chitnis T., Weiner H.L., Berry J.D., Gandhi R. Correlating serum microRNAs and clinical parameters in amyotrophic lateral sclerosis. *Muscle Nerve.* 2018;58(2):261-269. DOI 10.1002/mus.26106
- Rahman M.R., Islam T., Turanlı B., Zaman T., Faruquee H.M., Rahman M.M., Mollah M.N.H., Nanda R.K., Arga K.Y., Gov E., Moni M.A. Network-based approach to identify molecular signatures and therapeutic agents in Alzheimer's disease. *Comput. Biol. Chem.* 2019;78:431-439. DOI 10.1016/j.compbiolchem.2018.12.011
- Rahman M.R., Islam T., Zaman T., Shahjaman M., Karim M.R., Huq F., Quinn J.M.W., Holsinger R.M.D., Gov E., Moni M.A. Identification of molecular signatures and pathways to identify novel therapeutic targets in Alzheimer's disease: Insights from a systems biomedicine perspective. *Genomics.* 2020;112(2):1290-1299. DOI 10.1016/j.ygeno.2019.07.018
- Raihan O., Brishti A., Molla M.R., Li W., Zhang Q., Xu P., Khan M.I., Zhang J., Liu Q. The age-dependent elevation of miR-335-3p leads to reduced cholesterol and impaired memory in brain. *Neuroscience.* 2018;390:160-173. DOI 10.1016/j.neuroscience.2018.08.003
- Ramirez P., Zuniga G., Sun W., Beckmann A., Ochoa E., DeVos S., Hyman B., Chiu G., Roy E.R., Cao W., Orr M., Buggia-Prevot V., Ray W.J., Frost B. Pathogenic tau accelerates aging-associated activation of transposable elements in the mouse central nervous system. *Prog. Neurobiol.* 2022;208:102181. DOI 10.1016/j.pneurobio.2021.102181
- Ravel-Godreuil C., Zhaidi R., Bonnifet T., Joshi R.L., Fuchs J. Transposable elements as new players in neurodegenerative diseases. *FEBS Lett.* 2021;595(22):2733-2755. DOI 10.1002/1873-3468.14205
- Robinson M., Lee B.Y., Hane F.T. Recent progress in Alzheimer's disease research. Part 2: genetics and epidemiology. *J. Alzheimers Dis.* 2017;57(2):317-330. DOI 10.3233/JAD-161149
- Rogaev E.I., Lukiw W.J., Lavrushina O., Rogaeva E.A., St. George-Hyslop P.H. The upstream promoter of the β -amyloid precursor protein gene (*APP*) shows differential patterns of methylation in human brain. *Genomics.* 1994;22(2):340-347. DOI 10.1006/geno.1994.1393
- Samadian M., Gholipour M., Hajiesmaeili M., Taheri M., Ghafouri-Fard S. The eminent role of microRNAs in the pathogenesis of Alzheimer's disease. *Front. Aging Neurosci.* 2021;13:641080. DOI 10.3389/fnagi.2021.641080
- Sataranatarajan K., Feliels D., Mariappan M.M., Lee H.J., Lee M.J., Day R.T., Bindu H., Yalamanchili H.B., Choudhury G.G., Barnes J.L., Remmen H.V., Richardson A., Kasinath B.S. Molecular events in matrix protein metabolism in the aging kidney. *Aging Cell.* 2012;11(6):1065-1073. DOI 10.1111/accel.12008
- Satoh J.I., Kino Y., Niida S. MicroRNA-seq data analysis pipeline to identify blood biomarkers for Alzheimer's disease from public data. *Biomark. Insights.* 2015;10:21-31. DOI 10.4137/BMI.S25132
- Schwartzentruber J., Cooper S., Liu J.Z., Barrio-Hernandez I., Bello E., Kumasaka N., Young A.M.H., Franklin R.J.M., Johnson T., Estrada K., Gaffney D.J., Beltrao P., Bassett A. Genome-wide meta-analysis, fine-mapping and integrative prioritization implicate new Alzheimer's disease risk genes. *Nat. Genet.* 2021;53(3):392-402. DOI 10.1038/s41588-020-00776-w
- Serrano-Pozo A., Das S., Hyman B.T. APOE and Alzheimer's disease: advances in genetics, pathophysiology, and therapeutic approaches. *Lancet Neurol.* 2021;20(1):68-80. DOI 10.1016/S1474-4422(20)30412-9
- Sherrington R., Rogaev E.I., Liang Y., Rogaeva E.A., Levesque G., Ikeda M., Chi H., Lin C., Li G., Holman K., Tsuda T., Mar L., Foncin J.F., Bruni A.C., Montesi M.P., Sorbi S., Rainero I., Pinessi L., Nee L., Chumakov I., Pollen D., Brookes A., Sanseau P., Polinsky R.J., Wasco W., Da Silva H.A., Haines J.L., Perikicak-Vance M.A., Tanzi R.E., Roses A.D., Fraser P.E., Rommens J.M., St. George-Hyslop P.H. Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease. *Nature.* 1995;375(6534):754-760. DOI 10.1038/375754a0
- Sierksma A., Lu A., Salta E., Eynden E.V., Callaerts-Vegh Z., D'Hooge R., Blum D., Buee L., Fiers M., Stooper B.D. Dereglulation of neuronal miRNAs induced by amyloid- β or TAU pathology. *Mol. Neurodegener.* 2018;13(1):54. DOI 10.1186/s13024-018-0285-1
- Smith R.G., Pishva E., Shireby G., Smith A.R., Roubroeks J.A.Y., Hannon E., Wheildon G., Mastroeni D., Gasparoni G., Riemschneider M., Giese A., Sharp A.J., Schalkwyk L., Haroutunian V., Viechtbauer W., van den Hove D.L.A., Weedon M., Brokaw D., Francis P.T., Thomas A.J., Love S., Morgan K., Walter J., Coleman P.D., Bennett D.A., De Jager P.L., Mill J., Lunnon K. A meta-analysis of epigenome-wide association studies in Alzheimer's disease highlights novel differentially methylated loci across cortex. *Nat. Commun.* 2021;12(1):3517. DOI 10.1038/s41467-021-23243-4
- Smith-Vikos T., Liu Z., Parsons C., Gorospe M., Ferrucci L., Gill T.M., Slack F.J. A serum miRNA profile of human longevity: findings from the Baltimore Longitudinal Study of Aging (BLSA). *Aging (Albany N.Y.).* 2016;8(11):2971-2987. DOI 10.18632/aging.101106

- Sun W., Samimi H., Gamez M., Zare H., Frost B. Pathogenic tau-induced piRNA depletion promotes neuronal death through transposable element dysregulation in neurodegenerative tauopathies. *Nat. Neurosci.* 2018;21(8):1038-1048. DOI 10.1038/s41593-018-0194-1
- Swarbrick S., Wragg N., Ghosh S., Stolzing A. Systematic review of miRNA as biomarkers in Alzheimer's disease. *Mol. Neurobiol.* 2019;56(9):6156-6167. DOI 10.1007/s12035-019-1500-y
- Tan L., Yu J.T., Tan M.S., Liu Q.Y., Wang H.F., Zhang W., Jiang T., Tan L. Genome-wide serum microRNA expression profiling identifies serum biomarkers for Alzheimer's disease. *J. Alzheimers Dis.* 2014;40(4):1017-1027. DOI 10.3233/JAD-132144
- Tan X., Luo Y., Pi D., Xia L., Li Z., Tu Q. MiR-340 reduces the accumulation of amyloid- β through targeting BACE1 (β -site amyloid precursor protein cleaving enzyme 1) in Alzheimer's disease. *Curr. Neurovasc. Res.* 2020;17(1):86-92. DOI 10.2174/1567202617666200117103931
- Ukai T., Sato M., Akutsu H., Umezawa A., Mochida J. MicroRNA-199a-3p, microRNA-193b, and microRNA-320c are correlated to aging and regulate human cartilage metabolism. *J. Orthop. Res.* 2012;30(12):1915-1922. DOI 10.1002/jor.22157
- Van Meter M., Kashyap M., Rezazadeh S., Geneva A.J., Morello T.D., Seluanov A., Gorbunova V. SIRT6 represses LINE1 retrotransposons by ribosylating KAP1 but this repression fails with stress and age. *Nat. Commun.* 2014;5:5011. DOI 10.1038/ncomms6011
- Wang D., Fei Z., Wang H. MiR-335-5p inhibits β -amyloid (A β) accumulation to attenuate cognitive deficits through targeting c-jun-N-terminal kinase 3 in Alzheimer's disease. *Curr. Neurovasc. Res.* 2020;17(1):93-101. DOI 10.2174/1567202617666200128141938
- Watcharanurak P., Mutirangura A. Human RNA-directed DNA-methylation methylates high-mobility group box 1 protein-produced DNA gaps. *Epigenomics.* 2022;14(12):741-756. DOI 10.2217/epi-2022-0022
- Wei G., Qin S., Li W., Chen L., Ma F. MDTE DB: a database for microRNAs derived from Transposable element. *IEEE/ACM Trans. Comput. Biol. Bioinform.* 2016;13:1155-1160. DOI 10.1109/TCBB.2015.2511767
- Wong N.W., Chen Y., Chen S., Wang X. OncomiR: an online resource for exploring pan-cancer microRNA dysregulation. *Bioinformatics.* 2018;34(4):713-715. DOI 10.1093/bioinformatics/btx627
- Wood J.G., Helfand S.L. Chromatin structure and transposable elements in organismal aging. *Front. Genet.* 2013;4:274. DOI 10.3389/fgene.2013.00274
- Xu X., Gu D., Xu B., Yang C., Wang L. Circular RNA circ_0005835 promotes neural stem cells proliferation and differentiates to neuron and inhibits inflammatory cytokines levels through miR-576-ep in Alzheimer's disease. *Environ. Sci. Pollut. Res. Int.* 2022;29(24):35934-35943. DOI 10.1007/s11356-021-17478-3
- Yurov Y.B., Vorsanova S.G., Iourov I.Y. FISHing for chromosome instability and aneuploidy in the Alzheimer's disease brain. *Methods Mol. Biol.* 2023;2561:191-204. DOI 10.1007/978-1-0716-2655-9_10
- Zhang H., Yang H., Zhang C., Jing Y., Wang C., Liu C., Zhang R., Wang J., Zhang J., Zen K., Zhang C., Li D. Investigation of microRNA expression in human serum during the aging process. *J. Gerontol. A Biol. Sci. Med. Sci.* 2015;70(1):102-109. DOI 10.1093/gerona/glu145
- Zhang T., Brinkley T.E., Liu K., Feng X., Marsh A.P., Kritchevsky S., Zhou X., Nicklas B.J. Circulating miRNAs as biomarkers of gait speed responses to aerobic exercise training in obese older adults. *Aging (Albany N.Y.)*. 2017;9(3):900-913. DOI 10.18632/aging.101199
- Zhao X., Wang S., Sun W. Expression of miR-28-3p in patients with Alzheimer's disease before and after treatment and its clinical value. *Exp. Ther. Med.* 2020;20(3):2218-2226. DOI 10.3892/etm.2020.8920
- Zheng D., Sabbagh J.J., Blair L.J., Darling A.L., Wen X., Dickey C.A. MicroRNA-511 binds to FKBP5 mRNA, which encodes a chaperone protein, and regulates neuronal differentiation. *J. Biol. Chem.* 2016;291(34):17897-17906. DOI 10.1074/jbc.M116.727941

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

Received November 9, 2022. Revised October 30, 2023. Accepted November 2, 2023.

DOI 10.18699/vjgb-24-28

Non-viral systems for intracellular delivery of genome editing tools

I.H. Shaikhutdinov , P.V. Ilyasov  , O.V. Gribkova , L.V. Limareva 

Samara State Medical University of the Ministry of Healthcare of the Russian Federation, Samara, Russia

 p.v.ilyasov@samsmu.ru

Abstract. A hallmark of the last decades is an extensive development of genome editing systems and technologies propelling genetic engineering to the next level. Specific and efficient delivery of genome editing tools to target cells is one of the key elements of such technologies. Conventional vectors are not always suitable for this purpose due to a limited cargo volume, risks related to cancer and immune reactions, toxicity, a need for high-purity viral material and quality control, as well as a possibility of integration of the virus into the host genome leading to overexpression of the vector components and safety problems. Therefore, the search for novel approaches to delivering proteins and nucleic acids into cells is a relevant priority. This work reviews abiotic vectors and systems for delivering genome editing tools into target cells, including liposomes and solid lipid particles, other membrane-based vesicles, cell-penetrating peptides, micelles, dendrimers, carbon nanotubes, inorganic, polymer, metal and other nanoparticles. It considers advantages, drawbacks and preferred applications of such systems as well as suitability thereof for the delivery of genome editing systems. A particular emphasis is placed on metal-organic frameworks (MOFs) and their potential in the targeted intracellular delivery of proteins and polynucleotides. It has been concluded that further development of MOF-based vectors and technologies, as well as combining MOFs with other carriers can result in safe and efficient delivery systems, which would be able to circulate in the body for a long time while recognizing target cells and ensuring cell-specific delivery and release of intact cargoes and, thereby, improving the genome editing outcome.

Key words: metal-organic frameworks; vesicles; nanoparticles; viral vectors; gene editing.

For citation: Shaikhutdinov I.H., Ilyasov P.V., Gribkova O.V., Limareva L.V. Non-viral systems for intracellular delivery of genome editing tools. *Vavilovskii Zhurnal Genetiki i Seleksii = Vavilov Journal of Genetics and Breeding*. 2024;28(2): 239-248. DOI 10.18699/vjgb-24-28

Невирусные системы внутриклеточной доставки инструментов редактирования генома

И.Х. Шайхутдинов , П.В. Ильясов  , О.В. Грибкова , Л.В. Лимарева 

Самарский государственный медицинский университет Министерства здравоохранения Российской Федерации, Самара, Россия

 p.v.ilyasov@samsmu.ru

Аннотация. Последние десятилетия отмечены интенсивным развитием технологий и систем редактирования генов, которое вывело генную инженерию на новый уровень. Важным звеном этих технологий является специфичная и эффективная доставка компонентов таких систем в клетки-мишени. Традиционные векторы не всегда подходят для этой цели ввиду ограниченного объема полезной нагрузки, рисков, связанных с канцерогенезом и иммуногенностью, токсичности, необходимости высокой степени очистки и оценки качества полученных вирусных носителей, а также возможности встраивания вируса в геном хозяина, что может приводить к сверхэкспрессии компонентов вируса и проблемам с безопасностью. Это обуславливает актуальность поиска новых средств внутриклеточной доставки белков и нуклеиновых кислот. В данной работе приведен обзор абиотических векторов и систем доставки инструментов для редактирования генома, включая липосомы и твердые липидные наночастицы, мембранные везикулы иной природы, пептиды, проникающие в клетки, мицеллы, дендримеры, углеродные нанотрубки, неорганические, полимерные и другие наночастицы, металл-органические каркасные полимеры. Рассмотрены их преимущества, недостатки и предпочтительные области применения, а также возможность их использования для доставки систем редактирования генов. Особое внимание уделено металл-органическим каркасным полимерам и их потенциалу в качестве средств избирательной внутриклеточной доставки белков и полинуклеотидов. Сделан вывод о том, что дальнейшее развитие таких векторов и технологий на их основе может привести к появлению безопасных и эффективных систем доставки, способных длительно циркулировать в крови и распознавать клетки-мишени, обеспечивая адресное высвобождение полезной нагрузки в неизменном состоянии и тем самым улучшая результаты редактирования генов.

Ключевые слова: металл-органические каркасные полимеры; везикулы; наночастицы; вирусные векторы; редактирование генов.

Introduction

The last decades were marked by the development of novel strategies and genome editing tools for treatment of hereditary and acquired diseases. Such tools include but are not limited to specific synthetic oligonucleotides, recombinant zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), genome editing systems based on clustered regularly interspaced short palindromic repeats and associated enzymes (CRISPR/Cas), and genomic DNA base editors. Their efficacy strongly depends on the methods for delivery thereof into the target cells and tissues. Currently, various approaches and vector systems, having their specific advantages and drawbacks, are being used for these purposes.

The major challenges of such delivery inherent for genome editing tools include a large size of CRISPR/Cas or TALEN components, a large negative charge of RNAs, immunogenic potential, low efficacy, and off-target side effects (Singh D. et al., 2016). Transfection of such tools is also complicated by multiple factors impeding the cell and nucleus penetration by nucleic acids and proteins, with additional issues and limitations often conferred by the delivery methods and systems which were supposedly designed to facilitate the passing of the barriers. All these reduce the efficacy of genetic manipulations with the target cells.

The above tools have been delivered into the cells using a variety of techniques including electroporation, mechanoporation, microinjection, hydrodynamic injection, sonoporation, etc. (Moscoso, Steer, 2020). Among them, the most common ones are electroporation, due to its ease of use, high efficiency of *in vivo* transfection and genome editing, and microinjection, which allows to inject DNA directly to the nucleus. Particularly, when using CRISPR/Cas, microinjection allows to control the amount of Cas/sgRNA complex to be injected and to overcome the molecular weight limitations (Wang H.X. et al., 2017). The main limitations of the electroporation are low cell viability after the manipulations and a need to adjust the technique to the particular cells and vectors. In the case of microinjection, the limitations include relatively high complexity, labor intensity and cost of the procedure. Moreover, these methods are not suitable for all tissues of the body *in vivo*, and they have generally been used for small animal genome editing.

Another approach to the delivery of nucleic acids and proteins into the cells is based on vectors, which are able to penetrate the cells without using any ancillary tools. Conventionally, this assumes the use of viral vectors as they have an evolutionarily optimized machinery for introducing their genetic material into host cells. They are highly stable, can readily penetrate biological barriers, drive efficient transfection and induce long-term gene expression, and are able to infect both proliferating and nonproliferating cells (Huang et al., 2011). At the same time, serious disadvantages of the viral vectors include restricted cargo volume, cancer risk, immunogenic properties, toxicity, and a need for high purification and quality control of the vector used. Moreover, many viral vectors integrate themselves into the target cell genome, which may result in the overexpression of the genome editing system components and potentially cause safety issues (Hanlon et al., 2019).

Therefore, search for and development of alternative non-viral vector systems that would be able to bind nucleic acids and proteins and release them in a controlled manner is a relevant priority. Such delivery systems should have a number of advantages, particularly, an ability to load and deliver large molecules, an ease of preparation, low toxicity, minimal immune reactivity, and a possibility of customization of the properties defining their practical implementation. Almost all abiotic vectors have a positive charge required for electrostatic DNA complexing (Mintzer, Simanek, 2009). In contrast with other delivery systems, they are able to transfer the editing complexes in various forms including DNA, ribonucleoproteins and mRNA (Liu C. et al., 2019; Niggemann et al., 2020). Notably, non-viral vectors perform transient delivery, which is preferred in some cases of genome editing. Genome editing components are degraded shortly after cell penetration, thereby reducing the off-target effects (Mout et al., 2017a). In addition, many non-viral vectors can be commercially manufactured with the defined parameters.

Genome editing tool delivery systems such as liposomes and solid lipid particles, other membrane-based vesicles, cell-penetrating peptides, micelles, dendrimers, carbon nanotubes, inorganic, polymer, metal and other nanoparticles, and metal-organic frameworks (MOFs) are especially noteworthy. The Table shows advantages and drawbacks for some of them, with more details provided in the following sections.

Lipid-based nanoparticles

Liposome-mediated gene transfer was one of the first strategies for introducing foreign genetic material into target cells (Mintzer et al., 2009). Currently, composition of the liposomes used for this purpose widely varies and may include, e. g., cationic lipids, polyethylene glycol, cholesterol, phospholipids, dioleoylphosphatide acid, etc. (Kim et al., 2020; Patel et al., 2020). They readily penetrate target cells and ensure specific delivery, which significantly reduces effects in off-target tissues and organs *vs.* DNA vector-based delivery of CRISPR (Yeh et al., 2018). This supports the significance of studies of lipid carriers as the delivery systems for genome editing tools. For example, in a study (Andey et al., 2019), a lipoplex was synthesized incorporating small interfering RNA (siRNA) of the SOX2 transcription factor for target therapy of SOX2-enriched lung tumors in CB-17 nude mice. 85 % of animals administered with the lipoplex showed a reduction in their tumor weight and volume, which was associated with the reduction in SOX2 protein expression. A review (Lu Z.R. et al., 2021) provides a summary on the use of DODAP and other ionizable pH-sensitive lipids, which can also respond to other changes in the environment to produce the nanoparticles incorporating siRNA molecules for cancer therapy and targeted oncogene silencing. The use of lipid nanoparticles for the delivery of regulatory RNAs into the cells is also described in (Wang C. et al., 2021; Eygeris et al., 2022), as well as in an integrated review devoted to the use of lipids and their derivatives for the RNA delivery (Zhang Y. et al., 2021). These works suggest different approaches and advantages of lipid nanoparticles, including the relative ease of the targeted delivery, possibility of controlled release, protection from aggregation and elimination from the blood stream by the host immune system due to the use of PEG and other protective molecules, endosomal escape, etc.

Summary of advantages and drawbacks of non-viral systems for delivery of genome editing tools

Delivery system	Advantages	Drawbacks
Liposomes and lipid particles	Ease of preparation Efficient targeted delivery Possibility of adaptation for particular purposes Possibility of controlled release	Propensity for aggregation Issues with size control and stability
Extracellular vesicles	Biocompatibility Efficient and specific delivery Cargo protection	Relatively low stability No standardized production methods
Cell-penetrating peptides	Low toxicity Efficient transfection Possibility to control vector structure	Risk of immunization Low specificity
Dendrimers	Possibility to control vector properties Cargo protection Efficient transfection	Toxicity
Polymer nanoparticles	Ease of preparation Safety Possibility to control vector properties	Relatively low efficiency of delivery
Metal nanoparticles	Biocompatibility Broad spectrum of binding components Efficient delivery	Reported cases of toxicity and immunogenic activity
Metal-organic frameworks	Ease of preparation Biocompatibility Possibility to control pore parameters Cargo protection Possibility of controlled release	Relatively low efficacy and specificity of targeted delivery Reported cases of toxicity and immunogenic activity

Commercially available lipid nanovehicles include, e. g., Lipofectamine 2000, Lipofectamine 3000, RNAiMAX, which are used to deliver CRISPR/Cas9 components as a mix of Cas9 mRNA, gRNA and ribonucleoproteins into various cells (Yu X. et al., 2016). Lipid nanoparticles allow simultaneous encapsulation and delivery of several RNA types (mRNA and siRNA) into the target cells (Ball et al., 2018). In addition, they can be adapted to a particular way of administration, cell type and genome editing tool (Liu J. et al., 2019; Lokugamage et al., 2021).

However, the issues with controlling the size, uniformity and stability of the lipid nanoparticles restrict their use, particularly for *in vivo* gene therapies. Sometimes, these issues may be addressed by modification of the nanoparticle surface with PEG and other polymers, or by using nanoparticle cores made of a different material (for example, gold or polystyrene) and forming lipid layers with the incorporated cargoes over the core (Yan et al., 2022); however, this generally hampers preparation and use of such carriers and, in a number of cases, it would be practical to omit such systems in favor of other biotic and abiotic vectors.

Extracellular vesicles: exosomes and microvesicles

A number of scientists propose natural cell membrane-derived vesicles, including exosomes, microvesicles and apoptotic bodies, as the carriers to deliver genome editing tools *in vitro*

and *in vivo* while protecting them in the biological fluids and extracellular matrix.

Exosomes are extracellular vesicles produced by all cells. They were initially considered as drug carriers due to their small size, perfect biocompatibility, ability to transfer biomolecules into the cells and specific expression of the cell surface receptors. Further studies have shown that the exosomes carrying siRNAs can protect their cargo from enzymatic cleavage (half-life > 48 h), while naked siRNAs have half-lives of less than 6 h (Yang Z. et al., 2016). Moreover, encapsulation of siRNA into the exosomes improved its absorption by the cells.

Kamerkar S. et al. have constructed exosomes carrying a siRNA that targeted proto-oncogenic KRAS GTPase. These exosomes inhibited tumor development in various mouse pancreatic cancer models and significantly increased overall survival (Kamerkar et al., 2017).

To reduce immunogenic potential of exosomes carrying siRNAs and proteins in mice with Alzheimer’s disease, mouse dendritic cell-derived vesicles were used (Alvarez-Erviti et al., 2011). In this case, the proteins characteristic for target cells were fused to Lamp2b, which is abundant in exosome membranes. Such modification resulted in efficient cell type-specific gene knockdown while minimizing host immune response.

Moreover, the studies have shown that modified exosomes can transfer guide RNA and Cas9 protein between

HuH7 line cells (Chen R. et al., 2019). This work describes intercellular delivery of CRISPR/Cas9 components ensuring cleavage of hepatitis B virus and papilloma virus DNA in the infected cells.

Although the *in vivo* transfer of genome editing tools using exosomes showed no apparent side effects, the loading and targeting efficacy of such delivery systems is understudied. Another limitation of the clinical use of exosomes is the lack of standardized methods for their isolation and analysis (Doyle, Wang, 2019). Therefore, detailed studies of mechanisms and consequences of the vesicle-mediated delivery are needed as the result of such delivery may strongly depend on the cargo and cells used.

Microvesicles, another type of extracellular vesicles, are also of interest as potential delivery means. In contrast to exosomes, which are derived from endosomes, microvesicles are formed directly from the plasma membrane. They are larger than exosomes, which allows increasing their actual payload (Kanada et al., 2015). The potential of epithelium-derived microvesicles as a delivery system for CRISPR/Cas9 and sorafenib was assessed in the hepatocellular carcinoma model (Samuel et al., 2020), which showed enhanced microvesicle homing towards the tumor cells and a synergy of the agents loaded. A number of works also describes the use of microvesicles for the delivery of siRNAs and miRNAs into the cells to regulate intracellular and tissue processes, such as fibrosis, tumor growth inhibition, and the like (Vader et al., 2017; Stolzenburg, Harris, 2018).

Cell-penetrating peptides

An efficient delivery system must perform in a variety of tissues, ensuring rapid cargo release, be functional with low payload doses, non-toxic and easy to use in clinical practice. These properties, among others, are common for cell-penetrating peptides (CPPs). These peptides can bind to different molecules, interact with membrane structures, penetrate cells and deliver their cargo into the cytoplasm or nucleus. There are a lot of such peptides that can bind the molecules of interest in a covalent or non-covalent manner and translocate into the cells by means of direct membrane crossing, endocytosis, or formation of a transport channel in the membrane. Due to a number of their advantages, CPPs are widely used in studies to transfer small RNAs/DNAs, plasmids, antibodies, and nanoparticles into the cells. Their beneficial properties include controllable low toxicity, high transfection efficacy, and structural flexibility (Lopez-Vidal et al., 2021).

In a study (Ramakrishna et al., 2014), CPP was conjugated with a modified Cas9 protein and gRNA to induce gene disruptions in the target site in embryonic stem cells, dermal fibroblasts, HEK293T, HeLa, and human embryonic carcinoma cells. This genome editing tool delivery system efficiently changed target gene expression with the reduction in off-target mutation rate *vs.* the plasmid-based transfection.

Lopez-Vidal E.M. et al. successfully used a conjugate of a short synthetic peptide with low arginine content and anti-sense oligonucleotides for the transfection of HeLa654 cells and cardiac tissue of transgenic mice *in vivo* (Lopez-Vidal et al., 2021).

The efficacy of CPPs as vectors for gene delivery was shown for their complexes with modified viruses, plasmid

DNA, small interfering RNAs, oligonucleotides, DNA origami platforms, full-length genes, etc. (Taylor, Zahid, 2020). Features limiting their use include their high molecular weight, risk of host immunization, and insufficient delivery specificity.

Dendrimers

Dendrimers are another example of abiotic vectors. They are generally characterized by advantageous safety, lack of immunogenic potential, high efficacy, reproducibility, controllable size, broad range of possible modifications, an ability to form stable complexes with different molecules and deliver several molecule types (e. g., drug and gene) at once (Abedi-Gaballu et al., 2018). They can also promote release from endosomes after cell penetration due to the proton sponge effect. Dendrimer molecules can associate with different moieties and ligands, including antibodies, signaling molecules, imaging probes, photosensitisers, etc. (Kim et al., 2020). A unique property of dendrimers is their chemical and physical stability inherent to their chemical structure (Kalomiraki et al., 2016).

Dendrimers are extensively branched synthetic macromolecules having a well-defined structure and composition. These molecules are produced by the repeated assembly of polymer layers over the core. There are many dendrimer types, including peptide, poly(L-lysine), polyamideamine (PAMAM), silicone, polyethyleneimine, and other dendrimers. PAMAM dendrimers are the most extensively studied ones as drug and gene delivery systems. They form stable complexes with DNAs, siRNAs, and miRNAs referred to as dendriplexes. These complexes show high transfection efficacy and ability to protect nucleic acids from damage (Fant et al., 2008). Their modifications make it possible to create the derivatives possessing reduced toxicity, increased gene delivery specificity and efficacy (Abedi-Gaballu et al., 2018; Liu C. et al., 2019).

In the recent decade, three common strategies for dendrimer modification are used: (1) surface modification with different moieties (Yang J. et al., 2015); (2) hybrid vector formation (Biswas et al., 2013); (3) creation of self-assembling supramolecular nanoparticles (Yadav et al., 2020).

The first strategy is exemplified by the studies by Nagasaki T. et al. summarized in a review (Nagasaki, Shinkai, 2007), which used a cationic polyazobenzene dendrimer modified with L-lysine (Lys-G2). The dendrimer complex with a plasmid DNA ensured increased transfection efficacy when administered to cytoplasm and UV-irradiated.

The second strategy implies conjugation of ligands, polymers, inorganic nanoparticles, etc. with the dendrimer complex surface (Lin et al., 2018), which improves the dendrimer carrier properties. Mbatha L.S. et al. (Mbatha et al., 2021) have developed hybrid carriers by means of derivatization of gold nanoparticles with folic acid and 5th generation polyamidoamines. Their cytotoxicity and transgene expression efficacy were assessed *in vitro* using a luciferase reporter gene. The hybrid vectors ensured an increased luciferase expression *vs.* PAMAM dendrimers with folic acid or unbound dendrimers.

An example of the third strategy is a study aimed at building supramolecular nanoparticles of variable size (30–450 nm) from three different units, PAMAM dendrimer with adamantane, branched polyethyleneimine conjugated

with cyclodextrin, and polyethylene glycol with adamantane (Lu S. et al., 2020). These nanoparticles were used as a vector for the anti-cancer RNA interference agents, which resulted in reduced vascularization and inhibition of the lung tumor xenograft growth in a mouse model.

A study (Zarebkohan et al., 2015) yielded a PAMAM-PEG-dendrimer coated by serine-arginine-leucine (SRL) tripeptide for the delivery of genes into C6 glioma line cells. The results showed that such nanoparticles efficiently transfected the brain tumor cells.

Thus, dendrimers are a promising means for the delivery of genetic materials and genome editing tools. However, one of their critical limitations is related to their toxicity, with 3th to 5th generation dendrimers being less toxic than higher generations (Shcharbin et al., 2013). Moreover, dendrimer cytotoxicity depends on their branch elasticity (Tang et al., 1996), hydrophobic properties, the number and nature of the surface and core modifications (Somani et al., 2018). A broad range of modifications changing these parameters allows selection of the most suitable ones in order to minimize adverse effects of dendrimer-based vectors.

Polymer nanoparticles

Polymer nanoparticles possess chemical variety and have great potential due to their flexible structural modifications. They are widely used to deliver nucleic acids and other substances into cells and tissues.

These carriers are built from various natural and synthetic polymers. Natural macromolecules have a number of advantages over synthetic ones, which are generally consigned to the lack of toxicity, relatively low cost and ease of preparation. They include celluloses, starches, gelatin, collagen, chitosan, agar, pectin, inulin, dextrin, etc. These biopolymers can be modified to create delivery systems addressing particular tasks (Yadav et al., 2020; Basinska et al., 2021). For example, chitosan is the natural polymer that is most commonly used for CRISPR/Cas9 delivery. Its main advantages are biocompatibility, biodegradability, and lack of cytotoxicity. Qiao J. et al. encapsulated red fluorescent protein and Cas9/ribonucleoprotein fused to a polyglutamate peptide tag together with donor DNA into the chitosan nanoparticles. The polymer carrier ensured simultaneous delivery of both the genome editing tool and the single-strand DNA matrix while showing highly efficient transfection of HeLa cells with no cytotoxicity (Qiao et al., 2019).

The list of synthetic polymers is also large enough. Among them, the most explored delivery means include polylactic and polyglycolic acids, their copolymers, polycaprolactam, polyhydroxybutyrate, etc. They possess good biocompatibility and biodegradability, which support their wide use in medicine, biotechnology, agriculture and other fields (Singh A.V., 2011; Zhang S. et al., 2021).

There are reports of the ongoing development of complex carriers comprising several polymers at once, which allows to overcome the drawbacks of particular components owing to the advantages of others. Thus, in (Luo et al., 2018), a block copolymer of polyethylene glycol, β -poly(lactic-glycolic) acid and cationic lipids was used to obtain specific nanoparticles for the delivery of Cas9 mRNA and CRISPR/Cas9 plasmids into the macrophages. The resulting carriers induced specific

Cas9 expression in the macrophages and monocytes both *in vitro* and *in vivo*.

Rui Y. et al. synthesized polymers from carboxylated branched poly(β -amino esters) by stepwise copolymerization. Their results showed that C5-capped polymer ensured maximum cargo release efficacy after absorption by the cells. Furthermore, it was used to produce the nanoparticles for CRISPR-Cas9 ribonucleoprotein encapsulation. The authors found that the delivery of genome editing tools led to 77 % and 47 % knockout of the target gene in HEK-293T and GL261 mouse glioma cells, respectively (Rui et al., 2019).

Therefore, polymeric nanoparticles are generally safe, easy to produce and customizable. Moreover, they undergo degradation in the host body and are suitable for all strategies of CRISPR-Cas9 delivery. However, the efficacy of delivery using the polymeric carriers is thought to be insufficient (Liu C. et al., 2019).

Gold nanoparticles

A number of studies propose gold nanoparticles as a vector base to address the issues with *in vitro* and *in vivo* delivery of genome editing tools. It was shown that small (<3 nm) gold nanoparticles are biocompatible but possess cytotoxicity and immunogenic potential (Shukla et al., 2005). Gold nanoparticles can be bound with various ligands, drug molecules, genome editing tools, which expands their applications.

Gold nanoparticles used to transfer ribonucleoproteins for genome editing into the brain cells showed no cytotoxicity or adverse effects on the neuron function (Lee et al., 2018). A paper (Glass et al., 2017) describes efficient elimination of a DNA mutation leading to Duchene muscular dystrophy in a mouse model using gold nanoparticles carrying CRISPR components, with minimum off-target effects. In another study (Jia et al., 2017), gold nanoparticles covalently conjugated with a siRNA successfully delivered their cargo into the macrophages, which resulted in the inhibition of inflammation and restoration of the heart function in a laboratory animal cardiomyopathy model.

In a study (Mout et al., 2017b), arginine-coated gold nanoparticles were conjugated with the synthetic constructs of ribonucleoproteins and Cas9 oligoglutamate-tagged protein. These complexes were incubated with HeLa, HEK-293T, and Raw 264.7 cell cultures. The delivery system ensured highly efficient (about 90 %) transfer of Cas9 and ribonucleoproteins into the cytoplasm and nucleus, with 23 to 30 % genome editing efficacy. Tao Y. et al. (Tao et al., 2021) have shown the suitability of surface-modified gold nanoparticles for real-time monitoring of the biological effects during genome editing.

The limitations of gold nanoparticles include a lack of knowledge on the correlation of their immunogenic potential and toxicity with appropriate physicochemical properties, such as size, shape, charge, and surface modifications (Dykman, Khlebtsov, 2017). The approaches to reduce toxicity of such carriers and improve the delivery efficacy include the use of the complex nanoparticles comprising polyethyleneimine, polyethylene glycol, and other components promoting reduction in the immunogenic properties of the particles and preventing their binding to off-target receptors (Li Y. et al., 2017).

Metal-organic frameworks

Current studies in the field of abiotic vectors include extensive development of the carriers derived from metal-organic frameworks (MOFs) as the non-viral vehicles to deliver nucleic acids into target cells. MOFs are a novel class of porous materials. Their crystal lattice is formed by coordinate bonds between the central alkaline-earth or transition metal ions (Ca, Mg, Zn, Ti, Zr, Mn, Pd, Cu, Cr, Cd, etc.) and organic ligands having chelating moieties (Cheetham et al., 1999; Valtchev et al., 2009; Farha et al., 2012; Paz et al., 2012; Furukawa et al., 2013; Yu Y. et al., 2013; Li H. et al., 2018; Corella-Ochoa et al., 2019). MOF synthesis produces high-ordered porous crystal structures with strictly defined pore parameters (Wang Z., Cohen, 2009). Moreover, the MOF technology allows controlling the porosity and pore size in accordance with the cargo properties.

In addition, particular MOFs (e. g., based on zinc, calcium, magnesium, titan, zirconium, iron ions and biocompatible organic ligands, including polycarbonates, imidazolates, amines, phosphates, etc.) are biodegradable and low-toxic (Horcajada et al., 2012; Lyu et al., 2021). Therefore, such MOFs are widely used in experimental medicine as controlled-release drug carriers (Su et al., 2015; Ranjbar et al., 2018; Chen G. et al., 2019; Osorio-Toribio et al., 2020). A nanosized zeolite-like framework based on imidazole and zinc salts (ZIF) is particularly useful for these purposes. It has low toxicity, broad controllability of pore parameters, buffer properties and endosomal escape ability (Alsaiani et al., 2018). In a number of studies, MOFs are used for encapsulation of biologically active compounds such as insulin (Chen Y. et al., 2018), heparin (Vinogradov et al., 2018), hemoglobin (Peng et al., 2019). Moreover, a research team (Liang et al., 2016) successfully encapsulated living cells into a MOF, which ensured their preservation and physical protection.

Encapsulation of genome editing tools in the pores of such materials prevents their degradation in the physiological conditions until they reach their targets (Peng et al., 2018). There are two mechanisms of encapsulation of genome editing tools into the MOFs. The first one is the encapsulation by direct absorption into the pores. For example, a paper (Teplensky et al., 2019) describes the encapsulation of an RNA molecule into the pores of NU-1000, a zirconium-based MOF. The second mechanism is the biomineralization, i. e. the building of a metal-organic framework over the material to be encapsulated (Li Y. et al., 2019).

In a study (Alsaiani et al., 2018), the encapsulation of CRISPR/Cas9 into ZIF-8 was described. The cargo weight reached 1.2 % of the total polymer weight, with 17 % pore loading efficacy, which the authors considered a good result in contrast to previously reported values for the MOF-based delivery systems. The polymer showed no cytotoxic properties in concentrations up to 200 mg/mL, was stable in physiological conditions but was rapidly destroyed at pH of 5–6, which creates the potential for controlled cargo release *in vivo*. This complex also had an enhanced endosomal escape ability over the cationic lipid-based vehicles and reduced target gene expression twofold when incubated for 2 days and threefold when incubated for 4 days, which was two times higher than the efficacy of the target gene knockdown with lipofectamine-mediated CRISPR/Cas9 delivery.

Specific delivery to the target cells is critical for improving the genome editing efficacy and safety. Alyami M.Z. et al. proposed a coating for ZIF-8 with encapsulated CRISPR/Cas9, which was based on MCF-7 human breast adenocarcinoma cell membrane. Incubation of such modified MOF with MCF-7, HeLa, HDFn, and aTC cells showed that MCF-7 possessed the maximum carrier absorption efficacy while the other cell lines absorbed the agent to a small extent. Moreover, such a composite, when transfected to the MCF-7 cells, inhibited EGFP expression threefold vs. the HeLa membrane-coated ZIF (Alyami et al., 2020).

Currently, there is an ongoing discussion of particular chemistries useful for the controlled delivery of Cas9/gRNA into the cells using MOFs in the presence of endogenic or external signals (Yang X. et al., 2019; Lyu et al., 2021). Thus, the carrier systems that penetrate the cells by endocytosis come to the organelles with an acid content, such as endosomes or lysosomes. Considering intracellular pH levels, the pH-sensitive hybrid carriers were created from silicon dioxide and ZIF (SMOFs) for efficient encapsulation and delivery of hydrophilic compounds (Wang Y. et al., 2020). SMOF nanoparticles with encapsulated ribonucleoproteins ensured efficient genome editing *in vivo* in the mouse retinal pigment epithelium after subretinal injection.

In addition, abnormal cells and tissues often have a unique microenvironment with specific levels of pH and other active substances such as enzymes and ATP which could be used for MOF-mediated targeted delivery. Yang X. et al., relying on the activation of ATP production in some disorders, created an ATP-sensitive zeolite-like framework based on imidazole and zinc ions (ZIF-90). This material efficiently encapsulated CRISPR/Cas9 and ensured delivery of a large amount of protein payload into the cell matrix, regardless of the particle size and molecular weight. In the presence of ATP, ZIF-90/protein conjugates were destroyed, releasing the protein due to competitive coordination between ATP and Zn²⁺ in ZIF-90. After transfection, target gene expression in HeLa cells was inhibited by up to 35 % (Yang X. et al., 2019).

The study by Chen T.T. et al. also showed that ZIF-8 nanoparticles were able to release encapsulated proteins rapidly in acid media but not at pH 7.4 (Chen T.T. et al., 2018), which may be preferred in some disorders.

Despite certain advances and potential of MOFs as vectors for genome editing tools, there are also issues yet to be addressed. Particularly, there is a need for the following studies: (1) to improve the specificity and efficacy of targeted effects of MOF nanoparticles; (2) to increase MOF/biomolecule conjugate stability in the bloodstream with intravenous administration; (3) to find ways for reducing the immunogenic potential and toxicity of MOFs; (4) to estimate the long-term safety of the carriers; (5) to finalize the large-scale production of the carriers with defined parameters (Lyu et al., 2021; Zheng et al., 2021).

Conclusion

Currently, there is a variety of methods and systems for the delivery of genome editing tools. They have both unique advantages and drawbacks. At the same time, it is worth acknowledging that a single universal carrier for delivery of

all types of the agents cannot be developed. It seems clear that the choice and use of certain viral or non-viral vectors must primarily be defined by the specific aspects of the problem to be solved. In particular, synthetic carriers are preferred for the simultaneous loading of several substances and components, which is especially relevant for genome editing. Therefore, there is an increasing number of proposals to combine different non-viral delivery systems. For example, in a number of cases, it makes sense to deliver the genes and therapies using cell-penetrating peptides combined with nanoparticles, micelles, liposomes, or polymers. In this context, MOF-based carriers, which allow the implementation of a broad spectrum of capabilities, have great potential. Further development of such vectors and technologies can result in safe and efficient delivery systems that would be able to circulate in the body for a long time while recognizing target cells and ensuring cell-specific delivery and release of intact cargoes.

References

- Abedi-Gaballu F., Dehghan G., Ghaffari M., Yekta R., Abbaspour-Ravasjani S., Baradaran B., Dolatabadi J.E.N., Hamblin M.R. PAMAM dendrimers as efficient drug and gene delivery nano-systems for cancer therapy. *Appl. Mater. Today*. 2018;12:177-190. DOI 10.1016/j.apmt.2018.05.002
- Alsaiani S.K., Patil S., Alyami M., Alamoudi K.O., Aleisa F.A., Merzaban J.S., Li M., Khashab N.M. Endosomal escape and delivery of CRISPR/Cas9 genome editing machinery enabled by nanoscale zeolitic imidazolate framework. *J. Am. Chem. Soc.* 2018;140(1):143-146. DOI 10.1021/jacs.7b11754
- Alvarez-Erviti L., Seow Y., Yin H., Betts C., Likhachev S., Wood M.J. Delivery of siRNA to the mouse brain by systemic injection of targeted exosomes. *Nat. Biotechnol.* 2011;29(4):341-345. DOI 10.1038/nbt.1807
- Alyami M.Z., Alsaiani S.K., Li Y., Qutub S.S., Aleisa F.A., Sougra R., Merzaban J.S., Khashab N.M. Cell-type-specific CRISPR/Cas9 delivery by biomimetic metal organic frameworks. *J. Am. Chem. Soc.* 2020;142(4):1715-1720. DOI 10.1021/jacs.9b11638
- Andey T., Bora-Singhal N., Chellappan S.P., Singh M. Cationic lipoplexes for treatment of cancer stem cell-derived murine lung tumors. *Nanomedicine*. 2019;18:31-43. DOI 10.1016/j.nano.2019.02.007
- Ball R.L., Hajj K.A., Vizelman J., Bajaj P., Whitehead K.A. Lipid nanoparticle formulations for enhanced co-delivery of siRNA and mRNA. *Nano Lett.* 2018;18(6):3814-3822. DOI 10.1021/acs.nanolett.8b01101
- Basinska T., Gadzinowski M., Mickiewicz D., Slomkowski S. Functionalized particles designed for targeted delivery. *Polymers (Basel)*. 2021;13(12):2022. DOI 10.3390/polym13122022
- Biswas S., Deshpande P.P., Navarro G., Dodwadkar N.S., Torchilin V.P. Lipid modified triblock PAMAM-based nanocarriers for siRNA drug co-delivery. *Biomaterials*. 2013;34(4):1289-1301. DOI 10.1016/j.biomaterials.2012.10.024
- Cheetham A.K., Ferey G., Loiseau T. Open-framework inorganic materials. *Angew. Chem. Int. Ed. Engl.* 1999;38(22):3268-3292. DOI 10.1002/(SICI)1521-3773(19991115)38:22<3268::AID-ANIE3268>3.0.CO;2-U
- Chen G., Luo J., Cai M., Qin L., Wang Y., Gao L., Huang P., Yu Y., Ding Y., Dong X., Yin X., Ni J. Investigation of metal-organic framework-5 (MOF-5) as an antitumor drug oridonin sustained release carrier. *Molecules*. 2019;24(18):3369. DOI 10.3390/molecules24183369
- Chen R., Huang H., Liu H., Xi J., Ning J., Zeng W., Shen C., Zhang T., Yu G., Xu Q., Chen X., Wang J., Lu F. Friend or foe? Evidence indicates endogenous exosomes can deliver functional gRNA and Cas9 protein. *Small*. 2019;15(38):e1902686. DOI 10.1002/smll.201902686
- Chen T.T., Yi J.T., Zhao Y.Y., Chu X. Biomimetic metal-organic framework nanoparticles enable intracellular delivery and endo-lysosomal release of native active proteins. *J. Am. Chem. Soc.* 2018;140(31):9912-9920. DOI 10.1021/jacs.8b04457
- Chen Y., Li P., Modica J.A., Drou R.J., Farha O.K. Acid-resistant mesoporous metal-organic framework toward oral insulin delivery: protein encapsulation, protection, and release. *J. Am. Chem. Soc.* 2018;140(17):5678-5681. DOI 10.1021/jacs.8b02089
- Corella-Ochoa M.N., Tapia J.B., Rubín H.N., Lillo V., Gonzalez-Cobos J., Nunez-Rico J.L., Balestra S.R.G., Almora-Barrios N., Lledos M., Guell-Bara A., Cabezas-Gimenez J., Escudero-Adan E.C., Vidal-Ferran A., Calero S., Reynolds M., Marti-Gastaldo C., Galan-Mascaros J.R. Homochiral metal-organic frameworks for enantioselective separations in liquid chromatography. *J. Am. Chem. Soc.* 2019;141(36):14306-14316. DOI 10.1021/jacs.9b06500
- Doyle L.M., Wang M.Z. Overview of extracellular vesicles, their origin, composition, purpose, and methods for exosome isolation and analysis. *Cells*. 2019;8(7):727. DOI 10.3390/cells8070727
- Dykman L.A., Khlebtsov N.G. Immunological properties of gold nanoparticles. *Chem. Sci.* 2017;8(3):1719-1735. DOI 10.1039/c6sc03631g
- Eygeris Y., Gupta M., Kim J., Sahay G. Chemistry of lipid nanoparticles for RNA delivery. *Acc. Chem. Res.* 2022;55(1):2-12. DOI 10.1021/acs.accounts.1c00544
- Fant K., Esbjörner E.K., Lincoln P., Nordén B. DNA condensation by PAMAM dendrimers: self-assembly characteristics and effect on transcription. *Biochemistry*. 2008;47(6):1732-1740. DOI 10.1021/bi7017199
- Farha O.K., Eryazici I., Jeong N.C., Hauser B.G., Wilmer C.E., Sarjeant A.A., Snurr R.Q., Nguyen S.T., Yazaydin A.O., Hupp J.T. Metal-organic framework materials with ultrahigh surface areas: is the sky the limit? *J. Am. Chem. Soc.* 2012;134(36):15016-15021. DOI 10.1021/ja3055639
- Furukawa H., Cordova K.E., O'Keeffe M., Yaghi O.M. The chemistry and applications of metal-organic frameworks. *Science*. 2013;341(6149):1230444. DOI 10.1126/science.1230444
- Glass Z., Li Y., Xu Q. Nanoparticles for CRISPR-Cas9 delivery. *Nat. Biomed. Eng.* 2017;1(11):854-855. DOI 10.1038/s41551-017-0158-x
- Hanlon K.S., Kleinstiver B.P., Garcia S.P., Zaborowski M.P., Volak A., Spirig S.E., Muller A., Sousa A.A., Tsai S.Q., Bengtsson N.E., Loov C., Ingelsson M., Chamberlain J.S., Corey D.P., Aryee M.J., Joung J.K., Breakefield X.O., Maguire C.A., Gyorgy B. High levels of AAV vector integration into CRISPR-induced DNA breaks. *Nat. Commun.* 2019;10(1):4439. DOI 10.1038/s41467-019-12449-2
- Horcjada P., Gref R., Baati T., Allan P.K., Maurin G., Couvreur P., Ferey G., Morris R.E., Serre C. Metal-organic frameworks in biomedicine. *Chem. Rev.* 2012;112(2):1232-1268. DOI 10.1021/cr200256v
- Huang Y., Liu X., Dong L., Liu Z., He X., Liu W. Development of viral vectors for gene therapy for chronic pain. *Pain Res. Treat.* 2011;2011:968218. DOI 10.1155/2011/968218
- Jia C., Chen H., Wei M., Chen X., Zhang Y., Cao L., Yuan P., Wang F., Yang G., Ma J. Gold nanoparticle-based miR155 antagonist macrophage delivery restores the cardiac function in ovariectomized diabetic mouse model. *Int. J. Nanomedicine*. 2017;12:4963-4979. DOI 10.2147/IJN.S138400

- Kalomiraki M., Thermos K., Chaniotakis N.A. Dendrimers as tunable vectors of drug delivery systems and biomedical and ocular applications. *Int. J. Nanomedicine*. 2016;11:1-12. DOI 10.2147/IJN.S93069
- Kamerkar S., LeBleu V.S., Sugimoto H., Yang S., Ruivo C.F., Melo S.A., Lee J.J., Kalluri R. Exosomes facilitate therapeutic targeting of oncogenic KRAS in pancreatic cancer. *Nature*. 2017;546(7659):498-503. DOI 10.1038/nature22341
- Kanada M., Bachmann M.H., Hardy J.W., Frimansson D.O., Bronsart L., Wang A., Sylvester M.D., Schmidt T.L., Kaspar R.L., Butte M.J., Matin A.C., Contag C.H. Differential fates of biomolecules delivered to target cells via extracellular vesicles. *Proc. Natl. Acad. Sci. USA*. 2015;112(12):E1433-1442. DOI 10.1073/pnas.1418401112
- Kim D., Le Q.V., Wu Y., Park J., Oh Y.K. Nanovesicle-mediated delivery systems for CRISPR/Cas genome editing. *Pharmaceutics*. 2020;12(12):1233. DOI 10.3390/pharmaceutics12121233
- Lee B., Lee K., Panda S., Gonzales-Rojas R., Chong A., Bugay V., Park H.M., Brenner R., Murthy N., Lee H.Y. Nanoparticle delivery of CRISPR into the brain rescues a mouse model of fragile X syndrome from exaggerated repetitive behaviours. *Nat. Biomed. Eng.* 2018;2(7):497-507. DOI 10.1038/s41551-018-0252-8
- Li H., Wang K., Sun Y., Lollar C.T., Li J., Zhou H.-C. Recent advances in gas storage and separation using metal-organic frameworks. *Materials Today*. 2018;21(2):108-121. DOI 10.1016/j.mattod.2017.07.006
- Li Y., Chen Y., Li J., Zhang Z., Huang C., Lian G., Yang K., Chen S., Lin Y., Wang L., Huang K., Zeng L. Co-delivery of microRNA-21 antisense oligonucleotides and gemcitabine using nanomedicine for pancreatic cancer therapy. *Cancer Sci.* 2017;108(7):1493-1503. DOI 10.1111/cas.13267
- Li Y., Zhang K., Liu P., Chen M., Zhong Y., Ye Q., Wei M.Q., Zhao H., Tang Z. Encapsulation of plasmid DNA by nanoscale metal-organic frameworks for efficient gene transportation and expression. *Adv. Mater.* 2019;31(29):e1901570. DOI 10.1002/adma.201901570
- Liang K., Richardson J.J., Cui J., Caruso F., Doonan C.J., Falcaro P. Metal-organic framework coatings as cytoprotective exoskeletons for living cells. *Adv. Mater.* 2016;28(36):7910-7914. DOI 10.1002/adma.201602335
- Lin L., Fan Y., Gao F., Jin L., Li D., Sun W., Li F., Qin P., Shi Q., Shi X., Du L. UTMD-promoted co-delivery of gemcitabine and miR-21 inhibitor by dendrimer-entrapped gold nanoparticles for pancreatic cancer therapy. *Theranostics*. 2018;8(7):1923-1939. DOI 10.7150/thno.22834
- Liu C., Wan T., Wang H., Zhang S., Ping Y., Cheng Y. A boronic acid-rich dendrimer with robust and unprecedented efficiency for cytosolic protein delivery and CRISPR-Cas9 gene editing. *Sci. Adv.* 2019;5(6):eaaw8922. DOI 10.1126/sciadv.aaw8922
- Liu J., Chang J., Jiang Y., Meng X., Sun T., Mao L., Xu Q., Wang M. Fast and efficient CRISPR/Cas9 genome editing *in vivo* enabled by bioreducible lipid and messenger RNA nanoparticles. *Adv. Mater.* 2019;31(33):e1902575. DOI 10.1002/adma.201902575
- Lokugamage M.P., Vanover D., Beyersdorf J., Hatit M.Z.C., Roto-lo L., Echeverri E.S., Peck H.E., Ni H., Yoon J.K., Kim Y., Santangelo P.J., Dahlman J.E. Optimization of lipid nanoparticles for the delivery of nebulized therapeutic mRNA to the lungs. *Nat. Biomed. Eng.* 2021;5(9):1059-1068. DOI 10.1038/s41551-021-00786-x
- Lopez-Vidal E.M., Schisse C.K., Mohapatra S., Bellovoda K., Wu C.L., Woo J.A., Malmberg A.B., Loas A., Gomez-Bombarelli R., Pentelute B.L. Deep learning enables discovery of a short nuclear targeting peptide for efficient delivery of antisense oligomers. *JACS Au*. 2021;1(11):2009-2020. DOI 10.1021/jacsau.1c00327
- Lu S., Bao X., Hai W., Shi S., Chen Y., Yu Q., Zhang M., Xu Y., Peng J. Multi-functional self-assembled nanoparticles for pVEGF-shRNA loading and anti-tumor targeted therapy. *Int. J. Pharm.* 2020;575:118898. DOI 10.1016/j.ijpharm.2019.118898
- Lu Z.R., Laney V.E.A., Hall R., Ayat N. Environment-responsive lipid/siRNA nanoparticles for cancer therapy. *Adv. Healthc. Mater.* 2021;10(5):e2001294. DOI 10.1002/adhm.202001294
- Luo Y.L., Xu C.F., Li H.J., Cao Z.T., Liu J., Wang J.L., Du X.J., Yang X.Z., Gu Z., Wang J. Macrophage-specific *in vivo* gene editing using cationic lipid-assisted polymeric nanoparticles. *ACS Nano*. 2018;12(2):994-1005. DOI 10.1021/acsnano.7b07874
- Lyu Y., Yang C., Lyu X., Pu K. Active delivery of CRISPR system using targetable or controllable nanocarriers. *Small*. 2021;17(24):e2005222. DOI 10.1002/sml.202005222
- Mbatha L.S., Maiyo F., Daniels A., Singh M. Dendrimer-coated gold nanoparticles for efficient folate-targeted mRNA delivery *in vitro*. *Pharmaceutics*. 2021;13(6):900. DOI 10.3390/pharmaceutics13060900
- Mintzer M.A., Simanek E.E. Nonviral vectors for gene delivery. *Chem. Rev.* 2009;109(2):259-302. DOI 10.1021/cr800409e
- Moscoco C.G., Steer C.J. The evolution of gene therapy in the treatment of metabolic liver diseases. *Genes (Basel)*. 2020;11(8):915. DOI 10.3390/genes11080915
- Mout R., Ray M., Lee Y.W., Scaletti F., Rotello V.M. *In vivo* delivery of CRISPR/Cas9 for therapeutic gene editing: progress and challenges. *Bioconjug. Chem.* 2017a;28(4):880-884. DOI 10.1021/acs.bioconjchem.7b00057
- Mout R., Ray M., Yesilbag Tonga G., Lee Y.W., Tay T., Sasaki K., Rotello V.M. Direct cytosolic delivery of CRISPR/Cas9-ribonucleoprotein for efficient gene editing. *ACS Nano*. 2017b;11(3):2452-2458. DOI 10.1021/acsnano.6b07600
- Nagasaki T., Shinkai S. The concept of molecular machinery is useful for design of stimuli-responsive gene delivery systems in the mammalian cell. *J. Incl. Phenom. Macrocycl. Chem.* 2007;58(3-4):205-219. DOI 10.1007/s10847-007-9303-6
- Niggemann P., Gyorgy B., Chen Z.Y. Genome and base editing for genetic hearing loss. *Hear. Res.* 2020;394:107958. DOI 10.1016/j.heares.2020.107958
- Osorio-Toribio G., Velasquez-Hernandez M.J., Mileo P.G.M., Zarate J.A., Aguila-Rosas J., Leyva-Gomez G., Sanchez-Sanchez R., Magana J.J., Perez-Diaz M.A., Lazaro I.A., Forgan R.S., Maurin G., Lima E., Ibarra I.A. Controlled transdermal release of antioxidant ferulate by a porous Sc(III) MOF. *iScience*. 2020;23(6):101156. DOI 10.1016/j.isci.2020.101156
- Patel S., Ashwanikumar N., Robinson E., Xia Y., Mihai C., Griffith J.P., Hou S., Esposito A.A., Ketova T., Welsher K., Joyal J.L., Almarsson Ö., Sahay G. Naturally-occurring cholesterol analogues in lipid nanoparticles induce polymorphic shape and enhance intracellular delivery of mRNA. *Nat. Commun.* 2020;11(1):983. DOI 10.1038/s41467-020-14527-2
- Paz F.A., Klinowski J., Vilela S.M., Tomé J.P., Cavaleiro J.A., Rocha J. Ligand design for functional metal-organic frameworks. *Chem. Soc. Rev.* 2012;41(3):1088-1110. DOI 10.1039/C1CS15055C
- Peng S., Bie B., Sun Y., Liu M., Cong H., Zhou W., Xia Y., Tang H., Deng H., Zhou X. Metal-organic frameworks for precise inclusion of single-stranded DNA and transfection in immune cells. *Nat. Commun.* 2018;9(1):1293. DOI 10.1038/s41467-018-03650-w

- Peng S., Liu J., Qin Y., Wang H., Cao B., Lu L., Yu X. Metal-organic framework encapsulating hemoglobin as a high-stable and long-circulating oxygen carriers to treat hemorrhagic shock. *ACS Appl. Mater. Interfaces*. 2019;11(39):35604-35612. DOI 10.1021/acsami.9b15037
- Qiao J., Sun W., Lin S., Jin R., Ma L., Liu Y. Cytosolic delivery of CRISPR/Cas9 ribonucleoproteins for genome editing using chitosan-coated red fluorescent protein. *Chem. Commun. (Camb)*. 2019;55(32):4707-4710. DOI 10.1039/c9cc00010k
- Ramakrishna S., Kwaku Dad A.B., Beloor J., Gopalappa R., Lee S.K., Kim H. Gene disruption by cell-penetrating peptide-mediated delivery of Cas9 protein and guide RNA. *Genome Res*. 2014;24(6):1020-1027. DOI 10.1101/gr.171264.113
- Ranjbar M., Pardakhty A., Amanatfard A., Asadipour A. Efficient drug delivery of beta-estradiol encapsulated in Zn-metal-organic framework nanostructures by microwave-assisted coprecipitation method. *Drug Des. Devel. Ther*. 2018;12:2635-2643. DOI 10.2147/DDDT.S173324
- Rui Y., Wilson D.R., Choi J., Varanasi M., Sanders K., Karlsson J., Lim M., Green J.J. Carboxylated branched poly(beta-amino ester) nanoparticles enable robust cytosolic protein delivery and CRISPR-Cas9 gene editing. *Sci. Adv*. 2019;5(12):eaay3255. DOI 10.1126/sciadv.aay3255
- Samuel M.S., Suman S., Venkateshkannan, Selvarajan E., Mathmani T., Pugazhendhi A. Immobilization of Cu₃(btc)₂ on graphene oxide-chitosan hybrid composite for the adsorption and photocatalytic degradation of methylene blue. *J. Photochem. Photobiol. B*. 2020;204:111809. DOI 10.1016/j.jphotobiol.2020.111809
- Shcharbin D., Shakhbazov A., Bryszewska M. Poly(amidoamine) dendrimer complexes as a platform for gene delivery. *Expert Opin. Drug Deliv*. 2013;10(12):1687-1698. DOI 10.1517/17425247.2013.853661
- Shukla R., Bansal V., Chaudhary M., Basu A., Bhonde R.R., Sasstry M. Biocompatibility of gold nanoparticles and their endocytotic fate inside the cellular compartment: a microscopic overview. *Langmuir*. 2005;21(23):10644-10654. DOI 10.1021/la0513712
- Singh A.V. Biopolymers in drug delivery: a review. *Pharmacology-online*. 2011;1:666-674
- Singh D., Sternberg S.H., Fei J., Doudna J.A., Ha T. Real-time observation of DNA recognition and rejection by the RNA-guided endonuclease Cas9. *Nat. Commun*. 2016;7:12778. DOI 10.1038/ncomms12778
- Somani S., Laskar P., Altwaijry N., Kewcharoenvong P., Irving C., Robb G., Pickard B.S., Dufès C. PEGylation of polypropyleneimine dendrimers: effects on cytotoxicity, DNA condensation, gene delivery and expression in cancer cells. *Sci. Rep*. 2018; 8(1):9410. DOI 10.1038/s41598-018-27400-6
- Stolzenburg L.R., Harris A. Microvesicle-mediated delivery of miR-1343: impact on markers of fibrosis. *Cell Tissue Res*. 2018; 371(2):325-338. DOI 10.1007/s00441-017-2697-6
- Su H., Sun F., Jia J., He H., Wang A., Zhu G. A highly porous medical metal-organic framework constructed from bioactive curcumin. *Chem. Commun*. 2015;51(26):5774-5777. DOI 10.1039/c4cc10159f
- Tang M.X., Redemann C.T., Szoka F.C., Jr. *In vitro* gene delivery by degraded polyamidoamine dendrimers. *Bioconjug. Chem*. 1996;7(6):703-714. DOI 10.1021/bc9600630
- Tao Y., Yi K., Hu H., Shao D., Li M. Coassembly of nucleus-targeting gold nanoclusters with CRISPR/Cas9 for simultaneous bioimaging and therapeutic genome editing. *J. Mater. Chem. B*. 2021;9(1):94-100. DOI 10.1039/d0tb01925a
- Taylor R.E., Zahid M. Cell penetrating peptides, novel vectors for gene therapy. *Pharmaceutics*. 2020;12(3):225. DOI 10.3390/pharmaceutics12030225
- Teplensky M.H., Fantham M., Poudel C., Hockings C., Lu M., Guna A., Aragones-Anglada M., Moghadam P.Z., Li P., Farha O.K., Fernández S.B.Q., Richards F.M., Jodrell D.I., Kaminski Schierle G., Kaminski C.F., Fairen-Jimenez D. A highly porous metal-organic framework system to deliver payloads for gene knockdown. *Chem*. 2019;5(11):2926-2941. DOI 10.1016/j.chempr.2019.08.015
- Vader P., Mager I., Lee Y., Nordin J.Z., Andaloussi S.E., Wood M.J. Preparation and isolation of siRNA-loaded extracellular vesicles. *Methods Mol. Biol*. 2017;1545:197-204. DOI 10.1007/978-1-4939-6728-5_14
- Valtchev V., Mintova S., Tsapatsis M. (Eds.). Ordered Porous Solids. Recent Advances and Prospects. Oxford, Amsterdam: Elsevier, 2009. DOI 10.1016/B978-0-444-53189-6.X0001-7
- Vinogradov V.V., Drozdov A.S., Mingabudinova L.R., Shabanova E.M., Kolchina N.O., Anastasova E.I., Markova A.A., Shtil A.A., Milichko V.A., Starova G.L., Precker R.L.M., Vinogradov A.V., Hey-Hawkins E., Pidko E.A. Composites based on heparin and MIL-101(Fe): the drug releasing depot for anticoagulant therapy and advanced medical nanofabrication. *J. Mater. Chem. B*. 2018;6(16):2450-2459. DOI 10.1039/c8tb00072g
- Wang C., Zhang Y., Dong Y. Lipid nanoparticle-mRNA formulations for therapeutic applications. *Acc. Chem. Res*. 2021;54(23): 4283-4293. DOI 10.1021/acs.accounts.1c00550
- Wang H.X., Li M., Lee C.M., Chakraborty S., Kim H.W., Bao G., Leong K.W. CRISPR/Cas9-based genome editing for disease modeling and therapy: challenges and opportunities for nonviral delivery. *Chem. Rev*. 2017;117(15):9874-9906. DOI 10.1021/acs.chemrev.6b00799
- Wang Y., Shahi P.K., Xie R., Zhang H., Abdeen A.A., Yodsanit N., Ma Z., Saha K., Pattnaik B.R., Gong S. A pH-responsive silica-metal-organic framework hybrid nanoparticle for the delivery of hydrophilic drugs, nucleic acids, and CRISPR-Cas9 genome-editing machineries. *J. Control. Release*. 2020;324:194-203. DOI 10.1016/j.jconrel.2020.04.052
- Wang Z., Cohen S.M. Postsynthetic modification of metal-organic frameworks. *Chem. Soc. Rev*. 2009;38(5):1315-1329. DOI 10.1039/b802258p
- Yadav S., Sharma A.K., Kumar P. Nanoscale self-assembly for therapeutic delivery. *Front. Bioeng. Biotechnol*. 2020;8:127. DOI 10.3389/fbioe.2020.00127
- Yan Y., Liu X.Y., Lu A., Wang X.Y., Jiang L.X., Wang J.C. Non-viral vectors for RNA delivery. *J. Control. Release*. 2022;342: 241-279. DOI 10.1016/j.jconrel.2022.01.008
- Yang J., Zhang Q., Chang H., Cheng Y. Surface-engineered dendrimers in gene delivery. *Chem. Rev*. 2015;115(11):5274-5300. DOI 10.1021/cr500542t
- Yang X., Tang Q., Jiang Y., Zhang M., Wang M., Mao L. Nanoscale ATP-responsive zeolitic imidazole Framework-90 as a general platform for cytosolic protein delivery and genome editing. *J. Am. Chem. Soc*. 2019;141(9):3782-3786. DOI 10.1021/jacs.8b11996
- Yang Z., Xie J., Zhu J., Kang C., Chiang C., Wang X., Wang X., Kuang T., Chen F., Chen Z., Zhang A., Yu B., Lee R.J., Teng L., Lee L.J. Functional exosome-mimic for delivery of siRNA to cancer: *in vitro* and *in vivo* evaluation. *J. Control. Release*. 2016; 243:160-171. DOI 10.1016/j.jconrel.2016.10.008
- Yeh W.H., Chiang H., Rees H.A., Edge A.S.B., Liu D.R. *In vivo* base editing of post-mitotic sensory cells. *Nat. Commun*. 2018;9(1): 2184. DOI 10.1038/s41467-018-04580-3

- Yu X., Liang X., Xie H., Kumar S., Ravinder N., Potter J., de Mollerat du Jeu X., Chesnut J.D. Improved delivery of Cas9 protein/gRNA complexes using lipofectamine CRISPRMAX. *Bio-technol. Lett.* 2016;38(6):919-929. DOI 10.1007/s10529-016-2064-9
- Yu Y., Ren Y., Shen W., Deng H., Gao Z. Applications of metal-organic frameworks as stationary phases in chromatography. *Trends Anal. Chem.* 2013;50:33-41. DOI 10.1016/j.trac.2013.04.014
- Zarebkohan A., Najafi F., Moghimi H.R., Hemmati M., Deevband M.R., Kazemi B. Synthesis and characterization of a PAMAM dendrimer nanocarrier functionalized by SRL peptide for targeted gene delivery to the brain. *Eur. J. Pharm. Sci.* 2015; 78:19-30. DOI 10.1016/j.ejps.2015.06.024
- Zhang S., Shen J., Li D., Cheng Y. Strategies in the delivery of Cas9 ribonucleoprotein for CRISPR/Cas9 genome editing. *Theranostics.* 2021;11(2):614-648. DOI 10.7150/thno.47007
- Zhang Y., Sun C., Wang C., Jankovic K.E., Dong Y. Lipids and lipid derivatives for RNA delivery. *Chem. Rev.* 2021;121(20): 12181-12277. DOI 10.1021/acs.chemrev.1c00244
- Zheng Q., Li W., Mao L., Wang M. Nanoscale metal-organic frameworks for the intracellular delivery of CRISPR/Cas9 genome editing machinery. *Biomater. Sci.* 2021;9(21):7024-7033. DOI 10.1039/d1bm00790d

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

Received September 26, 2022. Revised July 21, 2023. Accepted September 12, 2023.

DOI 10.18699/vjgb-24-29

Problems of creating antibody phage libraries and their solutions

V.S. Aripov , N.V. Volkova , A.A. Ilyichev , D.N. Shcherbakov 

State Research Center of Virology and Biotechnology "Vector", Koltsovo, Novosibirsk Region, Russia

 aripov_vs@vector.nsc.ru

Abstract. Phage display has become an efficient, reliable and popular molecular technique for generating libraries encompassing millions or even billions of clones of divergent peptides or proteins. The method is based on the correspondence between phage genotype and phenotype, which ensures the presentation of recombinant proteins of known amino acid composition on the surface of phage particles. The use of affinity selection allows one to choose variants with affinity for different targets from phage libraries. The implementation of the antibody phage display technique has revolutionized the field of clinical immunology, both for developing tools to diagnose infectious diseases and for producing therapeutic agents. It has also become the basis for efficient and relatively inexpensive methods for studying protein-protein interactions, receptor binding sites, as well as epitope and mimotope identification. The antibody phage display technique involves a number of steps, and the final result depends on their successful implementation. The diversity, whether natural or obtained by combinatorial chemistry, is the basis of any library. The choice of molecular techniques is critical to ensure that this diversity is maintained during the phage library preparation step and during the transformation of *E. coli* cells. After a helper phage is added to the suspension of transformed *E. coli* cells, a bacteriophage library is formed, which is a working tool for performing the affinity selection procedure and searching for individual molecules. Despite the apparent simplicity of generating phage antibody libraries, a number of subtleties need to be taken into account. First, there are the features of phage vector preparation. Currently, a large number of phagemid vectors have been developed, and their selection is also of great importance. The key step is preparing competent *E. coli* cells and the technology of their transformation. The choice of a helper phage and the method used to generate it is also important. This article discusses the key challenges faced by researchers in constructing phage antibody libraries.

Key words: phage library; bacteriophage replicative form; monoclonal antibodies; helper phage; competent cells.

For citation: Aripov V.S., Volkova N.V., Ilyichev A.A., Shcherbakov D.N. Problems of creating antibody phage libraries and their solutions. *Vavilovskii Zhurnal Genetiki i Seleksii = Vavilov Journal of Genetics and Breeding*. 2024;28(2):249-257. DOI 10.18699/vjgb-24-29

Проблемы создания фаговых библиотек антител и пути их решения

В.С. Арипов , Н.В. Волкова , А.А. Ильичев , Д.Н. Щербаков 

Государственный научный центр вирусологии и биотехнологии «Вектор» Роспотребнадзора, р. п. Кольцово, Новосибирская область, Россия

 aripov_vs@vector.nsc.ru

Аннотация. Фаговый дисплей стал эффективной, надежной и востребованной молекулярной техникой для создания библиотек, содержащих миллионы или даже миллиарды клонов, экспонирующих различающиеся пептиды или белки. В основе этого метода лежит соответствие между генотипом и фенотипом фага, обеспечивающее презентацию на поверхности фаговых частиц рекомбинантных белков с известным аминокислотным составом. Использование процедуры аффинной селекции позволяет проводить отбор из фаговых библиотек вариантов, обладающих сродством к различным мишеням. Внедрение технологии фагового дисплея антител имело революционное значение в области клинической иммунологии, как для создания инструментов диагностики инфекционных заболеваний, так и для получения терапевтических агентов. Она стала также основой эффективных и относительно недорогих методов исследования белок-белковых взаимодействий, сайтов связывания рецепторов, идентификации эпитопов и мимотопов. Технология фагового дисплея антител включает в себя ряд этапов, от успешной реализации которых зависит финальный результат. Основа любой библиотеки – разнообразие, природное или полученное при помощи методов комбинаторной химии. Критически важным является подбор молекулярных техник, обеспечивающих сохранение этого разнообразия на этапе получения библиотек фагмидов и на этапе трансформации клеток *E. coli*. После добавления фага-помощника к суспензии трансформированных клеток *E. coli* происходит формирование библиотеки бактериофагов, которая служит рабочим инструментом для проведения процедуры аффинной селекции и поиска индивидуальных молекул. Несмотря на кажущуюся простоту создания фаговых библиотек антител, существует ряд тонкостей, которые необходимо учитывать. В первую очередь это особенности подготовки фагмидного вектора.

В настоящее время разработано большое количество фагмидных векторов, и их выбор также имеет большое значение. Ключевым этапом считается подготовка компетентных клеток *E. coli* и технология их трансформации. Немаловажен выбор фага-помощника и способа его подготовки. Статья посвящена основным проблемам, с которыми сталкиваются исследователи при создании фаговых библиотек антител. Ключевые слова: фаговая библиотека; репликативная форма бактериофага; моноклональные антитела; фаг-помощник; компетентные клетки.

Introduction

In 1985, Smith was the first to describe the phage display technique and show that filamentous phages are able to display a peptide of interest on their surface after insertion of a foreign DNA fragment into the major coat protein gene (Smith, 1985). Subsequently, Parmley and Smith described the process for selecting and enriching phage libraries based on affinity binding called “biopanning” (Parmley, Smith, 1988). And in 1991, McCafferty and Winter were the first to use phage display technology for antibody production, having generated their combinatorial libraries using filamentous bacteriophages to screen for antigen-specific monoclonal antibodies (McCafferty et al., 1990).

The filamentous bacteriophages M13, f1, and fd were key tools of phage display technology. They are stable in solutions under various conditions, including extreme pH and high temperatures (Brigati, Petrenko, 2005), and insensitive to lytic enzymes (including DNases and proteases) (Larocca et al., 2002).

The filamentous bacteriophage M13 (Fig. 1) is one of the most commonly used bacteriophages for antibody phage display (Sokullu et al., 2019). M13 phage has a high replication capacity, and large foreign DNA fragments can be inserted into its genome by genetic engineering methods. This bacteriophage is non-lytic, infects and replicates in *Escherichia coli* strains carrying F⁺ that form F-pili (O’Callaghan et al., 1973). The greatest portion of the virion consists of single-stranded DNA covered with approximately 2,700 copies of the major coat protein pVIII. Each end of the phage consists of five copies of two different proteins: pVII and pIX at one end, pIII and pVI at the other end (see Fig. 1). Virion length

depends on the length of the genome packed into a phage particle. Up to 12,000 nucleotides can be added to the wild-type genome without disrupting phage packaging (Rasched, Oberer, 1986).

During the replication cycle of filamentous bacteriophages, their genome is simultaneously present in the bacterial cell as a large number of copies of the replicative form (double-stranded circular DNA) and single-stranded DNA (ssDNA). The presence of ssDNA in the DNA sample interferes with cloning as it creates a “background” of clones without insertion of the target gene. Using the purified replicative form, it is possible to clone DNA fragments within the sites encoding the bacteriophage protein of interest, mainly the pIII and pVIII proteins. Such vector systems are referred to as systems 3 and 8, respectively (Kay et al., 1996).

The main differences in using the pIII and pVIII proteins consist in the length of foreign peptides, presentation of chimeric proteins on the surface of the phage progeny, and the effect on viability (Ilyichev et al., 1989; Bass et al., 1990; Barbas et al., 1991). Compared to the pVIII protein, the pIII protein can ensure representation of longer amino acid sequences (Näkelä et al., 1978). The pVIII protein system is commonly used for small peptides. Larger polypeptides interfere with the function of the pVIII protein, so a system with the pIII protein is used for them. Due to the structural properties of Ff phages, application of the pIII protein in the phage display system results in less than five copies of the fusion protein being present in the phage progeny. It is only in the absence of the pIII gene in the helper phage genome that five copies of this protein will be present in the phage progeny (Smith, 1993).

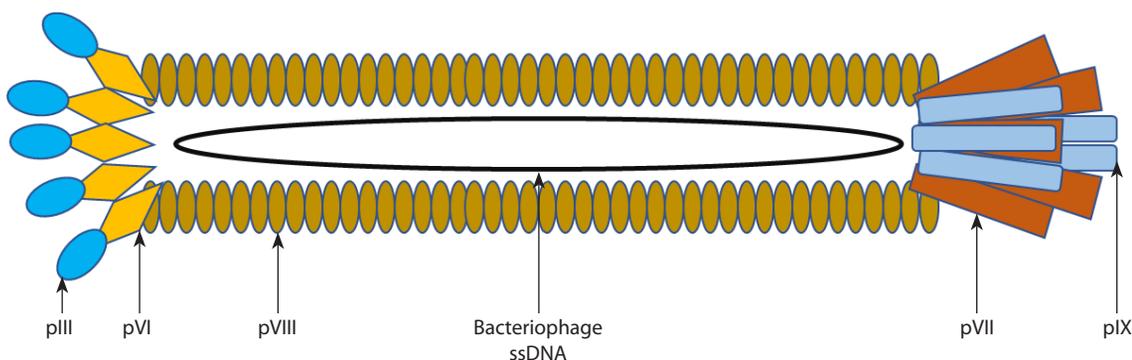


Fig. 1. Schematic representation of the filamentous bacteriophage M13.

The filamentous phage M13 is approximately 880 nm long and 6.5 nm wide and is surrounded by single-stranded DNA. The capsid of bacteriophage M13 is formed by five copies of different minor coat proteins (pIII, pVI, pVII, and pIX), but the major coat protein pVIII, represented by ~2,700 copies, makes up the most part of the capsid (Chung et al., 2011).

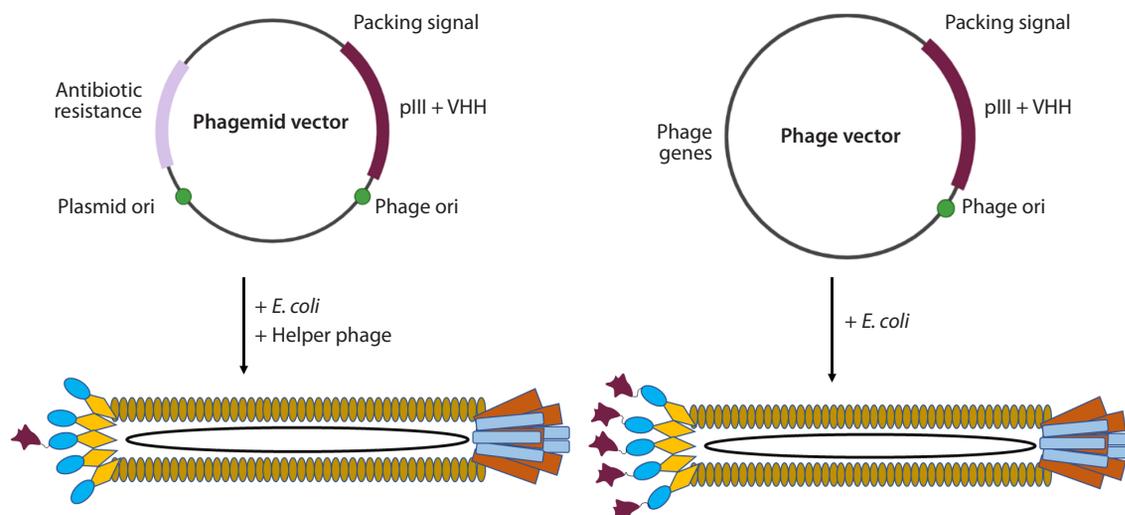


Fig. 2. Distinguishing between a phagemid and a phage vector.

The use of pVIII allows production of phage particles containing much more than five copies of chimeric proteins: hundreds or even thousands of them (Veronese et al., 1994). However, because of the multiple copies of chimeric proteins, the avidity effect begins to play a major role, which can impede the selection of protein variants with different affinities. Therefore, pVIII is used in phage display to expand the range of potential ligands, and pIII is used to reduce or eliminate the avidity effect to select high-affinity proteins. In addition, the N-terminal domain of the major coat protein pIII is involved in host cell infection. To infect *E. coli*, filamentous phages use filamentous bacterial structures known as F-pili. The end of the phage with the pIII protein interacts with the TolA protein on the bacterial surface. Although the pIII protein is relatively large, excessive perturbation of its structure can lead to disruption of its interaction with F-pili. This can have a significant effect on progeny phage viability; therefore, applications of the pIII protein for protein fragment display are limited. Such problems do not occur when the VIII protein is used.

As for the application of other bacteriophage proteins, very few papers have so far reported the feasibility of using the pVII and pIX proteins (Gao et al., 1999, 2002). The pVII and pIX proteins form a complex at the end of the phage particle opposite to the end carrying the widely used pIII protein. In order to demonstrate the feasibility of using these proteins for display, a phagemid variant was developed in which the variable regions of the heavy and light chains of the antibody were fused to the N-termini of pVII and pIX, respectively. Remarkably, the fusion proteins interact to form a functional Fv-binding domain on the phage surface. This approach looks promising for displaying complex libraries of peptides and proteins that could form combinatorial heterodimeric structures of the Fv domain of an antibody. However, this format has not yet been widely used.

In addition to phage vectors, phagemids are also used for library construction. Phagemids are vector molecules combining the properties of a plasmid and a phage vector (Kay et al.,

1996). The main difference between phagemids and phage vectors is that phagemids are vectors derived from Ff-phages that contain the plasmid origin of replication (Ori) for double-stranded replication and the fl Ori to allow single-stranded replication and packaging into phage particles (Fig. 2). They usually either do not encode at all, or encode only one type of major coat protein, and other structural and functional proteins required to complete the bacteriophage life cycle are provided by the helper phage (Ledsgaard et al., 2018).

An alternative approach is proposed by Chasteen and his team, who have developed “bacterial packaging cell lines” containing M13-based helper plasmids (Chasteen et al., 2006). The use of such cells ensures the synthesis of phage packaging proteins that assemble phage particles as efficiently as a helper phage infection, thereby eliminating the helper phage stage of bacterial cell infection.

Other elements, such as molecular tags and selective markers, are also inserted into the phagemids to facilitate subsequent operations, such as protein purification. Tagging can also be used to facilitate library screening.

The relatively small size of phagemids allows one to clone larger gene fragments encoding fusion proteins. The efficiency of phagemid transformation is higher than that of phage vectors, allowing the construction of large repertoire libraries (Qi et al., 2012).

There are a large number of phagemid variants. Type III and VIII phagemids can be reviewed in more detail in (Qi et al., 2012). The most common phagemids are pHEN, pComb, pSEX, and pADL. Each of them has several derivatives. Each phagemid has an original design, the choice of which depends on the researcher’s goal. Thus, the pADL-10b phagemid has fl ori to enable single-stranded replication and packaging into phage particles, a copy of LacI to provide lac promoter repression, and a strong transcription terminator to prevent unwanted and toxic expression during cloning. This design ensures reliable cloning of variant antibody genes and increased library stability during screening (Kreber et al., 1997). This vector is recommended for scFv display and is best suited

when a re-cloning step is envisaged to synthesize the protein in a soluble form.

Incorporation of the amber stop codon (TAG) between the target protein and the phage coat protein initiates expression of the fusion protein in *E. coli* suppressor strains such as TG1. In this case, the TAG codon is translated as a glutamine residue. However, in non-suppressor strains, such as HB2151, a soluble form of the recombinant protein will be produced because such strains do not contain glutamine in the amber stop codon. In this case, TAG would be the normal stop codon (Hoogenboom et al., 1991).

The expression level of the pADL series phagemids is equivalent to that of the pComb3 phagemid and significantly lower than that of the pHEN phagemid. Overexpression leads to toxicity, which makes antibody libraries genetically unstable and contributes to a strong tendency towards loss of diversity. In this respect, the pADL series phagemids provide a good balance between toxicity and synthesis of target antibodies in the periplasmic space of bacterial strains (Ishina et al., 2020).

The most popular derivative of the pComb vector is pComb3XSS. It contains two hydrolysis sites of the SfiI restrictase, which recognizes the GGCCNNNNNGGCC sequence. The lack of specificity for the five central nucleotide pairs allows PCR products encoding antibody fragments to be inserted in the desired orientation. The orientation is determined by the design of the forward and reverse primers, which also contain the SfiI site. The Comb3XSS vector also contains 6×His and HA tags, which allow protein purification and detection. It also contains an amber stop codon to switch off expression of the pIII fusion protein by switching to a non-suppressor *E. coli* strain, allowing production of soluble protein without subcloning.

The key points in working with phage libraries are vector preparation, preparation of well-competent cells, electroporation and amplification of the helper phage. Each of these points is discussed below and some recommendations are provided.

Phagemid preparation

Preparing a vector for insertion of target DNA sequences is one of the key aspects in phage library construction. The main difficulty consists in purifying the phagemid DNA preparation from single-stranded and linear forms of DNA. These DNA forms constitute the “background” on the control dish during electroporation, making it difficult to assess the “representation” of the phage library. Most importantly, they can have a negative impact on the level of library diversity. Various phage DNA purification protocols are used to overcome this problem. Once DNA has been isolated, it is possible to purify the preparation from the single-stranded and linear form using the phenol method. This is the “classic” method, but it has a number of disadvantages, such as significant losses of the replicative form of DNA, and handling phenol requires special working conditions. DNA isolation using the phenol method often depends on the skill of the person performing this procedure (Maniatis et al., 1984).

There is also a method of isolating the replicative form of phage DNA using ammonium acetate. Unlike the phenol

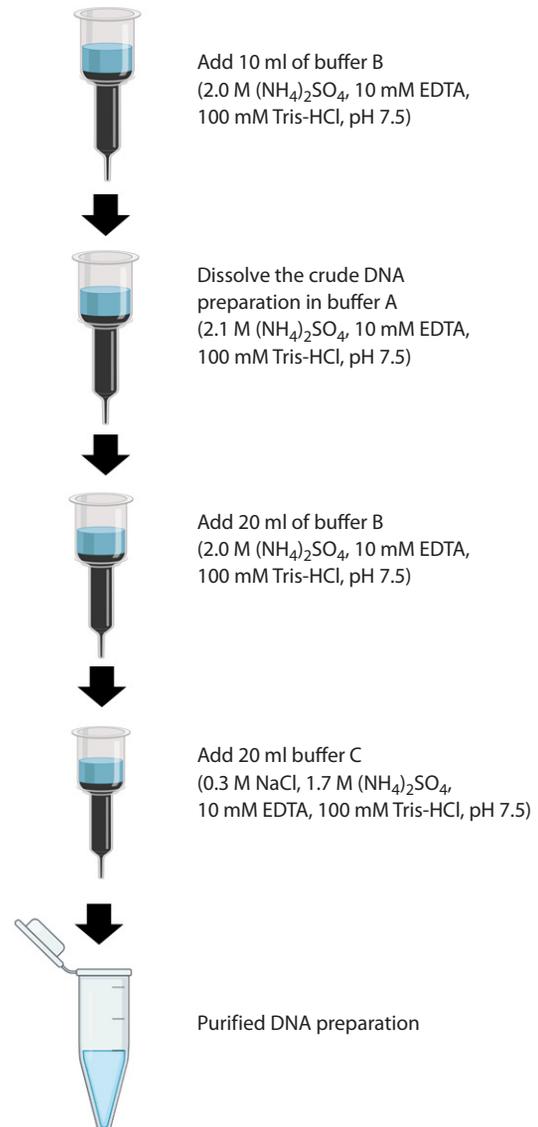


Fig. 3. Scheme of phage vector purification to remove single-stranded and linear forms of DNA.

method, it does not require special working conditions but is rather time-consuming (Maniatis et al., 1984).

An alternative is to purify the DNA preparation using sorbents (e. g., HiTrap PlasmidSelect Xtra Sorbent). The main steps of DNA purification using a sorbent are shown in Figure 3.

This method is not ideal either, but it does not require any special conditions to carry out the purification procedure. After purification on the sorbent, one should be prepared for significant DNA loss. In the best-case scenario, the total DNA yield will be 10 %. Meanwhile, this amount of DNA purified from satellite forms is quite sufficient to obtain a phage library.

Competent cells

Highly competent cells are a critical step in obtaining a phage library with high representation. TG1 cells are considered to be among the best cell lines for obtaining large phage libra-

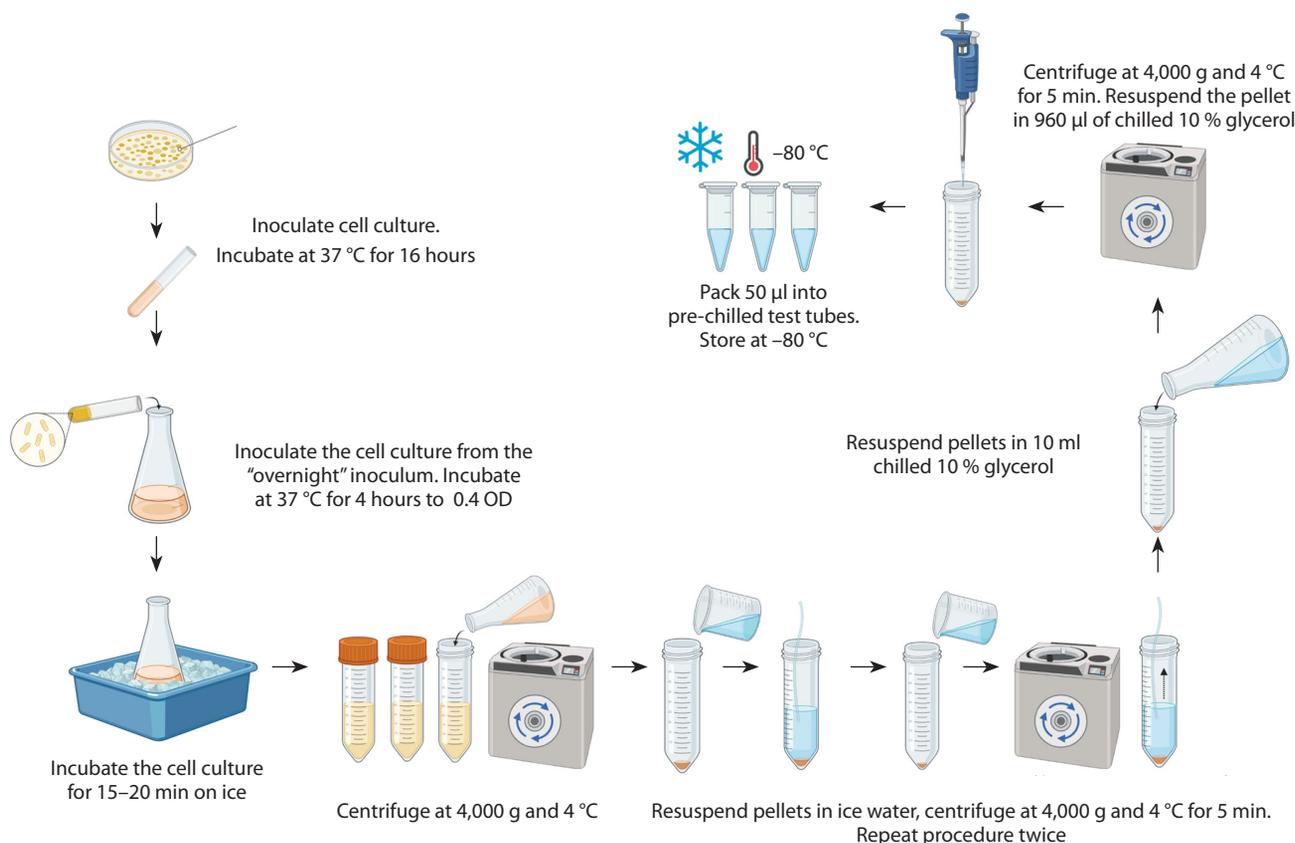


Fig. 4. Typical scheme for obtaining competent cells.

ries (Tu et al., 2005; Chai et al., 2020). TG1 is an unmodified derivative of the *E. coli* strain JM101. Currently, TG1 is one of the fastest growing *E. coli* clones, which undoubtedly plays an important role in setting up the experiment.

There are commercial companies offering ready-made cells with a competence of $\geq 4 \cdot 10^{10}$ CFU/ μ g DNA. However, the supply of such cells is not always possible, and the transport process leads to temperature fluctuations that reduce cellular competence. Therefore, the development of a protocol for obtaining highly competent cells for the generation of phage antibody libraries is an extremely important task (Fig. 4).

To obtain a large diversity of the phage library (10^{10-11} CFU), it is necessary to observe simple but very important conditions in the preparation of competent cells. It is important to ensure proper culturing conditions, and temperature in particular. The recommended temperature for culturing is +37 °C. It is also important to ensure that the cell culture density does not exceed 0.4 OD. Cells should be in the early to mid logarithmic phase of growth. Using cells in the late logarithmic or stationary phase will significantly reduce electroporation efficiency. It is also very important to avoid temperature extremes. Once your culture has grown to the required density, all manipulations need to be carried out on "wet" ice.

It is also necessary to resuspend the cells very carefully at each stage, since they are very fragile and easily destroyed. It is recommended to store the obtained cells at -80 °C (Chai et al.,

2020). At this temperature, the cells retain their competence for a longer time. However, cellular competence deteriorates over time, so freshly prepared cells should be used to maintain the diversity of phage libraries.

Electroporation

Prior to electroporation (Fig. 5), its efficiency was assessed using pUC19 plasmid DNA at a known concentration when a phage library was obtained (Chai et al., 2020).

Before starting electroporation, it is necessary to thaw the competent cells in ice. This manipulation preserves the viability of the cells. Simultaneously with thawing the cells, any nutrient medium needs to be warmed to +37 °C.

During electroporation, it is important to carry out the manipulations clearly and quickly. First, 1 μ l (10 pg/ μ l) of DNA needs to be added. This is the amount needed to deliver the control plasmid pUC19 to the thawed electrocompetent cells. If electroporation is performed with a ligase mixture that has been pre-cleared from the ligation buffer, it is necessary to take much larger (15–25-fold) amounts of DNA (Pardon et al., 2014).

It is also important to control the voltage according to the cuvette size. Immediately after the electrical pulse, 1 ml of SOC recovery medium must be added and the cells need to be incubated for 1 hour at +37 °C in a thermostat without stirring. After incubation, the transformed cells are centrifuged, 950 μ l of supernatant is removed, the precipitate is

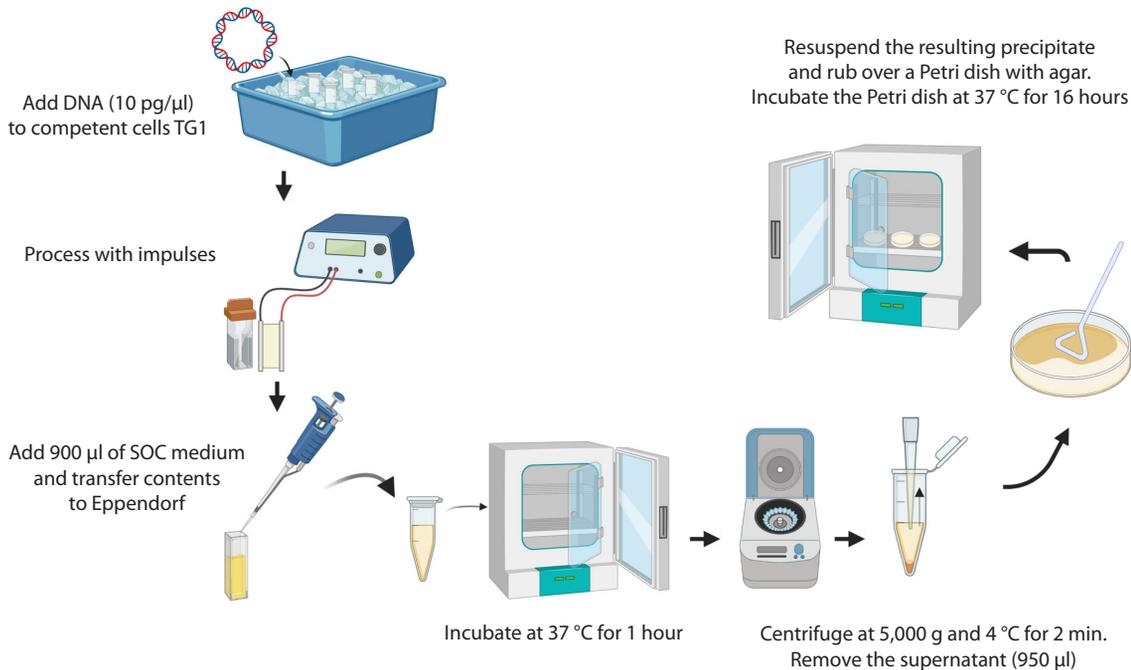


Fig. 5. Scheme of electroporation.

resuspended and spread on a Petri dish containing agar, the required antibiotic and 2 % glucose. Incubation is performed at +37 °C for 16 hours.

On the following day, the electroporation efficiency is assessed. Electroporation efficiency is defined as the number of colony-forming units obtained by electroporation of 1 μg of plasmid per given volume of competent cells. The higher the competence of the cells obtained, the greater the diversity of the phage library. Cells above 10¹⁰ CFU are considered highly competent.

Helper phage

Together with the phagemid, the helper phage is an important component for successful production of diverse libraries. It is required for the assembly of phage particles by infecting cells with phagemid. The helper phage carries all the genes required for infection, replication, assembly and budding, thus providing the phagemid, which carries the gene encoding the pIII-scFv fusion protein, with the proteins required for amplification.

The helper phage infects the bacterium by first attaching to its F-pili and then, once attached, transporting its genome into the cytoplasm of the host cell. Inside the cell, the phage genome initiates the production of phagemid single-stranded DNA in the cytoplasm. The phagemid DNA is then packaged into phage particles. Phage particles containing single-stranded DNA are released from the bacterial cell into the extracellular environment. Filamentous phages inhibit bacterial growth but, unlike phage λ and phage T7, do not lyse host cells. Helper phages are usually designed so that their DNA packaging is less efficient than that of phagemids, and the resulting phage particles preferentially contain phagemid

DNA. This is due to a defective replication start site (Lund et al., 2010).

M13KO7 is the main helper phage currently used in laboratories (Russel et al., 1986; Vieira, Messing, 1987; Du, Zhang, 2009). All other helper phages are derived from it. The M13KO7 helper phage is a derivative of phage M13 and carries a heterologous, low copy number p15A replication start site. When the M13KO7 phage is present alone in the host bacterium, replication is sufficient to produce its high titers. However, if a large number of copies of the phagemid are present, it will displace the helper phage genome during packaging, meaning that most of the resulting phage particles will carry phagemid DNA (Vieira, Messing, 1987). The capsid of the resulting phage particles will contain pIII encoded by the helper phage genome as well as the chimeric pIII-scFv protein encoded by the phagemid. In practice, there is preferential incorporation of wild-type pIII from the helper phage. This means that most phage particles will be “bald” (lacking the chimeric pIII-scFv). The preferred display of pIII over fused pIII-scFv is probably due to a combination of differences in expression/translation levels and the fact that some of the fused pIII-scFv appear to be degraded.

Choosing a specific helper phage to work with is quite difficult. Despite the stated differences in properties, in practice some derivatives of the M13KO7 helper phage do not differ from each other. For example, if we compare all the derivatives of the helper phage CM13, it gives larger plaques and higher titers (twice as many CFU/ml), so it is easier to prepare. The CM13 helper phage gives consistently good yields, even at low multiplicity of infection, but has a lower packing factor. At the same time, the M13KO7 helper phage gives similar phage yields, requires precise adherence to bacterial culture

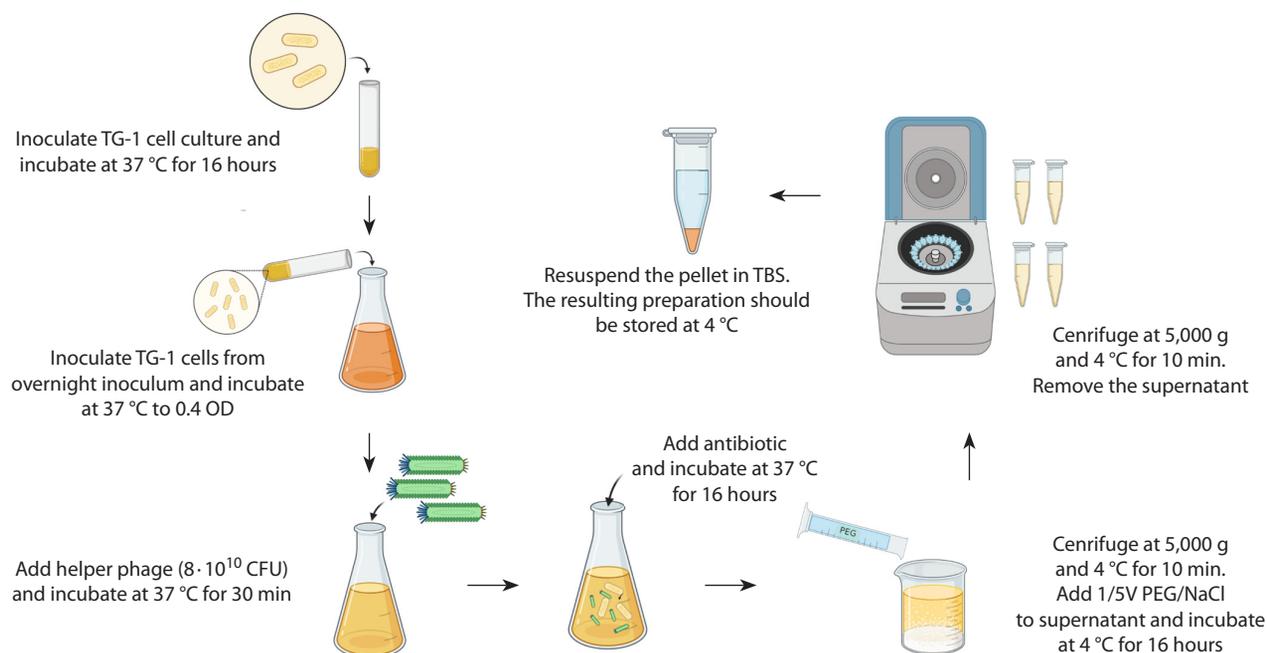


Fig. 6. Scheme of helper phage amplification.

density (about 0.5 OD₆₀₀) for better cell contamination, and ensures a high packing factor. The M13KO7 helper phage is recommended for producing single-stranded DNA, e. g. for Kunkel mutagenesis, due to its better packing ratio. The helper phage R408 is an fl derivative, does not carry antibiotic resistance and may improve packaging of single-stranded DNA (Dotto et al., 1981).

Hyperphages are a very attractive replacement for conventional helper phages (Rondot et al., 2001). Hyperphages contain a genome with a deletion of the pIII gene. To obtain them, packaging cells from *E. coli* are used to synthesize functional pIII, which is necessary for the formation of a complete particle and packaging of the phage genome. The resulting hyperphages carry functional pIII on their surface, but their genome is defective in the pIII gene. Each of the resulting phages carries multiple copies of the antibody or peptide on its surface, which greatly increases panning efficiency.

Although the helper phage is commercially available, for convenience it is useful to amplify a helper phage that is as good as the commercial phage in terms of biological titer. One of the important criteria for helper phage amplification is the optical density of the infected *E. coli* cell culture; the optimum value is usually 0.4 OD (Fig. 6). Higher densities result in a large percentage of non-transduced bacteria not producing virions, while lower densities can increase mismatches caused by differences in phage growth rate and lower phage production, as well as the expression of toxic fusion proteins that limit bacterial growth. The infectious dose of the helper phage is also critical. The optimal dose is $8 \cdot 10^{10}$ CFU/ml; the excess usually leads to overconsumption of the preparation.

When infecting cells with a helper phage, incubation should be performed without agitation. Agitation promotes cell growth; therefore, not all the cells will be infected with a helper phage. It is also important to respect the time intervals: incubation of the cell culture with a helper phage at +37 °C should not exceed 30 minutes. Extraordinary growth of helper phage-infected cells is achieved by adding an antibiotic (depending on the antibiotic embedded in the helper phage genome), in most cases kanamycin.

If the helper phage is precipitated with PEG/NaCl solution, the preparation should be incubated for 16 hours at +4 °C. This promotes better precipitation of the helper phage. Precipitation should be performed twice, since some preparation will remain in the supernatant after the first precipitation. The resulting preparation should be dissolved in TBS buffer. The amount of TBS buffer used is chosen individually and depends on the precipitate size. The resulting preparation must be stored at +4 °C.

Conclusion

In this article, we have attempted to reflect on the main problems faced by researchers dealing with phage library construction and the ways of solving them. In our opinion, one of the key problems in the construction of phage antibody libraries is the preparation of the phagemid vector for the insertion of antibody gene diversity, on which the library diversity depends critically. Equally important is amplification and obtaining a purified helper phage preparation, as well as obtaining highly competent cells for electroporation.

Despite these methodological challenges, the phage display technique continues to be successfully developed. This tech-

nology is attractive for a number of reasons, including its relative simplicity, the ability to screen large numbers of samples in a short time, and the ability to select antibodies specific to a wide range of antigens. Phage display has opened up new prospects for antibody discovery and development. Application of phage libraries of human antibody fragments makes it possible to avoid the use of chimeric or humanized antibodies for therapy, thus avoiding problems associated with the immunogenicity of drugs.

We would like to point out that this article is not an ideal guide but may help one solve some problems in obtaining a phage library of antibodies with a large repertoire.

References

- Barbas C.F., Kang A.S., Lerner R.A., Benkovic S.J. Assembly of combinatorial antibody libraries on phage surfaces: the gene-III site. *Proc. Natl. Acad. Sci. USA*. 1991;88(18):7978-7982. DOI 10.1073/pnas.88.18.7978
- Bass S., Greene R., Wells J.A. Hormone phage: an enrichment method for variant proteins with altered binding-properties. *Proteins*. 1990;8(4):309-314. DOI 10.1002/prot.340080405
- Brigati J.R., Petrenko V.A. Thermostability of landscape phage probes. *Anal. Bioanal. Chem*. 2005;382(6):1346-1350. DOI 10.1007/s00216-005-3289-y
- Chai D., Wang G., Fang L., Li H., Liu S., Zhu H., Zheng J. The optimization system for preparation of TG1 competent cells and electrotransformation. *MicrobiologyOpen*. 2020;9(7):e1043. DOI 10.1002/mbo3.1043
- Chasteen L., Ayriess J., Pavlik P., Bradbury A.R. Eliminating helper phage from phage display. *Nucleic Acids Res*. 2006;34(21):e145. DOI 10.1093/nar/gkl1772
- Chung W.J., Oh J.W., Kwak K., Lee B.Y., Meyer J., Wang E., Hexemer A., Lee S.W. Biomimetic self-templating supramolecular structures. *Nature*. 2011;478(7369):364-368. DOI 10.1038/nature10513
- Dotto G.F., Enea V., Zinder N.D. Functional analysis of bacteriophage fl intergenic region. *Virology*. 1981;114(2):463-473. DOI 10.1016/0042-6822(81)90226-9
- Du D., Zhang R. Application and progress of helper phage in phage display. *Microbiology*. 2009;36(2):261-266
- Gao C., Mao S., Lo C.H., Wirsching P., Lerner R.A., Janda K.D. Making artificial antibodies: a format for phage display of combinatorial heterodimeric arrays. *Proc. Natl. Acad. Sci. USA*. 1999;96(11):6025-6030. DOI 10.1073/pnas.96.11.6025
- Gao C.S., Mao S.L., Kaufmann G., Wirsching P., Lerner R.A., Janda K.D. A method for the generation of combinatorial antibody libraries using pIX phage display. *Proc. Natl. Acad. Sci. USA*. 2002;99(20):12612-12616. DOI 10.1073/pnas.192467999
- Hoogenboom H.R., Griffiths A.D., Johnson K.S., Chiswell D.J., Hudson P., Winter G. Multi-subunit proteins on the surface of filamentous phage: methodologies for displaying antibody (Fab) heavy and light chains. *Nucleic Acids Res*. 1991;19(15):4133-4137. DOI 10.1093/nar/19.15.4133
- Ilyichev A.A., Minenkova O.O., Tatkov S.I., Karpishev N.N., Eroshkin A.M., Petrenko V.A., Sandakchiev L.S. Creation of a viable variant of phage M13 by inserting foreign peptide in major coat protein. *Doklady AN SSSR = Doklady Biological Sciences*. 1989;307(2):481-483 (in Russian)
- Ishina I.A., Filimonova I.N., Zakharova M.Y., Ovchinnikova L.A., Mamedov A.E., Lomakin Y.A., Belogurov A.A., Jr. Exhaustive search of the receptor ligands by the CyCLOPS (Cytometry Cell-Labeling Operable Phage Screening) technique. *Int. J. Mol. Sci*. 2020;21(17):6258. DOI 10.3390/ijms 21176258
- Kay B., Winter J., McCafferty J. (Eds.). Phage Display of Peptides and Proteins: A Laboratory Manual. N.Y.: Acad. Press, 1996. DOI 10.1016/B978-0-12-402380-2.X5000-X
- Krebber A., Bornhauser S., Burmester J., Honegger A., Willuda J., Bosshard H.R., Pluckthun A. Reliable cloning of functional antibody variable domains from hybridomas and spleen cell repertoires employing a reengineered phage display system. *J. Immunol. Methods*. 1997;201(1):35-55. DOI 10.1016/S0022-1759(96)00208-6
- Larocca D., Burg M.A., Jensen-Pergakes K., Ravey E.P., Gonzalez A.M., Baird A. Evolving phage vectors for cell targeted gene delivery. *Curr. Pharm. Biotechnol*. 2002;3(1):45-57. DOI 10.2174/1389201023378490
- Ledsgaard L., Kilstrup M., Karatt-Vellatt A., McCafferty J., Laustsen A.H. Basics of antibody phage display technology. *Toxins*. 2018;10(6):236. DOI 10.3390/toxins10060236
- Lund P.E., Hunt R.C., Gottesman M.M., Kimchi-Sarfaty C. Pseudovirions as vehicles for the delivery of siRNA. *Pharm Res*. 2010;27(3):400-420. DOI 10.1007/s11095-009-0012-2
- Maniatis T., Fritsch E., Sambrook J. Molecular Cloning. Cold Spring Harbor Laboratory Press, 1982
- McCafferty J., Griffiths A.D., Winter G., Chiswell D.J. Phage antibodies: filamentous phage displaying antibody variable domains. *Nature*. 1990;348(6301):552-554. DOI 10.1038/348552a0
- Näkelä O., Kaartinen M., Pelkonen J.L., Karjalainen K. Inheritance of antibody specificity V. Anti-2-phenylloxazalone in the mouse. *J. Exp. Med*. 1978;148(6):1644-1660. DOI 10.1084/jem.148.6.1644
- O'Callaghan R., Bradley R., Paranchych W. The effect of M13 phage infection upon the F pili of *E. coli*. *Virology*. 1973;54(1):220-229. DOI 10.1016/0042-6822(73)90131-1
- Pardon E., Laeremans T., Triest S., Rasmussen S.G.F., Wohlkönig A., Ruf A., Muyltermans S., Hol W.G.J., Kobilka B.K., Steyaert J. A general protocol for the generation of Nanobodies for structural biology. *Nat. Protoc*. 2014;9(3):674-693. DOI 10.1038/nprot.2014.039
- Parmley S.F., Smith G.P. Antibody-selectable filamentous FD phage vectors: affinity purification of target genes. *Gene*. 1988;73(2):305-318. DOI 10.1016/0378-1119(88)90495-7
- Qi H., Lu H., Qiu H.J., Petrenko V., Liu A. Phagemid vectors for phage display: properties, characteristics and construction. *J. Mol. Biol*. 2012;417(3):129-143. DOI 10.1016/j.jmb.2012.01.038
- Rasched I., Oberer E. Ff coliphages: structural and functional relationships. *Microbiol. Rev*. 1986;50(4):401-427. DOI 10.1128/mr.50.4.401-427.1986
- Rondot S., Koch J., Breitling F., Dübel S. A helper phage to improve single-chain antibody presentation in phage display. *Nat. Biotechnol*. 2001;19(1):75-78. DOI 10.1038/83567
- Russel M., Kidd S., Kelley M.R. An improved filamentous helper phage for generating single-stranded plasmid DNA. *Gene*. 1986;45(3):333-338. DOI 10.1016/0378-1119(86)90032-6
- Smith G.P. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science*. 1985;228(4705):1315-1317. DOI 10.1126/science.4001944
- Smith G.P. Preface (Surface display and peptide libraries. Cold Spring Harbor Laboratory, April 4-7, 1992. Proceedings). *Gene*. 1993;128(1):1-2. DOI 10.1016/0378-1119(93)90145-s
- Sokullu E., Soleymani Abyaneh H., Gauthier A.M. Plant/bacterial virus-based drug discovery, drug delivery, and therapeutics.

- Pharmaceutics*. 2019;11(5):211. DOI 10.3390/pharmaceutics11050211
- Tu Z., He G., Li K.X., Chen M.J., Chang J., Chen L., Yao Q., Dongping P., Ye H., Shi J., Wu X. An improved system for competent cell preparation and high efficiency plasmid transformation using different *Escherichia coli* strains. *Electron. J. Biotechnol.* 2005;8(1):113-120. DOI 10.2225/vol8-issue1-fulltext-8
- Veronese F.D.M., Willis A.E., Boyerthompson C., Appella E., Perham R.N. Structural mimicry and enhanced immunogenicity of peptide epitopes displayed on filamentous bacteriophage. The V3 loop of HIV-1 gp120. *J. Mol. Biol.* 1994;243(2):167-172. DOI 10.1006/jmbi.1994.1643
- Vieira J., Messing J. Production of single-stranded plasmid DNA. *Methods Enzymol.* 1987;153:3-11. DOI 10.1016/0076-6879(87)53044-0

Funding. The work was carried out under the Thematic Government Order 23/21 of the Plan of Main Activities of the State Research Center of Virology and Biotechnology "VECTOR" for 2024.

Acknowledgements. The authors would like to express their sincere gratitude to the Head of the Department of Molecular Immunology, Doctor of Sciences in Biology, Alexander Vladimirovich Taranin, Senior Researchers, PhD Nikolay Andreevich Chikaev and PhD Alexander Matveevich Nayakshin, for their consultations during the work with phage libraries.

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

Received July 27, 2023. Revised January 18, 2024. Accepted January 19, 2024.

Прием статей через электронную редакцию на сайте <http://vavilov.elpub.ru/index.php/jour>
Предварительно нужно зарегистрироваться как автору, затем в правом верхнем углу страницы выбрать «Отправить рукопись». После завершения загрузки материалов обязательно выбрать опцию «Отправить письмо», в этом случае редакция автоматически будет уведомлена о получении новой рукописи.

«Вавиловский журнал генетики и селекции»/“Vavilov Journal of Genetics and Breeding”
до 2011 г. выходил под названием «Информационный вестник ВОГиС»/
“The Herald of Vavilov Society for Geneticists and Breeding Scientists”.

Сетевое издание «Вавиловский журнал генетики и селекции» – 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
на платформе Web of Science, 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:

русскаяязычная 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 28.03.2024. Формат 60 × 84 1/8. Уч.-изд. л. 18.1.