

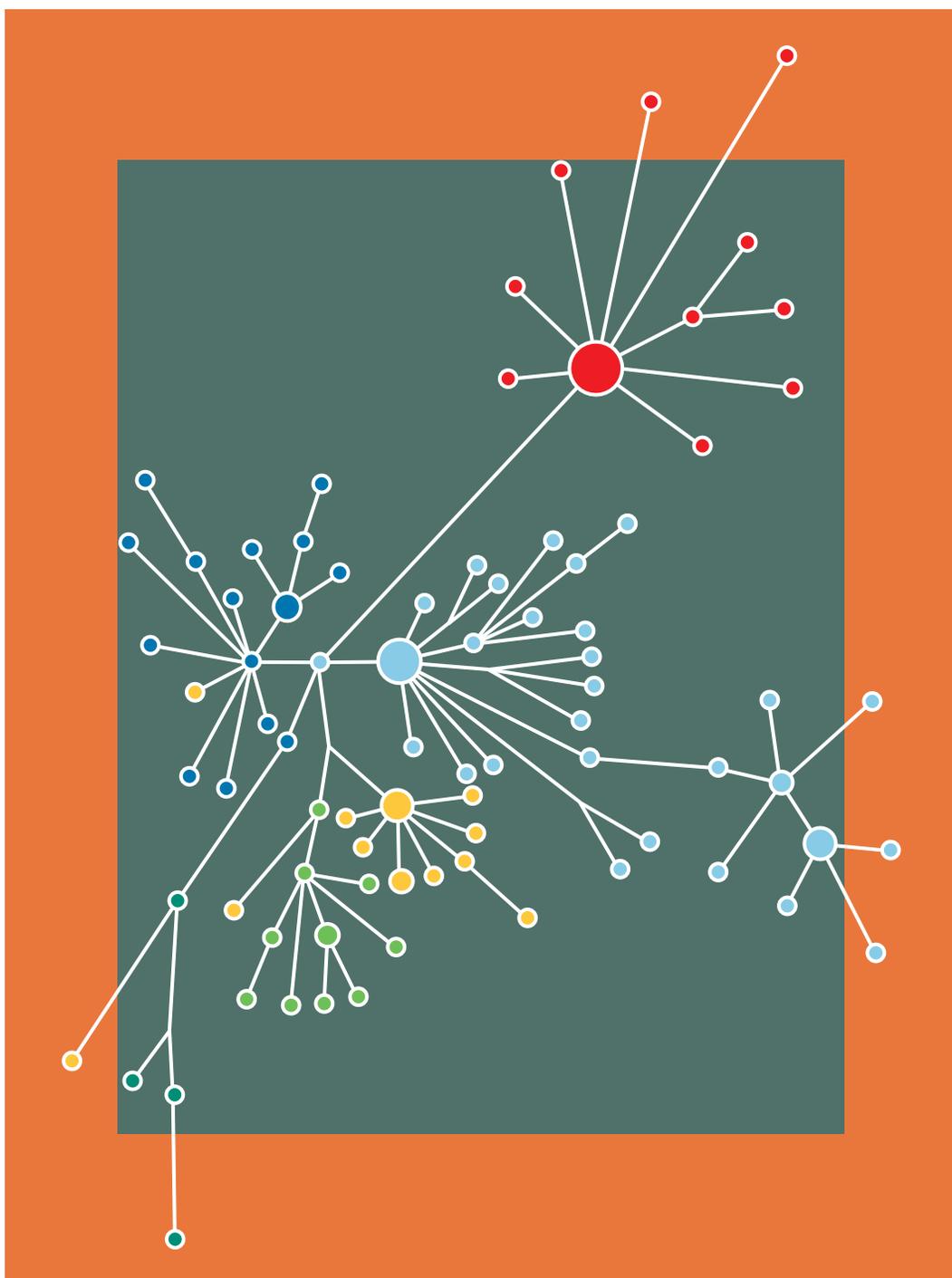
ВАВИЛОВСКИЙ

ЖУРНАЛ ГЕНЕТИКИ И СЕЛЕКЦИИ

VAVILOV JOURNAL OF GENETICS AND BREEDING

2023 • 27 • 1

Медицинская генетика / Популяционная генетика человека / Эпигенетика и регуляция активности генов /
Актуальные технологии



vavilov.ejpub.ru/
biomet.nsc.ru/vogjs
vavilov_journal@biomet.nsc.ru
Индекс издания 42153



Научный рецензируемый журнал

ВАВИЛОВСКИЙ ЖУРНАЛ ГЕНЕТИКИ И СЕЛЕКЦИИ

Основан в 1997 г.

Периодичность 8 выпусков в год

DOI 10.18699/VJGB-23-01

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)

Scientific Peer Reviewed Journal

VAVILOV JOURNAL OF GENETICS AND BREEDING

VAVILOVSKII ZHURNAL GENETIKI I SELEKTSII

*Founded in 1997**Published 8 times annually*

DOI 10.18699/VJGB-23-01

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. Bespalova, Full Member of the RAS, Dr. Sci. (Agricul.), Russia

N.P. Bondar, Cand. Sci. (Biology), Russia

S.A. Borinskaya, Dr. Sci. (Biology), Russia

P.M. Borodin, Professor, Dr. Sci. (Biology), Russia

A. Börner, Dr. Sci., Germany

M. Chen, Dr. Sci. (Biology), People's Republic of China

S.A. Demakov, Dr. Sci. (Biology), Russia

T.A. Gavrilenko, Dr. Sci. (Biology), Russia

I. Grosse, Professor, Dr. Sci., Germany

N.E. Gruntenko, Dr. Sci. (Biology), Russia

S.G. Inge-Vechtomov, Full Member of the RAS, Dr. Sci. (Biology), Russia

E.K. Khlestkina, 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),
RussiaD.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.), BelarusS.V. Shestakov, Full Member of the RAS, Dr. Sci. (Biology),
RussiaV.A. Stepanov, Full Member of the RAS, Dr. Sci. (Biology),
RussiaI.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),
RussiaN.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

- 5 **ОТ РЕДАКТОРА**
К 40-летию Научно-исследовательского института медицинской генетики Томского национального исследовательского медицинского центра Российской академии наук.
В.А. Степанов, И.Н. Лебедев

Медицинская генетика

- 7 **ОБЗОР**
Генетическая канва герменевтики феномена сочетания болезней человека.
Е.Ю. Брагина, В.П. Пузырёв
- 18 **ОБЗОР**
Молекулярно-генетические основы буллезного эпидермолиза.
Ю.Ю. Коталевская, В.А. Степанов
- 28 **ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ**
Сравнительная цитогенетика анэмбрионии и неразвивающейся беременности у человека.
Т.В. Никитина, Е.А. Саженова, Е.Н. Толмачева, Н.Н. Суханова, С.А. Васильев, И.Н. Лебедев

Популяционная генетика человека

- 36 **ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ**
Структура и происхождение генофонда тувинцев по данным аутосомных SNP и гаплогрупп Y-хромосомы.
В.А. Степанов, Н.А. Колесников, Л.В. Валихова, А.А. Зарубин, И.Ю. Хитринская, В.Н. Харьков
- 46 **ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ**
Связь генофонда хантов с народами Западной Сибири, Предуралья и Алтая-Саян по данным о полиморфизме аутосомных локусов и Y-хромосомы.
В.Н. Харьков, Н.А. Колесников, Л.В. Валихова, А.А. Зарубин, М.Г. Сваровская, А.В. Марусин, И.Ю. Хитринская, В.А. Степанов

- 55 **ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ**
Идентичные по происхождению блоки в геномах коренного населения Сибири демонстрируют генетические связи между популяциями.
Н.А. Колесников, В.Н. Харьков, К.В. Вагайцева, А.А. Зарубин, В.А. Степанов

Эпигенетика и регуляция активности генов

- 63 **ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ**
Экспрессия генов *NUP153* и *YWHAB* с их канонических промоторов и альтернативных промоторов ретротранспозона LINE-1 в плаценте первого триместра беременности.
В.В. Деменова, Е.Н. Толмачева, Т.В. Никитина, Е.А. Саженова, С.Ю. Юрьев, А.Ш. Махмутходжаев, А.С. Зуев, С.А. Филатова, А.Е. Дмитриев, Я.А. Даркова, Л.П. Назаренко, И.Н. Лебедев, С.А. Васильев
- 72 **ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ**
Изменение профиля метилирования ДНК в ткани печени при прогрессировании HCV-индуцированного фиброза до гепатоцеллюлярной карциномы.
И.А. Гончарова, А.А. Зарубин, Н.П. Бабушкина, Ю.А. Королева, М.С. Назаренко

Актуальные технологии

- 83 **МЕТОДЫ И ПРОТОКОЛЫ**
Влияние предварительной обработки образцов периферической крови человека на качество Hi-C библиотек.
М.М. Гридина, Э. Весна, М.Е. Миньженкова, Н.В. Шилова, О.П. Рыжкова, Л.П. Назаренко, Е.О. Беляева, И.Н. Лебедев, В.С. Фишман

VAVILOV JOURNAL OF GENETICS AND BREEDING

CONTENTS • 2023 • 27 • 1

- 5 **FROM THE EDITOR**
On the 40th anniversary of the
Research Institute of Medical Genetics,
Tomsk National Research Medical Center.
V.A. Stepanov, I.N. Lebedev
- Medical genetics**
- 7 **REVIEW**
Genetic outline of the hermeneutics
of the diseases connection phenomenon
in human. *E.Yu. Bragina, V.P. Puzyrev*
- 18 **REVIEW**
Molecular genetic basis of epidermolysis
bullosa. *Yu.Yu. Kotalevskaya, V.A. Stepanov*
- 28 **ORIGINAL ARTICLE**
Comparative cytogenetics of anembryonic
pregnancies and missed abortions in human.
*T.V. Nikitina, E.A. Sazhenova, E.N. Tolmacheva, N.N. Sukhanova,
S.A. Vasilyev, I.N. Lebedev*
- Human population genetics**
- 36 **ORIGINAL ARTICLE**
Structure and origin of Tuvan gene pool
according to autosome SNP and
Y-chromosome haplogroups. *V.A. Stepanov,
N.A. Kolesnikov, L.V. Valikhova, A.A. Zarubin,
I.Yu. Khitrinskaya, V.N. Kharkov*
- 46 **ORIGINAL ARTICLE**
Relationship of the gene pool of the Khants
with the peoples of Western Siberia,
Cis-Urals and the Altai-Sayan Region
according to the data on the polymorphism
of autosomic locus and the Y-chromosome.
*V.N. Kharkov, N.A. Kolesnikov, L.V. Valikhova, A.A. Zarubin,
M.G. Svarovskaya, A.V. Marusin, I.Yu. Khitrinskaya, V.A. Stepanov*
- 55 **ORIGINAL ARTICLE**
Blocks identical by descent in the genomes
of the indigenous population of Siberia
demonstrate genetic links between
populations. *N.A. Kolesnikov, V.N. Kharkov,
K.V. Vagaitseva, A.A. Zarubin, V.A. Stepanov*
- Epigenetics and gene regulation**
- 63 **ORIGINAL ARTICLE**
Expression of the *NUP153* and *YWHAB*
genes from their canonical promoters
and alternative promoters of the LINE-1
retrotransposon in the placenta
of the first trimester of pregnancy.
*V.V. Demeneva, E.N. Tolmacheva, T.V. Nikitina, E.A. Sazhenova,
S.Yu. Yuriev, A.Sh. Makhmutkhodzhaev, A.S. Zuev, S.A. Filatova,
A.E. Dmitriev, Ya.A. Darkova, L.P. Nazarenko, I.N. Lebedev,
S.A. Vasilyev*
- 72 **ORIGINAL ARTICLE**
Changes in DNA methylation profile in liver
tissue during progression of HCV-induced
fibrosis to hepatocellular carcinoma.
*I.A. Goncharova, A.A. Zarubin, N.P. Babushkina,
I.A. Koroleva, M.S. Nazarenko*
- Mainstream technologies**
- 83 **METHODS AND PROTOCOLS**
Influence of human peripheral blood
samples preprocessing on the quality
of Hi-C libraries. *M.M. Gridina, E. Vesna,
M.E. Minzhenkova, N.V. Shilova, O.P. Ryzhkova,
L.P. Nazarenko, E.O. Belyaeva, I.N. Lebedev, V.S. Fishman*

К 40-летию Научно-исследовательского института медицинской генетики Томского национального исследовательского медицинского центра Российской академии наук

Уважаемые читатели!
Предлагаем вашему вниманию выпуск, посвященный 40-летию Научно-исследовательского института медицинской генетики Томского национального исследовательского медицинского центра Российской академии наук. На страницах текущего номера – статьи, подготовленные сотрудниками института и их коллегами по материалам докладов, прозвучавших на XIII науч-

ной конференции «Генетика человека и патология», посвященной юбилею института, проведенной в Томске 20–22 ноября 2022 г.

История института началась 6 июля 1982 г. В этот день в Томске состоялось торжественное открытие Отдела медицинской генетики московского Института медицинской генетики Академии медицинских наук СССР. Спустя 5 лет, в 1987 г., Научно-исследовательский институт ме-



Научно-исследовательский институт медицинской генетики Томского НИМЦ РАН.



Генетическая клиника Научно-исследовательского института медицинской генетики Томского НИМЦ РАН.

дицинской генетики был преобразован в самостоятельное научное учреждение и вошел в состав Томского научного центра АМН СССР.

В 1993 г. НИИ медицинской генетики становится Федеральным центром медико-генетической службы Министерства здравоохранения РФ, а в 1994 г. в структуре института открывается Генетическая клиника – первое и единственное в России специализированное медицинское учреждение для пациентов с наследственными заболеваниями. С момента создания и по настоящее время в фокусе внимания института находятся ключевые направления современной медицинской генетики и генетики человека – клиническая генетика, популяционная генетика и геномика, генетика многофакторных болезней, молекулярная цитогенетика, онтогенетика, персонализированная медицина. Важной частью деятельности института стала подготовка научных и медицинских кадров. Институт готовит аспирантов и ординаторов, с 1998 г. работает диссертационный совет по защите диссертаций по специальности «Генетика». В 1989 г. в Сибирском государственном медицинском университете на базе НИИ медицинской генетики был открыт курс медицинской генетики, а в 1999 г. – первая в Сибири кафедра медицинской генетики. В 2016 г. НИИ медицинской генетики вошел в состав Томского национального исследовательского медицинского центра Российской академии наук – крупнейшего в современной России академического медицинского научно-исследовательского учреждения, объединившего научные и клинические базы шести научно-исследовательских институтов. Сегодня НИИ медицинской генетики Томского НИМЦ – динамично развивающийся институт, один из ведущих медико-генетических центров в России. Исследования института направлены на выявление фундаментальных основ наследственной патологии человека, разработку и внедрение технологий диагностики, лечения и профилактики наследственных заболеваний, развитие медицинской генетики в России как важного направления, интегрирующего практически все отрасли современной медицинской науки.

Открывает выпуск аналитический обзор работ, посвященных генетическим аспектам феномена неслучайного сочетания разных болезней, подготовленный Е.Ю. Брагиной и В.П. Пузырёвым. В работе обобщены и систематизированы современные представления о генетических основах синтропных и дистропных заболеваний, генетической архитектуре многофакторных болезней человека – одного из магистральных направлений научных исследований коллектива лаборатории популяционной генетики.

В обзоре Ю.Ю. Коталевской и В.А. Степанова «Молекулярно-генетические основы буллезного эпидермолиза»

рассмотрены новые патогенетические механизмы и гены, ответственные за развитие классических и синдромальных форм буллезного эпидермолиза – тяжелого наследственного заболевания, сопровождающегося хрупкостью кожи и мультисистемными поражениями.

В статье Т.В. Никитиной с соавторами «Сравнительная цитогенетика анэмбрионии и неразвивающейся беременности у человека» обобщены результаты многолетнего цитогенетического скрининга эмбриолетальных мутаций, проводимого в лаборатории цитогенетики института. Накопленный уникальный материал позволил выделить специфические хромосомные аномалии в группах эмбрионов, различающихся по степени тяжести нарушений внутриутробного развития.

Развитие технологий полногеномного анализа вывело на новый уровень работы института в области популяционной и эволюционной генетики. В трех статьях, подготовленных под первым авторством В.А. Степановым, В.Н. Харьковским и Н.А. Колесниковым с сотрудниками лаборатории эволюционной генетики, представлены новые данные о структуре и происхождении генофонда ряда коренных сибирских этносов, продемонстрирована информативность идентичных по происхождению блоков генома в установлении генетических связей между популяциями и определении эволюционных механизмов адаптации человека к факторам окружающей среды.

Эпигенетика и молекулярные механизмы регуляции активности генов при патологии эмбрионального развития и многофакторных заболеваниях также находятся в фокусе внимания коллектива. В работе В.В. Деменевой с соавторами приведены данные об особенностях экспрессии генов как с их канонических, так и с альтернативных промоторов ретротранспозона LINE-1 в зависимости от уровня его метилирования в тканях плаценты. В статье И.А. Гончаровой с соавторами прослежена динамика изменений характера метилирования ДНК в клетках печени при прогрессировании фиброза, индуцированного вирусом гепатита С, до гепатоцеллюлярной карциномы.

Завершает выпуск работа наших коллег из Института цитологии и генетики СО РАН (г. Новосибирск), в которой обозначены ключевые требования к пробоподготовке биологического материала для создания Hi-C библиотек ДНК в целях диагностики хромосомных перестроек современными методами 3D геномики.

Редакция «Вавиловского журнала генетики и селекции» поздравляет коллектив НИИ медицинской генетики Томского НИМЦ с юбилеем, желает дальнейшего развития, творческих успехов и новых достижений на благо российской медицинской генетики!

Научные редакторы выпуска:

академик РАН В.А. Степанов

д-р биол. наук, профессор РАН И.Н. Лебедев

Перевод на английский язык <https://vavilov.elpub.ru/jour>

Генетическая канва герменевтики феномена сочетания болезней человека

Е.Ю. Брагина¹✉, В.П. Пузырёв^{1, 2}

¹ Научно-исследовательский институт медицинской генетики, Томский национальный исследовательский медицинский центр Российской академии наук, Томск, Россия

² Сибирский государственный медицинский университет Министерства здравоохранения Российской Федерации, Томск, Россия
✉ elena.bragina@medgenetics.ru

Аннотация. Структура заболеваний у человека неоднородна, характеризуется различными вариантами сочетаний болезней, включая сопутствующие патологии, связанные общим патогенетическим механизмом, а также болезни, редко проявляющиеся совместно на фенотипическом уровне. В последнее время отмечается рост интереса к изучению закономерностей развития не отдельных болезней, а целых семейств, связанных общими патогенетическими механизмами и общими генами, вовлеченными в их развитие. В результате установлен существенный генетический компонент, контролирующий образование конгломератов болезней сложным образом, через функционально взаимодействующие модули отдельных генов в генных сетях. Аналитический обзор исследований по проблематике разных аспектов сочетания болезней и является целью настоящей работы. В обзоре использована метафора герменевтического круга для познания структуры закономерных связей между болезнями, приведены концептуальные рамки, связанные с множественностью заболеваний у индивида. Рассмотрена существующая терминология применительно к ним, среди которых мультиморбидность, полипатии, коморбидность, конгломераты, семейства, «вторые болезни», синтропия и другие. Приведены ключевые результаты, чрезвычайно полезные, прежде всего, для описания генетической архитектуры болезней многофакторной природы. Обобщения по проблеме исследования феномена сочетания болезней позволяют приблизиться к систематизации и естественной классификации болезней. С точки зрения практического здравоохранения описание феномена сочетания болезней имеет решающее значение для расширения интерпретационного горизонта клинициста и выхода за пределы узких, ориентированных на конкретную болезнь терапевтических решений.

Ключевые слова: феномен сочетания болезней; синтропия; дистропия; коморбидность; герменевтика.

Для цитирования: Брагина Е.Ю., Пузырёв В.П. Генетическая канва герменевтики феномена сочетания болезней человека. *Вавиловский журнал генетики и селекции*. 2023;27(1):7-17. DOI 10.18699/VJGB-23-03

Genetic outline of the hermeneutics of the diseases connection phenomenon in human

Е.Ю. Брагина¹✉, V.P. Puzyrev^{1, 2}

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

² Siberian State Medical University, Tomsk, Russia
✉ elena.bragina@medgenetics.ru

Abstract. The structure of diseases in humans is heterogeneous, which is manifested by various combinations of diseases, including comorbidities associated with a common pathogenetic mechanism, as well as diseases that rarely manifest together. Recently, there has been a growing interest in studying the patterns of development of not individual diseases, but entire families associated with common pathogenetic mechanisms and common genes involved in their development. Studies of this problem make it possible to isolate an essential genetic component that controls the formation of disease conglomerates in a complex way through functionally interacting modules of individual genes in gene networks. An analytical review of studies on the problems of various aspects of the combination of diseases is the purpose of this study. The review uses the metaphor of a hermeneutic circle to understand the structure of regular relationships between diseases, and provides a conceptual framework related to the study of multiple diseases in an individual. The existing terminology is considered in relation to them, including multimorbidity, polyopathies, comorbidity, conglomerates, families, "second diseases", syntropy and others. Here we summarize the key results that are extremely useful, primarily for describing the genetic architecture of diseases of a multifactorial nature. Summaries of the research problem of the disease connection phenomenon allow us to approach the systematization and natural classification of diseases. From practical healthcare perspective, the description of the disease connection phenomenon is crucial for expanding the clinician's interpretive horizon and moving beyond narrow, disease-specific therapeutic decisions.

Key words: diseases connection phenomenon; syntropy; dystropy; comorbidity; hermeneutics.

For citation: Bragina E.Yu., Puzyrev V.P. Genetic outline of the hermeneutics of the diseases connection phenomenon in human. *Vavilovskii Zhurnal Genetiki i Seleksii = Vavilov Journal of Genetics and Breeding*. 2023;27(1):7-17. DOI 10.18699/VJGB-23-03

Введение

Мы живем в «мире многих миров» (Виленкин, 2010), который поражает нас таинственностью и всеобщностью связей явлений, многообразием эволюционно-исторических событий, как в космическом масштабе, так и в планетарном – земляне (человечество) являют собой такую же всеобщность связей между собой и окружающим миром. Эти связи складываются закономерно или случайно, имеют долгую филогенетическую историю в 4 млрд лет и всего столетнюю онтогенетическую историю каждого индивида. Структура «людских» связей, проявляющаяся в метаболической и морфофизиологической вариабельности, составляет основу медицинских оценок – «норма» или «болезнь». С начала века в биологии и медицине обозначился новый подход к исследованию данных вопросов – сетевой анализ как попытка разобраться в законах, управляющих разного рода сетями, от социальных до сложной паутины генов, властвующих над всеми клетками и признаками, определяя здоровье или болезнь (Barabási et al., 2011).

Геном человека как совокупность всех генов вида *Homo sapiens* находится в сложных и не до конца изученных отношениях с окружающей природой и социумом. Особенностью таких геномно-феномных отношений является различие, отмечаемое теперь часто: геном ограничен (приблизительно 3 млрд пар оснований у человека), феном – нет (его предел зависит от того, как далеко мы хотим двигаться) (Raigen, Erpig, 2000). В 1930-х гг., за столетие до свершившейся «геномной революции» выдающийся российский генетик А.С. Серебровский, обсуждая проблему органической эволюции, обозначил ее как «бесконечно-конечное противоречие» в «единстве бесконечного числа признаков и конечного числа генов» (Серебровский, 1973).

В таком бесконечном мире среди признаков всегда, хотя это и не просто, можно наблюдать и выявлять связанные друг с другом признаки, в том числе имеющие отношение к патологии. В клинике это явление составляет основу диагностики и врачевания, а устойчивые сочетания определенных болезненных признаков представляют самостоятельный предмет исследования – феномен сочетания болезней (ФСБ).

В 1970 г. американский врач и специалист в области эпидемиологии неинфекционных заболеваний А. Файнштейн для сочетаний болезней у отдельных индивидов предложил термин «коморбидность», означающий проявление дополнительного клинического состояния, которое существует или возникает на фоне текущего заболевания (Feinstein, 1970). Таким клиническим состоянием может быть болезнь, патологический синдром, беременность, длительная «строгая» диета или осложнение после лекарственной терапии. Коморбидность представляет собой комплексы нескольких болезней (мегаформы, конгломераты), одновременно существующих у отдельных пациентов и наблюдающихся значительно чаще, чем это можно ожидать при случайном распределении.

Популярность термина «коморбидность» поразительна, особенно среди врачей-клиницистов: создано международное научное общество коморбидности (IRCMo, The International Research Community on Multimorbidity), с 2010 г. издается журнал 'The Journal of Multimorbidity and Co-

morbidity' (<https://journals.sagepub.com/description/COB>), функционирует медицинская онлайн-площадка для обсуждения диагностики и лечения пациентов с коморбидными диагнозами (<https://nexusacademy.ru/about>). Автору термина приписывается открытие, «внесшее ясность» в толкование коморбидной патологии (Верткин, 2015). И все же не оставляет ощущение переоценки «ясности» в понимании феномена и термина. Оно похоже на ситуацию, описанную в романе известного нобелиата У. Фолкнера: «Все сразу заговорили – горячо, наперебой, запальчиво, нереальное обращая в возможное, затем в вероятное, затем в неоспоримый факт, как это у людей всегда выходит, когда они желания облекают в слова»¹.

Тем не менее надо согласиться, что термин «коморбидность» оказался особенно удачным для клиницистов. Он стал «зонтиком», под которым разместились многочисленные названия сочетаний болезней, вариантов двух и более форм патологии у пациентов и, нередко, у их ближайших родственников. Порой такие заболевания называются фоновыми, или сопутствующими. Вообще же терминология названий таких сочетаний болезней, по нашим подсчетам, превышает 30 единиц. Среди них: мультиморбидность, полипатии, соболезность, конгломераты, семейства, «вторые болезни» и др. Чаще они являются болезнями, имеющими «общий корень» (родство патогенеза, транссиндромальная коморбидность), хотя в отношении других сочетаний между ними патогенетической общности не обнаруживается (транснозологическая коморбидность). Заметим, что специальных терминоведческих исследований немного и в результате их не сложились консенсусные отношения (Azaïs et al., 2016; Navickas et al., 2016). Однако в текущей ситуации определился объект исследования – «коморбидный пациент» (Верткин, 2015), и накапливается доброкачественный клинический и клинико-эпидемиологический материал, «подоспевший» стать основой для осуществления «омиксных» подходов к изучению проблемы ФСБ. И как показывают исследования, сочетание болезней имеет довольно серьезный генетический аспект, по поводу которого и подготовлена настоящая статья.

Концептуальный инструментарий в генетическом исследовании ФСБ

Приведем подбор (совокупность) взглядов (принципов, концепций), связанных между собой и образующих единую систему, полезную, на наш взгляд, для понимания (трактовки, толкования) ФСБ. Воспользуемся метафорой «герменевтический круг», которая описывает взаимосогласие отдельного (части) и целого, подобно герменевтическому правилу: целое надлежит понимать на основании отдельного, а отдельное – на основании целого (Гадамер, 1991). В отношении ФСБ как целого нам представляются обоснованными версии «отдельностей» для введения в герменевтический круг фрагментов концепций (учений, принципов) выдающихся клинических генетиков: советского невропатолога С.Н. Давиденкова (1880–1961), американского врача-генетика В.А. МакКьюсика (1921–2008), немецкого педиатра М. Пфаундлера (1872–1947) и ныне здравствующего немецко-американского клини-

¹ Фолкнер У. Шум и ярость (Пер. с англ. О. Сорока). СПб.: Азбука-Классика, 2010. С. 130.

циста Дж.М. Опитца. Все они ученые-генетики и одновременно, что особенно важно, практикующие врачи, исследовавшие полиморфизм проявления болезней и загадочный феномен сочетания нескольких патологий у одного больного.

Не придерживаясь хронологического порядка в появлении их трудов, последуем задуманной логике в изложении структуры герменевтического круга, т. е. тех «отдельностей», которые могут быть полезными в толковании явления как целого – феномена сочетания болезней.

Объединители» (lumpers) и «разделители» (splitters) (McKusick, 1969). В 1960-х гг. в медико-генетическом сообществе была открыта дискуссия: что такое нозология генетических болезней? Обсуждались главным образом менделевские болезни, но также болезни с наследственным предрасположением (многофакторные заболевания – МФЗ). Фенотипически больные представляют огромное клиническое разнообразие, а возможности уточнения этиологии заболеваний молекулярно-генетическими или цитогенетическими методами в те годы были существенно ограничены. Поэтому врачи-исследователи были вольны классифицировать больных, объединяя или разделяя их. Однако в ходе обсуждения проблемы было предложено важное обобщение – принципы медицинской генетики: плейотропизм, вариабельность (полиморфизм) и генетическая гетерогенность (McKusick, 1968). Именно эти принципы можно считать стабилизирующими смысловой контекст понимания ФСБ. Сегодняшние систематики патологии человека тоже опираются на данные принципы (Biesecker, 1998; Brunner, van Driel, 2004). Более того, с развитием геномной медицины появилась возможность описания генетической архитектуры болезней многофакторной природы, под которой понимают число генетических полиморфизмов, оказывающих влияние на риск болезней, распределение их аллельных частот и силу эффектов, а также генетическую манеру их поведения (аддитивность, доминантность и/или эпистаз, плейотропия) (Wray et al., 2008).

Синдром как плейотропия, гипотеза условных тропизмов (Давиденков, 1947; Opitz, Neri, 2013). Первое употребление слова «синдром» зафиксировано в 1541 г. и, как отмечают J.M. Opitz, G. Neri (2013), до сих пор используется для указания на общую причину, а не просто на совокупность симптомов. Эти же авторы обсуждают и другое словарное определение – синдром как совпадение проявлений, «характеризующих специфическое заболевание», неслучайное совпадение одинаковых или очень похожих комплексов проявлений у двух или более индивидов, позволяющих предположить сходный патогенез, подлежащий проверке причинно-следственной связи посредством обнаружения физических, инфекционных, токсикологических или генетических факторов.

Сегодня биохимические и сложные молекулярные/цитогенетические методы позволяют с высокой точностью выявлять генетические причины, эпигенетические модификации в сочетанных фенотипах или в синдромах. Объяснение таких сочетаний, их стойкости или «рассыпания» у потомков, тяжести проявлений сходного сочетания, толкование взаимоотношений многочисленных вариаций нормы или мелких аномалий с их развитыми формами

патологии было предложено С.Н. Давиденковым в гипотезе условных тропизмов (1947). Использованный им эволюционно-генетический подход к анализу более сотни нозологических форм нервных болезней человека показал, что нередкость совместного появления у одного больного или в одной семье болезней нервной системы объясняется условным тропизмом: помимо своего собственного влияния на развитие нервной системы, патологический задаток (ген) обладает способностью резко усиливать фенотипическое проявление остальных идущих в том же направлении особенностей генотипа, включающих многочисленные варианты. Так, например, легкая экскавация стопы может принять форму тяжелого фридрейховского уродства.

Ассоциации, синтропии и дистропии, гипотеза транзитивных генетических ассоциаций (Pfaundler, Seht, 1921; Blair et al., 2013). В знаменитом руководстве по диагностике врожденных болезней (Джонс, 2011) ассоциации определяются как сочетания пороков развития неясной этиологии, встречающиеся чаще, чем можно было бы ожидать от простой случайности. С момента введения термина понимание «ассоциации» носило оттенок беспокойства и неясности, как заметили J.M. Opitz и G. Neri (2013), согласившись на два варианта в определении этого термина: случайное совпадение (простая встреча, соседство) и сочетание аномалий (тесная связь, политопный дефект области тела). В 1900-х гг. появились новые обозначения по сути тех же ассоциаций, но для наследственных болезней и врожденных пороков развития был предложен термин «множественные абарты» (abart, нем. – уродство), а для широко распространенных заболеваний многофакторной природы, встречающихся у пациентов одновременно, – «синтропии» (Pfaundler, Seht, 1921). Кроме того, обозначив «взаимную склонность, притяжение» (attraction) двух болезней термином «синтропия», на основе богатого клинического материала и десятков тысяч вскрытий умерших больных ими зафиксировано другое, противоположное синтропии патологическое состояние – «взаимное отталкивание» (repulsion), несочетаемость (несовместимость, диссоциация) – «дистропии». Вместе с тем свое название получили и промежуточные, в известной степени случайные и «нейтральные» состояния – «нейтропии». По мнению этих же исследователей, термин «синдром» тоже можно считать синтропией, поскольку он означает «избирательное родство» составляющих его признаков. Другое свойство единения патологических состояний – появление как минимум двух болезней одновременно у одного пациента (синхрония). Таким образом, синтропия, синдром, синхрония (три «с») – понятия родственные, и главным фактором, их объединяющим, является сходный патогенез – так называемый общий корень в отношении, например, атеросклероза, диабета и ожирения (Stein O., Stein Y., 1995).

В современном определении синтропия – это природно-видовое явление сочетания двух и более патологических состояний (нозологий или синдромов) у индивидуума и его ближайших родственников, неслучайное и имеющее эволюционно-генетическую основу; это часть (выборка) фенома человека, представляющая собой ландшафт взаимодействующих признаков и болезней и отражающая

Феномен сочетания болезней у человека
<ul style="list-style-type: none"> • Синтропии (син.: ассоциации, коморбидность) • Дистропии (син.: контрассоциации, обратная коморбидность) • Транзитивные генетические ассоциации (син.: коморбидность менделевских и многофакторных болезней)

Рис. 1. Классификация форм сочетания болезней у человека.

непрерывающуюся молекулярно-генетическую причинность (Пузырев, 2002; Puzyrev et al., 2010). Гены, вовлеченные в развитие синтропий, названы синтропными генами. Более точно: синтропные гены – набор функционально взаимодействующих генов, локализованных во всем пространстве генома, коррегируемых генов, вовлеченных в общий для данной синтропии метаболический путь. В случае, когда регуляторные связи приводят к взаимоисключению на клиническом уровне отдельных фенотипов (дистропии), такие гены называют дистропными в отношении соответствующих фенотипов. Есть некоторое смысловое сходство понятий «синтропные и дистропные гены» с термином «коровые гены», которые обсуждались в недавно предложенной модели омнигенной наследственности МФЗ (Boyle et al., 2017).

Наконец о гипотезе транзитивных генетических ассоциаций, которые являются еще одной формой ассоциаций из двух описанных выше – синтропии (ассоциации в общепринятом понимании и наиболее частая форма) и дистропии (диссоциация). D.R. Blair с коллегами (Blair et al., 2013) выдвинули гипотезу: статистически значимая коморбидность между МФЗ и менделевским заболеванием представляет собой тип генетической ассоциации, в которой не-менделевский фенотип картируется на генетические локусы, вызывающие это менделевское заболевание. По сути, транзитивные ассоциации являются разновидностью синтропии, но фенотип есть результат сочетания МФЗ и менделевской болезни. Таких состояний, по мнению авторов гипотезы, примерно половина из всех коморбидных болезней – 54 % (Blair et al., 2013).

Классификация вариантов сочетания болезней у человека. Общепринятой классификации ФСБ не существует. Более того, задачи систематизации, уяснения общих свойств, фиксирующих закономерные связи во всем многообразии таких сочетаний, не сформулированы. Имеющиеся попытки классификации таких патологических явлений пока разрозненны и условны, чаще они носят дескриптивный характер. Особенно это относится к клинической классификации сочетаний, обозначенных термином «коморбидность», «коморбидный пациент» (Верткин и др., 2012). Сейчас можно констатировать, что предпринимались попытки систематизировать и понятие «синтропия» (Крылов, 2000): по механизмам формирования (этиологические, патогенетические, возрастные, ятрогенные, случайные), по времени возникновения (врожденные, отсроченные, симультанные, сукцессивные) и по клиническому значению (инертная, интерференции).

Ранее нами (Пузырев, 2015) было предложено выделение следующих форм сочетаний болезней у отдельных пациентов (рис. 1). Предлагаемая систематика форм ФСБ также является описательной (дескриптивной), но в ней просматриваются и элементы сущностных классификаций, и это связано в том числе с обозначением ключевых терминов характеристики сочетаний: ассоциации и синтропии. Помимо медицинской, известно несколько предметных областей в научных исследованиях, в которых используется термин «синтропия». В.Б. Вяткин (2016) выделяет три области наук, в которых понятие «синтропия» занимает важное место, предлагая классификацию синтропий (в порядке начала их использования) на медицинскую (синтропия Пфаундлера–Зехт), биофизическую (синтропия Фантаппи–Сент-Джерди–Фуллера) и информационную (синтропия Вяткина). Приведенные дополнительно два вида синтропии, на наш взгляд, не только имеют самостоятельное значение, но и важны для сущностного понимания биологических процессов, в том числе как в общей патологии, так и в частной – патогенезе ФСБ.

Заметим, что множественность болезней у индивида – проблема давняя, привлекавшая внимание исследователей до широкого использования коморбидности. Общность механизмов развития неслучайных патологических сочетаний отражена в названиях соответствующих концепций: «сумма болезней гомеостаза» (Дильман, 1968), «болезни адаптации» (Казначеев, 1980), «болезни сердечно-сосудистого континуума» (Dzau et al., 2006), «метаболический синдром» (Reaven, 1988). Важно рассмотрение данной проблемы с генетических позиций, концепций дизиосома (Goh et al., 2007) и сетевой медицины (Barabási et al., 2011; Колчанов и др., 2013).

Обобщения по проблеме исследования ФСБ позволяют приблизиться к сущностным классификациям феномена. Это важно. Как заметил М.Д. Голубовский (2006), хорошая система – это событие в науке, концептуальное открытие, новое видение гармонии в хаосе фактов. Именно поэтому включение классификаций в герменевтический круг представляется полезным.

Фактические данные по исследованию ФСБ

Синтропии (коморбидность)

Синтропия широко распространена, встречается чаще, чем мы себе представляем. Так, 438 распространенных заболеваний, зарегистрированных в историях болезней у пациентов биобанка Великобритании (UK Biobank; <https://www.ukbiobank.ac.uk/>), образуют более 11 тыс. возможных сочетаний (Dong et al., 2021). Глобальность проблемы инициировала огромное число исследований преимущественно эпидемиологического характера. Только в 2021 г. по запросу ‘comorbidity’ обнаружено 34 185 статей медицинского и биологического профиля в базе данных Национального центра биотехнологической информации США (<https://pubmed.ncbi.nlm.nih.gov/>). В настоящее время более 50 млн человек в возрасте 65 лет и старше, что составляет практически половину населения Европы, страдает одновременно двумя и более заболеваниями (Rijken et al., 2018). Прогнозируется неуклонный рост комор-

бидных пациентов, затрагивающий до 68 % населения, в последующие 20 лет (Kingston et al., 2018).

Молекулярные причины фенотипических связей в основном неизвестны, несмотря на высокую активность исследований в этом направлении (Reynolds et al., 2021; Jia et al., 2022; Quick et al., 2022; Shnyder et al., 2022; Wang et al., 2022). Благодаря этим работам стало понятно, что существенная доля (46 %) сопутствующих состояний обусловлена общим компонентом на уровне генов, SNP, взаимодействий в генных сетях (Dong et al., 2021), отражающим их патогенетически родственный характер. Например, гены *HLA-DQB1*, *TLR1*, *WDR36*, *LRRC32*, *IL1RL1*, *GSDMA*, *TSLP*, *IL33*, *SMAD3*, участвующие в патогенезе отдельных аллергических болезней, являются критическими для фенотипа, развивающегося по сценарию «атопического марша» (Ferreira et al., 2014). Между тем с точки зрения патогенеза выявляются неочевидные на первый взгляд связи между болезнями, существование которых ранее и не предполагали. Варикозная болезнь, согласно оценке генетических корреляций, связана с такими признаками, как интеллект, память, уровень образования (Shadrina et al., 2019), тогда как аутизм положительно коррелирует с аллергическим ринитом и аутоиммунными заболеваниями (Rzhetsky et al., 2007). Существенным дополнением к выявлению общих генов для сопутствующих заболеваний стало изучение биологических процессов, в которые вовлечены эти гены (Rubio-Perez et al., 2017). Применение такого рода подходов предоставляет более полную картину взаимосвязей заболеваний и общих патогенетических путей, знания о которых можно широко использовать, в том числе для лечения коморбидных пациентов.

С использованием наших данных исследования генетической компоненты аллергических болезней (Фрейдлин и др., 2015), с одной стороны, установлена молекулярная связь большинства аллергических заболеваний, а с другой стороны, в пространстве их молекулярных взаимоотношений с другими болезнями отмечается их близость к инфекционным заболеваниям и выраженная дистанционность от аутоиммунных болезней (рис. 2).

Наибольший вклад в коморбидность бронхиальной астмы и гипертонической болезни могут вносить гены *TLR4*, *CAT*, *ANG/RNASE4*, свидетельствуя о важности воспаления, процессов неоваскуляризации и окислительного стресса для патогенеза обоих заболеваний (Bragina et al., 2018). Развитие фенотипов бронхиальной астмы в сочетании с сердечно-сосудистыми/метаболическими нарушениями ассоциировано с отдельными генетическими вариантами, влияющими на экспрессию генов, в том числе *CAT*, *TLR4*, *ELF5*, *ABTB2*, *UTP25*, *TRAF3IP3*, *NFKB1*, *LOC105377347*, *C1orf74*, *IRF6* и др., в органах-мишенях исследуемого профиля заболеваний (Брагина и др., 2022).

Синтропные гены участвуют в патогенезе посредством сложных взаимодействий с другими генами, белками, факторами окружающей среды, которые в совокупности влияют на клинические проявления сопутствующих заболеваний. В большинстве случаев нарушения в синтропных генах локализованы преимущественно в некодирующих РНК и межгенных областях, функционально связанных с регуляцией транскрипции генов (Dong et al., 2021). В свою очередь транскрипция синтропных генов зависима от эпигенетических механизмов, в частности метилиро-

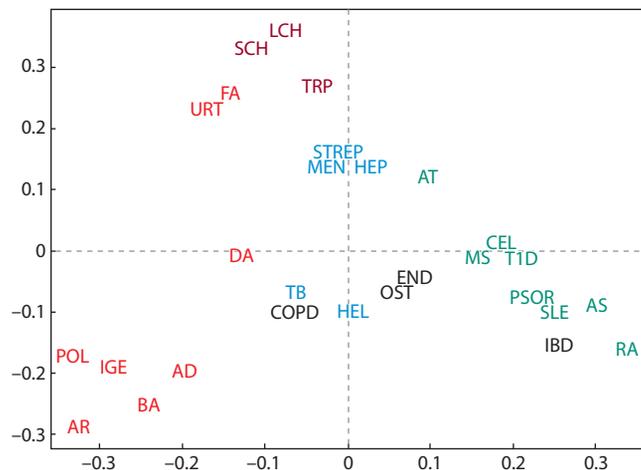


Рис. 2. Результаты многомерного шкалирования многофакторных заболеваний на основании общности ассоциированных с ними генов (использована иллюстрация из работы (Фрейдлин и др., 2015)).

AD – атопический дерматит; AR – аллергический ринит; AS – анкилозирующей спондилит; AT – аутоиммунный тиреоидит; BA – бронхиальная астма; CEL – целиакия; COPD – хроническая обструктивная болезнь легких; DA – лекарственная аллергия; END – эндометриоз; FA – пищевая аллергия; HEL – хеликобактерная инфекция; HEP – вирусный гепатит; IBD – воспалительное заболевание кишечника; IGE – уровень иммуноглобулина E; LCH – лейшманиоз; MEN – менингококковая инфекция; MS – рассеянный склероз; OST – остеопороз; POL – поллиноз; PSOR – псориаз; RA – ревматоидный артрит; SCH – шистосомоз; SLE – системная красная волчанка; STREP – стрептококковая инфекция; T1D – сахарный диабет 1-го типа; TB – туберкулез; TRP – трипаносомоз; URT – крапивница.

вания ДНК (Ferreira et al., 2017), что свидетельствует о модифицирующей роли внешнесредового воздействия на развитие комплексного фенотипа.

Многие синтропные гены являются известными лекарственными мишенями для терапии, в частности, аллергических (*FLG*, *IL13*, *IL1RL1*, *IL6R*, *INPP5D*, *NDFIP1*, *PTGER4*, *TSLP*, *STAT6*) (Ferreira et al., 2017), бронхолегочных и сердечно-сосудистых (*EDNRA*, *ADRB1*, *ADRB2*) заболеваний (Zolotareva et al., 2019; Dong et al., 2021). Более восьми тысяч лекарств нацелены на гены-мишени, действующие в развитии коморбидных состояний (Dong et al., 2021). Теоретически такие результаты не только подчеркивают важный вклад генов для фенотипических корреляций, но и предоставляют возможность перепрофилирования лекарств, нацеленных на общие генетические компоненты синтропных болезней.

Дистропии («диаметральные болезни»)

Контраст для синтропии составляют болезни, которые проявляются фенотипической конкуренцией одного патологического состояния относительно другого (дистропия). Дистропия затрагивает болезни разных категорий, включая иммунологические, онкологические, нейродегенеративные, сердечно-сосудистые, аутоиммунные и др. Спектр молекулярных механизмов, лежащих в основе данного феномена, представляется также весьма разнообразным. Исследования в отношении дистропии сфокусированы на поиске молекулярно-генетических различий между заболеваниями. В результате установлены различия в транскрипции одних и тех же генов при разных заболеваниях. На примере дистропии онкологических и

нейродегенеративных заболеваний (Catalá-López et al., 2014) выявлено, что дифференциально экспрессирующиеся гены преимущественно связаны с репарацией ДНК, функционированием митохондрий, стабилизацией p53, контролируют ангиогенез, клеточный цикл, транспорт ионов металлов, глюкозы, регуляцию апоптотических процессов, активируют миелоидные лейкоциты и фагоцитоз, вовлечены в передачу сигналов mTORC1 и KRAS (Forés-Martos et al., 2021; Pepe et al., 2021). Транскрипционные изменения неоднозначны для канцерогенеза; некоторые гены могут быть активированы при одних формах рака, но подавлены при других, что, вероятно, связано с особенностями сложных генетических и эпигенетических нарушений (Zhao et al., 2016). В то же время регистрируются общие закономерности. В частности, в работе (Ibáñez et al., 2014) идентифицированы гены *MT2A*, *MT1X*, *NFKB1A*, *AC009469.1*, *DHRS3*, *CDKN1A*, *TNFRSF1A*, *CRYBG3*, *IL4R*, *MT1M*, *FAM107A*, *ITPKC*, *MID1*, *IL11RA*, *AHNAK*, *KAT2B*, *BCL2*, *PTH1R*, *NFASC*, которые одновременно активируются при нескольких расстройствах ЦНС (болезнь Альцгеймера, болезнь Паркинсона, шизофрения) и подавляются при онкологических заболеваниях.

Вышеприведенные примеры свидетельствуют о том, что фенотипическая супрессия опосредована генетическими факторами. Потенциально «вредные» аллели могут приносить пользу, создавая некоторого рода компромисс повышенного риска развития одних болезней и низкого риска развития других. Компромиссы неизбежны, потому что вследствие сложной интегрированной работы целого организма несколько взаимодействующих частей должны работать вместе для выполнения определенных функций. Такая интеграция может привести к дилемме, которую часто называют «ценой сложности» (Wagner et al., 2008), возникающей в результате работы множества взаимодействующих частей, кооперирующихся для успешного выполнения функции. Изменение любой одной части неизбежно негативно повлияет на другие признаки, изменяя функцию и снижая общую производительность или приспособленность. Таким образом, механистическая основа компромиссов может быть сосредоточена в плейотропных генах, вовлеченных в общие биологические пути для разных признаков (Mauro, Ghalambor, 2020). В соответствии с этим предположением вполне закономерно можно считать наблюдаемые различия в транскрипции важных для дистропии генов, что демонстрирует двунаправленный характер биологических процессов, благодаря которому экспрессия или активация может отклоняться от некоторого оптимального значения (Crespi, Go, 2015).

Существенным образом формирует дистропию лекарственная терапия, поскольку лекарства могут быть связаны с регуляцией общих молекулярных процессов фенотипически полярных заболеваний. Например, использование антихолинэстеразного средства галантамина и избирательного ингибитора моноаминоксидазы селегилина при нейродегенеративных заболеваниях оказывает противораковый эффект (Lazarevic-Pasti et al., 2017; Ryu et al., 2018). Два препарата для терапии рака молочной железы (экземестан и эстрадиол) снижают риск болезни Альцгеймера и других деменций (Branigan et al., 2020; Guglielmotto et al., 2020).

Транзитивные генетические ассоциации

Гены, мутационные нарушения в которых лежат в основе редких и высокопенетрантных менделевских болезней, затрагивают развитие более распространенных форм заболеваний. Эффект мутаций может быть как предполагающим к развитию заболевания фактором, так и, наоборот, проявляющим супрессирующие свойства в отношении фенотипов. Приводятся разные оценки участия генов менделевских заболеваний в фенотипической экспансии многофакторной патологии. Около 300 генов, ассоциированных в полногеномных исследованиях с распространенными болезнями, лежат в основе ряда менделевских заболеваний (Lupski et al., 2011). По некоторым оценкам, доля менделевских генов в структуре многофакторных болезней составляет приблизительно 23 % (Spataro et al., 2017), однако с ростом данных полногеномного секвенирования это значение, вероятно, заметно возрастет. Если говорить о конкретной патологии, то из 30 генов, ассоциированных с уровнем липидов в сыворотке крови, одиннадцать (*ABCG8*, *LCAT*, *APOB*, *APOE*, *LDLR*, *PCSK9*, *CETP*, *LPL*, *LIPC*, *APOA5* и *ABCA1*) вовлечены в моногенные нарушения метаболизма липидов (Kathiresan et al., 2009). Эти гены, являющиеся причинными вариантами как менделевских расстройств, так и риска многофакторных заболеваний, как правило, обладают более высокой функциональной значимостью и более высокими уровнями экспрессии, чем гены, связанные только с распространенными болезнями. Кроме того, генетические варианты в условно «менделевских» генах являются в большей степени рисковыми для многофакторных болезней, чем варианты в генах, не связанных с менделевскими заболеваниями (Spataro et al., 2017).

Идея реализации мутационного бремени на распространенную патологию не нова. Экспериментальной основой этого феномена стала публикация Майкла Брауна и Джозефа Гольдштейна (Brown, Goldstein, 1986), показавших, что у пациентов с гетерозиготными мутациями в гене рецептора липопротеинов низкой плотности (*LDLR*) наряду с семейной гиперхолестеринемией отмечается коронарный атеросклероз и инфаркт миокарда. В 2013 г. David R. Blair (Blair et al., 2013) сформулировал гипотезу о транзитивности редких менделевских вариантов в патологический «аллельный континуум» в широком диапазоне конечных фенотипических эффектов от моногенных до сложных многофакторных болезней. На сегодняшний день накоплен обширный фактический материал в поддержку этой гипотезы. У носителей мутаций гена *FLG*, связанных с потерей функции филагрина, возрастает риск развития атопического дерматита (Sandilands et al., 2007), бронхиальной астмы в контексте атопического дерматита, в то же время снижается риск астмы без атопического дерматита (Palmer et al., 2006). Это позволяет заключить, что мутации гена *FLG* являются важным фактором риска реализации атопии в целом, но с разными шансами для конкретного фенотипа. Носители мутаций болезни Гоше, преимущественно L444P и N370S в гене глюкоцереброзидазы (*GBA*), подвержены повышенному риску болезни Паркинсона (Sidransky et al., 2009). Гетерозиготные носители мутаций в гене трансмембранного регулятора муковисцидоза (*CFTR*) предрасположены к

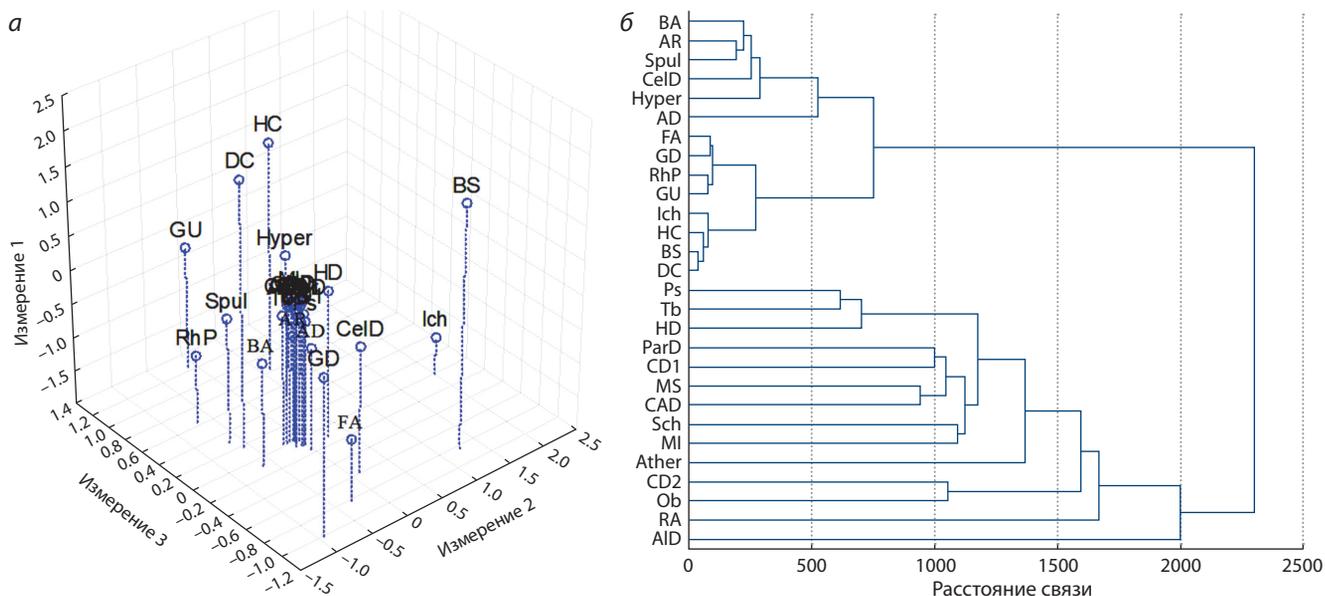


Рис. 3. Моделирование отношений между многофакторными/моногенными болезнями по общности ассоциированных с ними генов, по результатам многомерного шкалирования (а) и иерархического кластерного анализа (б).

AD – экзема (атопический дерматит); AID – болезнь Альцгеймера; AR – аллергический ринит; Ather – атеросклероз; BA – бронхиальная астма аллергическая; BS – синдром Бругада; CAD – ишемическая болезнь сердца; CD1 – сахарный диабет 1-го типа; CD2 – сахарный диабет 2-го типа; CelD – целиакия; DC – дилатационная кардиомиопатия; FA – пищевая аллергия; GD – болезнь Гоше; GU – язвенная болезнь желудка; HC – гипертрофическая кардиомиопатия; HD – болезнь Гентингтона; Hyper – артериальная гипертензия; Ich – икhtiоз; MI – инфаркт миокарда; MS – рассеянный склероз; Ob – ожирение; ParD – болезнь Паркинсона; Ps – псориаз; RA – ревматоидный артрит; RhP – полипозный синусит; Sch – шизофрения; Spul – саркоидоз; Tb – туберкулез.

идиопатическому панкреатиту (Weiss et al., 2005) и хронической обструктивной болезни легких (Divac et al., 2004).

Различные подходы используются для получения знаний об участии генов менделевских болезней в качестве причинных генов многофакторных заболеваний. Например, базируясь на приоритизации данных полногеномных ассоциативных исследований разных форм кардиомиопатий, установлено, что 70 % генов гипертрофической и 56 % генов дилатационной кардиомиопатий связаны с различными менделевскими заболеваниями. Это наводит на мысль, что существующая дихотомическая классификация заболеваний (моногенные и многофакторные) стала неактуальной и требуется переосмысление с учетом новых знаний о генетической структуре подверженности (Назаренко и др., 2022).

Потенциал отдельных мутаций генов оценивается в качестве протективных факторов в отношении онкологических заболеваний. В частности, активация апоптоза и аутофагии мутантным гентинггином (Гомбоева и др., 2020), а также онкотоксичная функция CAG-повторов (Mittmann et al., 2018), экспансия которых является причиной развития болезни Гентингтона, могут предотвращать развитие большинства видов рака у пациентов с этим наследственным заболеванием (Catalá-López et al., 2014). Молекулярный механизм онкопротекции мутации карликовости Ларона (OMIM #262500) (NM_000163.5(GHR):c.594A>G (p.Glu198=)) в гене рецептора гормона роста опосредован влиянием на активность генов, участвующих в контроле клеточного цикла, подвижности, роста и онкогенной трансформации (Werner et al., 2020).

Потеря функции отдельных белков вследствие мутаций (loss-of-function mutations) обеспечивает специфическую

резистентность в отношении некоторых распространенных фенотипов. Защита от сахарного диабета 2-го типа связана с носительством мутации в гене трансмембранного переносчика цинка типа 8 (*SLC30A8*), приводящей к синтезу укороченного белка (Flannick et al., 2014). В результате возникающего дефицита функции гена *SLC30A8* по механизму гаплонедостаточности у носителей мутантных аллелей наблюдается лучшая секреция инсулина из-за повышенной чувствительности к глюкозе и конверсии проинсулина в β -клетках поджелудочной железы. Другой пример относится к нонсенс-мутациям (Y142X, C679X и R46L), которые приводят к снижению холестерина липопротеинов низкой плотности, в гене пропротеиновой конвертазы субтилизин-кексинового типа 9 (*PCSK9*), лежащего в основе семейной гиперхолестеринемии (OMIM #603776) (Cohen et al., 2005). Гетерозиготные носители делеции F508del в гене трансмембранного регуляторного белка муковисцидоза (*CFTR*), являющейся причиной муковисцидоза, более устойчивы к инфекционным заболеваниям – холере, брюшному тифу и туберкулезу, поэтому некоторые авторы связывают высокую распространенность муковисцидоза в современной популяции человека с адаптивным преимуществом носителей мутации (Bosch et al., 2017).

Результаты классификации некоторых многофакторных и менделевских заболеваний, основанные на связанных с ними генах, позволили идентифицировать большую общую генетическую компоненту многофакторных болезней, о чем свидетельствует их ориентированность ближе к центру на рис. 3, а. Моногенные заболевания закономерно дистанцированы от них, за исключением болезни Гентингтона, которая по степени общности генов не только близка

к другим нейродегенеративным заболеваниям, но и имеет молекулярное сходство с инфекционными, аутоиммунными и кардиометаболическими болезнями (см. рис. 3, б). В целом по степени генетической «общности» и кластеризации большинство исследуемых заболеваний отражает общепринятую классификацию заболеваний. Однако подобное моделирование имеет ограничение, поскольку зависит от изученности генов, поэтому стоит ожидать смещение расположения моногенных заболеваний. Заполним некоторые пробелы в изученности ассоциированных с заболеваниями генов (причем в настоящее время речь идет о стремительно пополняющихся массивах геномной информации), остается приступить к более сложному этапу исследований, заключающемуся в понимании механизмов проявления эффекта мутаций, которые комбинируются определенным образом (Diss, Lehner, 2018), способствуя фенотипическому разнообразию.

Заключение

Последние десятилетия для геномных исследований стали важным этапом развития благодаря возможностям высокопроизводительных технологий и колоссальному количеству полученных данных. Ожидается, что к 2025 г. может быть секвенировано от 100 миллионов до 2 миллиардов человеческих геномов, что намного превышает рост в других динамично развивающихся областях, оперирующих большими данными, включая, например, астрономию, YouTube и Twitter (Stephens et al., 2015). Авторы упомянутой работы сравнивают геномные исследования с «четырёхглавым монстром», основываясь на четырех главных потребностях в геномике на протяжении всего жизненного цикла, получаемых в ходе секвенирования больших данных: сбор, хранение, распространение и анализ. Наибольших усилий из этих четырех потребностей требуют анализ и осмысление полученных результатов, распутывание сложной связи между генетическими вариантами и фенотипами, представляющей собой преимущественно стохастический процесс, с одной стороны ограниченный геномом, с другой – факторами окружающей среды. Соответственно рациональные способы постижения сложных с точки зрения биологии объектов в мире больших данных по-прежнему остаются актуальными.

Накапливающиеся в научной литературе результаты исследования феномена сочетания болезней (коморбидность, синтропия/дистропия) подводят к необходимости и возможности приблизиться к такому видению обобщения, которое обозначил выдающийся Карл Вёзе в своей работе: «...предназначение биологии не в доскональном описании объектов как моментальных снимков, а в философско-эволюционном понимании феномена их существования» (Woese, Goldenfeld, 2009). В этом контексте в нашей статье рассмотрен феномен сочетания болезней в рамках метафоры «герменевтического круга». Важно отметить историческую преемственность научных знаний по данному вопросу, в истоке которых изначально лежало холистическое представление о развитии живых организмов, начиная от жоффруизма, отраженного в принципах коннексии, единства элементарности и целостности (Холлоковский, 1915), до проявления сложного тропизма наследственных факторов (Давиденков, 1947) и принципов

систематизации в медицинской генетике (McKusick, 1968) и, наконец, организации структурно-функциональных генных сетей в рамках современных концепций сетевой биологии и медицины (Barabási et al., 2011; Колчанов и др., 2013).

Необходимо отметить недостаточно исчерпывающий характер существующей терминологии относительно сопутствующих болезней. Например, в отличие от термина «коморбидность», ставшего привычным в медицинской практике, генетический дискурс близости сопутствующих заболеваний наиболее полно истолковывается терминами «синтропия» и «дистропия», отражающими особенности патогенетических связей между болезнями. Патогенетический принцип вовлеченности генов в развитие сопутствующих болезней позволил классифицировать их как синтропные и дистропные гены (Пузырёв, 2015). В этом контексте важна классификация генов на механистической основе на ядерные/коровые (*англ.* core) и периферические гены, чьи омнигенные эффекты на развитие патологического фенотипа осуществляется через транс- и цис-регуляцию (Boyle et al., 2017; Liu et al., 2019). Очевидно, что, наряду с ядерными, периферические гены являются важными объектами для изучения коморбидности МФЗ, поскольку их глобальная активность в конкретных типах клеток определяет клеточную функцию и риск заболевания.

Молекулярная природа сопутствующих заболеваний, которая позволяет им быть связанными вместе во многих, нередко не смертельных и даже полезных сочетаниях, остается сложной для объяснения в силу некоторой «вольности генома», определяемой динамическим и нелинейным характером функционирования системы, регулируемой обратными связями, которые могут нарушаться предсказуемым, но индивидуальным образом. Степень пользы или вреда таких комбинаций болезней условного «адаптивного фенотипа» зависит от компромиссов, которые наиболее очевидны вследствие конкуренции за ограниченные ресурсы организма. Вероятно, уязвимость индивида для одних заболеваний относительно небольшого риска развития других сводится к определению некоторой «цены сложности», основой которой является плейотропное действие генов.

С одной стороны, давно описанный в клинической практике ФСБ представляет самостоятельный интерес для исследований фундаментального характера, становится дополнительным путем выяснения этиологии и патогенеза сложных заболеваний, к изучению которого привлекаются современные методологические и концептуальные подходы. С другой стороны, это важно для практического здравоохранения, поскольку описание феномена сопутствующих болезней имеет решающее значение для расширения интерпретационного горизонта клинициста и выхода за пределы узких, ориентированных на конкретную болезнь терапевтических решений. Расширяя наши знания о молекулярном разнообразии фенома человека, можно стимулировать пересмотр текущих классификаций болезней (Piro, 2012), выделение в таких классификациях подтипов с различным прогнозом для пациента и членов семьи, индивидуальных ответов на лечение (Manolio, 2013).

Список литературы / References

- Брагина Е.Ю., Гончарова И.А., Жалсанова И.Ж., Немеров Е.В., Назаренко М.С., Фрейдин М.Б., Пузырев В.П. Генетическая коморбидность гипертонической болезни и бронхиальной астмы. *Артериальная гипертензия*. 2022;28(1):87-95. DOI 10.18705/1607-419X-2022-28-1-87-95.
- [Bragina E.Yu., Goncharova I.A., Zhalsanova I.Z., Nemerov E.V., Nazarenko M.S., Freidin M.B., Puzyrev V.P. Genetic comorbidity of hypertension and bronchial asthma. *Arterial'naya Gipertenziya = Arterial Hypertension*. 2022;28(1):87-95. DOI 10.18705/1607-419X-2022-28-1-87-95. (in Russian)]
- Верткин А.Л. Коморбидный пациент. М., 2015.
- [Vertkin A.L. Comorbid Patient. Moscow, 2015. (in Russian)]
- Верткин А.Л., Румянцев М.А., Скотников А.С. Коморбидность. *Клин. медицина*. 2012;90(10):4-11.
- [Vertkin A.L., Rumyantsev M.A., Skotnikov A.S. Comorbidity. *Klinicheskaya Meditsyna = Clinical Medicine*. 2012;90(10):4-11. (in Russian)]
- Виленкин А. Мир многих миров. Физики в поисках параллельных вселенных. М.: Астрель, 2010.
- [Vilenkin A. Many Worlds in One. The Search for Other Universes. Moscow: Astrel Publ., 2010. (in Russian)]
- Вяткин В.Б. Об использовании термина «синтропия» в научных исследованиях. *Науч. обозрение. Реферат. журн.* 2016;3:81-84.
- [Vyatkin V.B. About application of the term "syntropy" in scientific research. *Nauchnoye Obzreniye. Referativnyy Zhurnal = Scientific Review. Abstract Journal*. 2016;3:81-84. (in Russian)]
- Гадамер Г.-Г. О круге понимания. Актуальность прекрасного. М., 1991.
- [Gadamer G.-G. On the Circle of Understanding. The Relevance of Beauty. Moscow, 2010. (in Russian)]
- Голубовский М.Д. Комментарий к «Диалог о систематике». Надежда Манделштам и Любимцев. *Природа*. 2006;6:77-80.
- [Golubovsky M.D. Commentaries on the Dialogue on Systematics. Nadezhda Mandelstam and Lyubishchev. *Priroda = Nature*. 2006;6:77-80. (in Russian)]
- Гомбоева Д.Е., Брагина Е.Ю., Назаренко М.С., Пузырев В.П. Обратная коморбидность между онкологическими заболеваниями и болезнью Гентингтона: обзор эпидемиологических и биологических доказательств. *Генетика*. 2020;56(3):260-271. DOI 10.31857/S0016675820030054.
- [Gombueva D.E., Bragina E.Y., Nazarenko M.S., Puzyrev V.P. The inverse comorbidity between oncological diseases and Huntington's disease: review of epidemiological and biological evidence. *Russ. J. Genet.* 2020;56(3):269-279. DOI 10.1134/S1022795420030059.]
- Давиденков С.Н. Эволюционно-генетические проблемы в невропатологии. Л.: ГИДУВ, 1947.
- [Davidenkov S.N. Evolutionary Genetic Problems in Neuropathology. Leningrad: GIDUV Publ., 1947. (in Russian)]
- Джонс К.Л. Наследственные синдромы по Дэвиду Смит. Атлас-справочник. М., 2011.
- [Jones K.L. Hereditary Syndromes According to David Smith. Atlas-reference book. Moscow: Praktika Publ., 2011. (in Russian)]
- Дильман В.М. Старение, климакс, рак. М., 1968.
- [Dilman V.M. Aging, Menopause, Cancer. Moscow, 1968. (in Russian)]
- Казначеев В.П. Современные аспекты адаптации. Новосибирск, 1980.
- [Kaznacheev V.P. Modern Aspects of Adaptation. Novosibirsk, 1980. (in Russian)]
- Колчанов Н.А., Игнатьева Е.В., Подколотная О.А., Лиховшвай В.А., Матушкин Ю.Г. Генные сети. *Вавиловский журнал генетики и селекции*. 2013;17(4-2):833-850.
- [Kolchanov N.A., Ignatieva E.V., Podkolodnaya O.A., Lihovshvai V.A. Gene networks. *Vavilovskii Zhurnal Genetiki i Selekcii = Vavilov Journal of Genetics and Breeding*. 2013;17(4-2):833-850. (in Russian)]
- Крылов А.А. К проблеме сочетаемости заболеваний. *Клин. медицина*. 2000;78(1):56-58.
- [Krylov A.A. To the problem of compatibility of diseases. *Klinicheskaya Meditsyna = Clinical Medicine*. 2000;78(1):56-58. (in Russian)]
- Назаренко М.С., Слепцов А.А., Пузырев В.П. «Менделевский код» в генетической структуре широко распространенных многофакторных заболеваний. *Генетика*. 2022;58(10):1101-1111. DOI 10.31857/S0016675822100058.
- [Nazarenko M.S., Slepcev A.A., Puzyrev V.P. "Mendelian code" in the genetic structure of complex diseases. *Genetics*. 2022;58(10):1101-1111. DOI 10.31857/S0016675822100058. (in Russian)]
- Пузырев В.П. Вольности генома и медицинская патогенетика. *Бюл. сиб. медицины*. 2002;1(2):16-29. DOI 10.20538/1682-0363-2002-2-16-29.
- [Puzyryov V.P. Liberties of genome and medical pathogenetics. *Byulleten Sibirskoy Meditsiny = Bulletin of Siberian Medicine*. 2002;1(2):16-29. DOI 10.20538/1682-0363-2002-2-16-29. (in Russian)]
- Пузырев В.П. Генетические основы коморбидности у человека. *Генетика*. 2015;51(4):491-502. DOI 10.7868/S0016675815040098.
- [Puzyrev V.P. Genetic bases of human comorbidity. *Russ. J. Genet.* 2015;51(4):408-417. DOI 10.1134/S1022795415040092.]
- Серебровский А.С. Некоторые проблемы органической эволюции. М., 1973.
- [Serebrovskiy A.S. Some Problems of Organic Evolution. Moscow, 1973. (in Russian)]
- Фрейдин М.Б., Огородова Л.М., Пузырев В.П. Патогенетика аллергических болезней. Новосибирск, 2015.
- [Freydin M.B., Ogorodova L.M., Puzyrev V.P. Pathogenetics of Allergic Diseases. Novosibirsk, 2015. (in Russian)]
- Холодковский Н.А. Ламаркизм и жоффруизм. *Природа*. 1915;4:533-542.
- [Holodkovsky N.A. Lamarckism and Geofreyism. *Priroda = Nature*. 1915;4:533-542. (in Russian)]
- Azaïs B., Bowis J., Wismar M. Facing the challenge of multimorbidity. *J. Comorb.* 2016;6(1):1-3. DOI 10.15256/joc.2016.6.71.
- Barabási A.L., Gulbahce N., Loscalzo J. Network medicine: a network-based approach to human disease. *Nat. Rev. Genet.* 2011;12(1):56-68. DOI 10.1038/nrg2918.
- Biesecker L.G. Lumping and splitting: molecular biology in the genetics clinic. *Clin. Genet.* 1998;53(1):3-7. DOI 10.1034/j.1399-0004.1998.531530102.x.
- Blair D.R., Lyttle C.S., Mortensen J.M., Bearden C.F., Jensen A.B., Khiabani H., Melamed R., Rabadan R., Bernstam E.V., Brunak S., Jensen L.J., Nicolae D., Shah N.H., Grossman R.L., Cox N.J., White K.P., Rzhetsky A. A nondegenerate code of deleterious variants in Mendelian loci contributes to complex disease risk. *Cell*. 2013;155(1):70-80. DOI 10.1016/j.cell.2013.08.030.
- Bosch L., Bosch B., De Boeck K., Nawrot T., Meyts I., Vanneste D., Le Bourlegat C.A., Croda J., da Silva Filho L.V.R.F. Cystic fibrosis carriership and tuberculosis: hints toward an evolutionary selective advantage based on data from the Brazilian territory. *BMC Infect. Dis.* 2017;17(1):340. DOI 10.1186/s12879-017-2448-z.
- Boyle E.A., Li Y.I., Pritchard J.K. An expanded view of complex traits: from polygenic to omnigenic. *Cell*. 2017;169(7):1177-1186. DOI 10.1016/j.cell.2017.05.038.
- Bragina E.Y., Goncharova I.A., Garaeva A.F., Nemerov E.V., Babovskaya A.A., Karpov A.B., Semenova Y.V., Zhalsanova I.Z., Gombueva D.E., Saik O.V., Zolotareva O.I., Ivanisenko V.A., Dosenko V.E., Hofstaedt R., Freidin M.B. Molecular relationships between bronchial asthma and hypertension as comorbid diseases. *J. Integr. Bioinform.* 2018;15(4):20180052. DOI 10.1515/jib-2018-0052.
- Branigan G.L., Soto M., Neumayer L., Rodgers K., Brinton R.D. Association between hormone-modulating breast cancer therapies and incidence of neurodegenerative outcomes for women with breast cancer. *JAMA Netw. Open*. 2020;3(3):e201541. DOI 10.1001/jamanetworkopen.2020.1541.
- Brown M.S., Goldstein J.L. A receptor-mediated pathway for cholesterol homeostasis. *Science*. 1986;232(4746):34-47. DOI 10.1126/science.3513311.

- Brunner H.G., van Driel M.A. From syndrome families to functional genomics. *Nat. Rev. Genet.* 2004;5(7):545-551. DOI 10.1038/nrg1383.
- Catalá-López F., Suárez-Pinilla M., Suárez-Pinilla P., Valderas J.M., Gómez-Beneyto M., Martínez S., Balanzá-Martínez V., Climent J., Valencia A., McGrath J., Crespo-Facorro B., Sanchez-Moreno J., Vieta E., Tabarés-Seisdedos R. Inverse and direct cancer comorbidity in people with central nervous system disorders: a meta-analysis of cancer incidence in 577,013 participants of 50 observational studies. *Psychother. Psychosom.* 2014;83(2):89-105. DOI 10.1159/000356498.
- Cohen J., Pertsemlidis A., Kotowski I.K., Graham R., Garcia C.K., Hobbs H.H. Low LDL cholesterol in individuals of African descent resulting from frequent nonsense mutations in *PCSK9*. *Nat. Genet.* 2005;37(2):161-165. DOI 10.1038/ng1509.
- Crespi B.J., Go M.C. Diametrical diseases reflect evolutionary-genetic tradeoffs: evidence from psychiatry, neurology, rheumatology, oncology and immunology. *Evol. Med. Public. Health.* 2015;2015(1):216-253. DOI 10.1093/emph/eov021.
- Diss G., Lehner B. The genetic landscape of a physical interaction. *eLife.* 2018;7:e32472. DOI 10.7554/eLife.32472.
- Divac A., Nikolic A., Mitic-Milicic M., Nagorni-Obradovic L., Petrovic-Stanojevic N., Dopudja-Pantic V., Nadaskic R., Savic A., Radjokovic D. High frequency of the R75Q CFTR variation in patients with chronic obstructive pulmonary disease. *J. Cyst. Fibros.* 2004;3(3):189-191. DOI 10.1016/j.jcf.2004.05.049.
- Dong G., Feng J., Sun F., Chen J., Zhao X.M. A global overview of genetically interpretable multimorbidities among common diseases in the UK Biobank. *Genome Med.* 2021;13(1):110. DOI 10.1186/s13073-021-00927-6.
- Dzau V.J., Antman E.M., Black H.R., Hayes D.L., Manson J.E., Plutzky J., Popma J.J., Stevenson W. The cardiovascular disease continuum validated: clinical evidence of improved patient outcomes. Part I: Pathophysiology and clinical trial evidence (risk factors through stable coronary artery disease). *Circulation.* 2006;114(25):2850-2870. DOI 10.1161/CIRCULATIONAHA.106.655688.
- Feinstein A.R. The pre-therapeutic classification of co-morbidity in chronic disease. *J. Chronic Dis.* 1970;23(7):455-468. DOI 10.1016/0021-9681(70)90054-8.
- Ferreira M.A., Matheson M.C., Tang C.S., Granell R., Ang W., Hui J., Kiefer A.K., Duffy D.L., Baltic S., Danoy P., Bui M., Price L., Sly P.D., Eriksson N., Madden P.A., Abramson M.J., Holt P.G., Heath A.C., Hunter M., Musk B., Robertson C.F., Le Souëf P., Montgomery G.W., Henderson A.J., Tung J.Y., Dharmage S.C., Brown M.A., James A., Thompson P.J., Pennell C., Martin N.G., Evans D.M., Hinds D.A., Hopper J.L., Australian Asthma Genetics Consortium Collaborators. Genome-wide association analysis identifies 11 risk variants associated with the asthma with hay fever phenotype. *J. Allergy Clin. Immunol.* 2014;133(6):1564-1571. DOI 10.1016/j.jaci.2013.10.030.
- Ferreira M.A., Vonk J.M., Baurecht H., Marenholz I., Tian C., Hoffman J.D., Helmer Q., Tillander A., Ullemer V., van Dongen J., ... Jorgensen E., Lee Y.A., Boomsma D.I., Almqvist C., Karlsson R., Koppelman G.H., Paternoster L. Shared genetic origin of asthma, hay fever and eczema elucidates allergic disease biology. *Nat. Genet.* 2017;49(12):1752-1757. DOI 10.1038/ng.3985.
- Flannick J., Thorleifsson G., Beer N.L., Jacobs S.B., Grarup N., Burt N.P., Mahajan A., Fuchsberger C., Atzmon G., Benediktsson R., ... Pedersen O., Go-T2D Consortium, T2D-GENES Consortium, Groop L., Cox D.R., Stefansson K., Altshuler D. Loss-of-function mutations in *SLC30A8* protect against type 2 diabetes. *Nat. Genet.* 2014;46(4):357-363. DOI 10.1038/ng.2915.
- Forés-Martos J., Boullosa C., Rodrigo-Domínguez D., Sánchez-Valle J., Suay-García B., Climent J., Falcó A., Valencia A., Puig-Butillé J.A., Puig S., Tabarés-Seisdedos R. Transcriptomic and genetic associations between Alzheimer's disease, Parkinson's disease, and cancer. *Cancers (Basel).* 2021;13(12):2990. DOI 10.3390/cancers13122990.
- Goh K.I., Cusick M.E., Valle D., Childs B., Vidal M., Barabási A.L. The human disease network. *Proc. Natl. Acad. Sci. USA.* 2007;104(21):8685-8690. DOI 10.1073/pnas.0701361104.
- Guglielmotto M., Manassero G., Vasciaveo V., Venezia M., Tabaton M., Tamagno E. Estrogens inhibit amyloid- β -mediated paired helical filament-like conformation of tau through antioxidant activity and miRNA 218 regulation in hTau mice. *J. Alzheimers Dis.* 2020;77(3):1339-1351. DOI 10.3233/JAD-200707.
- Ibáñez K., Boullosa C., Tabarés-Seisdedos R., Baudot A., Valencia A. Molecular evidence for the inverse comorbidity between central nervous system disorders and cancers detected by transcriptomic meta-analyses. *PLoS Genet.* 2014;10(2):e1004173. DOI 10.1371/journal.pgen.1004173.
- Jia G., Zhong X., Im H.K., Schoettler N., Pividori M., Hogarth D.K., Sperling A.L., White S.R., Naureckas E.T., Lytle C.S., Terao C., Kamatani Y., Akiyama M., Matsuda K., Kubo M., Cox N.J., Ober C., Rzhetsky A., Solway J. Discerning asthma endotypes through comorbidity mapping. *Nat. Commun.* 2022;13(1):6712. DOI 10.1038/s41467-022-33628-8.
- Kathiresan S., Willer C.J., Peloso G.M., Demissie S., Musunuru K., Schadt E.E., Kaplan L., Bennett D., Li Y., Tanaka T., ... Peltonen L., Orho-Melander M., Ordovas J.M., Boehnke M., Abecasis G.R., Mohlke K.L., Cupples L.A. Common variants at 30 loci contribute to polygenic dyslipidemia. *Nat. Genet.* 2009;41(1):56-65. DOI 10.1038/ng.291.
- Kingston A., Robinson L., Booth H., Knapp M., Jagger C., MODEM project. Projections of multi-morbidity in the older population in England to 2035: estimates from the Population Ageing and Care Simulation (PACSim) model. *Age Ageing.* 2018;47(3):374-380. DOI 10.1093/ageing/afx201.
- Lazarevic-Pasti T., Leskovic A., Momic T., Petrovic S., Vasic V. Modulators of acetylcholinesterase activity: from Alzheimer's disease to anti-cancer drugs. *Curr. Med. Chem.* 2017;24(30):3283-3309. DOI 10.2174/0929867324666170705123509.
- Liu X., Li Y.L., Pritchard J.K. Trans effects on gene expression can drive omnigenic inheritance. *Cell.* 2019;177(4):1022-1034.e6. DOI 10.1016/j.cell.2019.04.014.
- Lupski J.R., Belmont J.W., Boerwinkle E., Gibbs R.A. Clan genomics and the complex architecture of human disease. *Cell.* 2011;147(1):32-43. DOI 10.1016/j.cell.2011.09.008.
- Manolio T.A. Bringing genome-wide association findings into clinical use. *Nat. Rev. Genet.* 2013;14(8):549-558. DOI 10.1038/nrg3523.
- Mauro A.A., Ghalambor C.K. Trade-offs, pleiotropy, and shared molecular pathways: a unified view of constraints on adaptation. *Integr. Comp. Biol.* 2020;60(2):332-347. DOI 10.1093/icb/icaa056.
- McKusick V.A. Some principles of medical genetics. In: Bartalos M. (Ed.) *Genetics in Medical Practice*. London: Pitman Medical, 1968;43-54.
- McKusick V.A. On lumpers and splitters, or the nosology of genetic disease. *Perspect. Biol. Med.* 1969;12(2):298-312. DOI 10.1353/pbm.1969.0039.
- Murmann A.E., Gao Q.Q., Putzbach W.E., Patel M., Bartom E.T., Law C.Y., Bridgeman B., Chen S., McMahon K.M., Thaxton C.S., Peter M.E. Small interfering RNAs based on huntingtin trinucleotide repeats are highly toxic to cancer cells. *EMBO Rep.* 2018;19(3):e45336. DOI 10.15252/embr.201745336.
- Navickas R., Petric V.K., Feigl A.B., Seychell M. Multimorbidity: what do we know? What should we do? *J. Comorb.* 2016;6(1):4-11. DOI 10.15256/joc.2016.6.72.
- Opitz J.M., Neri G. Historical perspective on developmental concepts and terminology. *Am. J. Med. Genet. A.* 2013;161A(11):2711-2725. DOI 10.1002/ajmg.a.36244.
- Paigen K., Eppig J.T. A mouse phenome project. *Mamm. Genome.* 2000;11(9):715-717. DOI 10.1007/s003350010152.
- Palmer C.N., Irvine A.D., Terron-Kwiatkowski A., Zhao Y., Liao H., Lee S.G., Goudie D.R., Sandilands A., Campbell L.E., Smith F.J., O'Regan G.M., Watson R.M., Cecil J.E., Bale S.J., Compton J.G., DiGiovanna J.J., Fleckman P., Lewis-Jones S., Arseculeratne G., Sergeant A., Munro C.S., El Houate B., McElreavey K., Halk-

- jaer L.B., Bisgaard H., Mukhopadhyay S., McLean W.H. Common loss-of-function variants of the epidermal barrier protein filaggrin are a major predisposing factor for atopic dermatitis. *Nat. Genet.* 2006;38(4):441-446. DOI 10.1038/ng1767.
- Pepe P., Vatrano S., Cannarella R., Calogero A.E., Marchese G., Ravo M., Frassetto F., Pepe L., Pennisi M., Romano C., Ferri R., Salemi M. A study of gene expression by RNA-seq in patients with prostate cancer and in patients with Parkinson disease: an example of inverse comorbidity. *Mol. Biol. Rep.* 2021;48(11):7627-7631. DOI 10.1007/s11033-021-06723-0.
- Pfaundler M., Seht L.V. Über Syntropie von Krankheitszuständen. *Z. Kinder-Heilk.* 1921;30:100-120. DOI 10.1007/BF02222706.
- Piro R.M. Network medicine: linking disorders. *Hum. Genet.* 2012; 131(12):1811-1820. DOI 10.1007/s00439-012-1206-y.
- Puzyrev V.P., Makeeva O.A., Freidin M.B. Syntropy, genetic testing and personalized medicine. *Per. Med.* 2010;7(4):399-405. DOI 10.2217/pme.10.35.
- Quick C.R., Conway K.P., Swendsen J., Stapp E.K., Cui L., Merikangas K.R. Comorbidity and coaggregation of major depressive disorder and bipolar disorder and cannabis use disorder in a controlled family study. *JAMA Psychiatry.* 2022;79(7):727-735. DOI 10.1001/jamapsychiatry.2022.1338.
- Reaven G.M. Banting lecture 1988. Role of insulin resistance in human disease. *Diabetes.* 1988;37(12):1595-1607. DOI 10.2337/diab.37.12.1595.
- Reynolds R.J., Irvin M.R., Bridges S.L., Kim H., Merriman T.R., Arnett D.K., Singh J.A., Sumpter N.A., Lupi A.S., Vazquez A.I. Genetic correlations between traits associated with hyperuricemia, gout, and comorbidities. *Eur. J. Hum. Genet.* 2021;29(9):1438-1445. DOI 10.1038/s41431-021-00830-z.
- Rijken M., Hujala A., van Ginneken E., Melchiorre M.G., Groenewegen P., Schellevis F. Managing multimorbidity: profiles of integrated care approaches targeting people with multiple chronic conditions in Europe. *Health Policy.* 2018;122(1):44-52. DOI 10.1016/j.healthpol.2017.10.002.
- Rubio-Perez C., Guney E., Aguilar D., Piñero J., Garcia-Garcia J., Iadrola B., Sanz F., Fernandez-Fuentes N., Furlong L.I., Oliva B. Genetic and functional characterization of disease associations explains comorbidity. *Sci. Rep.* 2017;7(1):6207. DOI 10.1038/s41598-017-04939-4.
- Ryu I., Ryu M.J., Han J., Kim S.J., Lee M.J., Ju X., Yoo B.H., Lee Y.L., Jang Y., Song I.C., Chung W., Oh E., Heo J.Y., Kweon G.R. L-Deprenyl exerts cytotoxicity towards acute myeloid leukemia through inhibition of mitochondrial respiration. *Oncol. Rep.* 2018;40(6):3869-3878. DOI 10.3892/or.2018.6753.
- Rzhetsky A., Wajngurt D., Park N., Zheng T. Probing genetic overlap among complex human phenotypes. *Proc. Natl. Acad. Sci. USA.* 2007;104(28):11694-11699. DOI 10.1073/pnas.0704820104.
- Sandilands A., Terron-Kwiatkowski A., Hull P.R., O'Regan G.M., Clayton T.H., Watson R.M., Carrick T., Evans A.T., Liao H., Zhao Y., Campbell L.E., Schmutz M., Gruber R., Janecke A.R., Elias P.M., van Steensel M.A., Nagtzaam I., van Geel M., Steijlen P.M., Munro C.S., Bradley D.G., Palmer C.N., Smith F.J., McLean W.H., Irvine A.D. Comprehensive analysis of the gene encoding filaggrin uncovers prevalent and rare mutations in ichthyosis vulgaris and atopic eczema. *Nat. Genet.* 2007;39(5):650-654. DOI 10.1038/ng2020.
- Shadrina A.S., Sharapov S.Z., Shashkova T.I., Tsepilov Y.A. Varicose veins of lower extremities: insights from the first large-scale genetic study. *PLoS Genet.* 2019;15(4):e1008110. DOI 10.1371/journal.pgen.1008110.
- Shnyder N.A., Novitsky M.A., Neznanov N.G., Limankin O.V., Asadullin A.R., Petrov A.V., Dmitrenko D.V., Narodova E.A., Popenko N.V., Nasyrova R.F. Genetic predisposition to schizophrenia and depressive disorder comorbidity. *Genes (Basel).* 2022;13(3):457. DOI 10.3390/genes13030457.
- Sidransky E., Nalls M.A., Aasly J.O., Aharon-Peretz J., Annesi G., Barbosa E.R., Bar-Shira A., Berg D., Bras J., Brice A., ... Tsuji S., Wittstock M., Wolfsberg T.G., Wu Y.R., Zabetian C.P., Zhao Y., Ziegler S.G. Multicenter analysis of glucocerebrosidase mutations in Parkinson's disease. *N. Engl. J. Med.* 2009; 361(17):1651-1661. DOI 10.1056/NEJMoa0901281.
- Spataro N., Rodríguez J.A., Navarro A., Bosch E. Properties of human disease genes and the role of genes linked to Mendelian disorders in complex disease aetiology. *Hum. Mol. Genet.* 2017;26(3):489-500. DOI 10.1093/hmg/ddw405.
- Stein O., Stein Y. Smooth muscle cells and atherosclerosis. *Curr. Opin. Lipidol.* 1995;6(5):269-274. DOI 10.1097/00041433-199510000-00005.
- Stephens Z.D., Lee S.Y., Faghri F., Campbell R.H., Zhai C., Efron M.J., Iyer R., Schatz M.C., Sinha S., Robinson G.E. Big data: astronomical or genomics? *PLoS Biol.* 2015;13(7):e1002195. DOI 10.1371/journal.pbio.1002195.
- Wagner G.P., Kenney-Hunt J.P., Pavlicev M., Peck J.R., Waxman D., Cheverud J.M. Pleiotropic scaling of gene effects and the 'cost of complexity'. *Nature.* 2008;452(7186):470-472. DOI 10.1038/nature06756.
- Wang M., Tang S., Yang X., Xie X., Luo Y., He S., Li X., Feng X. Identification of key genes and pathways in chronic rhinosinusitis with nasal polyps and asthma comorbidity using bioinformatics approaches. *Front. Immunol.* 2022;13:941547. DOI 10.3389/fimmu.2022.941547.
- Weiss F.U., Simon P., Bogdanova N., Mayerle J., Dworniczak B., Horst J., Lerch M.M. Complete cystic fibrosis transmembrane conductance regulator gene sequencing in patients with idiopathic chronic pancreatitis and controls. *Gut.* 2005;54(10):1456-1460. DOI 10.1136/gut.2005.064808.
- Werner H., Sarfstein R., Nagaraj K., Laron Z. Laron syndrome research paves the way for new insights in oncological investigation. *Cells.* 2020;9(11):2446. DOI 10.3390/cells9112446.
- Woese C.R., Goldenfeld N. How the microbial world saved evolution from the scylla of molecular biology and the charybdis of the modern synthesis. *Microbiol. Mol. Biol. Rev.* 2009;73(1):14-21. DOI 10.1128/MMBR.00002-09.
- Wray N.R., Goddard M.E., Visscher P.M. Prediction of individual genetic risk of complex disease. *Curr. Opin. Genet. Dev.* 2008;18(3): 257-263. DOI 10.1016/j.gde.2008.07.006.
- Zhao R., Choi B.Y., Lee M.H., Bode A.M., Dong Z. Implications of genetic and epigenetic alterations of *CDKN2A* (p16^{INK4a}) in cancer. *EBioMedicine.* 2016;8:30-39. DOI 10.1016/j.ebiom.2016.04.017.
- Zolotareva O., Saik O.V., Königs C., Bragina E.Y., Goncharova I.A., Freidin M.B., Dosenko V.E., Ivanisenko V.A., Hofestädt R. Comorbidity of asthma and hypertension may be mediated by shared genetic dysregulation and drug side effects. *Sci. Rep.* 2019;9(1):16302. DOI 10.1038/s41598-019-52762-w.

ORCID ID

E.Yu. Bragina orcid.org/0000-0002-1103-3073
V.P. Puzyrev orcid.org/0000-0002-2113-4556

Благодарности. Исследование выполнено в рамках государственного задания Министерства науки и высшего образования, № 122020300041-7.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию 17.10.2022. После доработки 25.12.2022. Принята к публикации 26.12.2022.

Перевод на английский язык <https://vavilov.elpub.ru/jour>

Молекулярно-генетические основы буллезного эпидермолиза

Ю.Ю. Коталевская^{1, 2}✉, В.А. Степанов³

¹ Московский областной научно-исследовательский клинический институт им. М.Ф. Владимирского, Москва, Россия

² Благотворительный фонд «БЭЛА. Дети-бабочки», Москва, Россия

³ Научно-исследовательский институт медицинской генетики, Томский национальный исследовательский медицинский центр Российской академии наук, Томск, Россия

✉ kotalevskaya@mail.ru

Аннотация. Буллезный эпидермолиз (БЭ) – наследственное нарушение, вызывающее хрупкость кожи, обусловленную изменениями генов, отвечающих за целостность кожи и дермо-эпидермальную адгезию. Хрупкость кожи проявляется снижением устойчивости к внешним механическим воздействиям, клинические признаки которой – образование пузырей, эрозий и ран на коже и слизистых оболочках. Для БЭ характерен широкий фенотипический спектр, при тяжелых типах, кроме кожи и слизистых, отмечаются мультисистемность поражения и развитие внекожных осложнений, высокая летальность. Выделено более 30 клинических подтипов БЭ, сгруппированных в четыре основных типа: простой, пограничный, дистрофический БЭ и синдром Киндлера. На сегодняшний день БЭ обуславливают патогенные варианты в 16 различных генах, которые кодируют белки, входящие в состав крепящих структур кожи, и сигнальные белки. Генетические дефекты в этих генах служат причиной нарушения функции клеточных структур, процессов дифференцировки, пролиферации и апоптоза клеток, приводя к механической неустойчивости кожи. Образование укороченных белков или уменьшение их количества обуславливает в основном функциональные нарушения, формируя легкие или среднетяжелые фенотипы. При нулевых генетических вариантах, вследствие которых экспрессия белка утрачивается полностью, возникают структурные нарушения, влекущие тяжелую клиническую картину. Для большинства вовлеченных в патогенез БЭ генов обнаружены определенные связи между характером и локализацией генетических дефектов с тяжестью клинических проявлений заболевания. Установление точного диагноза зависит от корреляции клинических, генеалогических и иммуногистологических данных в сочетании с молекулярно-генетическим исследованием. В целом изучение клинических, генетических и ультраструктурных изменений при БЭ значительно расширяет понимание естественного течения заболевания и пополняет данные о корреляциях генотип-фенотип, способствует поиску и изучению эпигенетических и негенетических факторов-модификаторов заболевания, а также разработке подходов к радикальному лечению заболевания. Новые возможности технологий секвенирования позволили описать новые фенотипы и изучить их генетические и молекулярные механизмы. В настоящей статье описаны патогенетические аспекты и гены, вызывающие классические и редкие синдромальные подтипы БЭ.

Ключевые слова: буллезный эпидермолиз; патогенез; корреляции генотип-фенотип; гетерогенность.

Для цитирования: Коталевская Ю.Ю., Степанов В.А. Молекулярно-генетические основы буллезного эпидермолиза. *Вавиловский журнал генетики и селекции*. 2023;27(1):18-27. DOI 10.18699/VJGB-23-04

Molecular genetic basis of epidermolysis bullosa

Yu.Yu. Kotalevskaya^{1, 2}✉, V.A. Stepanov³

¹ Moscow Regional Research and Clinical Institute, Moscow, Russia

² Charitable Foundation "BELA. Butterfly Children", Moscow, Russia

³ Research Institute of Medical Genetics, Tomsk National Research Medical Center of the Russian Academy of Sciences, Tomsk, Russia

✉ kotalevskaya@mail.ru

Abstract. Epidermolysis bullosa (EB) is an inherited disorder of skin fragility, caused by mutations in a large number of genes associated with skin integrity and dermal-epidermal adhesion. Skin fragility is manifested by a decrease in resistance to external mechanical influences, the clinical signs of which are the formation of blisters, erosions and wounds on the skin and mucous membranes. EB is a multisystemic disease and characterized by a wide phenotypic spectrum with extracutaneous complications in severe types, besides the skin and mucous membranes, with high mortality. More than 30 clinical subtypes have been identified, which are grouped into four main types: simplex EB, junctional EB, dystrophic EB and Kindler syndrome. To date, pathogenic variants in 16 different genes are associated with EB and encode proteins that are part of the skin anchoring structures or are signaling proteins. Genetic mutations cause dysfunction of cellular structures, differentiation, proliferation and apoptosis of cells, leading to mechanical instability of the skin. The formation of reduced proteins or decrease in their level leads mainly to functional disorders, forming mild or intermediate severe phenotypes. Absent protein expression is a result of null genetic variants and leads to

structural abnormalities, causing a severe clinical phenotype. For most of the genes involved in the pathogenesis of EB, certain relationships have been established between the type and position of genetic variant and the severity of the clinical manifestations of the disease. Establishing an accurate diagnosis depends on the correlation of clinical, genealogical and immunohistological data in combination with molecular genetic testing. In general, the study of clinical, genetic and ultrastructural changes in EB has significantly expanded the understanding of the natural history of the disease and supplemented the data on genotype-phenotype correlations, promotes the search and study of epigenetic and non-genetic disease modifier factors, and also allows developing approaches to radical treatment of the disease. New advances of sequencing technologies have made it possible to describe new phenotypes and study their genetic and molecular mechanisms. This article describes the pathogenetic aspects and genes that cause main and rare syndromic subtypes of EB.

Key words: epidermolysis bullosa; pathogenesis; genotype phenotype correlations; heterogeneity.

For citation: Kotalevskaya Yu.Yu., Stepanov V.A. Molecular genetic basis of epidermolysis bullosa. *Vavilovskii Zhurnal Genetiki i Selektii = Vavilov Journal of Genetics and Breeding*. 2023;27(1):18-27. DOI 10.18699/VJGB-23-04

Введение

Буллезный эпидермолиз (БЭ) представляет собой группу редких и в настоящее время неизлечимых, генетически детерминированных наследственных заболеваний кожи. Буллезный эпидермолиз характеризуется хрупкостью кожи и слизистых оболочек, возникающей при механической травме, кажущейся незначительной по сдвигающей силе, часто сопровождается поражением ногтей, зубов и волос (Pânzaru et al., 2022). Спектр характерных проявлений со стороны кожи широк и включает пузыри, эрозии, раны, которые могут стать хроническими, рубцы, корки, милиумы, атрофию кожи и диспигментацию. При редких подтипах возможно поражение не только кожи, но и мышц, желудочно-кишечного тракта, почек и др., что обусловлено характером экспрессии дефектного белка.

Степень тяжести заболевания варьирует от фенотипически легких до тяжелых инвалидизирующих или летальных вариантов течения, что и определяет ожидаемый прогноз продолжительности жизни. Тяжелые подтипы БЭ развиваются как системные заболевания с вторичным полиорганным поражением и задержкой развития, анемией, заболеваниями сердца и костей, двигательными нарушениями, ранней предрасположенностью к раку кожи и преждевременной смертью. Лечение БЭ исключительно симптоматическое и направлено на профилактику механических травм, уход за ранами, лечение инфекционных осложнений и внекожных проявлений заболевания. К настоящему времени никакие подходы к терапии не могут излечить пациентов от БЭ (Pânzaru et al., 2022).

Буллезный эпидермолиз – наглядная модель механо-буллезного заболевания, а исследование приводящих к нему механизмов позволило значительно продвинуться в понимании основ физиологии и патофизиологии кожи. Полученные знания нашли свое отражение в классификации БЭ, которая пересматривалась несколько раз за последнее десятилетие международной консенсусной группой (Has et al., 2020a). Буллезный эпидермолиз подразделяется на четыре основных типа БЭ: простой (ПБЭ), пограничный (ПоБЭ), дистрофический (ДБЭ) и синдром Киндлера (СК), которые базируются на ультраструктурных изменениях и уровне образования пузырей в слоях кожи и отражают последствия генетических дефектов на функцию белков. Буллезный эпидермолиз клинически и генетически очень гетерогенен, наследуется по аутосомно-доминантному (АД) или аутосомно-рецессивному (АР) типу (Has et al., 2020a). Успехи в понимании патогене-

за БЭ способствуют разработке потенциально эффективных белковых, клеточных и генотерапевтических методов лечения (Has et al., 2020b).

Базальный слой эпидермиса, зона базальной мембраны (ЗБМ) и внеклеточный матрикс – это ключевые субрегионы, которые занимают центральное место в патофизиологии БЭ (Uitto et al., 2017), а генетические нарушения изменяют структуру или функцию их белков (Mariath et al., 2020). Патогенные варианты в 16 различных генах обуславливают генетическую и аллельную гетерогенность БЭ с формированием четырех основных типов БЭ, включающих более 30 клинических подтипов. Гены, связанные с БЭ, кодируют внутриклеточные, трансмембранные или внеклеточные белки, главным образом структурные компоненты цитоскелета (кератин 5 и 14), ЗБМ (интегрин $\alpha 6 \beta 4$, коллаген типа XVII, ламинин-332, коллаген типа VII, альфа-субъединица интегрина $\alpha 3$, киндин-1) или белки межклеточной адгезии – десмоплакин, плакофилин, плакоглобин (см. таблицу) (Has, Bruckner-Tuderman, 2014). В таблице представлены ключевые процессы патогенеза, приводящие к определенному фенотипу.

Основные типы буллезного эпидермолиза

Простой БЭ – наиболее распространенный тип, на долю которого приходится около 70 % всех пациентов с БЭ (Has, Fischer, 2019), согласно последней классификации, включает 14 клинических подтипов. Простой БЭ имеет широкий спектр степени тяжести: от легкой, с образованием пузырей на руках и ногах, до генерализованных форм с внекожными поражениями, иногда с летальным исходом (Fine, 2010). Простой БЭ чаще всего вызывается дефектами кератиновых филаментов базальных кератиноцитов, имеет разную генетическую основу: связан с изменениями как минимум в семи генах и представляет наибольшее клиническое разнообразие.

Большинство подтипов ПБЭ наследуются по АД типу, хотя в некоторых регионах мира встречается АР тип наследования (Gostyńska et al., 2015; Vahidnezhad et al., 2019). Наиболее распространенные подтипы ПБЭ, наблюдаемые в клинической практике, вызваны мутациями в генах кератина 5 или кератина 14 (70–80 % случаев), при этом, по литературным данным, не менее чем у 17 % больных ПБЭ мутации возникали *de novo* (Bolling et al., 2011; Wertheim-Tysarowska et al., 2016). Кроме того, ПБЭ с АД типом наследования может быть связан с гетерозиготными вариантами в генах *PLEC* или *KLHL24* (Grilletta,

Классификация буллезного эпидермолиза (БЭ) и основные механизмы патогенеза

Подтип	Тип	Дефектный ген	Механизм
Простой БЭ – внутриэпидермальный			
Локализованный Средней тяжести	АД	<i>KRT5, KRT14</i>	Аномальная кератиновая цитоскелетная сеть и базальный цитолиз
Тяжелый	АД	<i>KRT5, KRT14</i>	Аномальная кератиновая цитоскелетная сеть, комкование кератиновых тонофиламентов с последующим базальным цитолизом
С пятнистой пигментацией	АД	Преимущественно <i>KRT5</i> , реже <i>KRT14</i>	Разрыв кератиновых филаментов, базальный цитолиз и дополнительная агрегация плотно упакованных сложных меланосом в околоядерной цитоплазме базальных кератиноцитов
С мигрирующей кольцевидной эритемой	АД	<i>KRT5</i>	Удлинение кератина 5 вследствие позднего образования терминирующего кодона вызывает воспаление, опосредованное Т-клетками
Средней тяжести с кардиомиопатией	АД	<i>KLHL24</i>	Патогенные варианты приводят к образованию усеченного и более стабильного белка KLHL24 с последующей повышенной деградацией KRT14
Средней тяжести, вызванный мутацией в гене <i>PLEC</i>	АД и AP	<i>PLEC</i>	Уменьшенные ПД вследствие нарушения внутренней бляшки, к которой прикрепляется кератиновый цитоскелет, с последующим базальным цитолизом
Средней тяжести с мышечной дистрофией	AP	<i>PLEC</i>	Расслоение максимально близко к ЗБМ; ПД значительно уменьшены в размере; нарушение взаимодействия саркомеров вследствие бесстержневой изоформы плектина внутри Z-дисков; дефектное крепление между собранными десминовыми филаментами запускает образование агрегатов десминовых белков, а также вторичную митохондриальную недостаточность
Тяжелый с атрезией привратника	AP	<i>PLEC</i>	Отсутствует плектин
Простой БЭ	AP	<i>KRT5, KRT14</i>	Отсутствие либо значительное снижение пучков промежуточных филаментов в базальных кератиноцитах
Локализованный или средней тяжести с дефицитом BP230	AP	<i>DST</i>	Отсутствие внутренних бляшек ПД, компенсаторное увеличение KRT14 и плектина, что может объяснять мягкий фенотип
Локализованный или средней тяжести с дефицитом экзофилина 5	AP	<i>EXPH5</i>	Нарушение внутриклеточного транспорта везикул по актиновой и тубулиновой сетям; увеличение перинуклеарных везикул с аномальным кератином; утрата адгезии базального кератиноцита
Локализованный с нефропатией (дефицит CD151)	AP	<i>CD151</i>	Патогенные варианты приводят к снижению адгезии кератиноцитов, опосредованной комплексами ламинин-332-интегрин $\alpha 3\beta 1$ в эпидермисе и подоцитах
Пограничный БЭ – внутри светлой пластинки			
Тяжелый	AP	<i>LAMA3, LAMB3, LAMC2</i>	Ламинин-332 обычно отсутствует; уменьшенные ПД; аномальные или отсутствующие суббазальные плотные пластинки; уменьшение якорных фибрилл
Средней тяжести	AP	<i>LAMA3, LAMB3, LAMC2, COL17A</i>	Снижение уровня ламинина-332; отсутствие или уменьшение уровня коллагена типа XVII
С атрезией привратника	AP	<i>ITGA6, ITGB4</i>	Отсутствует или значимо снижен интегрин $\alpha 6\beta 4$; патогенные варианты в гене <i>ITGB4</i> , приводящие к частичной экспрессии интегрин $\beta 4$, могут вызывать более мягкий фенотип
Локализованный	AP	<i>LAMA3, LAMB3, LAMC2, COL17A, ITGB4, ITGA3</i>	Различные аномалии и уровни экспрессии в дефектных белках
Инверсный	AP	<i>LAMA3, LAMB3, LAMC2</i>	Снижение экспрессии ламинина-332
С поздним началом	AP	<i>COL17A</i>	Снижение или патологическая экспрессия коллагена типа XVII

Окончание таблицы

Подтип	Тип	Дефектный ген	Механизм
Ларинго-онихо-кожный синдром	AP	LAMA3	Аномально укороченная субъединица α 3A ламинина-332
С интерстициальным заболеванием легких и нефротическим синдромом	AP	IGT3	Частыми являются варианты с потерей функции субъединицы интегрин α 3; миссенс-варианты могут вызывать более легкое течение заболевания и улучшать выживаемость
Дистрофический БЭ – плотная пластинка			
ДДБЭ, средней тяжести	AD	COL7A1	Уменьшенный или аномальный коллаген типа VII; обычно возникает из-за миссенс-мутаций, вызывающих замену глицина в шарнирной области тройной спирали коллагена типа VII
ДДБЭ, локализованный	AD	COL7A1	Уменьшенный или аномальный коллаген типа VII, возникающий в результате моноаллельных делеций, миссенс-вариантов или мутаций сайта сплайсинга
ДДБЭ, пруригинозный	AD	COL7A1	Точный патогенетический механизм неизвестен
ДДБЭ, самоизлечивающийся	AD	COL7A1	Внутриклеточное накопление несекретируемого проколлагена VII; сохранение коллагена типа VII в базальных кератиноцитах; постепенное улучшение образования коллагена типа VII и якорных фибрилл по неизвестным причинам
РДБЭ, средней тяжести	AP	COL7A1	Комбинации биаллельных патогенных вариантов в COL7A1 (миссенс-, нонсенс-, инсерции, делеции и варианты сайта сплайсинга) приводят к уменьшению или образованию аномального коллагена типа VII
РДБЭ, тяжелый	AP	COL7A1	Биаллельные нулевые варианты в COL7A1, которые приводят к значительному снижению или отсутствию коллагена типа VII и, следовательно, к отсутствию функциональных якорных фибрилл
РДБЭ, инверсный	AP	COL7A1	Предполагается, что специфические замены аргинина и глицина в тройной спирали коллагена типа VII снижают термостабильность белка, вызывая клинические проявления на участках тела с более высокой температурой, в том числе на слизистых оболочках
РДБЭ, локализованный	AP	COL7A1	Уменьшенный или аномальный коллаген типа VII
РДБЭ, пруригинозный	AP	COL7A1	То же, что и при ДДБЭ, пруригинозный
РДБЭ, самоизлечивающийся	AP	COL7A1	То же, что и при ДДБЭ, самоизлечивающийся
ДБЭ, тяжелый	AD и AP	COL7A1	Патогенетические механизмы неизвестны, фенотип возникает у компаунд-гетерозигот по доминантной замене глицина в COL7A1 в одном аллеле и рецессивного варианта во втором аллеле, что изменяет белковое окружение в области ЗБМ, усиливая тяжесть клинических проявлений
Синдром Киндлера – вариативный и смешанный			
Синдром Киндлера	AP	FERMT1	Патогенные варианты способствуют разрушению сетей цитоскелета кератиноцитов, аномальной активации интегрин и потере адгезии кератиноцитов к подлежащей базальной мембране

Примечание. AD – аутосомно-доминантный тип наследования; AP – аутосомно-рецессивный тип наследования; ЗБМ – зона базальной мембраны; ПД – полудесмосома; ДДБЭ – доминантный дистрофический буллезный эпидермолиз; РДБЭ – рецессивный дистрофический буллезный эпидермолиз.

2019; Kiritsi et al., 2021). У больных с ПБЭ также описаны редкие дигенные мутации в генах *KRT5* и *KRT14* (Sathishkumar et al., 2016).

Кератин 5 и кератин 14 имеют сходную белковую структуру, состоящую из центрального α -спирального стержневого домена, который отвечает за полимеризацию этих белков с образованием кератиновых тонофиламентов. Стержневой домен подразделяется на сегменты 1A, 1B, 2A и 2B с помощью гибких линкеров L1, L12 и L2, фланкирован вариативными доменами V1 и V2 в обоих белках. Кератин 5 имеет консервативный домен гомологии H1 и H2. Гены *KRT5* и *KRT14* экспрессируются в базальных

кератиноцитах эпидермиса, где их белковые продукты объединяются между собой и образуют гетеродимерные молекулы. Димеры K5 и K14 являются основными компонентами системы промежуточных филаментов кератиноцитов, которые собираются во внутриклеточную сеть (Bunick, Milstone, 2017).

Патогенные варианты в генах *KRT5* и *KRT14*, среди которых преобладают доминантные миссенс-варианты, влияют на способность кератинов взаимодействовать со своим партнером. Ключевое значение имеет расположение патогенного варианта в функциональных доменах генов *KRT5* или *KRT14* (Arin et al., 2010). Доминантно-нега-

тивные патогенные варианты сгруппированы в начале 1A или конце 2B сегментов спирального стержневого домена *KRT5* и *KRT14* и типичны для тяжелого генерализованного ПБЭ, так как эти домены очень консервативны и считаются критическими для сборки филаментов.

Одними из частых патогенных вариантов считаются варианты р.Glu477Lys – в гене *KRT5* и р.Arg125Cys, р.Arg125His, р.Asn123Ser – в гене *KRT14* (Bolling et al., 2011; Vahidnezhad et al., 2016). При ПБЭ средней тяжести патогенные варианты расположены во второй части сегментов 1A или 2B стержневого домена *KRT5* и *KRT14*. При данном подтипе они не меняют процесс удлинения кератинов во время сборки филаментов, но ослабляют их функцию (Has, Bruckner-Tuderman, 2014). При локализованном подтипе ПБЭ патогенные варианты расположены кластерами как в *KRT5*, так и в *KRT14*, обычно за пределами высоко консервативных граничных мотивов стержневого домена, а также в линкерах L12, кроме того, в гене *KRT5* в домене H1, вызывая структурную нестабильность филаментов (Bardhan et al., 2020). Более четкие корреляции с генотипом обнаружены при подтипе ПБЭ с пятнистой пигментацией, он ассоциирован с патогенными вариантами в домене V1 гена *KRT5*. Так, на вариант р.Pro25Leu приходится 90–95 % мутаций при этом подтипе (Arim et al., 2010).

Тяжелый и среднетяжелый ПБЭ с АР наследованием связаны с редкими патогенными биаллельными вариантами в *KRT14* и *KRT5*, которые обнаружены в кровнородственных семьях (Vahidnezhad et al., 2016). Гомозиготные мутации в гене *KRT5* приводят к тяжелому фенотипу, внеожным проявлениям и ранней смертности (Has et al., 2006).

В последнем пересмотре классификации БЭ охарактеризованы редкие синдромальные подтипы ПБЭ, связанные с мутациями в генах *PLEC*, *KLHL24*, *DST*, *EXPH5* и *CD151* (см. таблицу), рассмотрим их далее.

Белок плектин, кодируемый геном *PLEC*, представляет собой цитоскелетный белок, связывающий сеть промежуточных филаментов с ПД и, таким образом, действует как медиатор механической стабильности кератиноцитов в коже (Natsuga, 2015). Большое количество альтернативно сплайсированных первых экзонов гена плектина образуют множественные изоформы белка и определяют разную экспрессию в тканях, что обеспечивает клиническое разнообразие, формируя четыре редких фенотипа ПБЭ.

Патогенные варианты в гене *PLEC* в основном были обнаружены в экзонах 31 и 32; варианты с потерей функции белка обуславливают развитие более тяжелых фенотипов, таких как ПБЭ с атрезией привратника, а в результате нулевых вариантов гена *PLEC* – к ПБЭ с мышечной дистрофией, где волокна скелетных мышц теряют свою структурную целостность из-за дефектов десминовых филаментов (Natsuga, 2015). Простой БЭ средней тяжести с АР типом наследования вызван специфической гомозиготной нонсенс-мутацией р.Arg16X в первом экзоне, кодирующем изоформу плектина 1a, приводящей к отсутствию только этой специфической изоформы (Gostuńska et al., 2015). В экзоне 31 гена *PLEC* описана доминантная аминокислотная замена р.Arg2110Trp, вызывающая частичную потерю функции белка и фрагментацию ПД

(Kiritsi et al., 2021), что клинически проявляется как ПБЭ средней тяжести.

Белок *KLHL24* принадлежит к семейству высоко консервативных белков с VTB/kelch доменами, патогенные варианты в гене *KLHL24* приводят к нарушению регуляции аутоубиквитинирования, изменяют регуляцию деградации молекул кератина 14 и вызывают его фрагментацию (Dhanao et al., 2013). При подтипе ПБЭ, связанном с мутациями в гене *KLHL24*, во всех описанных случаях наблюдался гетерозиготный вариант в стартовом кодоне, наиболее распространенным был с.1A-G с доминантно-негативным эффектом (Bardhan et al., 2020). У 85 % пациентов при данном подтипе ПБЭ в молодом возрасте развивается дилатационная кардиомиопатия, обусловленная *KLHL24*-опосредованной деградацией десмина – основного белка промежуточных филаментов кардиомиоцитов (Grilletta, 2019).

Дистонин (BPAG1) относится к семейству белков плакинов (Ganani et al., 2021). Ген *DST* кодирует эпителиальную изоформу BPAG1-е, которая является структурным компонентом внутренних бляшек ПД, состоит из спирально-спирального стержневого домена и фланкирующих N- и C-концов. N-конец белка BPAG1-е участвует в его интеграции в ПД и имеет сайты связывания для коллагена типа XVII и интегрин $\beta 4$, тогда как C-конец – ключевая точка прикрепления кератиновых промежуточных филаментов (Kumar et al., 2015). Показано, что мутации в *DST* связаны с нарушением адгезии кератиноцитов, повышенной миграцией клеток при сниженной экспрессии $\beta 4$ -интегринов на клеточной поверхности (Ganani et al., 2021), клинически приводят к мягкому фенотипу.

Экзофилин-5, эффекторный белок RAB27b GTPase, кодируемый геном *EXPH5*, не является структурным компонентом промежуточных филаментов, десмосом или ПД. Хотя его роль до конца не известна, предполагается, что он способствует регуляции внутриклеточного транспорта везикул, включая контроль их образования и движения по актиновой и тубулиновой сетям, а также секреции экзосом (Natsuga et al., 2010). Описаны единичные семьи с гомозиготными вариантами в гене *EXPH5*, приводящими к сдвигу рамки считывания, а также в сочетании с нонсенс-вариантами с легкими клиническими проявлениями заболевания.

В эпидермисе экспрессия трансмембранного белка CD151 локализуется в ПД, связываясь с интегрином $\alpha \beta 4$ и стабилизируя его взаимодействие с ламинином-332, играет критическую роль в формировании комплекса ПД. Белок CD151 обеспечивает клеточную адгезию и внутриклеточный везикулярный транспорт интегринов. В почках он образует комплексы с интегрин $\alpha \beta 1$ и $\alpha \beta 1$ и необходим для правильной сборки базальных мембран клубочков и канальцев (Margadant et al., 2010). Дефект в белке CD151 определяет клинические проявления у лиц с CD151-ассоциированным ПБЭ, включая нефропатию с протеинурией (Karamatic Crew et al., 2004).

Пограничный БЭ (ПоБЭ) также представляет собой клинически и генетически гетерогенную группу нарушений хрупкости кожи, включает девять клинических подтипов, является редким типом БЭ (Has et al., 2020a). Подтипы ПоБЭ имеют патогномичные признаки, на-

пример, при тяжелом генерализованном подтипе быстро образуется грануляционная ткань в типичных местах, характерна высокая летальность (Kiritsi et al., 2011). Фенотипическая изменчивость при ПоБЭ крайне широкая – от только дистрофии ногтей до летального исхода на первом году жизни. Патогенные варианты в семи разных генах приводят к развитию ПоБЭ, все подтипы наследуются по АР типу. К наиболее частым подтипам ПоБЭ приводят патогенные варианты в генах *LAMA3*, *LAMB3* и *LAMC2*, кодирующих $\alpha 3$ -, $\beta 3$ - и $\gamma 2$ -цепи ламинина-332, а также в гене *COL17A1*, кодирующем коллаген типа XVII (Uitto et al., 2016). Редкие фенотипы ПоБЭ связаны с дефицитом интегринина $\alpha\beta 4$, приводя к развитию ПоБЭ с атрезией привратника и дефицитом $\alpha 3$ субъединицы интегринина $\alpha 3\beta 1$, вызывая ПоБЭ с поражением органов дыхания и почек (Kiritsi et al., 2013).

Белок ламинин-332 – это гетеротример, состоящий из $\alpha 3$ -, $\beta 3$ - и $\gamma 2$ -цепей, которые кодируются генами *LAMA3*, *LAMB3* и *LAMC2* соответственно. Вместе с внеклеточным доменом коллагена типа XVII они создают якорные филаменты. Белок ламинин-332 связывается С-концом, образуемым α -цепью, с интегрининами $\alpha 3\beta 1$ в участках фокальной адгезии и интегрининами $\alpha\beta 4$ в ПД, соединяя поверхность базальных кератиноцитов с дермо-эпидермальной БМ (Dogic et al., 1998). В дерме N-концы цепей ламинина-332 связываются с коллагеном типа VII так, что якорные филаменты и якорные фибриллы соединяются напрямую (Aumailley et al., 2003). Потеря экспрессии ламинина-332 вызывает крайнюю хрупкость кожи и образование избыточной грануляционной ткани при генерализованном тяжелом ПоБЭ. При ламинин-332-дефицитных подтипах ПоБЭ в 70 % случаев изменен ген *LAMB3*. Примерно по 9 % больных ПоБЭ имеют мутации в генах *LAMA3* и в *LAMC2* соответственно (Varki et al., 2006; Uitto et al., 2016). Наиболее частый патогенный вариант – это p.R635X как «горячая» мутационная точка, на долю которой приходится 45–63 % всех патогенных аллелей гена *LAMB3* при генерализованном тяжелом ПоБЭ, в результате отсутствует один из трех белков, которые собираются в ламинин-332.

К развитию более легких проявлений БЭ приводят миссенс-мутации, мутации сайта сплайсинга и делеции с сохранением рамки считывания, изменяя ключевые положения белковых субъединиц и влияя на способность ламинина $\alpha 3$, $\beta 3$ и $\gamma 2$ собираться в тримерную молекулу, ее вторичную структуру и ее способность формировать внутриклеточные якорные фибриллы (Kiritsi et al., 2011).

Особым фенотипом, ларинго-онихо-кожным синдромом (ЛОК-синдром), проявляются патогенные варианты, образующие стоп-кодон в экзоне 39, специфичном для $\alpha 3$ субъединицы гена *LAMA3*, где к настоящему времени описано три причинных варианта: p.V51fs, p.Gln157Ter и p.Trp16Ter (Wang et al., 2022). Недавно С. Prodingер с коллегами (2021) сообщили о трех новых мутациях в гене *LAMA3* за пределами экзона 39.

Белок коллаген типа XVII – гомотример, состоящий из трех идентичных субъединиц, является трансмембранным белком и представляет собой основной структурный компонент ПД, имеет как внутри-, так и внеклеточный домены. Коллаген типа XVII действует как рецептор кле-

точной поверхности для белков внеклеточного матрикса (van den Bergh, Giudice, 2003). Внеклеточный домен коллагена типа XVII связан с ламинином-332, в связи с этим он принимает участие в создании якорных филаментов, может контролировать подвижность клеток, определяет пространственную ориентацию ламинина-332 и расположение в коллаген-IV-содержащей плотной пластинке БМ (Tong, Xu, 2004).

Этот белок также регулирует дифференцировку амелобластов – эпителиальных клеток, принимающих участие в образовании зубной эмали (Asaka et al., 2009). Дефекты зубной эмали, от точечных до генерализованной гипоплазии, встречаются при всех подтипах ПоБЭ, возникающих из-за нарушения адгезии одонтогенного эпителия, из которого происходят амелобласты (Wright et al., 2015).

Коллаген типа XVII играет центральную роль в регуляции пролиферации межфолликулярного эпидермиса, участвуя в поддержании стволовых клеток волосяного фолликула, где управляет программой их старения, что может объяснить необратимое выпадение волос у людей с дефицитом коллагена типа XVII (Matsumura et al., 2016).

Патогенные варианты в гене *COL17A1* обычно приводят к ПоБЭ средней тяжести (Pasmoosj et al., 2004), хотя было описано несколько случаев со смертельным исходом с наличием патогенных вариантов *COL17A1* (Murrell et al., 2007). По данным D. Kiritsi с коллегами (2011), 69 % вариантов гена *COL17A1* были нонсенс-вариантами, инсерциями или делециями, 19 % – миссенс-вариантами и 12 % – вариантами сайта сплайсинга. Патогенные варианты, приводящие к пропуску экзона в гене *COL17A1*, оказывают смягчающее влияние на фенотип, позволяя продуцировать достаточно функциональный белок (Condrat et al., 2019).

В некоторых случаях нонсенс-мутации могут вызывать незначительно выраженные проявления генерализованного ПоБЭ средней тяжести из-за механизмов альтернативного сплайсинга. Показано, что у больных с гомозиготной нонсенс-мутацией p.R795X в экзоне 33 в результате альтернативного сплайсинга образуется мРНК *COL17A1*, что позволяет продуцировать незначительное количество коллагена типа XVII.

Интегрины – это гетеродимерные трансмембранные рецепторы, состоящие из α - и β -субъединиц, которые формируют функциональный рецептор (Masunaga et al., 2017). В эпидермисе наиболее распространены интегринины $\alpha 3\beta 1$, $\alpha\beta 4$ и $\alpha 2\beta 1$. Интегрин $\alpha\beta 4$ связывается с ламинином-332 и кератиновыми филаментами внутри клетки, что позволяет координировать клеточный ответ, регулировать адгезию, миграцию и пролиферацию кератиноцитов. Интегрин $\alpha\beta 4$ также участвует в формировании целостности и стабильности ПД и взаимодействует с коллагеном типа XVII, плектином и дистонином (Has, Nyström, 2015). Группа $\beta 1$ -интегринов связана в основном с базальной поверхностью кератиноцитов и участвует в формировании фокальных контактов. Интегрин $\alpha 3\beta 1$ обнаруживается как на базальной, так и на латеральной поверхности базальных кератиноцитов, где он вовлечен в межклеточные контакты.

Ген *ITGA6* кодирует субъединицу $\alpha 6$, ген *ITGB4* – субъединицу $\beta 4$ интегрина $\alpha 6\beta 4$. Патогенные варианты в этих генах, приводящие к преждевременной терминации трансляции, формируют тяжелый фенотип, часто с летальным исходом в неонатальном периоде. Большинство мутаций находится в гене *ITGB4*, описаны варианты сайта сплайсинга, малые делеции и вставки, аминокислотные замены, формирующие редкий подтип – ПоБЭ с атрезией привратника (Masunaga et al., 2017). Исследования корреляций генотипа и фенотипа указывают на то, что варианты, расположенные во внеклеточном домене *ITGB4*, обычно связаны с более тяжелым фенотипом по сравнению с находящимися в цитоплазматическом хвосте (Mariath et al., 2021). В гене *ITGA6* описаны единичные варианты с потерей функции у пациентов из кровнородственных семей, клинически проявляющиеся ранней манифестацией и часто с летальным исходом (Schumann et al., 2013; Masunaga et al., 2017).

Ген *ITGA3* кодирует субъединицу $\alpha 3$ интегрина, которая связана с субъединицей $\beta 1$ и образует интегрин $\alpha 3\beta 1$, участвующий во взаимодействиях с белками внеклеточного матрикса, включая ламинины. Субъединица интегрин $\alpha 3$ экспрессируется в базальных кератиноцитах, подоцитах, эпителиальных клетках канальцев, альвеолярных эпителиальных клетках и многих других тканях (Bardhan et al., 2020).

Описаны несколько случаев ПоБЭ с интерстициальным заболеванием легких и почечными аномалиями, связанными с патогенными вариантами в гене *ITGA3*, экспрессия которого в разных тканях объясняет полиорганное поражение, наблюдаемое у пациентов. Кроме того, отношения между субъединицей интегрин $\alpha 3$ и клеточной мембраной являются сложными, включая посттрансляционные модификации, расщепление, гетеродимеризацию с субъединицей интегрин $\beta 1$ и ассоциацию с CD151. Аминокислотные замены могут мешать этим событиям и действовать как нулевые мутации, приводя к тяжелому течению заболевания (Has et al., 2012). Варианты, обеспечивающие экспрессию остаточного, укороченного или дисфункционального белка, могут привести к более мягкому фенотипу и повышению выживаемости (Liu et al., 2021).

Дистрофический тип БЭ подразделяют на две основные группы: доминантный ДБЭ (ДДБЭ) и рецессивный ДБЭ (РДБЭ). Клиническое разнообразие включает 11 подтипов, при всех подтипах имеются кожные и внекожные проявления различной степени тяжести. Как правило, РДБЭ протекает тяжелее, чем ДДБЭ, спектр варьирует от тяжелых кожных симптомов с рубцеванием и фиброзом, вторичными осложнениями, внекожными проявлениями и высоким риском развития плоскоклеточного рака до легкой хрупкости кожи на конечностях или только дистрофии ногтей. Однако между АД и AP подтипами существует значительное фенотипическое совпадение, что часто клинически затрудняет установление типа наследования ДБЭ у пациента, особенно если пробанд – единственный больной в семье.

Дистрофический тип БЭ развивается вследствие мутаций только в одном гене, *COL7A1*, который кодирует коллаген типа VII – основной белок якорных фибрилл,

обеспечивающих прикрепление БМ к подлежащей дерме. Патогенные варианты в гене *COL7A1* обуславливают нарушение выработки и молекулярной структуры коллагена, вызывая расщепление верхних слоев дермы и разрушение якорных фибрилл. Природа и местонахождение патогенных вариантов – важные детерминанты фенотипа (Hovnanian et al., 1997), который определяется экспрессией и остаточной функцией коллагена VII (Mariath et al., 2020).

Коллаген типа VII – нефибриллярный коллаген, синтезируется как эпидермальными кератиноцитами, так и дермальными фибробластами и локализуется в зоне БМ ниже слоев эпителия, представляя собой гомотример, состоящий из трех идентичных $\alpha 1$ -полипептидных цепей (Uitto et al., 1992). Каждая полипептидная $\alpha 1$ -цепь содержит центральный тройной спиральный коллагеновый домен и концевые неколлагеновые домены NC-1 и NC-2 (Chung, Uitto, 2010). Тройной спиральный домен состоит из повторяющейся последовательности Gly-X-Y, которая прерывается неколлагеновыми регионами, самый большой из них состоит из 39 аминокислотных остатков и известен как «шарнирная» область.

Домен NC-1 опосредует прикрепление якорных фибрилл к базальной мембране и островкам коллагена IV в дерме (Bruckner-Tuderman et al., 2013). Домен NC-2 содержит консервативные цистеины, участвующие в образовании дисульфидных связей, которые обеспечивают взаимодействие между гомотримерами коллагена типа VII. Кроме того, образованные якорными фибриллами петли в сосочковом слое дермы захватывают и механически удерживают волокна интерстициального коллагена, представляющего главным образом коллагенами типов I, III и V.

Коллаген типа VII способствует миграции кератиноцитов, что является одним из этапов заживления ран, обеспечивающим их реэпителизацию (Woodley et al., 2008). Показано, что при ДБЭ уменьшены размеры или количество якорных фибрилл либо они отсутствуют (Uitto, Christiano, 1992), что определяет основной механизм и тяжесть развития клинических проявлений. Нарушение функции коллагена типа VII приводит к глубоким дефектам кожи, рубцеванию слизистых оболочек, образованию милиумов и фиброзу.

Известны сотни мутаций гена *COL7A1*, ассоциированных с ДБЭ (Sawamura et al., 2005; Has et al., 2020a). Так, большинство случаев ДДБЭ – результат доминантно-негативных мутаций. Около 75 % пациентов с ДДБЭ имеют варианты замены глицина в тройном спиральном домене Gly-X-Y, особенно в экзонах 73, 74 и 75 (Varki et al., 2007). В этой «горячей» точке замены остатков глицина могут приводить к более выраженной дестабилизации белка, чем замена глициновых остатков внутри длинного непрерывного коллагенового сегмента, а варианты, находящиеся рядом с «шарнирным» регионом, вызывают нарушение сворачивания молекул белка и их накопление внутри клеток (Chen et al., 2001). Предполагается, что экзон 73 может кодировать аминокислотные остатки, важные для способности коллагена типа VII обеспечивать подвижность кератиноцитов (Woodley et al., 2008).

Глициновые и другие аминокислотные замены и варианты нарушения сплайсинга за пределами Gly-X-Y об-

ласти, также обнаруживаются при ДДБЭ, а внутрисемейная фенотипическая изменчивость позволяет считать, что другие факторы могут влиять на устойчивость клеток к трению (Koss-Harnes et al., 2002).

Тяжелый генерализованный РДБЭ обычно возникает из-за отсутствия продукта гена *COL7A1* в результате нулевых генетических вариантов на обоих аллелях, около 30 % которых являются нонсенс-вариантами с образованием стоп-кодона или вариантов сплайсинга, приводящих к крупным делециям, определяя тяжесть заболевания (van den Akker et al., 2011). Многие пациенты с РДБЭ средней тяжести – компаунд-гетерозиготы по преждевременному стоп-кодону и замещению глицина в коллагеновом домене, другому миссенс-варианту или вариантам, нарушающим сплайсинг, приводят к дестабилизации тройной спирали или к конформационным изменениям белка, что отражается на его функциональности (Pânzaru et al., 2022).

Такое разнообразие комбинаций генетических вариантов объясняет широкий спектр клинических проявлений. Так, например, варианты p.Gly2049Glu и p.Arg2063Trp, примыкающие к «шарнирному» региону, снижают способность поддерживать адгезию фибробластов и приводят к значительно сниженной способности поддерживать миграцию кератиноцитов, что замедляет заживление эрозий у больных РДБЭ (Varki et al., 2007). Более легкие формы РДБЭ часто обусловлены комбинацией вариантов нарушения сплайсинга и миссенс-вариантов. Замены глицина также могут встречаться при РДБЭ.

Синдром Киндлера (СК) – редкий тип БЭ, характеризуется хрупкостью кожи и образованием акральных пузырей с рождения, развитием атрофии кожи, фоточувствительностью, пойкилодермией, диффузным ладонно-подошвенным гиперкератозом и псевдосиндактилиями (Lai-Cheong, McGrath, 2022). Морфологически СК отличается от других типов БЭ тем, что образование пузырей вариативно и может происходить на разных уровнях дермо-эпидермального соединения. Синдром Киндлера развивается в результате патогенных вариантов в гене *FERMT1*. Заболевание наследуется по АР типу.

Ген *FERMT1* кодирует белок киндлин-1, являющийся мультидоменным белком фокальной адгезии. Киндлин-1 участвует в соединении между актиновым цитоскелетом и внеклеточным матриксом посредством фокальной адгезии, а также в интегрин-ассоциированных сигнальных путях (Has et al., 2011). Отсутствие киндлина-1 приводит к дезорганизации кератиноцитов в результате некорректной интегрин-опосредованной клеточной адгезии и миграции (Rognoni et al., 2016). В гене *FERMT1* зарегистрировано более 90 патогенных вариантов с потерей функции, включая: миссенс-, нонсенс-варианты и варианты сплайсинга; инсерции и Alu-опосредованные генные перестройки, которые приводят к отсутствию белка киндлина-1 или продукции нефункционального белка (Lai-Cheong, McGrath, 2022). Факторы окружающей среды играют важную роль в фенотипическом разнообразии СК и обуславливают тяжесть течения заболевания. X. Zhang с коллегами (2017) предположили, что гомолог 1 семейства фермитина важен для подавления вызванного ультрафиолетовым излучением воспаления и репарации ДНК.

Заключение

Мультисистемность проявлений БЭ и вовлечение в патогенез значительного количества белков, обеспечивающих механическую устойчивость кожи, обусловлены его генетической гетерогенностью. Патогенные варианты в генах, кодирующих белки крепящих комплексов эпидермиса и дермы, а также сигнальных белков, определяющих целостность кожи, приводят к их структурным и функциональным дефектам. Буллезный эпидермолиз характеризуется выраженным клиническим разнообразием и одновременно схожими проявлениями при разных генотипах. Изучение и накопление данных о естественном течении заболевания, корреляциях генотип-фенотип вносят вклад в понимание патогенеза БЭ и определяют развитие методов симптоматической и этиопатогенетической, в частности генной, терапии.

Список литературы / References

- Arin M.J., Grimberg G., Schumann H., de Almeida H. Jr., Chang Y.-R., Tadini G., Kohlhasse J., Krieg T., Bruckner-Tuderman L., Has C. Identification of novel and known *KRT5* and *KRT14* mutations in 53 patients with epidermolysis bullosa simplex: correlation between genotype and phenotype. *Br. J. Dermatol.* 2010;162(6):1365-1369. DOI 10.1111/j.1365-2133.2010.09657.x.
- Asaka T., Akiyama M., Domon T., Nishie W., Natsuga K., Fujita Y., Abe R., Kitagawa Y., Shimizu H. Type XVII collagen is a key player in tooth enamel formation. *Am. J. Pathol.* 2009;174(1):91-100. DOI 10.2353/AJPATH.2009.080573.
- Aumailley M., el Khal A., Knöss N., Tunggal L. Laminin 5 processing and its integration into the ECM. *Matrix Biol.* 2003;22(1):49-54. DOI 10.1016/S0945-053X(03)00013-1.
- Bardhan A., Bruckner-Tuderman L., Chapple I.L.C., Fine J.-D., Harper N., Has C., Magin T.M., Marinkovich M.P., Marshall J.F., McGrath J.A., Mellerio J.E., Polson R., Heagerty A.H. Epidermolysis bullosa. *Nat. Rev. Dis. Primers.* 2020;6(1):78. DOI 10.1038/s41572-020-0210-0.
- Bolling M.C., Lemmink H.H., Jansen G.H.L., Jonkman M.F. Mutations in *KRT5* and *KRT14* cause epidermolysis bullosa simplex in 75 % of the patients. *Br. J. Dermatol.* 2011;164(3):637-644. DOI 10.1111/j.1365-2133.2010.10146.x.
- Bruckner-Tuderman L., McGrath J.A., Robinson E.C., Uitto J. Progress in Epidermolysis bullosa research: summary of DEBRA International Research Conference 2012. *J. Invest. Dermatol.* 2013;133(9):2121-2126. DOI 10.1038/jid.2013.127.
- Bunick C.G., Milstone L.M. The X-ray crystal structure of the keratin 1-keratin 10 helix 2B heterodimer reveals molecular surface properties and biochemical insights into human skin disease. *J. Invest. Dermatol.* 2017;137(1):142-150. DOI 10.1016/j.jid.2016.08.018.
- Chen M., Keene D.R., Costa F.K., Tahk S.H., Woodley D.T. The carboxyl terminus of type VII collagen mediates antiparallel dimer formation and constitutes a new antigenic epitope for epidermolysis bullosa acquisita autoantibodies. *J. Biol. Chem.* 2001;276(24):21649-21655. DOI 10.1074/JBC.M100180200.
- Chung H.J., Uitto J. Type VII collagen: the anchoring fibril protein at fault in dystrophic epidermolysis bullosa. *Dermatol. Clin.* 2010;28(1):93-105. DOI 10.1016/J.DET.2009.10.011.
- Condrat I., He Y., Cosgarea R., Has C. Junctional epidermolysis bullosa: allelic heterogeneity and mutation stratification for precision medicine. *Front. Med. (Lausanne).* 2019;5:363. DOI 10.3389/fmed.2018.00363.
- Dhanoa B.S., Cogliati T., Satish A.G., Bruford E.A., Friedman J.S. Update on the Kelch-like (KLHL) gene family. *Hum. Genomics.* 2013;7(1):13. DOI 10.1186/1479-7364-7-13.
- Dogic D., Rousselle P., Aumailley M. Cell adhesion to laminin 1 or 5 induces isoform-specific clustering of integrins and other focal

- adhesion components. *J. Cell Sci.* 1998;111(Pt. 6):793-802. DOI 10.1242/JCS.111.6.793.
- Fine J.-D. Inherited epidermolysis bullosa. *Orphanet J. Rare Dis.* 2010; 5:12. DOI 10.1186/1750-1172-5-12.
- Ganani D., Malovitski K., Sarig O., Gat A., Sprecher E., Samuelov L. Epidermolysis bullosa simplex due to bi-allelic *DST* mutations: Case series and review of the literature. *Pediatr. Dermatol.* 2021;38(2): 436-441. DOI 10.1111/pde.14477.
- Gostyńska K.B., Nijenhuis M., Lemmink H., Pas H.H., Pasmooij A.M.G., Lang K.K., Castañón M.J., Wiche G., Jonkman M.F. Mutation in exon 1a of *PLEC*, leading to disruption of plectin isoform 1a, causes autosomal-recessive skin-only epidermolysis bullosa simplex. *Hum. Mol. Genet.* 2015;24(11):3155-3162. DOI 10.1093/hmg/ddv066.
- Grilletta E.A. Cardiac transplant for epidermolysis bullosa simplex with KLHL24 mutation-associated cardiomyopathy. *JAAD Case Rep.* 2019;5(10):912-914. DOI 10.1016/j.jcdr.2019.08.009.
- Has C., Bauer J.W., Bodemer C., Bolling M.C., Bruckner-Tuderman L., Diem A., Fine J.-D., Heagerty A., Hovnanian A., Marinkovich M.P., Martinez A.E., McGrath J.A., Moss C., Murrell D.F., Palissson F., Schwieger-Briel A., Sprecher E., Tamai K., Uitto J., Woodley D.T., Zambruno G., Mellerio J.E. Consensus reclassification of inherited epidermolysis bullosa and other disorders with skin fragility. *Br. J. Dermatol.* 2020a;183(4):614-627. DOI 10.1111/bjd.18921.
- Has C., Bruckner-Tuderman L. The genetics of skin fragility. *Annu. Rev. Genomics Hum. Genet.* 2014;15(1):245-268. DOI 10.1146/annurev-genom-090413-025540.
- Has C., Castiglia D., del Rio M., Garcia Diez M., Piccinni E., Kiritsi D., Kohlhase J., Itin P., Martin L., Fischer J., Zambruno G., Bruckner-Tuderman L. Kindler syndrome: Extension of FERMT1 mutational spectrum and natural history. *Hum. Mutat.* 2011;32(11):1204-1212. DOI 10.1002/HUMU.21576.
- Has C., Chang Y.-R., Volz A., Hoeping D., Kohlhase J., Bruckner-Tuderman L. Novel keratin 14 mutations in patients with severe recessive epidermolysis bullosa simplex. *J. Invest. Dermatol.* 2006; 126(8):1912-1914. DOI 10.1038/sj.jid.5700312.
- Has C., Fischer J. Inherited epidermolysis bullosa: New diagnostics and new clinical phenotypes. *Exp. Dermatol.* 2019;28(10):1146-1152. DOI 10.1111/exd.13668.
- Has C., Nyström A. Epidermal basement membrane in health and disease. *Curr. Top. Membr.* 2015;76:117-170. DOI 10.1016/bs.ctm.2015.05.003.
- Has C., South A., Uitto J. Molecular therapeutics in development for epidermolysis bullosa: Update 2020. *Mol. Diagn. Ther.* 2020b; 24(3):299-309. DOI 10.1007/s40291-020-00466-7.
- Has C., Sparta G., Kiritsi D., Weibel L., Moeller A., Vega-Warner V., Waters A., He Y., Ankster Y., Esser P., Straub B.K., Hausser I., Bockenhauer D., Dekel B., Hildebrandt F., Bruckner-Tuderman L., Laube G.F. Integrin $\alpha 3$ mutations with, lung, and skin disease. *N. Engl. J. Med.* 2012;366(16):1508-1514. DOI 10.1056/NEJM0A1110813.
- Hovnanian A., Rochat A., Bodemer C., Petit E., Rivers C.A., Prost C., Fraitag S., Christiano A.M., Uitto J., Lathrop M., Barrandon Y., de Prost Y. Characterization of 18 new mutations in COL7A1 in recessive dystrophic epidermolysis bullosa provides evidence for distinct molecular mechanisms underlying defective anchoring fibril formation. *Am. J. Hum. Genet.* 1997;61(3):599-610. DOI 10.1086/515495.
- Karamatic Crew V., Burton N., Kagan A., Green C.A., Levene C., Flinter F., Brady R.L., Daniels G., Anstee D.J. CD151, the first member of the tetraspanin (TM4) superfamily detected on erythrocytes, is essential for the correct assembly of human basement membranes in kidney and skin. *Blood.* 2004;104(8):2217-2223. DOI 10.1182/blood-2004-04-1512.
- Kiritsi D., Has C., Bruckner-Tuderman L. Laminin 332 in junctional epidermolysis bullosa. *Cell Adh. Migr.* 2013;7(1):135-141. DOI 10.4161/CAM.22418.
- Kiritsi D., Kern J.S., Schumann H., Kohlhase J., Has C., Bruckner-Tuderman L. Molecular mechanisms of phenotypic variability in junctional epidermolysis bullosa. *J. Med. Genet.* 2011;48(7):450-457. DOI 10.1136/JMG.2010.086751.
- Kiritsi D., Tsakiris L., Schauer F. Plectin in skin fragility disorders. *Cells.* 2021;10(10):2738. DOI 10.3390/cells10102738.
- Koss-Harnes D., Høyheim B., Anton-Lamprecht I., Gjesti A., Jørgensen R.S., Jahnsen F.L., Olaisen B., Wiche G., Gedde-Dahl T. A site-specific plectin mutation causes dominant epidermolysis bullosa simplex Ogná: two identical *de novo* mutations. *J. Invest. Dermatol.* 2002;118(1):87-93. DOI 10.1046/j.0022-202x.2001.01591.x.
- Kumar V., Bouameur J.E., Bär J., Rice R.H., Hornig-Do H.T., Roop D.R., Schwarz N., Brodessaer S., Thiering S., Leube R.E., Wiesner R.J., Brazel C.B., Heller S., Binder H., Löffler-Wirth H., Seibel P., Margin T.M. A keratin scaffold regulates epidermal barrier formation, mitochondrial lipid composition, and activity. *J. Cell Biol.* 2015; 211(5):1057-1075. DOI 10.1083/JCB.201404147.
- Lai-Cheong J.E., McGrath J.A. Kindler syndrome. In: Murrell D. (Ed.). Blistering Diseases: Clinical Features, Pathogenesis, Treatment. Berlin; Heidelberg: Springer, 2022;433-439. DOI 10.1007/978-3-662-45698-9_43.
- Liu Y., Yue Z., Wang H., Li M., Wu X., Lin H., Han W., Lan S., Sun L. A novel *ITGA3* homozygous splice mutation in an ILNEB syndrome child with slow progression. *Clin. Chim. Acta.* 2021;523:430-436. DOI 10.1016/J.CCA.2021.10.027.
- Margadant C., Charafeddine R.A., Sonnenberg A. Unique and redundant functions of integrins in the epidermis. *FASEB J.* 2010;24(11): 4133-4152. DOI 10.1096/fj.09-151449.
- Mariath L.M., Santin J.T., Frantz J.A., Doriqi M.J.R., Schuler-Faccini L., Kiszewski A.E. Genotype-phenotype correlations on epidermolysis bullosa with congenital absence of skin: A comprehensive review. *Clin. Genet.* 2021;99(1):29-41. DOI 10.1111/cge.13792.
- Mariath L.M., Santin J.T., Schuler-Faccini L., Kiszewski A.E. Inherited epidermolysis bullosa: update on the clinical and genetic aspects. *An. Bras. Dermatol.* 2020;95(5):551-569. DOI 10.1016/j.abd.2020.05.001.
- Masunaga T., Ogawa J., Akiyama M., Nishikawa T., Shimizu H., Ishiko A. Compound heterozygosity for novel splice site mutations of *ITGA6* in lethal junctional epidermolysis bullosa with pyloric atresia. *J. Dermatol.* 2017;44(2):160-166. DOI 10.1111/1346-8138.13575.
- Matsumura H., Mohri Y., Thanh Binh N., Morinaga H., Fukuda M., Ito M., Kurata S., Hoeijmakers J., Nishimura E.K. Hair follicle aging is driven by transepidermal elimination of stem cells via COL17A1 proteolysis. *Science.* 2016;351(6273):aad4395. DOI 10.1126/science.aad4395.
- Murrell D.F., Pasmooij A.M.G., Pas H.H., Marr P., Klingberg S., Pfindner E., Uitto J., Sadowski S., Collins F., Widmer R., Jonkman M.F. Retrospective diagnosis of fatal BP180-deficient non-Herlitz junctional epidermolysis bullosa suggested by immunofluorescence (IF) antigen-mapping of parental carriers bearing enamel defects. *J. Invest. Dermatol.* 2007;127(7):1772-1775. DOI 10.1038/SJ.JID.5700766.
- Natsuga K. Plectin-related skin diseases. *J. Dermatol. Sci.* 2015;77(3): 139-145. DOI 10.1016/j.jdermsci.2014.11.005.
- Natsuga K., Nishie W., Shinkuma S., Arita K., Nakamura H., Ohyama M., Osaka H., Kambara T., Hirako Y., Shimizu H. Plectin deficiency leads to both muscular dystrophy and pyloric atresia in epidermolysis bullosa simplex. *Hum. Mutat.* 2010;31(10):E1687-E1698. DOI 10.1002/humu.21330.
- Pânzaru M.C., Caba L., Florea L., Braha E.E., Gorduză E.V. Epidermolysis bullosa – a different genetic approach in correlation with genetic heterogeneity. *Diagnostics.* 2022;12(6):1325. DOI 10.3390/diagnostics12061325.
- Pasmooij A.M.G., van der Steege G., Pas H.H., Sillevs Smitt J.H., Nijenhuis A.M., Zuiderveen J., Jonkman M.F. Features of epidermolysis bullosa simplex due to mutations in the ectodomain of type XVII collagen. *Br. J. Dermatol.* 2004;151(3):669-674. DOI 10.1111/J.1365-2133.2004.06041.X.
- Prodinger C., Chottianchaiwat S., Mellerio J.E., McGrath J.A., Ozoe-menă L., Liu L., Moore W., Laimer M., Petrof G., Martinez A.E.

- The natural history of laryngo-onycho-cutaneous syndrome: A case series of six pediatric patients and literature review. *Pediatr. Dermatol.* 2021;38(5):1094-1101. DOI 10.1111/PDE.14790.
- Rognoni E., Ruppert R., Fässler R. The kindlin family: functions, signaling properties and implications for human disease. *J. Cell Sci.* 2016;129(1):17-27. DOI 10.1242/JCS.161190.
- Sathishkumar D., Orrin E., Terron-Kwiatkowski A., Browne F., Martinez A.E., Mellerio J.E., Ogboli M., Hoey S., Ozoemena L., Liu L., Baty D., McGrath J.A., Moss C. The p.Glu477Lys mutation in keratin 5 is strongly associated with mortality in generalized severe epidermolysis bullosa simplex. *J. Invest. Dermatol.* 2016;136(3):719-721. DOI 10.1016/j.jid.2015.11.024.
- Sawamura D., Goto M., Yasukawa K., Sato-Matsumura K., Nakamura H., Ito K., Nakamura H., Tomita Y., Shimizu H. Genetic studies of 20 Japanese families of dystrophic epidermolysis bullosa. *J. Hum. Genet.* 2005;50(10):543-546. DOI 10.1007/S10038-005-0290-4.
- Schumann H., Kiritsi D., Pigors M., Hausser I., Kohlhase J., Peters J., Ott H., Hyla-Klekot L., Gacka E., Sieron A.L., Valari M., Bruckner-Tuderman L., Has C. Phenotypic spectrum of epidermolysis bullosa associated with $\alpha\beta 4$ integrin mutations. *Br. J. Dermatol.* 2013;169(1):115-124. DOI 10.1111/bjd.12317.
- Tong G., Xu R. The role of collagen XVII in regulating keratinocyte migration. *Lab. Invest.* 2004;84(10):1225-1226. DOI 10.1038/labinvest.3700168.
- Uitto J., Bruckner-Tuderman L., Christiano A.M., McGrath J.A., Has C., South A.P., Kopelan B., Robinson E.C. Progress toward treatment and cure of epidermolysis bullosa: Summary of the DEBRA international research symposium EB2015. *J. Invest. Dermatol.* 2016;136(2):352-358. DOI 10.1016/j.jid.2015.10.050.
- Uitto J., Christiano A.M. Molecular genetics of the cutaneous basement membrane zone. Perspectives on epidermolysis bullosa and other blistering skin diseases. *J. Clin. Invest.* 1992;90(3):687-692. DOI 10.1172/JCI115938.
- Uitto J., Chung-Honet L.C., Christiano A.M. Molecular biology and pathology of type VII collagen. *Exp. Dermatol.* 1992;1(1):2-11. DOI 10.1111/J.1600-0625.1992.TB00065.X.
- Uitto J., Has C., Vahidnezhad H., Youssefian L., Bruckner-Tuderman L. Molecular pathology of the basement membrane zone in heritable blistering diseases: The paradigm of epidermolysis bullosa. *Matrix Biol.* 2017;57-58:76-85. DOI 10.1016/j.matbio.2016.07.009.
- Vahidnezhad H., Youssefian L., Saeidian A.H., Mozafari N., Barzegar M., Sotoudeh S., Daneshpazhooh M., Isaian A., Zeinali S., Uitto J. *KRT5* and *KRT14* mutations in epidermolysis bullosa simplex with phenotypic heterogeneity, and evidence of semidominant inheritance in a multiplex family. *J. Invest. Dermatol.* 2016;136(9):1897-1901. DOI 10.1016/j.jid.2016.05.106.
- Vahidnezhad H., Youssefian L., Saeidian A.H., Uitto J. Phenotypic spectrum of epidermolysis bullosa: The paradigm of syndromic versus non-syndromic skin fragility disorders. *J. Invest. Dermatol.* 2019;139(3):522-527. DOI 10.1016/j.jid.2018.10.017.
- van den Akker P.C., Jonkman M.F., Rengaw T., Bruckner-Tuderman L., Has C., Bauer J.W., Klaussegger A., Zambruno G., Castiglia D., Mellerio J.E., McGrath J.A., van Essen A.J., Hofstra R.M.W., Swertz M.A. The international dystrophic epidermolysis bullosa patient registry: an online database of dystrophic epidermolysis bullosa patients and their *COL7A1* mutations. *Hum. Mutat.* 2011;32(10):1100-1107. DOI 10.1002/humu.21551.
- van den Bergh F., Giudice G.J. BP180 (type XVII collagen) and its role in cutaneous biology and disease. *Adv. Dermatol.* 2003;19:37-71.
- Varki R., Sadowski S., Pfindner E., Uitto J. Epidermolysis bullosa. I. Molecular genetics of the junctional and hemidesmosomal variants. *J. Med. Genet.* 2006;43(8):641-652. DOI 10.1136/JMG.2005.039685.
- Varki R., Sadowski S., Uitto J., Pfindner E. Epidermolysis bullosa. II. Type VII collagen mutations and phenotype-genotype correlations in the dystrophic subtypes. *J. Med. Genet.* 2007;44(3):181-192. DOI 10.1136/JMG.2006.045302.
- Wang R., Sun L., Habulieti X., Liu J., Guo K., Yang X., Ma D., Zhang X. Novel variants in *LAMA3* and *COL7A1* and recurrent variant in *KRT5* underlying epidermolysis bullosa in five Chinese families. *Front. Med.* 2022;16(5):808-814. DOI 10.1007/S11684-021-0878-X.
- Wertheim-Tysarowska K., Ołdak M., Giza A., Kutkowska-Kaźmierczak A., Sota J., Przybylska D., Woźniak K., Śniegórska D., Niepokój K., Sobczyńska-Tomaszewska A., Rygiel A.M., Płoski R., Bał J., Kowalewski C. Novel sporadic and recurrent mutations in *KRT5* and *KRT14* genes in Polish epidermolysis bullosa simplex patients: further insights into epidemiology and genotype-phenotype correlation. *J. Appl. Genet.* 2016;57(2):175-181. DOI 10.1007/s13353-015-0310-9.
- Woodley D.T., Hou Y., Martin S., Li W., Chen M. Characterization of molecular mechanisms underlying mutations in dystrophic epidermolysis bullosa using site-directed mutagenesis. *J. Biol. Chem.* 2008;283(26):17838-17845. DOI 10.1074/JBC.M709452200.
- Wright J.T., Carrion I.A., Morris C. The molecular basis of hereditary enamel defects in humans. *J. Dent. Res.* 2015;94(1):52-61. DOI 10.1177/0022034514556708.
- Zhang X., Luo S., Wu J., Zhang L., Wang W.-hui, Degan S., Erdmann D., Hall R., Zhang J.Y. *KIND1* loss sensitizes keratinocytes to UV-induced inflammatory response and DNA damage. *J. Invest. Dermatol.* 2017;137(2):475-483. DOI 10.1016/J.JID.2016.09.023.

ORCID ID

Yu.Yu. Kotalevskaya orcid.org/0000-0001-8405-8223
V.A. Stepanov orcid.org/0000-0002-5166-331X

Благодарности. Работа выполнена при поддержке Министерства науки и высшего образования Российской Федерации (Федеральная научно-техническая программа развития генетических технологий на 2019–2027 гг., соглашение № 075-15-2021-1061, РФ 193021X0029).

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию 17.10.2022. После доработки 10.12.2022. Принята к публикации 20.12.2022.

Перевод на английский язык <https://vavilov.elpub.ru/jour>

Сравнительная цитогенетика анэмбрионии и неразвивающейся беременности у человека

Т.В. Никитина , Е.А. Саженова, Е.Н. Толмачева, Н.Н. Суханова, С.А. Васильев, И.Н. Лебедев

Научно-исследовательский институт медицинской генетики, Томский национальный исследовательский медицинский центр Российской академии наук, Томск, Россия
 t.nikitina@medgenetics.ru

Аннотация. Невынашивание беременности является серьезной проблемой в репродукции человека, затрагивающей 10–15 % клинически распознаваемых беременностей. Среди случаев эмбриональной гибели можно выделить замершие (неразвивающиеся) беременности (НБ), при которых ультразвуковым признаком гибели эмбриона служит отсутствие сердцебиения, и анэмбрионии (АЭ) – отсутствие эмбриона в полости плодного мешка. Целью данного исследования было сравнение частоты хромосомных аномалий во внезародышевых тканях, выявляемых при стандартном цитогенетическом анализе материала спонтанных абортосов, в зависимости от наличия или отсутствия эмбриона. Проведено ретроспективное исследование 1551 спонтанного абортоса, проанализированного с помощью стандартного цитогенетического исследования с 1990 по 2022 г. (266 случаев АЭ и 1285 случаев НБ) в НИИ медицинской генетики Томского НИМЦ. Выполнен сравнительный анализ частоты хромосомных аномалий и распределения частот кариотипов в зависимости от наличия эмбриона в полости плодного мешка. Статистический анализ проводили с использованием критерия хи-квадрат с уровнем значимости $p < 0.05$. Суммарно частота хромосомных аномалий в исследованной выборке составила 53.6 % (832/1551). Доля аномальных кариотипов в группах АЭ и НБ значимо не различалась и составила 57.1 % (152/266) и 52.9 % (680/1285) для АЭ и НБ соответственно ($p = 0.209$). При НБ статистически значимо чаще встречались аномалии числа половых хромосом (6.8 % (88/1285) против 2.3 % (6/266), $p = 0.005$) и триплоидии (8.9 % (114/1285) против 4.9 % (13/266), $p = 0.031$). В то же время при отсутствии эмбриона статистически значимо чаще регистрировалась тетраплоидия (12.4 % (33/266) против 8.2 % (106/1285), $p = 0.031$). Соотношение полов (46,XY:46,XX) среди абортосов с нормальным кариотипом составило 0.54 и 0.74 для АЭ и НБ соответственно. Таким образом, хотя частоты некоторых типов хромосомных аномалий различаются между АЭ и НБ, суммарная частота хромосомных аномалий при АЭ не повышена по сравнению с НБ, что свидетельствует о необходимости поиска причин АЭ на других уровнях организации генома, включая микроструктурные перестройки хромосом, митохондриальные мутации, нарушения импринтинга и aberrантные эпигенетические модификации.
Ключевые слова: анэмбриония; неразвивающаяся беременность; невынашивание беременности; кариотип; хромосомные аномалии; половые хромосомы; триплоидия; тетраплоидия.

Для цитирования: Никитина Т.В., Саженова Е.А., Толмачева Е.Н., Суханова Н.Н., Васильев С.А., Лебедев И.Н. Сравнительная цитогенетика анэмбрионии и неразвивающейся беременности у человека. *Вавиловский журнал генетики и селекции*. 2023;27(1):28-35. DOI 10.18699/VJGB-23-05

Comparative cytogenetics of anembryonic pregnancies and missed abortions in human

T.V. Nikitina , E.A. Sazhenova, E.N. Tolmacheva, N.N. Sukhanova, S.A. Vasilyev, I.N. Lebedev

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

Abstract. Miscarriage is an important problem in human reproduction, affecting 10–15 % of clinically recognized pregnancies. The cases of embryonic death can be divided into missed abortion (MA), for which the ultrasound sign of the embryo death is the absence of cardiac activity, and anembryonic pregnancy (AP) without an embryo in the gestational sac. The aim of this study was to compare the frequency of chromosomal abnormalities in extraembryonic tissues detected by conventional cytogenetic analysis of spontaneous abortions depending on the presence or absence of an embryo. This is a retrospective study of 1551 spontaneous abortions analyzed using GTG-banding from 1990 to 2022 (266 cases of AP and 1285 cases of MA). A comparative analysis of the frequency of chromosomal abnormalities and the distribution of karyotype frequencies depending on the presence of an embryo in the gestational sac was carried out. Statistical analysis was performed using a chi-square test with a $p < 0.05$ significance level. The total frequency of chromosomal abnormalities in the study was 53.6 % (832/1551). The proportion of abnormal karyotypes in the AP and MA groups did not differ significantly and amounted to 57.1 % (152/266) and 52.9 % (680/1285) for AP and MA, respectively ($p = 0.209$). Sex chromosome aneuploidies and triploidies were significantly less common in the AP group than in the MA group (2.3 % (6/266) vs 6.8 % (88/1285), $p = 0.005$ and 4.9 % (13/266) vs 8.9 % (114/1285), $p = 0.031$, re-

spectively). Tetraploidies were registered more frequently in AP compared to MA (12.4 % (33/266) vs. 8.2 % (106/1285), $p=0.031$). The sex ratio among abortions with a normal karyotype was 0.54 and 0.74 for AP and MA, respectively. Thus, although the frequencies of some types of chromosomal pathology differ between AP and MA, the total frequency of chromosomal abnormalities in AP is not increased compared to MA, which indicates the need to search for the causes of AP at other levels of the genome organization, including microstructural chromosomal rearrangements, monogenic mutations, imprinting disorders, and epigenetic abnormalities.

Key words: anembryonic pregnancy; missed abortion; miscarriage; karyotype; chromosomal abnormalities; sex chromosomes; triploidy; tetraploidy.

For citation: Nikitina T.V., Sazhenova E.A., Tolmacheva E.N., Sukhanova N.N., Vasilyev S.A., Lebedev I.N. Comparative cytogenetics of anembryonic pregnancies and missed abortions in human. *Vavilovskii Zhurnal Genetiki i Seleksii = Vavilov Journal of Genetics and Breeding*. 2023;27(1):28-35. DOI 10.18699/VJGB-23-05

Введение

Одна из основных проблем в репродукции человека – невынашивание беременности, в результате которого от 10 до 15 % всех клинически распознаваемых беременностей заканчиваются гибелью эмбриона или плода (Larsen et al., 2013). Цитогенетические исследования обнаруживают хромосомные аномалии в среднем у 50–60 % абортусов первого триместра беременности (Menasha et al., 2005; van den Berg et al., 2012; Hardy et al., 2016; Soler et al., 2017; Wang et al., 2020; Wu et al., 2021), также в последние годы появляется все больше данных о связи невынашивания беременности с вариациями числа копий ДНК (CNV), генными мутациями, аномалиями метилирования и другими эпигенетическими aberrациями (Levy et al., 2014; Fu et al., 2018; Fan et al., 2020; Finley et al., 2022). Установление причины гибели эмбриона помогает оценить риск при последующих беременностях, кроме того, обнаружение патологического фактора важно для психологического состояния родителей.

Одной из самых ранних форм невынашивания беременности является анэмбриония – отсутствие эмбриона в полости плодного мешка. При анэмбрионии происходит имплантация бластоцисты в стенку матки, формируется плодный мешок, однако собственно эмбрион либо не развивается изначально, либо его формирование останавливается на самых ранних стадиях (не позднее 5-й недели гестации), и далее продолжают пролиферация и рост только внезародышевых структур.

В норме приблизительно с 6-й недели беременности с помощью вагинального УЗИ в плодном мешке обнаруживаются вторичный желточный мешок и первичные зародышевые листки, а в течение 7-й недели выявляется примитивная сердечная трубка. При ранней эмбриональной гибели отмечаются следующие ультразвуковые признаки: отсутствие сердцебиения эмбриона при диаметре плодного яйца ≥ 25 мм, копчиково-теменном размере ≥ 7 мм на сроке 6 недель и более; отсутствие эмбриона и его сердцебиения через 14 дней после выявления плодного яйца без желточного мешка; отсутствие эмбриона и его сердцебиения через 11 дней после выявления плодного яйца с желточным мешком (Doubilet et al., 2013). Таким образом, УЗИ позволяет дифференцировать две формы ранней эмбриональной гибели: анэмбрионию (АЭ) и неразвивающуюся беременность (НБ). Диагноз АЭ ставится при отсутствии эмбриона и желточного мешка в полости плодного яйца на сроке более 7 недель (Радзинский и др., 2015). Кроме того, ультразвуковыми критериями анэмбрионии являются плодное яйцо более 13 мм без

желточного мешка или более 18 мм без эмбриона. Отсутствие сердцебиения при наличии эмбриона служит признаком неразвивающейся беременности.

Следует иметь в виду, что в литературе существуют терминологические противоречия, которые затрудняют сравнение результатов исследований, проведенных в разных центрах. В МКБ-10 используются термины ‘blighted ovum’ и ‘missed abortion’, принятые много лет назад (Robinson, 1975), которые при современном широком применении ультразвуковой диагностики не вполне отражают клинические особенности (Farquharson et al., 2005). Специальная группа Европейского общества репродукции человека и эмбриологии (ESHRE) предложила использовать термины ‘anembryonic (empty sac) miscarriage’ для плодного яйца диаметром ≥ 8 мм и без желточного мешка или зародыша; ‘yolk sac miscarriage’ для плодного мешка с желточным мешком, но без зародыша; ‘embryonic miscarriage’ при КТР эмбриона более 7 мм и отсутствии сердцебиения (Kolte et al., 2015). Таким образом, диагноз АЭ включает в себя как пустой плодный мешок (empty sac), так и плодный мешок с наличием желточного мешка, но без эмбриона (yolk sac only).

Оценка частоты АЭ среди потерь беременностей первого триместра различается по данным разных авторов: от 16 % в ранних работах (Robinson, 1975), 22.6 % после ЭКО (Li et al., 2017) и до 30–40 % в большинстве исследований (Lathi et al., 2007; Cheng et al., 2014; Ouyang et al., 2016; Yoneda et al., 2018). Несмотря на распространенность АЭ, данные о частоте хромосомных аномалий при этой патологии противоречивы. Интуитивно представляется, что ранние и грубые нарушения, которые приводят к остановке развития собственно эмбриона на начальных этапах его формирования, должны быть связаны со значительно повышенной частотой и тяжестью хромосомных аномалий. Есть исследования, в которых такая зависимость обнаружена (Angiolucci et al., 2011). В то же время большинство работ, проведенных в последние годы, демонстрируют либо отсутствие значимых различий в частоте хромосомных аномалий между АЭ и НБ (Lathi et al., 2007; Muñoz et al., 2010; Ljunger et al., 2011; Liu et al., 2015), либо даже меньшую частоту аномальных кариотипов среди АЭ по сравнению с НБ (Ginsberg et al., 2001; Cheng et al., 2014; Li et al., 2017; Yoneda et al., 2018; Gu et al., 2021). Это противоречие обусловило актуальность проведения сравнительного исследования на значительной по размеру, в сравнении с опубликованными литературными данными, выборке случаев АЭ и НБ. В настоящей работе были исследованы частота и спектр хромосомных аномалий,

выявленных при стандартном цитогенетическом анализе 1551 случая эмбриональной гибели, в зависимости от наличия или отсутствия эмбриона.

Материалы и методы

Объектом исследования был 1551 спонтанный абортус, прокаротипированный в лаборатории цитогенетики НИИ медицинской генетики Томского НИМЦ. Материал поступал из гинекологических клиник городов Томска и Северска. Предоставлялась информация о возрасте супругов, акушерско-гинекологическом анамнезе женщины, числе и исходе предыдущих беременностей. Исследование одобрено Комитетом по биомедицинской этике НИИ медицинской генетики Томского НИМЦ.

В большинстве случаев кариотип был установлен с помощью стандартного цитогенетического исследования после культивирования экстраэмбриональных фибробластов (90.9 %, 1410 образцов) либо прямых препаратов ворсин хориона (1.9 %, 29 образцов). При невозможности провести кариотипирование применялись интерфазная FISH (флуоресцентная гибридизация *in situ*) (3.7 %, 58 образцов) и CGH (сравнительная геномная гибридизация) (3.5 %, 54 образца). С диагнозом АЭ (отсутствие эмбриона в полости плодного яйца на сроке более 7 недель, плодное яйцо более 13 мм без желточного мешка или более 18 мм без эмбриона) в лабораторию поступили 266 (17.2 %) абортусов, а 1285 абортусов (82.8 %) с наличием эмбриона были отнесены в группу НБ.

Операционный материал, обычно представлявший собой фрагменты плодного мешка, доставлялся в лабораторию в стерильном физиологическом растворе, тщательно отмывали от крови и отделяли от децидуальных тканей. Методы культивирования эмбриональных клеток, получения хромосомных препаратов, G-окраски метафазных хромосом, FISH и CGH описаны в (Лебедев, Никитина, 2013).

Статистическую значимость различий между частотами вычисляли с помощью критерия χ^2 . Проверку на нормаль-

ность распределения количественных показателей выполняли с использованием критерия Колмогорова–Смирнова. Вследствие выявленных отличий от нормального распределения сравнение групп проводили с применением непараметрического критерия Манна–Уитни. Отличия принимали статистически значимыми при $p < 0.05$. Соотношение полов (СП) вычислялось как отношение кариотипов 46,XY : 46,XX. Привычное невынашивание беременности (ПНБ) определялось как две и более потери беременности в анамнезе женщины.

Исследование выполнено на базе Центра коллективного пользования научно-исследовательским оборудованием и экспериментальным биологическим материалом «Медицинская геномика» Томского НИМЦ с использованием ресурсов биобанка «Биобанк населения Северной Евразии» НИИ медицинской генетики Томского НИМЦ.

Результаты

Сравнение демографических характеристик исследованных групп абортусов приведено в табл. 1. Выборки не различались по возрасту матерей и отцов, числу беременностей и спонтанных абортусов у женщин и доле супружеских пар с привычным невынашиванием беременности (ПНБ) в анамнезе женщины. Однако срок беременности (как по дате последней менструации, так и по УЗИ) в группе АЭ был значимо меньше, чем в группе НБ.

Суммарно частота хромосомных аномалий в исследованной выборке составила 53.6 % (832/1551). Частоты кариотипов по классам приведены в табл. 2. Частота хромосомных аномалий среди потерь беременности без эмбриона и беременностей с наличием эмбриональных структур составила 57.1 % (152/266) и 52.9 % (680/1285) соответственно и значимо не различалась между группами ($p = 0.209$). В структуре хромосомных аномалий также не наблюдалось значимых отличий между АЭ и НБ по частоте аутомных трисомий (27.8 и 22.5 %), аутомных моносомий (1.5 и 0.6 %), структурных aberrаций (2.3 и 1.0 %) и комбинированных аномалий, включающих соче-

Таблица 1. Сравнение демографических показателей исследуемых групп

Показатель	Всего	АЭ	НБ	<i>p</i>
Возраст матери, лет	28.6 ± 6.23 (24.0–33.0; 28.0)	28.3 ± 6.14 (24.0–32.0; 28.0)	28.5 ± 6.18 (24.0–33.0; 28.0)	0.491
Возраст отца, лет	30.9 ± 6.69 (26.0–35.0; 30.0)	30.5 ± 6.52 (26.0–35.0; 30.0)	30.9 ± 6.64 (26.0–35.0; 30.0)	0.405
Срок беременности, недель*	9.4 ± 2.23 (8.0–11.0; 9.1)	9.0 ± 2.13 (7.5–10.1; 9.0)	9.5 ± 2.24 (8.0–11.0; 9.3)	0.001
Срок беременности по УЗИ, недель	7.3 ± 1.84 (6.0–8.5; 7.0)	6.6 ± 1.62 (5.5–7.6; 6.5)	7.5 ± 1.84 (6.0–8.5; 7.3)	<0.001
Число беременностей	2.9 ± 2.20 (1.0–4.0; 2.0)	3.2 ± 2.42 (1.0–4.5; 2.0)	2.9 ± 2.18 (1.0–4.0; 2.0)	0.085
Число спонтанных абортусов	1.6 ± 1.05 (1.0–2.0; 1.0)	1.6 ± 1.13 (1.0–2.0; 1.0)	1.6 ± 1.04 (1.0–2.0; 1.0)	0.633
ПНБ в анамнезе, %	38.7	35.5	39.4	0.278

Примечание. Для количественных показателей указаны: среднее ± стандартное отклонение, в скобках – нижний и верхний квартиль; медиана; * срок беременности, рассчитанный по дате последней менструации; значение *p* указано для сравнения АЭ и НБ. Жирным шрифтом выделены статистически значимые различия.

Таблица 2. Частоты нормальных кариотипов и хромосомных аномалий, выявленных при АЭ и НБ

Кариотип	Всего <i>n</i> = 1551	АЭ <i>n</i> = 266	НБ <i>n</i> = 1285	<i>p</i>
46,XX	422 (27.2)	74 (27.8)	348 (27.1)	0.806
46,XY	297 (19.1)	40 (15.0)	257 (20.0)	0.062
Анеуплоидии по половым хромосомам	94 (6.1)	6 (2.3)	88 (6.8)	0.005
Трисомии аутосом	363 (23.4)	74 (27.8)	289 (22.5)	0.062
Триплоидии	127 (8.2)	13 (4.9)	114 (8.9)	0.031
Тетраплоидии	139 (9.0)	33 (12.4)	106 (8.2)	0.031
Структурные aberrации	19 (1.2)	6 (2.3)	13 (1.0)	0.094
Моносомии аутосом	12 (0.8)	4 (1.5)	8 (0.6)	0.136
Другие	11 (0.7)	3 (1.1)	8 (0.6)	0.372
Сочетанные аномалии	67 (4.3)	13 (4.9)	54 (4.2)	0.618

Примечание. В скобках указаны проценты; значение *p* указано для сравнения АЭ и НБ. Жирным шрифтом выделены статистически значимые различия.

тание различных типов хромосомных aberrаций у одного абортуса (4.9 и 4.2 %) (см. табл. 2). В то же время числовые аномалии половых хромосом при АЭ встречались втрое реже, чем при наличии эмбриона (2.3 и 6.8 %, $p = 0.005$). Эта разница оказалась еще больше выражена для моносомии X: 0.8 % (2/266) и 5.0 % (64/1285), $p < 0.001$.

Кроме того, отсутствие эмбриона реже ассоциировалось с триплоидным, но чаще – с тетраплоидным кариотипом: 4.9 и 8.9 % ($p = 0.031$); 12.4 и 8.2 % ($p = 0.031$) при АЭ и НБ соответственно. Поскольку часть тетраплоидных кариотипов в мозаичной форме может представлять собой культуральные артефакты (тетраплоидизацию в культуре), были уточнены частоты для тетраплоидий, подтвержденных FISH в некультивированных тканях; при этом различия оказались даже более статистически значимыми: 14/266 (5.3 %) при АЭ против 22/1285 (1.7 %) при НБ ($p < 0.001$).

Среди абортусов с нормальным кариотипом соотношение полов составило 0.54 для АЭ и 0.74 для НБ. Значимых отличий в распределении кариотипов 46,XX и 46,XY не выявлено ($p = 0.142$).

Обсуждение

Целью нашей работы была сравнительная оценка частоты хромосомных аномалий при потерях беременности с отсутствием эмбриона (АЭ) и с его наличием (НБ). Мы обследовали значительную по размеру (1551 абортус) выборку случаев ранней эмбриональной гибели и не обнаружили значимых отличий по частоте хромосомных аномалий между этими двумя группами (57.1 и 52.9 % соответственно, $p = 0.209$). Анализ ранее опубликованных сравнительных исследований показал, что относительно корреляции между наличием эмбриона в полости плодного мешка и кариотипом получены противоречивые результаты (табл. 3).

Частота аномальных кариотипов положительно коррелировала с диагнозом АЭ по данным (Angiolucci et al., 2011), однако эта корреляция вычислялась авторами относительно погибших эмбрионов с нормальными УЗИ-признаками. Мы пересчитали ее, сравнив АЭ с остальными

группами абортусов (с наличием эмбриона), и не обнаружили статистически достоверной зависимости частоты хромосомных аномалий от наличия/отсутствия эмбриона ($p = 0.381$) (см. табл. 3). В ряде публикаций (Lathi et al., 2007; Muñoz et al., 2010; Ljunger et al., 2011; Liu et al., 2015) ассоциации частоты аномальных кариотипов с наличием/отсутствием эмбриона не выявлено, как и в настоящей работе. Однако в большинстве исследований, проведенных в последние годы на относительно больших выборках, обнаружена отрицательная корреляция между отсутствием эмбриона и хромосомными aberrациями (Cheng et al., 2014; Li et al., 2017; Yoneda et al., 2018; Gu et al., 2021). Это расхождение может быть связано с различиями в методах установления кариотипа, размерах выборок и структуре популяций (см. табл. 3). Кроме того, некоторые исследования проведены на смещенных выборках: например, охватывают женщин с бесплодием (Li et al., 2017) или из групп высокого риска по анеуплоидиям (Muñoz et al., 2010), или средний возраст матерей в выборке был более 35 лет (Ginsberg et al., 2001; Muñoz et al., 2010; Angiolucci et al., 2011). Тем не менее результаты анализа совокупности имеющихся в литературе данных начиная с 2001 г., поскольку в более ранних работах обследованы небольшие выборки, свидетельствуют о том, что при АЭ частота хромосомных аномалий ниже, чем при замершей беременности с эмбрионом: 44.4 % (813/1833) против 59.3 % (2701/4558) соответственно, $p < 0.001$ (см. табл. 3). Сходные выводы были сделаны также по результатам метаанализа ($p = 0.03$, OR = 0.68, 95 % CI = 0.48–0.97) (Huang et al., 2019).

Возможным объяснением данного феномена может быть то, что на разных стадиях раннего эмбрионального развития критичны различные типы генетических нарушений. Косвенным подтверждением этого предположения является различие в частоте хромосомных аномалий для двух отличающихся патологических фенотипов, входящих в диагноз «анэмбриония»: пустой плодный мешок (empty sac) и плодный мешок с наличием желточного мешка, но без эмбриона (yolk sac only). Так, в работах (Ouyang et al., 2016; Li et al., 2017; Gu et al., 2021) в группе АЭ с пустым

Таблица 3. Частоты хромосомных аномалий при АЭ и НБ в различных исследованиях

Литературный источник	N	Метод	Частота хромосомных аномалий		p
			при АЭ	при НБ	
Настоящая работа	1551	Культуры, G-banding	57.1 % (152/266)	52.9 % (680/1285)	0.209
Романова, 2022	273	ППВХ, Q-banding	25.7 % (9/35)	67.2 % (160/238)	<0.050
Gu et al., 2021	1102 (887)	ХМА	47.1 % (64/136)	62.1 % (466/751)	0.001
Yoneda et al., 2018	151 (141)	Культуры, G-banding	50.8 % (32/53)	77.3 % (68/88)	<0.001
Li et al., 2017	2172 после ЭКО	CGH + FISH	28.1 % (138/491)	52.2 % (641/1227)	<0.001
			пустой плодный мешок		
			43.4 % (197/454)		0.002
			только желточный мешок		
			35.4 % (335/945)		<0.001
			суммарно АЭ		
Liu et al., 2015	183	Культуры, G-banding	56.1 % (32/57)	61.1 % (77/126)	0.526
Cheng et al., 2014	223	Культуры, G-banding	46.3 % (37/80)	61.5 % (88/143)	0.030
Angiolucci et al., 2011	156	G-banding	72.2 % (13/18)	33.8 % (23/68)*	0.006
				61.6 % (85/138)**	0.381
Ljunger et al., 2011	259 (239)	ППВХ, G-banding	54.5 % (48/88)	65.6 % (99/151)	0.092
Muñoz et al., 2010	185	ППВХ, G-banding	60.5 % (26/43)	67.6 % (96/142)	0.387
Lathi et al., 2007	272	Культуры, G-banding	58.2 % (53/91)	68.0 % (123/181)	>0.050
Ginsberg et al., 2001	129	G-banding	57.1 % (12/21)	90.7 % (98/108)	<0.001
Bcero	6411	–	44.4 % (813/1833)	59.3 % (2701/4558)	<0.001

Примечание. N – объем выборки, в скобках указано число случаев, для которых проведено сравнение АЭ и НБ. G-banding – стандартное кариотипирование; ППВХ – прямые препараты ворсин хориона; ХМА – хромосомный микроматричный анализ; CGH – сравнительная геномная гибридизация; FISH – флуоресцентная гибридизация *in situ*.

* Относительно абортусов с нормальным фенотипом по УЗИ; ** относительно абортусов с наличием эмбриона по УЗИ.

плодным мешком частота хромосомных аномалий была статистически значимо ниже, чем в группе АЭ с наличием только желточного мешка (см. табл. 3).

Наличие желточного мешка в плодном яйце без эмбриона означает, что имеющиеся нарушения проявляются после обособления гипобласта и эпибласта, происходящего на 6–7-й день после оплодотворения. Гипобласт, дающий начало энтодерме желточного мешка, продолжает развитие, тогда как клетки эпибласта, дающие начало трем зародышевым листкам (энтодерме, эктодерме и мезодерме), составляющим ткани собственно эмбриона, блокируются. Пустой плодный мешок означает, что нарушения проявились до разделения внутренней клеточной массы на гипобласт и эпибласт, т. е. во время имплантации (Boss et al., 2018). На такой ранней стадии влияние негенетических факторов вряд ли может быть существенным. Сниженная частота хромосомных аномалий у этих зародышей может быть объяснена тем, что на столь ранних этапах развития критичны не столько изменение дозы генов вследствие анеуплоидии, сколько, например, точковые мутации, CNV или эпигенетические аномалии, нарушающие активность генов, важных в раннем эмбриогенезе.

Учитывая тот факт, что отсутствие эмбриона в полости плодного мешка не связано с повышенной частотой выявляемых хромосомных аномалий, и то, что чем раньше происходит прекращение развития собственно эмбриона, тем больший вклад в этот феномен вносят генетические, а не материнские или средовые факторы, перспективным представляется поиск в материале АЭ эпигенетических аномалий и генетических нарушений субхромосомного уровня (Лебедев и др., 2013). Так, с помощью микроматричного анализа было выявлено большее число вариаций числа копий ДНК (CNV) при АЭ, чем при НБ (299 и 132 соответственно), причем при АЭ обнаружено 54.3 % патогенетически значимых делеций и 45.7 % дупликаций, а при НБ – только дупликации (Савченко и др., 2018). Интересно, что отличался и спектр входящих в CNV генов: при АЭ чаще затрагивались гены, отвечающие за базовые биологические процессы (миграция, клеточные контакты, адгезия), тогда как при НБ – гены, отвечающие за морфогенез.

Мы обнаружили различия в частотах различных типов хромосомных аномалий между абортусами с наличием и отсутствием эмбриона. В нашей выборке анеуплоидии

по половым хромосомам (в особенности кариотип 45,X) реже встречались при АЭ, чем при НБ (см. табл. 2). Статистически значимо более высокая частота кариотипа 45,X обнаружена при замерших беременностях с наличием эмбриона и в большинстве опубликованных сравнительных исследований (Minelli et al., 1993; Muñoz et al., 2010; Cheng et al., 2014; Веропотвелян, Кодунов, 2015; Liu et al., 2015; Li et al., 2017; Ozawa et al., 2019; Gu et al., 2021). Эти результаты показывают, что моносомия X не оказывает заметного негативного влияния на ранний рост собственно эмбриона, и такие зародыши погибают на более поздних этапах, возможно, вследствие недостатка функции трофобласта (Ahern et al., 2022).

Другой тип аномалий, чаще встречающийся при наличии эмбриона, – триплоидии. Предполагается, что триплоидии в основном возникают в результате ошибок оплодотворения, приводящих либо к диандрии (наличию двух отцовских наборов хромосом), либо к дигинии (наличию двух материнских наборов) (Thaker, 2005). Вследствие явления хромосомного импринтинга хромосомы, полученные от отца, способствуют пролиферации преимущественно тканей трофобласта. Возможно, именно диандрические триплоидии приводят к фенотипу АЭ, однако это предположение нуждается в проверке.

Общим для двух рассматриваемых типов хромосомных аномалий является то, что механизм их возникновения не связан с мейотическим нерасхождением в ооцитах. Случаи моносомии X-хромосомы чаще всего вызваны ошибками в мейозе отца, а не матери (Hassold et al., 1988; Segawa et al., 2017), а триплоидии – ошибками оплодотворения. Поэтому среди абортусов от молодых матерей удельный вес этих типов аномалий кариотипа повышен (Soler et al., 2017; Wang et al., 2020; Gu et al., 2021). Тот факт, что в наших выборках АЭ и НБ возраст матерей не различался, подтверждает, что более высокая частота встречаемости при НБ отражает лучшую выживаемость эмбрионов с моносомией X и триплоидией.

Выявленная нами повышенная частота тетраплоидий при АЭ согласуется с данными (Веропотвелян, Кодунов, 2015; Ozawa et al., 2019) и свидетельствует о критическом влиянии тетраплоидного кариотипа, приводящем к более раннему прекращению развития эмбриона.

Мы обнаружили, что СП среди абортусов с нормальным кариотипом отклоняется от нормального и составляет 0.54 для АЭ и 0.74 для НБ. Несмотря на то что различия между группами не достигали статистически достоверного уровня ($p = 0.142$), они согласуются с данными, полученными ранее в нашей лаборатории, хотя и на значительно меньших выборках. В работе (Евдокимова и др., 2000) было показано, что доля эмбрионов с кариотипом 46,XY снижалась с ростом тяжести нарушения развития: СП составляло 0.77 для самопроизвольных абортусов с хорошо развитыми эмбрионами без существенной внутриутробной задержки развития, 0.60 для НБ и 0.31 для АЭ (по сравнению со значением 1.10 у контрольной группы медицинских абортусов). Одной из причин смещенного СП может быть контаминация культур клетками материнского происхождения (материнская контаминация). Но так как обе наши выборки (АЭ и НБ) анализировались параллельно и частота материнской контаминации невы-

сока (Nikitina et al., 2005), это нивелирует ее возможное влияние на СП в данной работе. В недавно опубликованном масштабном исследовании соотношения полов в раннем онтогенезе у человека (от зачатия до рождения) было показано, что СП уменьшается примерно в первую неделю после зачатия (из-за избыточной смертности эмбрионов мужского пола), а затем увеличивается в течение как минимум 10–15 недель (из-за избыточной смертности эмбрионов женского пола) (Orzack et al., 2015). Таким образом, избыток погибших эмбрионов женского пола в первом триместре беременности, вероятно, отражает реально существующий феномен.

Развитие клеточных технологий предлагает уникальную возможность для изучения биологических механизмов, приводящих к нарушению эмбриогенеза. Так, индуцированные плюрипотентные стволовые клетки (ИПСК) воспроизводят характеристики эмбриональных стволовых клеток, включая неограниченную пролиферативную способность и плюрипотентность, т. е. способность к дифференцировке в производные трех зародышевых листков. Показано, что ИПСК можно получить в том числе из тканей трофобласта не только от эмбрионов с нормальным кариотипом, но и с некоторыми хромосомными анеуплоидиями (например, моносомия X и трисомия 13) (Parveen et al., 2017; Long et al., 2020). Если удастся репрограммировать клетки трофобласта от случаев анэмбрионии и получить линии ИПСК, это откроет возможность изучения процессов, происходящих в производных различных зародышевых листков и приводящих к ранней остановке развития собственно эмбриона.

Заключение

Мы обнаружили, что структура хромосомных аномалий отличается между АЭ и НБ и наличие эмбриона положительно ассоциировано с анеуплоидиями по половым хромосомам и триплоидиям, а отсутствие эмбриона – с тетраплоидиями. В то же время суммарная частота хромосомных аномалий при АЭ и НБ не различалась, что свидетельствует о необходимости поиска причин АЭ на других уровнях организации генома, включая микроструктурные перестройки хромосом, моногенные мутации, нарушения импринтинга и другие aberrантные эпигенетические модификации генома.

Список литературы / References

- Веропотвелян Н.П., Кодунов Л.О. Особенности соотношения хромосомных аномалий и терминов персистенции плодного яйца среди неразвивающихся беременностей в случаях анэмбрионии и при наличии эмбриона (анализ 1328 случаев). *Здоровье женщины*. 2015;5(101):74–80. DOI 10.15574/HW.2015.101.74.
[Veropotvelyan N.P., Kodunov L.O. The features ratio of chromosomal abnormalities and terms of persistence of gestational sacs among undeveloping pregnancies in cases without embryo (anembryonia) and with the existing embryo: analysis of 1328 cases. *Zdorov'e Zhenshchiny = Health of Woman*. 2015;5(101):74–80. DOI 10.15574/HW.2015.101.74. (in Russian)]
- Евдокимова В.Н., Никитина Т.В., Лебедев И.Н., Суханова Н.Н., Назаренко С.А. К вопросу о соотношении полов при ранней эмбриональной летальности у человека. *Онтогенез*. 2000;31(4): 251–257.
[Evdokimova V.N., Nikitina T.V., Lebedev I.N., Sukhanova N.N., Nazarenko S.A. About the sex ratio in connection with early embryo-

- nic mortality in man. *Russ. J. Developmental Biology*. 2000;31(4): 251-257.]
- Лебедев И.Н., Кашеварова А.А., Скрябин Н.А., Никитина Т.В., Лопаткина М.Е., Мельников А.А., Саженова Е.А., Иванова Т.В., Евтущенко И.Д. Матричная сравнительная геномная гибридизация (аггау-СГН) в диагностике хромосомного дисбаланса и CNV-полиморфизма при анэмбрионии. *Журн. акушерства и женских болезней*. 2013;62(2):117-125.
- [Lebedev I.N., Kashevarova A.A., Skryabin N.A., Nikitina T.V., Lopatkina M.E., Melnikov A.A., Sazhenova E.A., Ivanova T.V., Evtushenko I.D. Array-based comparative genomic hybridization (array-CGH) in analysis of chromosomal aberrations and CNV in blighted ovum pregnancies. *Zhurnal Akusherstva i Zhenskikh Bolezney = Journal of Obstetrics and Women's Diseases*. 2013;62(2): 117-125. (in Russian)]
- Лебедев И.Н., Никитина Т.В. Цитогенетика нарушений эмбрионального развития человека (Наследственность и здоровье). Томск: Печатная мануфактура, 2013.
- [Lebedev I.N., Nikitina T.V. Cytogenetics of Human Embryonic Development Disorders (Heredity and Health). Tomsk: Pechatnaya Manufaktura Publ., 2013. (in Russian)]
- Радзинский В.Е., Маклецова С.А., Алеев И.А., Руднева О.Д., Рябинкина Т.С. Неразвивающаяся беременность. Методические рекомендации МАРС (Междисциплинарной ассоциации специалистов репродуктивной медицины). М.: Редакция журнала StatusPraesens, 2015.
- [Radzinsky V.E., Makletsova S.A., Aleev I.A., Rudneva O.D., Ryabinkina T.S. Missed miscarriage. Guidelines by MARS Reproductive Health Professional Medical Association. Moscow: StatusPraesens Editorial Office, 2015. (in Russian)]
- Романова О.А. Особенности иммунореактивности эндометрия при неразвивающейся беременности, ассоциированной с хромосомными аномалиями хорiona: Дис. ... канд. мед. наук. СПб., 2022.
- [Romanova O.A. Endometrium immunoreactivity in missed miscarriage associated with chromosome aberrations in the chorion. Cand. Med. Sci. Diss. St. Petersburg, 2022. (in Russian)]
- Савченко Р.Р., Кашеварова А.А., Скрябин Н.А., Жигалина Д.И., Лопаткина М.Е., Никитина Т.В., Васильев С.А., Лебедев И.Н. Анализ CNVs при анэмбрионии и неразвивающейся беременности. *Мед. генетика*. 2018;17(3):49-54. DOI 10.25557/2073-7998.2018.03.49-54.
- [Savchenko R.R., Kashevarova A.A., Skryabin N.A., Zhigalina D.I., Lopatkina M.E., Nikitina T.V., Vasiliev S.A., Lebedev I.N. Analysis of CNVs in anembryonic pregnancy and missed abortions. *Meditsinskaya Genetika = Medical Genetics*. 2018;17(3):49-54. DOI 10.25557/2073-7998.2018.03.49-54. (in Russian)]
- Ahern D.T., Bansal P., Faustino I., Kondaveeti Y., Glat-Deeley H.R., Banda E.C., Pinter S.F. Monosomy X in isogenic human iPSC-derived trophoblast model impacts expression modules preserved in human placenta. *Proc. Natl. Acad. Sci. USA*. 2022;119(40): e2211073119. DOI 10.1073/pnas.2211073119.
- Angiolucci M., Murru R., Melis G., Carcassi C., Mais V. Association between different morphological types and abnormal karyotypes in early pregnancy loss. *Ultrasound Obstet. Gynecol.* 2011;37(2):219-225. DOI 10.1002/uog.7681.
- Boss A.L., Chamley L.W., James J.L. Placental formation in early pregnancy: how is the centre of the placenta made? *Hum. Reprod. Update*. 2018;24(6):750-760. DOI 10.1093/humupd/dmy030.
- Cheng H.H., Ou C.Y., Tsai C.C., Chang S.D., Hsiao P.Y., Lan K.C., Hsu T.Y. Chromosome distribution of early miscarriages with present or absent embryos: female predominance. *J. Assist. Reprod. Genet.* 2014;31(8):1059-1064. DOI 10.1007/s10815-014-0261-9.
- Doubilet P.M., Benson C.B., Bourne T., Blaivas M., Society of Radiologists in Ultrasound Multispecialty Panel on Early First Trimester Diagnosis of Miscarriage and Exclusion of a Viable Intrauterine Pregnancy, Barnhart K.T., Benacerraf B.R., Brown D.L., Filly R.A., Fox J.C., Goldstein S.R., Kendall J.L., Lyons E.A., Porter M.B., Pretorius D.H., Timor-Tritsch I.E. Diagnostic criteria for nonviable pregnancy early in the first trimester. *N. Engl. J. Med.* 2013;369(15): 1443-1451. DOI 10.1056/NEJMra1302417.
- Fan L., Wu J., Wu Y., Shi X., Xin X., Li S., Zeng W., Deng D., Feng L., Chen S., Xiao J. Analysis of chromosomal copy number in first-trimester pregnancy loss using next-generation sequencing. *Front. Genet.* 2020;11:545856. DOI 10.3389/fgene.2020.545856.
- Farquharson R.G., Jauniaux E., Exalto N., ESHRE Special Interest Group for Early Pregnancy (SIGEP). Updated and revised nomenclature for description of early pregnancy events. *Hum. Reprod.* 2005;20(11):3008-3011. DOI 10.1093/humrep/dei167.
- Finley J., Hay S., Oldzej J., Meredith M.M., Dzidic N., Slim R., Aradhya S., Hovanes K., Sahoo T. The genomic basis of sporadic and recurrent pregnancy loss: a comprehensive in-depth analysis of 24,900 miscarriages. *Reprod. Biomed. Online*. 2022;45(1):125-134. DOI 10.1016/j.rbmo.2022.03.014.
- Fu M., Mu S., Wen C., Jiang S., Li L., Meng Y., Peng H. Whole exome sequencing analysis of products of conception identifies novel mutations associated with missed abortion. *Mol. Med. Rep.* 2018;18(2): 2027-2032. DOI 10.3892/mmr.2018.9201.
- Ginsberg N.A., Strom C., Verlinsky Y. Crown-rump lengths in missed miscarriages and trisomy 21. *Ultrasound Obstet. Gynecol.* 2001; 18(5):488-490. DOI 10.1046/j.0960-7692.2001.00571.x.
- Gu C., Li K., Li R., Li L., Li X., Dai X., He Y. Chromosomal aneuploidy associated with clinical characteristics of pregnancy loss. *Front. Genet.* 2021;12:667697. DOI 10.3389/fgene.2021.667697.
- Hardy K., Hardy P.J., Jacobs P.A., Lewallen K., Hassold T.J. Temporal changes in chromosome abnormalities in human spontaneous abortions: results of 40 years of analysis. *Am. J. Med. Genet. A*. 2016; 170(10):2671-2680. DOI 10.1002/ajmg.a.37795.
- Hassold T., Benham F., Leppert M. Cytogenetic and molecular analysis of sex-chromosome monosomy. *Am. J. Hum. Genet.* 1988;42(4): 534-541.
- Huang J., Zhu W., Tang J., Saravelos S.H., Poon L.C.Y., Li T.C. Do specific ultrasonography features identified at the time of early pregnancy loss predict fetal chromosomal abnormality? – A systematic review and meta-analysis. *Genes Dis.* 2019;6(2):129-137. DOI 10.1016/j.gendis.2018.10.001.
- Kolte A.M., Bernardi L.A., Christiansen O.B., Quenby S., Farquharson R.G., Goddijn M., Stephenson M.D. ESHRE Special Interest Group, Early Pregnancy. Terminology for pregnancy loss prior to viability: a consensus statement from the ESHRE early pregnancy special interest group. *Hum. Reprod.* 2015;30(3):495-498. DOI 10.1093/humrep/deu299.
- Larsen E.C., Christiansen O.B., Kolte A.M., Macklon N. New insights into mechanisms behind miscarriage. *BMC Med.* 2013;11:154. DOI 10.1186/1741-7015-11-154.
- Lathi R.B., Mark S.D., Westphal L.M., Milki A.A. Cytogenetic testing of anembryonic pregnancies compared to embryonic missed abortions. *J. Assist. Reprod. Genet.* 2007;24(11):521-524. DOI 10.1007/s10815-007-9166-1.
- Levy B., Sigurjonsson S., Pettersen B., Maisenbacher M.K., Hall M.P., Demko Z., Lathi R.B., Tao R., Aggarwal V., Rabinowitz M. Genomic imbalance in products of conception: single-nucleotide polymorphism chromosomal microarray analysis. *Obstet. Gynecol.* 2014;124(2 Pt.1):202-209. DOI 10.1097/aog.0000000000000325.
- Li X., Ouyang Y., Yi Y., Tan Y., Lu G. Correlation analysis between ultrasound findings and abnormal karyotypes in the embryos from early pregnancy loss after in vitro fertilization-embryo transfer. *J. Assist. Reprod. Genet.* 2017;34(1):43-50. DOI 10.1007/s10815-016-0821-2.
- Liu Y., Liu Y., Chen H., Du T., Tan J., Zhang J. The frequencies of the presence of embryonic pole and cardiac activity in early miscarriages with abnormal karyotypes. *Clin. Exp. Obstet. Gynecol.* 2015; 42(4):490-494.
- Ljunger E., Stavreus-Evers A., Cnattingius S., Ekbohm A., Lundin C., Annéren G., Sundström-Poromaa I. Ultrasonographic findings in spontaneous miscarriage: relation to euploidy and aneuploidy. *Fertil. Steril.* 2011;95(1):221-224. DOI 10.1016/j.fertnstert.2010.06.018.

- Long P., Liu Z., Wu B., Chen J., Sun C., Wang F., Huang Y., Chen H., Li Q., Ma Y. Generation of an induced pluripotent stem cell line from chorionic villi of a Patau syndrome spontaneous abortion. *Stem Cell Res.* 2020;45:101789. DOI 10.1016/j.scr.2020.101789.
- Menasha J., Levy B., Hirschhorn K., Kardon N.B. Incidence and spectrum of chromosome abnormalities in spontaneous abortions: new insights from a 12-year study. *Genet. Med.* 2005;7(4):251-263. DOI 10.1097/01.gim.0000160075.96707.04.
- Minelli E., Buchi C., Granata P., Meroni E., Righi R., Portentosio P., Giudici A., Ercoli A., Sartor M.G., Rossi A. Cytogenetic findings in echographically defined blighted ovum abortions. *Ann. Genet.* 1993;36(2):107-110.
- Muñoz M., Arigita M., Bennasar M., Soler A., Sanchez A., Borrell A. Chromosomal anomaly spectrum in early pregnancy loss in relation to presence or absence of an embryonic pole. *Fertil. Steril.* 2010;94(7):2564-2568. DOI 10.1016/j.fertnstert.2010.04.011.
- Nikitina T.V., Lebedev I.N., Sukhanova N.N., Sazhenova E.A., Nazarenko S.A. A mathematical model for evaluation of maternal cell contamination in cultured cells from spontaneous abortions: significance for cytogenetic analysis of prenatal selection factors. *Fertil. Steril.* 2005;83(4):964-972. DOI 10.1016/j.fertnstert.2004.12.009.
- Orzack S.H., Stubblefield J.W., Akmaev V.R., Colls P., Munné S., Scholl T., Steinsaltz D., Zuckerman J.E. The human sex ratio from conception to birth. *Proc. Natl. Acad. Sci. USA.* 2015;112(16):E2102-E2111. DOI 10.1073/pnas.1416546112.
- Ouyang Y., Tan Y., Yi Y., Gong F., Lin G., Li X., Lu G. Correlation between chromosomal distribution and embryonic findings on ultrasound in early pregnancy loss after IVF-embryo transfer. *Hum. Reprod.* 2016;31(10):2212-2218. DOI 10.1093/humrep/dew201.
- Ozawa N., Ogawa K., Sasaki A., Mitsui M., Wada S., Sago H. Maternal age, history of miscarriage, and embryonic/fetal size are associated with cytogenetic results of spontaneous early miscarriages. *J. Assist. Reprod. Genet.* 2019;36(4):749-757. DOI 10.1007/s10815-019-01415-y.
- Parveen S., Panicker M.M., Gupta P.K. Generation of an induced pluripotent stem cell line from chorionic villi of a Turner syndrome spontaneous abortion. *Stem Cell Res.* 2017;19:12-16. DOI 10.1016/j.scr.2016.12.016.
- Robinson H.P. The diagnosis of early pregnancy failure by sonar. *Br. J. Obstet. Gynaecol.* 1975;82(11):849-857. DOI 10.1111/j.1471-0528.1975.tb00588.x.
- Segawa T., Kuroda T., Kato K., Kuroda M., Omi K., Miyauchi O., Watanabe Y., Okubo T., Osada H., Teramoto S. Cytogenetic analysis of the retained products of conception after missed abortion following blastocyst transfer: a retrospective, large-scale, single-centre study. *Reprod. Biomed. Online.* 2017;34(2):203-210. DOI 10.1016/j.rbmo.2016.11.005.
- Soler A., Morales C., Mademont-Soler I., Margarit E., Borrell A., Borro V., Muñoz M., Sánchez A. Overview of chromosome abnormalities in first trimester miscarriages: a series of 1,011 consecutive chorionic villi sample karyotypes. *Cytogenet. Genome Res.* 2017;152(2):81-89. DOI 10.1159/000477707.
- Thaker H.M. The partly molar pregnancy that is not a partial mole. *Pediatr. Dev. Pathol.* 2005;8(2):146-147. DOI 10.1007/s10024-005-2164-3.
- van den Berg M.M., van Maarle M.C., van Wely M., Goddijn M. Genetics of early miscarriage. *Biochim. Biophys. Acta.* 2012;1822(12):1951-1959. DOI 10.1016/j.bbdis.2012.07.001.
- Wang H., Yuan D., Wang S., Luo L., Zhang Y., Ye J., Zhu K. Cytogenetic and genetic investigation of miscarriage cases in Eastern China. *J. Matern. Fetal. Neonatal. Med.* 2020;33(20):3385-3390. DOI 10.1080/14767058.2019.1572738.
- Wu X., Su L., Xie X., He D., Chen X., Wang M., Wang L., Zheng L., Xu L. Comprehensive analysis of early pregnancy loss based on cytogenetic findings from a tertiary referral center. *Mol. Cytogenet.* 2021;14(1):56. DOI 10.1186/s13039-021-00577-8.
- Yoneda S., Shiozaki A., Yoneda N., Sameshima A., Ito M., Shima T., Nakashima A., Yoshino O., Kigawa M., Takamori R., Shinagawa Y., Saito S. A yolk sac larger than 5 mm suggests an abnormal fetal karyotype, whereas an absent embryo indicates a normal fetal karyotype. *J. Ultrasound Med.* 2018;37(5):1233-1241. DOI 10.1002/jum.14467.

ORCID ID

T.V. Nikitina orcid.org/0000-0002-4230-6855
E.A. Sazhenova orcid.org/0000-0003-3875-3932
E.N. Tolmacheva orcid.org/0000-0002-0716-4302
S.A. Vasilyev orcid.org/0000-0002-5301-070X
I.N. Lebedev orcid.org/0000-0002-0482-8046

Благодарности. Исследование проведено в рамках выполнения государственного задания, № 122020300041-7.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию 30.09.2022. После доработки 14.11.2022. Принята к публикации 18.11.2022.

Перевод на английский язык <https://vavilov.elpub.ru/jour>

Структура и происхождение генофонда тувинцев по данным аутосомных SNP и гаплогрупп Y-хромосомы

В.А. Степанов, Н.А. Колесников, Л.В. Валихова, А.А. Зарубин, И.Ю. Хитринская, В.Н. Харьков 

Научно-исследовательский институт медицинской генетики, Томский национальный исследовательский медицинский центр Российской академии наук, Томск, Россия
 Vladimir-kharkov@medgenetics.ru

Аннотация. Тувинцы – один из наиболее компактно проживающих народов Южной Сибири, расселенный в основном на территории Тывы. Генофонд тувинцев является достаточно обособленным за счет эндогамии и очень низкой частоты межнациональных браков. Исследована структура генофонда тувинцев и других сибирских популяций по полногеномной панели аутосомных однонуклеотидных полиморфных маркеров и маркерам Y-хромосомы. Результаты анализа частот аутосомных SNP различными методами, сходства по составу гаплогрупп Y-хромосомы и YSTR-гаплотипов показывают, что генофонд тувинцев очень гетерогенен по составу генетических компонентов. Он включает в себя древний автохтонный енисейский компонент, доминирующий у чулымских тюрков и кетов, восточносибирский, преобладающий у якутов и эвенков, и дальневосточный, частота которого максимальна у нивхов и удэгейцев. Анализ состава IBD-блоков на аутосомах демонстрирует максимальное генетическое родство тувинцев с южными алтайцами, хакасами и шорцами, которые формировались при расселении тюркских групп популяций на территории Алтае-Саянского региона. Выявлен очень разнообразный состав генофонда тувинцев по различным сублиниям Y-хромосомных гаплогрупп, большинство из которых показывают сильную этническую специфичность. Филогенетический анализ отдельных Y-хромосомных гаплогрупп демонстрирует максимальную близость генофонда тувинцев с алтайцами, хакасами и шорцами. Внутри тувинского этноса обнаружены значительные различия между выборками из западных, южных и восточных районов Тывы по доле монгольского и енисейского генетических компонентов. Генетическое разнообразие тувинцев по Y-хромосомным гаплогруппам и максимально разнородный состав генетических компонентов свидетельствуют о самом высоком разнообразии тувинского генофонда по сравнению со всеми коренными народами Сибири. Обнаружены различия по частотам гаплогрупп Y-хромосомы между туджинцами и тувинцами и изменение частот гаплогрупп с юга на север, связанных с восточноазиатским компонентом. Большинство наиболее частых гаплогрупп Y-хромосомы у тувинцев демонстрирует эффект основателя, возраст формирования которых полностью согласуется с данными об их этногенезе. Ключевые слова: генофонд; популяции человека; генетическое разнообразие; генетические компоненты; Y-хромосома; тувинцы.

Для цитирования: Степанов В.А., Колесников Н.А., Валихова Л.В., Зарубин А.А., Хитринская И.Ю., Харьков В.Н. Структура и происхождение генофонда тувинцев по данным аутосомных SNP и гаплогрупп Y-хромосомы. *Вавиловский журнал генетики и селекции*. 2023;27(1):36-45. DOI 10.18699/VJGB-23-06

Structure and origin of Tuvan gene pool according to autosome SNP and Y-chromosome haplogroups

V.A. Stepanov, N.A. Kolesnikov, L.V. Valikhova, A.A. Zarubin, I.Yu. Khitrinskaya, V.N. Kharkov 

Research Institute of Medical Genetics, Tomsk National Research Medical Center of the Russian Academy of Sciences, Tomsk, Russia
 Vladimir-kharkov@medgenetics.ru

Abstract. Tuvans are one of the most compactly living peoples of Southern Siberia, settled mainly in the territory of Tuva. The gene pool of the Tuvans is quite isolated, due to endogamy and a very low frequency of interethnic marriages. The structure of the gene pool of the Tuvans and other Siberian populations was studied using a genome-wide panel of autosomal single nucleotide polymorphic markers and Y-chromosome markers. The results of the analysis of the frequencies of autosomal SNPs by various methods, the similarities in the composition of the Y-chromosome haplogroups and YSTR haplotypes show that the gene pool of the Tuvans is very heterogeneous in terms of the composition of genetic components. It includes the ancient autochthonous Yeniseian component, which dominates among the Chulym Turks and Kets, the East Siberian component, which prevails among the Yakuts and Evenks, and the Far Eastern component, the frequency of which is maximum among the Nivkhs and Udeges. Analysis of the composition of IBD-blocks on autosomes shows the maximum genetic relationship of the Tuvans with the Southern Altaians, Khakas and Shors, who were formed during the settlement of the Turkic groups of populations on the territory of the Altai-Sayan region. A very diverse composition of the Tuvan gene pool is shown for various sublines of Y-chromosomal haplogroups, most of which show strong ethnic specificity. Phylogenetic analysis of individual

Y-chromosome haplogroups demonstrates the maximum proximity of the gene pool of the Tuvans with the Altaians, Khakas and Shors. Differences in frequencies of Y-chromosome haplogroups between the Todzhans and Tuvans and a change in the frequencies of haplogroups from south to north associated with the East Asian component were found. The majority of the most frequent Y-chromosome haplogroups in the Tuvans demonstrate the founder effect, the formation age of which is fully consistent with the data on their ethnogenesis.

Key words: gene pool; human population; genetic diversity; genetic components; Y-chromosome; Tuvans.

For citation: Stepanov V.A., Kolesnikov N.A., Valikhova L.V., Zarubin A.A., Khitrinskaya I.Yu., Kharkov V.N. Structure and origin of Tuvan gene pool according to autosome SNP and Y-chromosome haplogroups. *Vavilovskii Zhurnal Genetiki i Selektzii = Vavilov Journal of Genetics and Breeding*. 2023;27(1):36-45. DOI 10.18699/VJGB-23-06

Введение

Для исследования популяционно- и эволюционно-генетических процессов, анализа генетического разнообразия и реконструкции генетической истории популяций генофонд коренного населения Южной Сибири представляет собой уникальную систему. Анализ состава и соотношения различных субстратных компонентов у сибирских народов имеет ряд невыясненных вопросов. Генетика в связи с этим дает богатейшие возможности для исследования проблем, поскольку развитие новых подходов к анализу популяционного генофонда позволяет вывести этногенетические исследования на совершенно новый уровень. Современные методы, применяемые в молекулярно-генетических исследованиях, и новые биоинформатические наработки дают возможность достоверно выявлять различные предковые генетические компоненты в составе генофонда различных народов и отдельных людей.

Важнейшей проблемой этнологии и антропологии населения Южной Сибири остается формирование коренных этносов, в решении которой в настоящее время большую роль играют методики анализа геномных данных. Генофонд коренного населения этого региона складывался за счет длительного и многоэтапного смешения большого числа локальных генофондов различных племен европеоидного и монголоидного происхождения. Коренные этносы Южной Сибири характеризуются различными антропологическими типами, сложной этнической и демографической историей. Смешение многочисленных тюркских, монгольских, енисейских, самодийских и угорских групп на основе генетического субстрата древних индоевропейских племен и таежных монголоидов сформировало в результате пеструю картину генетического разнообразия населения этого региона (Генофонд..., 2003).

Процессы слияния и ассимиляции с участием различных миграционных потоков сыграли важную роль в формировании современных тюркоязычных популяций Южной Сибири, особенно тувинцев. В эпоху энеолита, бронзы и раннего железа территория Тывы входила в зону обитания древнего европеоидного населения, на которой позднее сложились культуры скифо-сибирского мира (Алексеев, 1984). Проникновение центральноазиатского монголоидного компонента на территорию Южной Сибири датируется VIII–VI вв. до н. э. Примерно к этому же времени относится и появление лесного, таежного монголоидного компонентов (Киселев, 1951). Постепенно происходило увеличение монголоидного компонента от преобладания европеоидного в скифское время до формирования в XIII–XIV вв. н. э. современного

центральноазиатского антропологического типа тувинцев (Дебец, 1948).

Тувинцы – это один из наиболее компактно проживающих народов России, расселенный в основном на территории Тывы. В России, по данным Всероссийской переписи населения 2010 г., численность тувинцев составляет 263 934 человека. При этом генофонд тувинцев является относительно обособленным за счет эндогамии и очень низкой частоты межнациональных браков (Пузырев и др., 1999; Кучер и др., 2003). Показана гетерогенность родоплеменного состава тувинцев (Потапов, 1969). Для некоторых групп тувинцев отмечена изолированность локальных популяций, как вызванная географическими факторами, так и исторически сложившаяся, особенно ярко выраженная для тувинцев-тоджинцев из северо-восточной горной части Тывы. Исследованию генофонда тувинцев в последние годы посвящен ряд научных публикаций, которые были акцентированы на изучении общего спектра линий мтДНК, гаплогрупп Y-хромосомы и детализации отдельных клад (Степанов, Пузырев, 2000; Степанов и др., 2001, 2006; Derenko et al., 2006; Харьков и др., 2013; Дамба и др., 2018а, б; Агджоян и др., 2021).

Цель настоящего исследования – комплексный анализ структуры генофонда тувинцев и реконструкция их происхождения в сравнении с другими популяциями коренного населения Сибири. Для решения вопросов генетической близости тувинцев с другими коренными народами было выполнено генотипирование широкого геномного набора аутосомных маркеров с помощью высокоплотных биочипов, а также расширенного набора SNP- и STR-маркеров Y-хромосомы у различных коренных народов Сибири.

Материалы и методы

Материал исследования составили образцы ДНК мужчин общей численностью 419 образцов, представляющих коренное население Республики Тыва. Выборки были собраны в пос. Тээли (западная часть Тывы) ($N = 44$), пос. Кунгуртуг (юго-восток республики) ($N = 48$), пос. Тоора-Хем (северо-восточная часть Тывы) ($N = 23$) и г. Кызыл ($N = 304$). Образцы из Кызыла относили к конкретной территориальной группе в соответствии с местами рождения доноров. Образцы были разделены на пять территориально дистанцированных групп: запад (Барун-Хемчигский, Бай-Тайгинский, Дзун-Хемчигский, Сут-Хольский, Монгун-Тайгинский районы) ($N = 169$), центр (Чаа-Хольский, Тандынский, Каа-Хемский, Кызыльский, Улуг-Хемский, Чеди-Хольский, Пий-Хемский, Тес-Хемский, Овьюрский, Эрзинский районы) ($N = 179$), восток ($N = 71$),

включающий северо-восток (Тоджинский район) ($N = 23$) и юго-восток (Тере-Хольский район) ($N = 48$).

Забор первичного биологического материала (венозной крови) у доноров производили с соблюдением процедуры письменного информированного согласия на проведение исследования. На каждого донора составляли анкету с краткой родословной, указанием этнической принадлежности и мест рождения предков. Индивида относили к данной этнической группе на основании этнической идентификации – его собственной, его родителей, и места рождения.

Для анализа Y-хромосомных гаплогрупп и гаплотипов тувинцев были использованы все 419 образцов ДНК мужчин. Для генотипирования на чипах были выбраны неродственные образцы с. Тээли Бай-Тайгинского кожууна ($N = 28$).

Другие популяции коренного населения Сибири представлены чулымцами, хакасами-сагайцами, хакасами-качинцами, южными алтайцами, кетами, хантами, томскими татарами, бурятами, якутами, эвенками, нивхами, удэгейцами, а также калмыками, дунганями и киргизами.

Данные широкогеномного генотипирования получены с использованием микрочипов Infinium Multi-Ethnic Global-8 (Illumina) для SNP-генотипирования, включающего свыше 1.7 млн маркеров. Материал депонирован в биоресурсной коллекции «Биобанк населения Северной Евразии». Для сравнительного анализа использовали данные генотипов по 1677114 аутосомным SNP (биочип Illumina Multi-Ethnic Global-8) 917 образцов и данные генотипирования более 3000 Y-хромосомных SNP и 36 YSTR у более 1600 образцов мужчин, представляющих коренное население Сибири и соседних регионов. Охарактеризовано свыше 30 популяционных выборок, которые подробно описаны в наших прошлых работах (Колесников и др., 2021, 2022). Для анализа компонентного состава и количества примесей у отдельных индивидов и популяций были использованы методика NGSadmix (Scotte, 2013) и программа ADMIXTURE (Alexander et al., 2009, 2011), а также проведен сравнительный анализ данных аутосомных SNP и гаплогрупп и гаплотипов Y-хромосомы.

Кластеризацию массива генотипов аутосомных SNP и контроль качества выполняли с помощью протокола, разработанного Y. Guo с коллегами (2014) с использованием GenomeStudio (Illumina. GenomeStudio, модуль генотипирования v2.0.3), программного пакета, который компания Illumina создала для различных геномных анализов. Для фильтрации, нормализации и расчета стандартных геномных статистик и показателей оптимальным стал стандартный набор программ, включающий vcftools, bcftools, и plink. Для анализа идентичных по происхождению блоков сцепления применяли алгоритм Refined IBD (Browning B.L., Browning S.R., 2013), показывающий более точные результаты по сравнению с встроенными в plink алгоритмами. Предварительно генотипы были фазированы с использованием программного обеспечения Beagle 5.1 (Browning S.R., Browning B.L., 2007). Для сравнения популяций были получены суммы средних длин идентичных по происхождению блоков (сегментов IBD – identical by descent) – между парами индивидов.

Для изучения состава и структуры гаплогрупп Y-хромосомы в исследование были включены две системы генетических маркеров: диаллельных локусов, представленных SNP, и полиаллельных высоковариабельных микросателлитов (YSTR). С помощью 156 SNP-маркеров определяли принадлежность образцов к различным гаплогруппам. Классификация гаплогрупп приведена в соответствии с данными Международного общества генетической генеалогии (сайт www.isogg.org). Анализ STR-гаплотипов внутри гаплогрупп осуществляли с использованием 44 STR-маркеров нерекombинирующей части Y-хромосомы (DYS19, 385a, 385b, 388, 389I, 389II, 390, 391, 392, 393, 426, 434, 435, 436, 437, 438, 439, 442, 444, 445, 448, 449, 456, 458, 460, 461, 481, 504, 505, 518, 525, 531, 533, 537, 552, 570, 576, 635, 643, YCAIIa, YCAIIb, GATA H4.1, Y-GATA-A10, GGAAT1B07).

STR-маркеры генотипировали посредством капиллярного электрофореза на генетическом анализаторе ABI Prism 3730. Генотипирование SNP-маркеров проводили с помощью полимеразной цепной реакции и последующего анализа фрагментов ДНК с использованием полиморфизма длин рестрикционных фрагментов методом гель-электрофореза. Экспериментальные исследования осуществлены на базе Центра коллективного пользования научно-исследовательским оборудованием «Медицинская геномика» (НИИ медицинской генетики Томского НИМЦ). Построение медианных сетей гаплотипов Y-хромосомы выполняли с использованием программы Network v10.2.0.0 (Fluxus Technology Ltd; www.fluxus-engineering.com) по методу медианных сетей Бандельта (Bandelt, 1999). Оценку возраста генерации наблюдаемого разнообразия гаплотипов в гаплогруппах проводили методом ASD (Zhivotovsky, 2004) на основании средних квадратичных отличий в числе повторов между всеми маркерами.

Результаты и обсуждение

Генотипирование большого массива SNP позволяет очень детально изучить паттерны гаплотипического разнообразия, маркирующие различные субстратные и суперстратные пласты генофонда популяций, степень метисации с пришлым населением на различных уровнях – от индивидуального до родового и этнического, провести детальный анализ демографической истории различных популяций и проанализировать молекулярно-филогенетическую и филогеографическую структуру гаплогрупп Y-хромосомы. Это дает возможность точнее реконструировать генетико-демографические события, произошедшие в прошлом. Применение современных биоинформационных подходов на широком массиве SNP и детальной филогении унитарительских линий позволяет точнее реконструировать формирование генофонда тувинцев.

После обработки данных по результатам микрочипового исследования для фильтрации прогенотипированных образцов и выполнения дальнейших расчетов сначала был проведен поиск метисов среди тувинцев с использованием программы NGSadmix. Метод NGSadmix при запуске на сформированном нами массиве данных показал, что все прогенотипированные образцы тувинцев не

имеют метисации, что полностью соответствует данным анкетирования ДНК-доноров. Полученные данные о частотах SNP в исследованных популяционных выборках были использованы для выяснения генетических взаимоотношений между различными этносами. Для понижения размерности и выявления генетических компонентов был применен алгоритм ADMIXTURE.

Компонентный состав генофонда тувинцев

Для выявления индивидуальных генетических компонентов в генофонде изучаемых популяций использована программа ADMIXTURE, позволяющая обнаруживать смешанность состава набора индивидов на основе данных о генотипах и тем самым строить предположения о происхождении популяции. Моделирование с помощью ADMIXTURE в последнее время является одним из основных методов анализа при исследовании генофондов современных и древних популяций человека, позволяя анализировать одни и те же данные на разных иерархических уровнях.

Тувинцы, по сравнению с большинством сибирских популяций, демонстрируют очень разнообразный состав генетических компонентов. Наиболее отчетливо их распределение проявляется при $K = 12$. Практически для всех сибирских популяций показано полное доминирование одного генетического компонента, характерного для отдельных выборок или близкородственных друг другу коренных народов. Кроме тувинцев, достаточно гетерогенный компонентный состав обнаружен у хакасов-качинцев. Спектр генетических компонентов у качинцев практически полностью совпадает с тувинским, но отличается по их соотношению.

Алтайский компонент. С максимальной частотой у тувинцев (53 %) представлен генетический компонент, доминирующий у южных алтайцев (до 90 %). С учетом того что анализируемая выборка тувинцев представляет самый западный район, граничащий с Республикой Горный Алтай, это вполне закономерно. С достаточной частотой он представлен у родственных южным алтайцам киргизов (9.8 %) и хакасов-качинцев (7.6 %). Вероятно, этот генетический компонент связан с влиянием носителей тюркских языков на формирование современных южносибирских народов. Ранее близость алтайцев и тувинцев была показана при анализе частот аллелей гена *ZFX* (Хитринская и др., 2010), X-сцепленных STR-маркеров (Вагайцева и др., 2014), ферментов и белков крови (Спицын и др., 1984), частот групп крови системы ABO и по их антропологическим параметрам (Богданова, 1978а, б; Алексеев, 1984; Алексеева, 1984).

Восточносибирский компонент. Вторым по частоте у тувинцев является восточносибирский генетический компонент (21 %), доминирующий у якутов (94 %), эвенков из Якутии (93 %) и Забайкалья (62 %). Это соответствует лингвистическим данным о южносибирском происхождении предков современных якутов. У бурятов он составляет 30 %, у качинцев – 12 %, южных алтайцев – 4 %. Распределение этого генетического компонента согласуется с классификацией расовых типов. Тувинцы, тофалары, якуты и долганы – носители черт североазиатской малой расы – одного из подразделений континентальной ветви

большой монголоидной расы. В составе североазиатских монголоидов выделяют два умеренно различающихся типа – байкальский и центральноазиатский. Первый тип характерен прежде всего для тунгусо-маньчжурских народов, второй – для тюркских и монгольских (Тюркские народы Восточной Сибири, 2008).

Восточноазиатский компонент. На третьем месте у тувинцев (11 %) находится компонент, доминирующий у дунган (91 %), бурятов (63 %) и калмыков (54 %). Наиболее четко он проявляется при $K = 12$. Он составляет более значительную долю у киргизов (49 %), казахов (46 %), узбеков (43 %), хакасов-качинцев (41 %) и томских татар (24 %) и южных алтайцев (4 %). Именно этот генетический компонент отражает вклад в генофонд населения Южной Сибири наиболее поздних групп переселенцев с территории Монголии. Почти все остальные исследованные популяции Сибири и Дальнего Востока – якуты, шорцы, хакасы-сагайцы и чулымцы – демонстрируют практически полное отсутствие этого компонента. Не обнаружен он и у эвенков, хантов, кетов, чулымцев, чукчей, коряков и нивхов. Общая картина распределения этого генетического компонента хорошо согласуется с антропологическими и этнографическими данными о влиянии монгольской экспансии на этногенез исследованных этносов.

Енисейский компонент. Наибольшая доля этого компонента характерна для чулымских тюрков (94 %) и кетов (65 %). У кетов его доля ниже по причине метисации и выявления методом NGSAdmix недавней европеоидной примеси у многих образцов. У тувинцев его частота составляет 6.9 %, а у качинцев – 20 %. Полученные результаты сочетаются с данными этнологии, антропологии и лингвистики о вкладе енисейского компонента в формирование различных народов Алтае-Саян и историческими ареалами енисейских языков.

Дальневосточный компонент. Последний генетический компонент у тувинцев, представленный с достоверной частотой (4.9 %), преобладает у нивхов (96 %) и удэгейцев (56 %). С небольшой частотой он присутствует у забайкальских эвенков (11 %), бурятов (10 %), калмыков (8 %) и дунган (6 %). Вероятно, его присутствие отражает вклад таежных монголоидов, расселявшихся в древности в западном направлении из Приморья и Забайкалья.

Можно предположить, что в генофонде тувинцев может быть и самодийский компонент, однако для его определения требуется анализ популяционных групп, в которых он доминирует (ненцев, энцев, нганасан и селькупов).

Идентичные по происхождению блоки сцепления. В результате биоинформатической обработки данных генотипирования высокоплотных биочипов различных сибирских популяций проведен анализ совпадения общих по происхождению фрагментов ДНК между популяциями и отдельными индивидами. Сегментом, имеющим идентичные нуклеотидные последовательности, является IBD у двух или более лиц, если они унаследовали его от общего предка без рекомбинации, т. е. у этих людей сегмент имеет общее происхождение. Ожидаемая длина сегмента IBD зависит от количества поколений с момента существования последнего общего предка. Одно из применений анализа общих по происхождению участков генома – количественная оценка степени родства между индивидами,

что также может дополнить информацию о генетических связях популяций (Gusev et al., 2011).

Образцы из выборки тувинцев показали максимальное совпадение по IBD-блокам между собой (10.07 %), далее с выборкой южных алтайцев (1.62 %), эвенков (0.81 %), якутов (0.77 %), чулымцев (0.70 %), хакасов-сагайцев (0.66 %), хакасов-качинцев (0.64 %), бурятов (0.58 %), калмыков (0.57 %), удэгейцев (0.39 %) и хантов (0.38 %). Степень совпадения IBD-блоков между тувинцами и другими популяционными выборками согласуется с результатами программы ADMIXTURE по распределению в этих популяциях частот аллелей и общих генетических компонентов. Также был рассчитан коэффициент инбридинга (FROH) для всех индивидов по блокам гомозиготности (ROH). У тувинцев его значение (0.0151) намного ниже, чем у чулымцев (0.0292), казымских (0.0280) и русскиных хантов (0.0266), кетов (0.0259) и хакасов-сагайцев из предгорного Таштыпского района. Практически равны тувинцам по значению FROH выборки южных алтайцев (0.0168) и хакасов-качинцев Ширинского района (0.0146). Это свидетельствует об отсутствии значимой роли инбридинга в формировании генофонда современных популяций тувинцев.

Гаплогруппы Y-хромосомы

Для основных по частоте азиатских по происхождению гаплогрупп Y-хромосомы у тувинцев были прогенотипированы дополнительные терминальные SNP, которые позволили точнее разделить образцы на отдельные специфичные сублинии. Частоты встречаемости указаны только для них (см. таблицу). Частоты других достаточно редких гаплогрупп, представленных отдельными образцами, описанные в более ранней статье (Харьков и др., 2013), здесь

Частота встречаемости
основных гаплогрупп Y-хромосомы у тувинцев

Гаплогруппа	% (N)
N1a2b1b2b1 (B178, PF3415, Z35147, Z35149, Z35152)	23.9 (100)
N1a2b1b2a1~ (B228, Z35125, Z35127, Z35128)	0.9 (4)
N1a2b1b (B169 xB175, Z35117, Z35118)	2.4 (10)
N1a1a2~ (B187 xB449)	6.4 (27)
N1a1a1a1a3a2 (B219 xB199)	11.9 (50)
N1a1a1a (L708, L839 xL392)	0.5 (2)
Q1b1a3b1a~ (B30/YP1691, YP1693, YP1694)	12.9 (54)
C2b1a1a1a1 (F3850)	1.4 (6)
C2b1a1a2a (F1756)	1.4 (6)
C2b1c (M504)	2.8 (12)
C2b1b1 (M77)	10.5 (44)
R1a1a1b2e1~ (YP1505, YP1507, YP1508, YP1509)	4.1 (17)
R1a1a1b2a2a (Z2123)	1.4 (6)
R1a1a1b2a2a3b1a1~ (YP1542-1556)	1.2 (5)
R1a1a1b2 (Y43109)	6.4 (27)

не приводятся, так как для них не отбирались дополнительные SNP.

Самой частой гаплогруппой Y-хромосомы у тувинцев является N1a2b1-B169, составляющая в общем массиве образцов мужчин 24 %. Она разделена на три сублинии, отличающиеся по терминальным SNP и кластерам гаплотипов. С максимальной частотой у тувинцев выявлен ее вариант N1a2b1b2b1 (B178, PF3415, Z35147, Z35149, Z35152). Кроме тувинцев, к нему относятся два образца южных алтайцев. По данным сайта YFull, эта линия также обнаружена у одного мужчины из Киргизии и двух – из Китая. Гаплотипы этой линии имеют звездообразную филогению, свидетельствующую о сильном эффекте основателя (рис. 1).

Возраст этой линии у тувинцев по YSTR составляет 1442 года (SD = 368 лет). Присутствие ее у алтайцев, киргизов и жителей Китая в виде единичных образцов, возможно, связано с вхождением в их состав отдельных мужчин, имеющих тувинское происхождение. Эта линия у тувинцев определяет для них общий генетический субстрат, который связан с наследием самодийского населения территории Южной Сибири. Присутствие у тувинцев, хакасов и шорцев разных этноспецифичных вариантов гаплогруппы N1a2b1 свидетельствует о значительной генетической дифференциации между ними. Это подтверждает отсутствие миграций носителей данной гаплогруппы и обмена генами на протяжении последних нескольких сотен лет. Основным фактором ее распространения на территории Тывы были генетическая изоляция местных самодийских групп и интенсивное увеличение численности их популяции. Четыре образца тувинцев принадлежат к очень редкой параллельной линии N1a2b1b2a1~ (B228, Z35125, Z35127, Z35128). Ранее она была обнаружена у монголов (Пумäe et al., 2016). К третьей тувинской сублинии (xB175, Z35117, Z35118) относятся 10 образцов.

Гаплогруппа N1a1 (19 %) – вторая по частоте у тувинцев, разделяется на три ветви. В суммарной выборке ее частота уступает N1a2b1 всего на 5 %, охватывая на западе Тывы чуть менее 30 % образцов. Первая линия, N1a1a2~ (B187 xB449), в общей выборке тувинцев имеет частоту 6.4 %. В восточных районах, Тоджиинском и Тере-Хольском, эта гаплогруппа не найдена. Данный вариант является очень этноспецифичным и не обнаружен в других популяциях. Параллельная ей сестринская линия (N1a1a2~ B499) с относительно недавним расхождением с тувинской также характерна для хакасов-сагайцев и шорцев. Она доминирует по частоте в хакаских сеоках хый и хобый. У шорцев к этой гаплогруппе относятся все мужчины сеоков кый и кобий (Харьков, 2020). На медианных сетях гаплотипы этих линий у тувинцев, хакасов и шорцев формируют три кластера, не пересекающихся с тувинцами, за исключением одного образца (рис. 2).

При этом гаплотипы хакасов-сагайцев Таштыпского района, граничащего с Шорией, очень близки к шорцам и демонстрируют сильный недавний эффект основателя. Общий возраст кластера гаплотипов тувинцев составил 1863 года (SD = 294 года). Это показывает давнее разделение этих линий между тувинцами, хакасами и шорцами и достаточно сильные эффекты основателя для отдельных сеоков хакасов и шорцев. Данная сублиния имеет очень

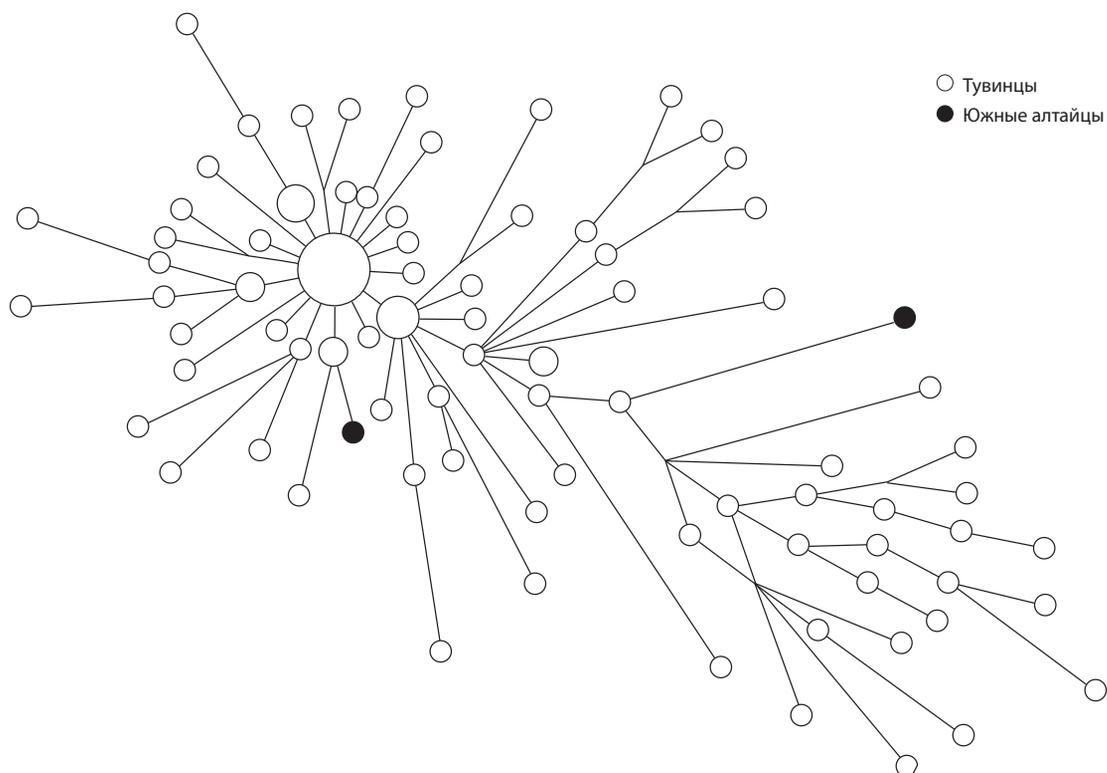


Рис. 1. Медианная сеть YSTR-гаплотипов гаплогруппы N1a2b1b2b1 у тувинцев и южных алтайцев.

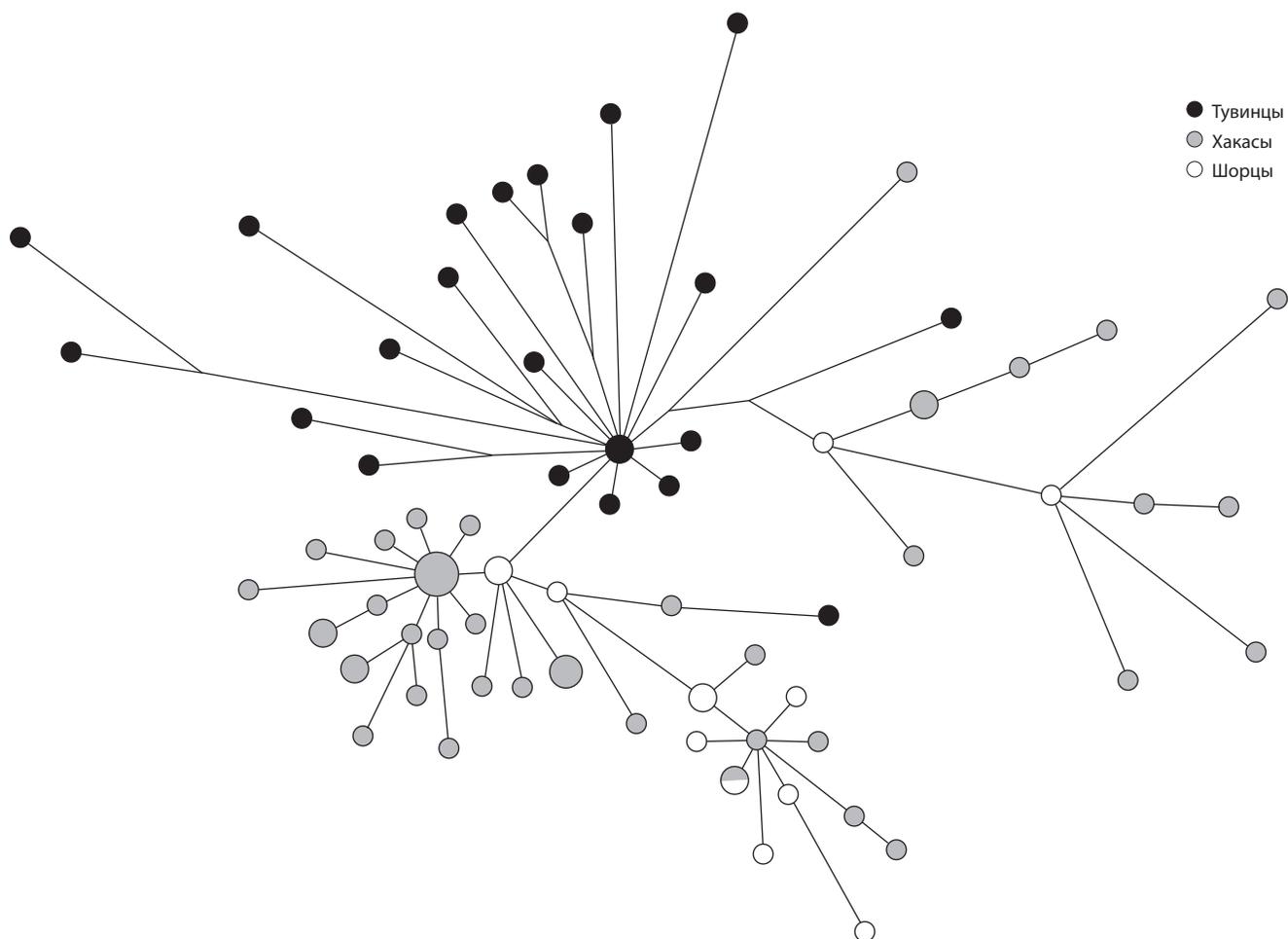


Рис. 2. Медианная сеть YSTR-гаплотипов гаплогруппы N1a1a2~ у тувинцев, хакасов и шорцев.

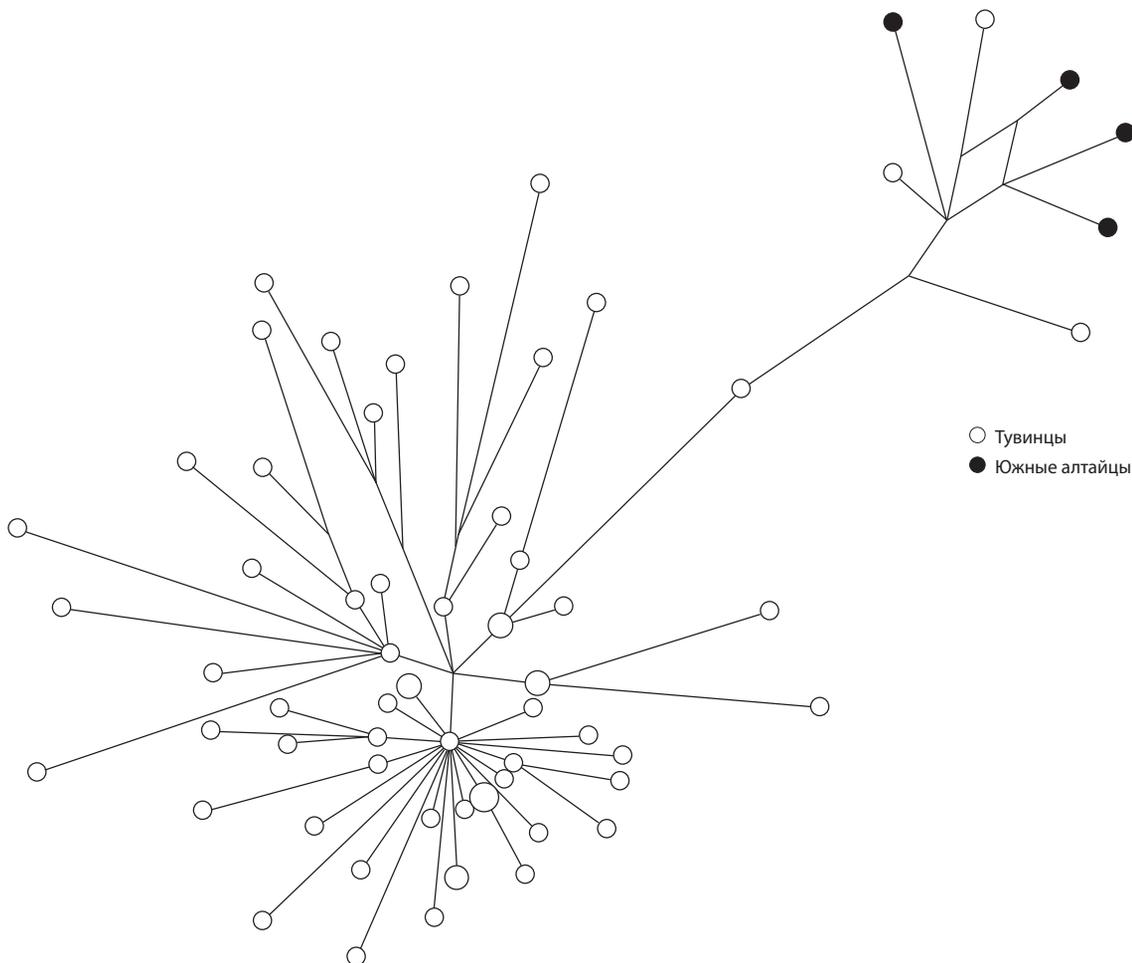


Рис. 3. Медианная сеть YSTR-гаплотипов гаплогруппы Q1b1a3b1a~ –B30 у тувинцев и южных алтайцев.

ограниченный географический ареал. Скорее всего, исходным местом ее распространения была территория Тывы, откуда она переместилась в Горную Шорию и затем в Хакасию. Она очень рано отделилась от основного ствола N1a1 и, как и многие другие редкие Y-хромосомные линии, сохранилась с достаточно высокой частотой лишь в относительно изолированных горных популяциях. Ее отделение от основного ствола гаплогруппы N1a1 произошло примерно 10 700 лет назад (данные с сайта YFull). В силу ее популяционной специфичности и изолированности неоднозначна связь этого варианта с самодийским, угорским или другими генетическими компонентами.

Вторая сублиния N1a1a у тувинцев – N1a1a1a1a3a2 (B219 xB199). Она представлена во всех районах и имеет частоту 11.9 %. К ней относятся и три образца алтайцев. У хакасов, шорцев и чулымцев она не обнаружена. Близкородственная ей линия N1a1a1a1a3a2c2-B199 доминирует у восточных бурятов и с достаточно высокой частотой представлена у западных бурятов. Появление этих линий связано с расселением в Тыве, Бурятии и на Алтае монгольских этнических групп. Возраст данной линии по гаплотипам у тувинцев составил 1500 лет (SD = 304 года). Распространение этой Y-хромосомной линии происходило немного позже, чем носителей тувинско-шорско-хакасской ветви N1a1a2~.

Всего два образца тувинцев-тоджинцев принадлежат к очень редкой линии N1a1a1a (L708, L839 xL392). По гаплотипам она очень близка к якутско-эвенкийской гаплогруппе N1a1a1a1a4a1a1, но не мутантна по ее терминальным SNP (M1979, M1984, M1988, M1991). Присутствие этого варианта Y-хромосомы у тоджинцев согласуется с распределением восточносибирского генетического компонента и совпадения по IBD-блокам с якутами и эвенками. К данной линии также относятся и четыре образца мужчин хакасов-сагайцев из Аскизского района с гаплотипами, близкими к тувинским.

Гаплогруппа Q1b1a3b1a~ (B30/YP1691, YP1693, YP1694) занимает 13 % в общей выборке тувинцев. Максимальная ее частота приходится на восточные выборки тоджинцев и тувинцев Тере-Хольского кожууна (25 %). К этой линии принадлежат и четыре образца южных алтайцев (рис. 3).

Градиент убывания этой гаплогруппы с востока на запад был показан на территории Тывы и ранее (Харьков и др., 2013; Дамба и др., 2018б; Агджоян и др., 2021). Наибольшая для тувинцев частота гаплогруппы Q1b1a3b1a~ в Тодже, видимо, является следствием их относительной генетической изолированности и сохранения большей доли местного автохтонного енисейского генетического компонента. Возраст этой линии по гаплотипам у тувинцев составил 2187 лет (SD = 446 лет). Частота распростра-

нения гаплогруппы Q1b1a3b1a~ и ее родственных линий Q1b1a3b1a2-B33 и Q1b1a3b4-B31 в популяциях коренных народов Южной и Западной Сибири отражает вклад в их генофонды достаточно древних по происхождению местных аборигенных групп населения, относящихся к енисейской языковой семье. Анализ Y-хромосомных сублиний Q1b1a3b показывает, что исходный центр происхождения и расселения ее носителей – территория современной Тывы.

У отдельных популяций с долей монгольского генетического компонента представлены отличающиеся гаплогруппы и сублинии, происхождение которых связано с расселением различных этнических групп, и разделенные по времени миграционными событиями. У тувинцев результатом монгольского вклада, помимо N1a1a1a1a3a2, являются гаплогруппы клад C2b1, O2 и O3. Все они очень близки вариантам, представленным с высокой частотой у монголов, бурят и калмыков. Доля C2b1c (M504) и C2b1a1a1a1 (F3850) максимальна в юго-восточной выборке (15 %). Линия C2b1a1a1a1 (F3850) найдена только в южных и юго-восточных районах. Более частая линия C2b1b1 (M77) показывает клинальное убывание частот с юго-востока на запад. Это характерно и для гаплогрупп O2 и O3. В генофонде практически всех исследованных популяций, у которых не обнаруживается монгольский генетический компонент, по аутосомным SNP отсутствуют и указанные гаплогруппы Y-хромосомы. Филогенетический анализ Y-хромосомных сублиний и гаплотипов показывает, что центр происхождения и расселения носителей монгольского компонента – территория Центральной Азии.

Указанные гаплогруппы у тувинцев являются наследием генетического вклада поздних монголоидных мигрантов, отражая вклад хуннских и монгольских переселенцев на территорию Тывы. Таким образом, генетические данные подтверждают, что проникновение монгольских кочевников на территорию Тывы шло с юга, постепенно распространяясь на северные области, и, соответственно, монголизация населения Тывы сильнее всего была выражена именно в южных районах. Это совпадает с данными палеоантропологии (Алексеев, 1984) и антропологии современного населения (Богданова, 1978а). Данные лингвистики, характеризующие юго-восточный диалект как сформировавшийся в результате значительного влияния монгольского языка, также полностью совпадают с полученным нами распределением этого компонента и частот гаплогрупп.

В состав гаплогруппы R1a1a (12 %) у тувинцев входят семь различных линий. Шесть мужчин тувинцев относятся к линии R1a1a1b2a2a (Z2123). К ней также принадлежат три алтайца и два киргиза. Еще пять тувинцев относятся к близкой ей линии R1a1a1b2a2a3b1a1~ (YP1542, YP1556). Она доминирует по частоте у южных алтайцев и телеутов. У 27 тувинцев обнаружена линия R1a1a1b2 (Y43109), которая разделена на три варианта, отличающихся по гаплотипам. К одному варианту принадлежат 16 тувинцев и 5 южных алтайцев; ко второму – четыре тувинца, хакасы из сеоков туран и хызыл хая и почти все шорцы сеоков тарткын, шор-кызай и кара-шор; к третьему – семь тувинцев, хакасы из различных сеоков бельтиров и бирюсинцев и шорцы сеоков челей и чедибер. Это подтверждает дан-

ные о том, что некоторые группы тувинцев, кочевавшие в Минусинской котловине и получившие позднее название «бельтиры», были полностью ассимилированы местными племенами, составив один из компонентов формирования этноса современных хакасов.

Гаплогруппа R1a1a1b2e1~ (YP1509) у тувинцев также разделена по гаплотипам на две линии. Девять образцов тувинцев первого варианта по гаплотипам очень близки к этому варианту у хакасов сеоков харга и шорцев сеока харга и челей и одному телеуту. Восемь образцов принадлежат к другому специфичному варианту, распространенному у теленгитов и северных алтайцев.

Показано очень большое разнообразие гаплогруппы R1a1a у коренного населения Алтае-Саянского региона. Различные ее сублинии разделились достаточно давно и не демонстрируют звездообразной филогении гаплотипов, пониженного разнообразия и следов эффекта основателя. Это свидетельствует о значительном размере эффективной численности популяций древних европеоидов и тюрков, которые привнесли эти компоненты в генофонд современных тувинцев, хакасов и шорцев. Эффекты основателя со значительным демографическим ростом обнаружены только у южных алтайцев, киргизов и телеутов по гаплогруппе R1a1a1b2a2a3b1a1~. Распространение различных обнаруженных сублиний гаплогруппы R1a1a на территории Тывы, Алтая, Хакасии и Шории связано, вероятнее всего, с тюрками и енисейскими киргизами.

Из других гаплогрупп у тувинцев единичными образцами представлены еще восемь (D, E, I1, I2a, J1, J2a, J2a1 и R1b). Скорее всего, их присутствие частично связано с недавней метисацией и более ранним расселением среднеазиатских популяций. Результаты исследования детальной филогении Y-хромосомных гаплогрупп позволили более точно провести анализ компонентного состава генофонда тувинцев. Это является дополнением к анализу аутосомных маркеров, что дает возможность подробно реконструировать формирование их генофонда. Данная информация важна для описания сходства и различия сравниваемых групп, а также уточнения их этногенеза. Различные гаплогруппы Y-хромосомы в составе тувинского генофонда демонстрируют их генетическую близость с алтайцами, хакасами, шорцами, бурятами, монголами, эвенками, кетами, чулымскими тюрками и телеутами. Это позволяет более детально охарактеризовать генофонд коренного южносибирского населения и генетические взаимосвязи и преемственность популяций, проживающих на этой территории.

Заключение

В настоящем исследовании проведено подробное изучение генофонда тувинцев на основании данных высокоплотных биочипов и широкого спектра SNP нерекombинирующей части Y-хромосомы. Обнаружен существенный гетерогенный состав генофонда тувинцев и хакасов как по аутосомным SNP, так и по различным сублиниям Y-хромосомных гаплогрупп. Показана максимальная близость генофонда тувинцев с алтайцами, хакасами и шорцами. Анализ IBD-блоков и отдельных редких вариантов мужских линий демонстрирует следы более ранних связей с древним аборигенным населением этого региона

и популяциями Восточной Сибири и Дальнего Востока. Внутри тувинского этноса выявлены значительные различия между выборками из западных, южных и восточных районов Тывы по доле монгольского и енисейского генетического компонентов. Генетическое разнообразие тувинцев по Y-хромосомным гаплогруппам и максимальное разнородный состав генетических компонентов свидетельствуют о самом высоком разнообразии тувинского генофонда по сравнению со всеми остальными коренными народами Сибири.

В дальнейшем мы планируем более детально проанализировать структуру генофондов южно- и западносибирских популяций, добавив популяционные выборки самодийских народов – ненцев и селькупов.

Список литературы / References

- Агджоян А.Т., Дамба Л.Д., Гурьянов В.М., Запорожченко В.В., Балановский О.П. Филогенетический анализ южносибирской гаплогруппы Q-YP1102 по данным о Y-SNP- и Y-STR-маркерах у тувинцев и окружающих популяций. *Генетика*. 2021;57(12): 1386-1395. DOI 10.31857/S001667582112002X.
- [Agdzhoyan A.T., Damba L.D., Gurianov V.M., Zaporozhchenko V.V., Balanovsky O.P. Phylogenetic analysis of the South Siberian Q-YP1102 haplogroup based on the data on Y-SNP and Y-STR markers in Tuvans and surrounding populations. *Russ. J. Genet.* 2021;57:1398-1407. DOI 10.1134/S1022795421120024.]
- Алексеев В.П. Краткое изложение палеоантропологии Тывы в связи с историческими вопросами. В: Антропо-экологические исследования в Тыве. М.: Наука, 1984;6-75.
- [Aleksseev V.P. Brief account of the paleoanthropology of Tuva in connection with historical issues. In: *Anthropoecological Research in Tuva*. Moscow: Nauka Publ., 1984;6-75. (in Russian)]
- Алексеева Т.И. Антропологические особенности современных тувинцев. Кефалометрия и кефалоскопия. В: Антропо-экологические исследования в Тыве. М.: Наука, 1984;75-114.
- [Aleksseeva T.I. Anthropological features of modern Tuvans. Cephalometry and cephaloscopy. In: *Anthropoecological Research in Tuva*. Moscow: Nauka Publ., 1984;75-114. (in Russian)]
- Богданова В.И. Антропологическое изучение современных тувинцев в 1972–1976 гг. В: Полевые исследования Института этнографии за 1976 г. М.: Наука, 1978а;187-198.
- [Bogdanova V.I. Anthropological study of modern Tuvans in 1972–1976. In: *Field Studies of the Institute of Ethnography in 1976*. Moscow: Nauka Publ., 1978а;187-198. (in Russian)]
- Богданова В.И. Некоторые вопросы формирования антропологического состава современных тувинцев. *Сов. этнография*. 1978б;6:46-58.
- [Bogdanova V.I. Some issues of the origins of the anthropological composition of present-day Tuvan people. *Sovetskaya Etnografiya = Soviet Ethnography*. 1978б;6:46-58. (in Russian)]
- Вагайцева К.В., Харьков В.Н., Черпинская К.В., Хитринская И.Ю., Степанов В.А. Генетическая вариабельность X-сцепленных STR-маркеров в популяциях Сибири. *Молекуляр. биология*. 2015;49(2):305-312. DOI 10.7868/S0026898415020147.
- [Vagaitseva K.V., Kharkov V.N., Cherpinskaya K.V., Khitrinskaya I.Yu., Stepanov V.A. Genetic variability of X-linked STR markers in Siberian populations. *Mol. Biol.* 2015;49(2):267-274. DOI 10.1134/S0026893315020132.]
- Генофонд населения Сибири. Новосибирск: Изд-во Ин-та археологии и этнографии СО РАН, 2003.
- [Gene Pool of the Population of Siberia. Novosibirsk: Publ. House of the Institute of Archeology and Ethnography SB RAS, 2003. (in Russian)]
- Дамба Л.Д., Айыжы Е.В., Монгуш Б.Б., Жабагин М.К., Юсупов Ю.М., Сабитов Ж.М., Агджоян А.Т., Маркина Н.В., Доржу Ч.М., Балановская Е.В., Балановский О.П. Комплексный подход в изучении родовой структуры тувинцев Республики Тыва на примере родов монгуш и ооржак. *Вестн. Тув. гос. ун-та. № 2. Естественные и сельскохозяйственные науки*. 2018а; 2(37):37-44.
- [Damba L.D., Aiyzhy E.V., Mongush B.B.O., Zhabagin M.K., Yusupov Yu.M., Sabitov Zh.M., Agdzhoyan A.T., Markina N.V., Dorzhu Ch.M., Balanovskaya E.V., Balanovsky O.P. Complex approach to the clan structure of Tuvans by the example of Mongush and Oorzhak clans. *Vestnik Tuvinskogo Gosudarstvennogo Universiteta. № 2. Estestvennye i Sel'skokhozyajstvennye Nauki = Bulletin of Tuva State University. No. 2. Natural and Agricultural Sciences*. 2018а;37(2):37-44. (in Russian)]
- Дамба Л.Д., Балановская Е.В., Жабагин М.К., Юсупов Ю.М., Богуннов Ю.В., Сабитов Ж.М., Агджоян А.Т., Короткова Н.А., Лавряшина М.Б., Монгуш Б.Б., Кавай-оол У.Н., Балановский О.П. Оценка вклада монгольской экспансии в генофонд тувинцев. *Вавиловский журнал генетики и селекции*. 2018б;22(5):611-619. DOI 10.18699/VJ18.402.
- [Damba L.D., Balanovskaya E.V., Zhabagin M.K., Yusupov Yu.M., Bogunov Yu.V., Sabitov Zh.M., Agdzhoyan A.T., Korotkova N.A., Lavryashina M.B., Mongush B.B., Kavai-ool U.N., Balanovsky O.P. Estimating the impact of Mongol expansion on gene pool of Tuvans. *Vavilovskii Zhurnal Genetiki i Selekcii = Vavilov Journal of Genetics and Breeding*. 2018б;22(5):611-619. DOI 10.18699/VJ18.402. (in Russian)]
- Дебес Г.Ф. Палеоантропология СССР. М.; Л.: Изд-во АН СССР, 1948.
- [Debets G.F. Paleoanthropology of the USSR. Moscow; Leningrad: Publishing House of the USSR Academy of Sciences, 1948. (in Russian)]
- Киселев С.В. История Южной Сибири. М.: АН СССР, 1951.
- [Kiselev S.V. History of South Siberia. Moscow: Publishing House of the USSR Academy of Sciences, 1951. (in Russian)]
- Колесников Н.А., Харьков В.Н., Зарубин А.А., Воевода М.И., Губина М.А., Штыгашева О.В., Максимова Н.Р., Сухомясова А.Л., Степанов В.А. Сигналы направленного отбора в популяциях коренного населения Сибири. *Генетика*. 2022;58(4):470-474. DOI 10.31857/S0016675822040075.
- [Kolesnikov N.A., Kharkov V.N., Zarubin A.A., Voevoda M.I., Gubina M.A., Shtygasheva O.V., Maksimova N.R., Sukhomyasova A.L., Stepanov V.A. Signals of directed selection in the Indigenous populations of Siberia. *Russ. J. Genet.* 2022;58(4):473-477. DOI 10.1134/S102279542204007X.]
- Колесников Н.А., Харьков В.Н., Зарубин А.А., Раджабов М.О., Воевода М.И., Губина М.А., Хуснутдинова Э.К., Литвинов С.С., Екомасова Н.В., Штыгашева О.В., Максимова Н.Р., Сухомясова А.Л., Степанов В.А. Особенности геномного распределения регионов высокой гомозиготности у коренного населения Северной Евразии на индивидуальном и популяционном уровнях на основе анализа SNP высокой плотности. *Генетика*. 2021;57(11): 1261-1275. DOI 10.31857/S0016675821110059.
- [Kolesnikov N.A., Kharkov V.N., Zarubin A.A., Radzhabov M.O., Voevoda M.I., Gubina M.A., Khusnutdinova E.K., Litvinov S.S., Ekomasova N.V., Shtygasheva O.V., Maksimova N.R., Sukhomyasova A.L., Stepanov V.A. Features of the genomic distribution of runs of homozygosity in the indigenous population of Northern Eurasia at the individual and population levels based on high density SNP analysis. *Russ. J. Genet.* 2021;57(11):1271-1284. DOI 10.1134/S1022795421110053.]
- Кучер А.Н., Ондар Э.А., Степанов В.А. Тувинцы: гены, демография, здоровье. Томск: Печатная мануфактура, 2003.
- [Kucher A.N., Ondar E.A., Stepanov V.A. Tuvinians: genes, demography, health. Tomsk: Pechatnaya Manufaktura Publ., 2003. (in Russian)]
- Потапов Л.П. Очерки народного быта тувинцев. М.: Наука, 1969.
- [Potapov L.P. Essays on the Folk Life of the Tuvans. Moscow: Nauka Publ., 1969. (in Russian)]

- Пузырев В.П., Эрдыниева Л.С., Кучер А.Н. Генетико-эпидемиологическое исследование населения Тувы. Томск: STT, 1999. [Puzyrev V.P., Erdynieva L.S., Kucher A.N. Genetic and Epidemiological Study of the Population of Tuva. Tomsk: STT Publ., 1999. (in Russian)]
- Спицын В.А., Боева С.Б., Филиппов И.К. Генетико-антропологическое изучение коренного населения Алтае-Саянского нагорья. В: Антропо-экологические исследования в Туве. М.: Наука, 1984;185-194. [Spitsyn V.A., Bueva S.B., Filippov I.K. Genetic and anthropological study of the indigenous population of the Altai-Sayan highland. In: Anthro-ecological Research in Tuva. Moscow: Nauka Publ., 1984;185-194. (in Russian)]
- Степанов В.А., Пузырев В.П. Анализ аллельных частот семи микросателлитных локусов Y-хромосомы в трех популяциях тувинцев. *Генетика*. 2000;36(2):241-248. [Stepanov V.A., Puzyrev V.P. Analysis of the allele frequencies of seven Y-chromosome microsatellite loci in three Tuvian populations. *Russ. J. Genet.* 2000;36(2):179-185.]
- Степанов В.А., Харьков В.Н., Пузырев В.П. Эволюция и филогеография линий Y-хромосомы человека. *Информационный вестник ВОВиС*. 2006;10(1):57-73. [Stepanov V.A., Kharkov V.N., Puzyrev V.P. Evolution and phylogeography of human Y-chromosomal lineages. *Informatsionny Vestnik VOGiS = The Herald of Vavilov Society for Geneticists and Breeding Scientists*. 2006;10(1):57-73. (in Russian)]
- Степанов В.А., Хитринская И.Ю., Пузырев В.П. Генетическая дифференциация населения Тувы по полиморфным Alu-инсерциям. *Генетика*. 2001;37(4):563-569. [Stepanov V.A., Khitrinskaya I.Yu., Puzyrev V.P. Genetic differentiation of the Tuva population with respect to the Alu-insertions. *Russ. J. Genet.* 2001;37(4):453-459. DOI 10.1023/A:1016623030663.]
- Тюркские народы Восточной Сибири. М.: Наука, 2008. [Turkic Peoples of East Siberia. Moscow: Nauka Publ., 2008. (in Russian)]
- Харьков В.Н., Новикова Л.М., Штыгашева О.В., Лузина Ф.А., Хитринская И.Ю., Волков В.Г., Степанов В.А. Генофонд хакасов и шорцев по маркерам Y-хромосомы: общие компоненты и генетическая структура родов. *Генетика*. 2020;56(7):826-833. DOI 10.31857/S0016675820070073. [Kharkov V.N., Novikova L.M., Shtygasheva O.V., Luzina F.A., Khitrinskaya I.Yu., Volkov V.G., Stepanov V.A. Gene pool of Khakass and Shors for Y chromosome markers: common components and tribal genetic structure. *Russ. J. Genet.* 2020;56(7):849-855. DOI 10.1134/S1022795420070073.]
- Харьков В.Н., Хамина К.В., Медведева О.Ф., Симонова К.В., Хитринская И.Ю., Степанов В.А. Структура генофонда тувинцев по маркерам Y-хромосомы. *Генетика*. 2013;49(12):1416-1425. DOI 10.7868/S0016675813120035. [Kharkov V.N., Khamina K.V., Medvedeva O.F., Simonova K.V., Khitrinskaya I.Yu., Stepanov V.A. Gene-pool structure of Tuvians inferred from Y-chromosome marker data. *Russ. J. Genet.* 2013;49(12):1236-1244. DOI 10.1134/S102279541312003X.]
- Хитринская И.Ю., Харьков В.Н., Степанов В.А. Генетическое разнообразие X-хромосомы в популяциях коренных этносов Сибири: структура неравновесия по сцеплению и филогеография гаплотипов локуса ZFX. *Молекуляр. биология*. 2010;44(5):804-815. [Khitrinskaya I.Yu., Khar'kov V.N., Stepanov V.A. Genetic diversity of the chromosome X in aboriginal Siberian populations: The structure of linkage disequilibrium and haplotype phylogeography of the ZFX locus. *Mol. Biol.* 2010;44(5):709-719. DOI 10.1134/S0026893310050055.]
- Alexander D.H., Lange K. Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics*. 2011;12:246. DOI 10.1186/1471-2105-12-246.
- Alexander D.H., Novembre J., Lange K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 2009;19(9):1655-1664. DOI 10.1101/gr.094052.
- Bandelt H.J. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* 1999;16(1):37-48. DOI 10.1093/oxfordjournals.molbev.a026036.
- Browning B.L., Browning S.R. Improving the accuracy and efficiency of identity-by-descent detection in population data. *Genetics*. 2013;194(2):459-471. DOI 10.1534/genetics.113.150029.
- Browning S.R., Browning B.L. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am. J. Hum. Genet.* 2007;81(5):1084-1097. DOI 10.1086/521987.
- Derenko M., Malyarchuk B., Denisova G., Wozniak M., Dambueva I., Dorzhu C., Luzina F., Miścicka-Sliwka D., Zakharov I. Contrasting patterns of Y-chromosome variation in South Siberian population from Baikal and Altai-Sayan regions. *Hum. Genet.* 2006;118(5):591-604. DOI 10.1007/s00439-005-0076-y.
- Guo Y., He J., Zhao S., Wu H., Zhong X., Sheng Q., Samuels D.C., Shyr Y., Long J. Illumina human exome genotyping array clustering and quality control. *Nat. Protoc.* 2014;9(11):2643-2662. DOI 10.1038/nprot.2014.174.
- Gusev A., Palamara P.F., Aponte G., Zhuang Z., Darvasi A., Gregeresen P., Pe'er I. The architecture of long-range haplotypes shared within and across populations. *Mol. Biol. Evol.* 2012;29(2):473-486. DOI 10.1093/molbev/msr133.
- Ilumäe A.-M., Reidla M., Chukhryaeva M., Järve M., Post H., Karmin M., Saag L., Agdzhoyan A., Kushniarevich A., Litvinov S., Ekomasova N., Tambets K., Metspalu E., Khusainova R., Yunusbayev B., Khusnutdinova E.K., Osipova L.P., Fedorova S., Utevska O., Koshel S., Balanovska E., Behar D.M., Balanovsky O., Kivisild T., Underhill P.A., Villems R., Rootsi S. Human Y chromosome haplogroup N: A non-trivial time-resolved phylogeography that cuts across language families. *Am. J. Hum. Genet.* 2016;99(1):163-173. DOI 10.1016/j.ajhg.2016.05.025.
- Skotte L., Korneliussen T.S., Albrechtsen A. Estimating individual admixture proportions from next generation sequencing data. *Genetics*. 2013;195(3):693-702. DOI 10.1534/genetics.113.154138.
- Zhivotovsky L.A., Underhill P.A., Cinnioglu C., Kayser M., Morar B., Kivisild T., Scozzari R., Cruciani F., Destro-Bisol G., Spedini G., Chambers G.K., Herrera R.J., Yong K.K., Gresham D., Tournev I., Feldman M.W., Kalaydjieva L. The effective mutation rate at Y-chromosome STRs with application to human population divergence time. *Am. J. Hum. Genet.* 2004;74(1):50-61. DOI 10.1086/380911.

ORCID ID

V.A. Stepanov orcid.org/0000-0002-5166-331X
N.A. Kolesnikov orcid.org/0000-0001-8855-577X
A.A. Zarubin orcid.org/0000-0001-6568-6339
V.N. Kharkov orcid.org/0000-0002-1679-2212

Благодарности. Исследование выполнено за счет гранта Российского научного фонда № 22-64-00060 (<https://rscf.ru/project/22-64-00060/>).

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию 14.10.2022. После доработки 28.12.2022. Принята к публикации 28.12.2022.

Перевод на английский язык <https://vavilov.elpub.ru/jour>

Связь генофонда хантов с народами Западной Сибири, Предуралья и Алтая-Саян по данным о полиморфизме ауtosомных локусов и Y-хромосомы

В.Н. Харьков , Н.А. Колесников, Л.В. Валихова, А.А. Зарубин, М.Г. Сваровская, А.В. Марусин, И.Ю. Хитринская, В.А. Степанов

Научно-исследовательский институт медицинской генетики, Томский национальный исследовательский медицинский центр Российской академии наук, Томск, Россия
 vladimir-kharkov@medgenetics.ru

Аннотация. Ханты – коренной сибирский народ, проживающий на территории Западной Сибири, в основном на территории Ханты-Мансийского и Ямало-Ненецкого автономных округов. Настоящее исследование направлено на комплексный анализ структуры генофонда хантов и их сравнение с другими популяциями коренного населения Южной и Западной Сибири. Для решения вопросов генетической близости хантов с другими коренными народами выполнено генотипирование широкого геномного набора ауtosомных маркеров с помощью высокоплотных биочипов, а также расширенного набора SNP- и STR-маркеров Y-хромосомы у различных этнических групп: хакасов, тувинцев, южных алтайцев, сибирских татар, чулымцев (тюркская языковая семья) и кетов (енисейская языковая семья). Результаты анализа частот ауtosомных SNP различными методами, сходства по составу гаплогрупп Y-хромосомы и YSTR-гаплотипов свидетельствуют, что генофонд хантов достаточно специфичен. При анализе ауtosомных SNP в обеих выборках полностью доминирует угорский генетический компонент (до 99–100 %). Выборки хантов показали максимальное совпадение по IBD-блокам между собой, с выборкой кетов, чулымцев, тувинцев, томских татар, хакасов-качинцев и южных алтайцев. Степень совпадения IBD-блоков между хантами, кетами и томскими татарами согласуется с результатами распределения в этих популяциях частот аллелей и общих генетических компонентов. По составу гаплогрупп Y-хромосомы две выборки хантов значительно различаются между собой. Детальный филогенетический анализ различных гаплогрупп Y-хромосомы позволил описать и уточнить различия в филогении и структуре отдельных этноспецифичных сублиний, определить их родство, следы экспансии численности в генофонде хантов. Варианты разных гаплогрупп Y-хромосомы у хантов, хакасов и тувинцев восходят к общим для них предковым линиям. Результаты сравнительного анализа образцов мужчин также свидетельствуют о близком генетическом родстве между хантами и ненцами, коми, удмуртами и кетами. Специфичность гаплотипов, обнаружение различных терминальных SNP подтверждают, что ханты достаточно долго не имели контактов с другими этносами, кроме ненцев, в состав которых вошло много хантыйских родов. Ключевые слова: генофонд; популяции человека; генетическое разнообразие; генетические компоненты; Y-хромосома; ханты.

Для цитирования: Харьков В.Н., Колесников Н.А., Валихова Л.В., Зарубин А.А., Сваровская М.Г., Марусин А.В., Хитринская И.Ю., Степанов В.А. Связь генофонда хантов с народами Западной Сибири, Предуралья и Алтая-Саян по данным о полиморфизме ауtosомных локусов и Y-хромосомы. *Вавиловский журнал генетики и селекции*. 2023;27(1):46-54. DOI 10.18699/VJGB-23-07

Relationship of the gene pool of the Khants with the peoples of Western Siberia, Cis-Urals and the Altai-Sayan Region according to the data on the polymorphism of autosomic locus and the Y-chromosome

V.N. Kharkov , N.A. Kolesnikov, L.V. Valikhova, A.A. Zarubin, M.G. Svarovskaya, A.V. Marusin, I.Yu. Khitrinskaya, V.A. Stepanov

Research Institute of Medical Genetics, Tomsk National Research Medical Center of the Russian Academy of Sciences, Tomsk, Russia
 vladimir-kharkov@medgenetics.ru

Abstract. Khanty are indigenous Siberian people living on the territory of Western Siberia, mainly on the territory of the Khanty-Mansiysk and Yamalo-Nenets Autonomous Okrugs. The present study is aimed at a comprehensive analysis of the structure of the Khanty gene pool and their comparison with other populations of the indigenous population of Southern and Western Siberia. To address the issues of genetic proximity of the Khanty with other indigenous peoples, we performed genotyping of a wide genomic set of autosomal markers using high-density biochips, as well

as an expanded set of SNP and STR markers of the Y-chromosome in various ethnic groups: Khakas, Tuvans, Southern Altaians, Siberian Tatars, Chulymys (Turkic language family) and Kets (Yeniseian language family). The structure of the gene pool of the Khanty and other West Siberian and South Siberian populations was studied using a genome-wide panel of autosomal single nucleotide polymorphic markers and Y-chromosome markers. The results of the analysis of autosomal SNPs frequencies by various methods, the similarities in the composition of the Y-chromosome haplogroups and YSTR haplotypes indicate that the Khanty gene pool is quite specific. When analyzing autosomal SNPs, the Ugrian genetic component completely dominates in both samples (up to 99–100 %). The samples of the Khanty showed the maximum match in IBD blocks with each other, with a sample of the Kets, Chulymys, Tuvans, Tomsk Tatars, Khakas, Kachins, and Southern Altaians. The degree of coincidence of IBD blocks between the Khanty, Kets, and Tomsk Tatars is consistent with the results of the distribution of allele frequencies and common genetic components in these populations. According to the composition of the Y-chromosome haplogroups, the two samples of the Khanty differ significantly from each other. A detailed phylogenetic analysis of various Y-chromosome haplogroups made it possible to describe and clarify the differences in the phylogeny and structure of individual ethnospecific sublines, to determine their relationship, traces of population expansion in the Khanty gene pool. Variants of different haplogroups of the Y-chromosome in the Khanty, Khakas and Tuvans go back to their common ancestral lines. The results of a comparative analysis of male samples indicate a close genetic relationship between the Khanty and Nenets, Komi, Udmurts and Kets. The specificity of haplotypes, the discovery of various terminal SNPs confirms that the Khanty did not come into contact with other ethnic groups for a long time, except for the Nenets, which included many Khanty clans. Key words: gene pool; human population; genetic diversity; genetic components; Y-chromosome; Khanty.

For citation: Kharkov V.N., Kolesnikov N.A., Valikhova L.V., Zarubin A.A., Svarovskaya M.G., Marusin A.V., Khitrinskaya I.Yu., Stepanov V.A. Relationship of the gene pool of the Khants with the peoples of Western Siberia, Cis-Urals and the Altai-Sayan Region according to the data on the polymorphism of autosomic locus and the Y-chromosome. *Vavilovskii Zhurnal Genetiki i Seleksii = Vavilov Journal of Genetics and Breeding*. 2023;27(1):46-54. DOI 10.18699/VJGB-23-07

Введение

Изучение структуры генофондов популяций различных сибирских регионов – одно из приоритетных направлений современной генетики человека, оно помогает раскрыть детально часть вопросов, связанных с их этногенезом.

Ханты – коренной народ, проживающий на территории Западной Сибири, в основном на территории Ханты-Мансийского (ХМАО) и Ямало-Ненецкого (ЯНАО) автономных округов, а также Тюменской области. Небольшие группы хантов проживают на севере Томской области и в Республике Коми. По данным Всероссийской переписи 2010 г., численность хантов составила 30943 человек, из которых 61.6 % проживают в ХМАО и 30.7 % – в ЯНАО. У хантов выделяются три большие этнографические группы, совпадающие с группами их языковых диалектов – северные, южные и восточные. Причем южные (прииртышские) ханты были тюркизированы и вошли в состав сибирских татар, смешавшись с ними, а также ассимилированы русскими переселенцами (Народы Западной Сибири, 2005).

Популяции хантов представляют значительный интерес для популяционно-генетических исследований как в силу относительно слабой их изученности с привлечением современных геномных технологий, так и по причине специфичности генофондов отдельных групп, развивавшихся в условиях длительной генетической изоляции.

Расселение хантов в древности было очень широким – от низовий Оби на севере до Барабинских степей на юге и от Енисея на Востоке до Зауралья, включая реки Северная Сосьва и Ляпин, а также часть рек Пелым и Конда на западе. С XIX в. за Урал из Прикамья и Приуралья стали переселяться манси, которых теснили коми-зыряне и русские. С более раннего времени к северу уходила и часть южных манси в связи с созданием в XIV–XV вв. Тюменского и Сибирского ханств – государств сибирских татар, а позднее (XVI–XVII вв.) и с освоением Сибири русскими. В XVII–XVIII вв. на Пелыме и Конде уже жили манси.

Часть хантов также переселялись из западных областей на восток и север (на Обь с ее левых притоков), это фиксируется статистическими данными архивов. Их места занимали манси. Так, к концу XIX в. на реках Северная Сосьва и Ляпин не осталось остяцкого населения, которое либо переселилось на Обь, либо слилось с новоприсельцами (Народы России, 1994).

На севере ханты вступали в контакт с ненцами, часть их была ассимилирована ими, что подтверждается этнографическими данными, а также нашим исследованием родовой структуры гыданских ненцев по маркерам Y-хромосомы (Харьков и др., 2021). Миграция хантов на север и восток продолжалась и в XX в. К XX в. южные ханты были почти полностью ассимилированы татарами и русскими.

Исторически сложилось так, что хантыйское население не было однородным ни по языку, ни по культуре. Одни ученые разделяют хантыйский язык на две крупные группы – западную и восточную, а другие еще подразделяют западные диалекты на южные и северные. В антропологическом отношении ханты – наиболее характерные представители уральского типа, к которому относятся также манси, селькупы, ненцы, барабинские татары, шорцы, северные алтайцы и хакасы. Самые близкие родственники хантов по происхождению, языку и культуре – манси (Брук, 1986).

Целью настоящего исследования стали комплексный анализ структуры генофонда хантов и реконструкция их происхождения в сравнении с другими популяциями коренного населения Южной и Западной Сибири. Для решения вопросов генетической близости хантов с другими коренными народами было выполнено генотипирование широкого геномного набора аутосомных маркеров с помощью высокоплотных биочипов, а также расширенного набора SNP- и STR-маркеров Y-хромосомы у различных этнических групп: хакасов, тувинцев, южных алтайцев, сибирских татар, чулымцев (тюркская языковая семья) и кетов (енисейская языковая семья).

Материалы и методы

Материалом исследования были образцы ДНК мужчин и женщин из двух популяций хантов деревни Русскинская, Сургутского района и села Казым, Белоярского района ХМАО. Забор первичного биологического материала (венозной крови) у доноров производили с соблюдением процедуры письменного информированного согласия на проведение исследования. На каждого донора составляли анкету с краткой родословной, указанием этнической принадлежности и места рождения предков. Индивиды относили к данной этнической группе на основании этнической идентификации – его собственной, его родителей, и места рождения.

Для анализа Y-хромосомных гаплогрупп и гаплотипов хантов использованы 120 образцов ДНК мужчин д. Русскинская ($N = 64$) и с. Казым ($N = 54$) ХМАО. Для генотипирования на высокоплотных микрочипах были выбраны неродственные образцы с. Казым ($N = 30$) и д. Русскинская ($N = 26$). Другие популяции коренного населения Сибири представлены: чулымцами ($N = 22$), хакасами (сагайцами Таштыпского района, $N = 29$, и качинцами Ширинского района, $N = 26$), южными алтайцами (с. Бешпельтир Чемальского района, $N = 24$, и с. Кулада Онгудайский района, $N = 25$), кетами (пос. Келлог Туруханского района Красноярского края, $N = 15$), томскими татарами (пос. Черная Речка, пос. Эушта и пос. Тахтамышево Томского района, $N = 20$), тувинцами (с. Тээли Бай-Тайгинского кожууна, $N = 28$).

Данные широкогеномного генотипирования получены с использованием микрочипов Infinium Multi-Ethnic Global-8 (Illumina) для SNP-генотипирования, включающего свыше 1.7 млн маркеров. Материал депонирован в биоресурсной коллекции «Биобанк населения Северной Евразии».

Кластеризацию массива генотипов аутосомных SNP (single nucleotide polymorphism) и контроль качества выполняли с помощью протокола, разработанного Y. Guo с коллегами (2014), с использованием GenomeStudio (Illumina. GenomeStudio, модуль генотипирования v2.0.3), программного пакета, который компания Illumina создала для различных геномных анализов. Для фильтрации, нормализации и расчета стандартных геномных статистик и показателей оптимальным был стандартный набор программ, включающий vcftools, bcftools, и plink.

Для анализа идентичных по происхождению блоков сцепления использовали алгоритм Refined IBD (Browning B.L., Browning S.R., 2013), показывающий более точные результаты по сравнению с встроенными в plink алгоритмами. Предварительно генотипы были фазированы с помощью программного обеспечения Beagle 5.1 (Browning S.R., Browning B.L., 2007). Для сравнения популяций были получены суммы средних длин идентичных по происхождению блоков (сегментов IBD – identical by descent) между парами индивидов.

Для анализа генетических взаимоотношений между популяциями применяли метод tSNE. Для определения компонентного состава и количества примесей у отдельных индивидов и популяций использованы методика NGSadmix (Skotte et al., 2013) и программа ADMIXTURE (Alexander et al., 2009; Alexander, Lange, 2011).

Для изучения состава и структуры гаплогрупп Y-хромосомы в исследование были включены две системы генетических маркеров: диаллельных локусов, представленных SNP, и полиаллельных высоковариабельных микросателлитов (YSTR). С помощью 138 SNP-маркеров определяли принадлежность образцов к различным гаплогруппам. Классификация гаплогрупп соответствует данным Международного общества генетической генеалогии (сайт www.isogg.org).

Анализ STR-гаплотипов внутри гаплогрупп проводили с использованием 44 STR-маркеров нерекombинирующей части Y-хромосомы (DYS19, 385a, 385b, 388, 389I, 389II, 390, 391, 392, 393, 426, 434, 435, 436, 437, 438, 439, 442, 444, 445, 448, 449, 456, 458, 460, 461, 481, 504, 505, 518, 525, 531, 533, 537, 552, 570, 576, 635, 643, YCAIIa, YCAIIb, GATA H4.1, Y-GATA-A10, GGAAT1B07). STR-маркеры генотипировали с помощью капиллярного электрофореза на генетическом анализаторе ABI Prism 3730. Генотипирование SNP-маркеров проводили с применением полимеразной цепной реакции и последующего анализа фрагментов ДНК с помощью ПДРФ-анализа (полиморфизм длин рестрикционных фрагментов).

Экспериментальные исследования осуществлены на базе Центра коллективного пользования научно-исследовательским оборудованием «Медицинская геномика» (НИИ медицинской генетики Томского НИМЦ). Построение медианных сетей гаплотипов Y-хромосомы выполняли с использованием программы Network v.10.2.0.0 (Fluxus Technology Ltd; www.fluxus-engineering.com) по методу медианных сетей Бандельта (Bandelt et al., 1999). Оценку возраста генерации наблюдаемого разнообразия гаплотипов в гаплогруппах проводили методом ASD (Zhitovskiy et al., 2004) на основании средних квадратичных отличий в числе повторов между всеми маркерами.

Результаты и обсуждение

Полученный в результате генотипирования высокоплотных микрочипов большой массив данных по аутосомным SNP в выборках хантов и других коренных сибирских народов позволяет максимально детально охарактеризовать генофонд исследованных выборок с применением различных методик. Генотипирование расширенного набора специфичных SNP Y-хромосомы из различных гаплогрупп дает возможность гораздо точнее описать молекулярно-филогенетическую и филогеографическую структуры отдельных гаплогрупп Y-хромосомы.

После обработки данных по результатам микрочипового исследования для фильтрации прогенотипированных образцов и выполнения дальнейших расчетов был проведен поиск среди хантов метисов с использованием программы NGSadmix. Алгоритм этой программы позволяет определять соотношение предковых компонентов по данным NGS с относительно малой глубиной покрытия. Принцип расчетов аналогичен другим программам, таким как FRAPPE и ADMIXTURE, но NGSadmix, в отличие от них, эффективно работает, когда имеется статистическая неточность в отдельных генотипах. Метод NGSadmix при запуске на сформированном нами массиве данных показал, что практически все образцы хантов не имеют метисации, что полностью соответствует данным анкетирования.

рования ДНК-доноров. Метисация с русскими (до 30 %) обнаружена только для одного мужчины из д. Русскинская. Его принадлежность к европейской Y-хромосомной линии R1b1a1b-L407 подтверждает метисацию с отцовской стороны. Этот образец был исключен из дальнейших расчетов.

Полученные данные о частотах SNP в исследованных образцах были использованы для выяснения генетических взаимоотношений между включенными в работу популяционными выборками. Для понижения размерности, пространственного анализа и выявления генетических компонентов мы остановились на двух алгоритмах: tSNE и ADMIXTURE. Метод tSNE позволяет более четко разделить массив данных на отдельные этноспецифичные группы образцов по сравнению с методом главных компонент (PCA).

Генетические взаимоотношения хантов с другими популяциями Западной и Южной Сибири

При анализе массива данных по частотам аутосомных SNP с помощью метода tSNE на уровне отдельных образцов показано, что две выборки хантов очень близки, при этом образцы казымских и русскинских хантов не пересекаются на графике и отделены друг от друга (рис. 1).

Ханты характеризуются специфическими особенностями генофонда и не кластеризуются с другими популяциями. По сравнению с субэтническими группами хакасов и южных алтайцев из разных поселков более удаленные географически выборки хантов демонстрируют максимальную генетическую близость. Ближе всего к хантам расположены выборки кетов и томских татар. Генетические расстояния между хантами и популяциями Южной Сибири значительно больше. Этнически и территориально близкие друг другу выборки располагаются на рис. 1 довольно близко, но каждый образец входит в отдельный этноспецифичный кластер. Исключение составляют лишь несколько единичных образцов хакасов.

Компонентный состав генофонда популяций. Современные методы, применяемые в геномных исследованиях, и новые биоинформационные подходы дают возможность достоверно выявлять предковые генетические компоненты разного происхождения в составе генофонда различных популяций и отдельных индивидов. Для обнаружения индивидуальных генетических компонентов в генофонде изучаемых популяций использована программа ADMIXTURE, позволяющая выявлять смешанность состава набора индивидов на основе данных о генотипах и тем самым строить предположения о происхождении популяции.

Моделирование с помощью ADMIXTURE в последнее время считается одним из основных методов анализа при исследовании генофондов современных и древних популяций человека, позволяет анализировать одни и те же данные на разных иерархических уровнях. При задании числа предковых компонентов больше двух в большинстве исследованных популяций обнаруживается специфичный для хантов генетический компонент, наиболее отчетливо проявляющийся на анализируемом массиве популяционных выборок при $K = 8$, который можно интерпретировать как «угорский» генетический пласт в генофонде современных популяций. Ханты характеризуются

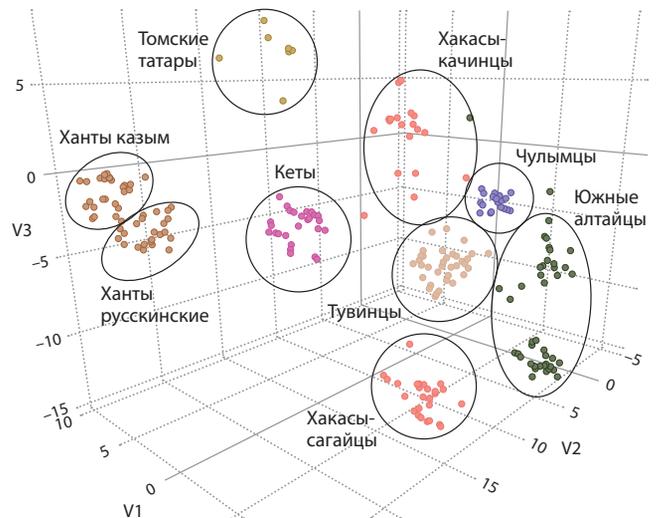


Рис. 1. Дифференциация геномов населения Южной и Западной Сибири по трем компонентам tSNE.

доминированием этого компонента, который является их генетической основой (до 99–100 % на уровне большинства индивидов). Значительна доля данного компонента также у кетов (до 45–50 % у отдельных индивидов) и томских татар (до 5–9 %). Ранее было показано, что существенную долю этот компонент занимает также в генофонде популяций Волго-Уральского региона – башкир (до 25 %), марийцев (до 20 %), коми, удмуртов и чувашей (до 15 %). С меньшей частотой он представлен практически во всех южносибирских выборках – у тувинцев, чулымцев, алтайцев и хакасов-сагайцев (от 5 до 10 %) (Харьков и др., 2020).

Доминирование угорского компонента у всех образцов хантов, начиная с $K = 3$, и практически полное отсутствие других генетических компонентов в их геномах на индивидуальном и популяционном уровне свидетельствуют, что их предковые популяции очень долгое время находились в генетической изоляции. Это говорит о том, что древнее угорское население современной территории расселения хантов не смешивалось с другими этносами, и подтверждает отсутствие других групп переселенцев с территории Южной Сибири и степной зоны.

Полученный результат показывает, что общая картина распределения компонентов хорошо согласуется с географическим местоположением исследуемых популяций, привязкой к конкретному региону, антропологическим и языковым различиям. Эта информация позволяет более точно судить о сходстве и различии сравниваемых популяций по составу предковых компонентов, а также о процессе формирования их генофонда.

Идентичные по происхождению блоки сцепления. В результате биоинформационной обработки данных генотипирования высокоплотных биочипов различных сибирских популяций проведен анализ совпадения общих по происхождению фрагментов ДНК между популяциями и отдельными индивидами. Сегментом, имеющим идентичные нуклеотидные последовательности, является IBD у двух или более лиц, если они унаследовали его от общего

предка без рекомбинации, т. е. у этих людей сегмент имеет общее происхождение. Ожидаемая длина сегмента IBD зависит от количества поколений с момента появления последнего общего предка. Одно из применений анализа общих по происхождению участков генома – количественная оценка степени родства между индивидами, что также может дополнить информацию о генетических связях популяций (Gusev et al., 2012).

Выборки хантов показали максимальное совпадение по IBD-блокам между собой (6 %), далее – с выборкой кетов (1.45 %), чулымцев (0.71 %), тувинцев (0.35 %), томских татар (0.33 %), хакасов-качинцев (0.32 %) и южных алтайцев (0.28 %). Среди хантов большее совпадение IBD-блоков наблюдается у русскиных хантов (23.5 %) по сравнению с казымскими (18.1 %).

Степень совпадения IBD-блоков между хантами, кетами и томскими татарами согласуется с результатами tSNE и ADMIXTURE по распределению в этих популяциях частот аллелей и общих генетических компонентов. При этом у русскиных хантов, имеющих наибольшую сумму средних длин сегментов IBD между парами индивидов, наибольший вклад вносят IBD длиной более 10 cM (42–46 %), что говорит о недавнем сильном инбридинге внутри популяции. Для подтверждения этого был рассчитан коэффициент инбридинга FROH для всех индивидов для трех классов блоков гомозиготности (ROH). Среди западно-сибирских популяций максимальные значения имеет популяция чулымцев (0.0292), близкие по значению – казымские (0.0280) и русскиные ханты (0.0266) и кеты (0.0259). Среди южносибирских популяций, включающих алтайцев, томских татар, тувинцев и хакасов, максимальное значение установлено и для выборки хакасов-сагайцев из предгорного Таштыпского района (0.0318), вдвое превосходящее у хакасов-качинцев равнинного Ширинского района. Минимальное значение характерно для томских татар (0.0071).

Наблюдается высокая корреляция для FROH > 1.5 с суммой средних длин сегментов IBD (IBD > 1.5 cM) между парами индивидов внутри сибирских популяций ($r = 0.9246$, $p < 5.612e-09$). Для расчета коэффициента

корреляции Спирмена использовали cor.test в программе R. Отношение суммы средних длин сегментов IBD (IBD > 1.5 cM) между парами индивидов к коэффициенту геномного инбридинга (FROH > 1.5) у русскиных хантов выше, чем у казымских. Эти показатели геномного инбридинга и распределения длин IBD внутри популяций хантов хорошо согласуются с их территориальной изолированностью и подтверждают отсутствие недавних потоков генов между популяциями на протяжении сотен лет.

Гаплогруппы Y-хромосомы. В результате анализа частоты встречаемости использованных SNP-маркеров в исследуемых выборках хантов было выявлено восемь гаплогрупп Y-хромосомы. По составу и частотам гаплогрупп выборки русскиных и казымских мужчин хантов существенно отличаются друг от друга. Только две гаплогруппы присутствуют в обеих выборках (см. таблицу).

К сублинии N1a2b1b1 у русскиных хантов принадлежат 39 образцов, а у казымских – всего 3. Терминальными для этой линии у хантов являются SNP Y68212, Y70717, Y70315, Y70327. Эта хантыйская сублиния близка к вариантам N1a2b1b1 у чулымцев (VL65, Z35095, Z35099, Z35102) и хакасов-качинцев (Z35093, Z35097, Z35103) (Валихова и др., 2022).

Гаплогруппа N1a2b1b1 у хантов – этноспецифичная, не совпадает по терминальным SNP и гаплотипам с доминирующей у ненцев N1a2b1b1a~ (B171, B170, Z35091, Z35092) (Харьков и др., 2021).

Особенность этнического состава большинства южно-сибирских народов – наличие родов (сеоков), где счет родства ведется по мужской линии. Такая родовая структура характерна для шорцев, хакасов, северных и южных алтайцев, телеутов. Все остальные образцы мужчин из различных западно- и южносибирских популяций (энцев, хакасов-сагайцев, шорцев, челканцев и тувинцев, а также хакаских сеоков, ранее входивших в состав бельтиров и бирюсинцев, ассимилированных в конце 19–начале 20 в.), относятся к другим сублиниям гаплогруппы N1a2b. Медианная сеть гаплотипов (рис. 2) демонстрирует у хантов звездообразную филогению с недавним эффектом основателя и преобладанием по частоте предкового гаплотипа.

Специфичный кластер гаплотипов хантов равноудален от всех сеоков хакасов-качинцев. Возраст этого кластера у хантов составил 858 лет (SD = 338 лет), что примерно в полтора-два раза превышает возраст кластеров качинских сеоков хасха – 487 лет (SD = 153 года), ызыр – 501 год (SD = 203 года), соххы – 585 лет (SD = 215 лет) (Харьков и др., 2020) и чулымских тюрок – 667 лет (SD = 194 года). Таким образом, ханты по этой гаплогруппе имеют прямую генетическую связь с качинцами, чулымцами и ненцами, предковые линии которых разошлись достаточно давно и отражают их связь с народами самодийской языковой группы.

Вторая общая для двух выборок хантов гаплогруппа – N1a2b2a1 (VL97, L1419, Y3185, Y3188, Y3189, Y3190, Y111190), ранее обозначавшаяся как европейская линия N1b-E. Эта сублиния обнаружена у башкир, казанских татар, коми, марийцев, карел, вепсов, финнов и русских (<https://www.yfull.com/>). Филогенетически ближе всего к хантам по этой линии расположены образцы коми. Этноспецифичные ветви хантов и коми объединяют SNP

Частота встречаемости гаплогрупп Y-хромосомы у хантов

Гаплогруппа	д. Русскиная % (N)	с. Казым % (N)
N1a2b1b1 (Y68212, Y70717, Y70315, Y70327)	60.9 (39)	5.5 (3)
N1a2b2a1 (VL97, L1419, Y3185, Y3188, Y3189, Y3190, Y111190)	9.4 (6)	16.6 (9)
N1a2b2b1~ (Z35076)	–	5.5 (3)
N1a1a1a2a1c1~ (Y13850, Y13852)	–	24.1 (13)
N1a2b1b1b1~ (B172, Z35108)	–	9.2 (5)
Q1b1a3b1a2~ (Z35974 xB32, B33, Z35993)	–	38.9 (21)
Q1a2b~ (M25, L716, YP1674, YP1676)	4.7 (3)	–
R1a1a1b2 (Y43850, S7280, FGC687, FGC38304)	25.0 (16)	–

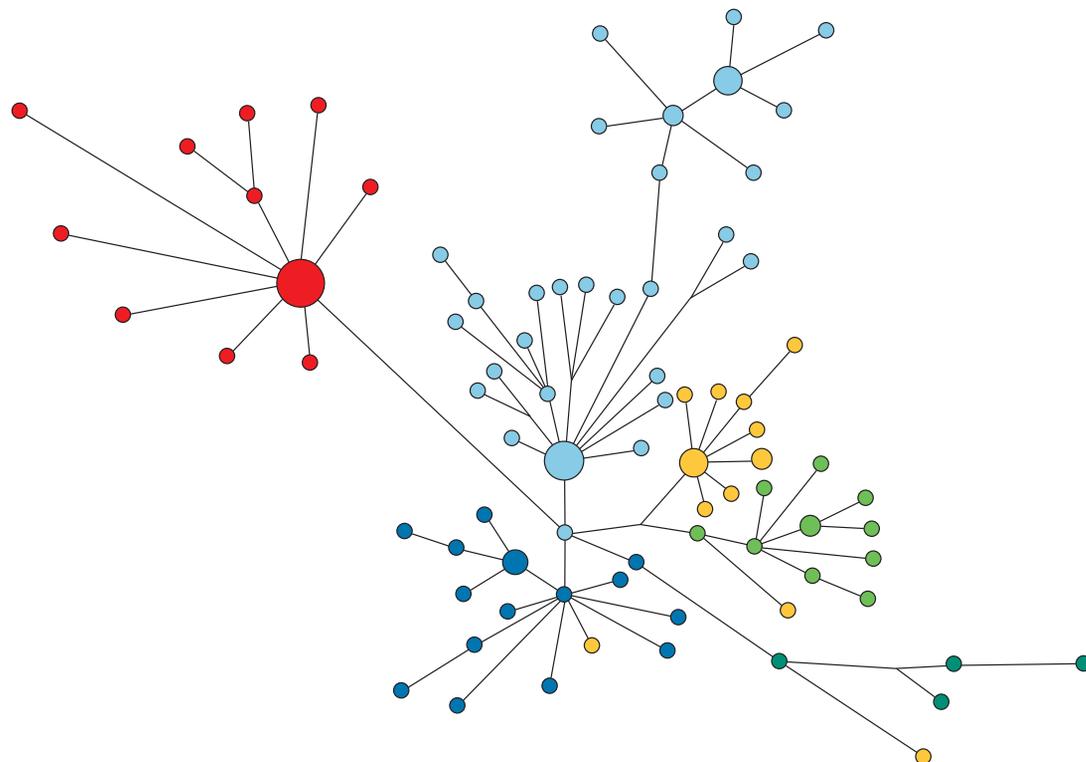


Рис. 2. Медианная сеть YSTR-гаплотипов гаплогруппы N1a2b1b1 у хантов, чулымцев и хакасов-качинцев.

Голубым цветом обозначены ханты, красным – чулымцы, синим – хакасы сеока соххы, зеленым – хакасы сеока ызыр, желтым – хакасы сеока хаса, темно-зеленым – сеока пурут.

Y65017 и Y89655, не обнаруженные у других популяций. Ханты и коми имеют наименее древнего общего предка по этой гаплогруппе по сравнению с остальными европейскими популяциями.

По данным сайта YFull, эта ветвь отделилась от предковой линии примерно 2800 лет назад. Теоретически возможны два варианта появления этой гаплогруппы у современных хантов и коми: 1) наследование от общей древней предковой группы угорских племен; 2) недавнее смешение хантов с этническими коми, переселенцами в Сибирь. Но результаты анализа геномных данных методами NGSadmix, ADMIXTURE, IBD-блоков и различия по терминальным SNP Y-хромосом не подтверждают второй вариант. YSTR-гаплотипы этой линии у хантов и коми также отличаются на несколько мутаций. Ранее V.N. Pimenoff с коллегами (2008) выдвинули предположение, что когда обско-угорские ханты и манси ушли на западные склоны Уральских гор и на северо-запад Сибири, и было сформировано уникальное объединение N1b-A и N1b-E. Это сочетание сублиний N1b у хантов и манси предполагает недавнее слияние западной и восточной линий в Северо-Западной Сибири. Наши новые данные не противоречат этой версии.

Все остальные гаплогруппы представлены только в отдельных выборках хантов. К гаплогруппе N1a2b2b1~ (Z35076) относятся три образца казымских хантов. Наиболее близкая ей линия N1a2b2b1~ (B528, Y24382, Z35076, Z35077) также распространена у коми. У удмуртов, татар, чувашей и башкир присутствует ее более современная линия (B226). YSTR-гаплотипы этой гаплогруппы у коми

и удмуртов ближе друг другу, чем к образцам хантов. Присутствие у хантов и коми двух гаплогрупп, N1a2b2a1 и N1a2b2b1~, с этноспецифичными терминальными SNP и отличающимися гаплотипами говорит об их наследовании от достаточно древних общих предков, скорее всего, входивших в состав раннего угорского населения этих территорий.

Тринадцать образцов казымских хантов относятся к гаплогруппе N1a1a1a1a2a1c1~ (Y13850, Y13852), из них семь имеют фамилию Пяк, ненецкую по происхождению, относящейся к лесным ненцам. Все семь образцов имеют очень близкие гаплотипы и являются потомками относительно недавнего общего предка ненца. В анкетах этих мужчин, считающих себя хантами по отцовской линии, с разной глубиной были указаны ненецкие предки. Остальные шесть мужчин этой гаплогруппы отличаются гаплотипами от рода Пяк.

В нашем исследовании тазовских ненцев (Харьков и др., 2021) обнаружено, что к данной гаплогруппе полностью принадлежат все мужчины, представляющие хантыйские по происхождению рода Салиндер, Лар и Тибичи. Представители этих родов сформировались в XVIII–XIX вв. в низовьях Оби в результате развития ненецкого крупнотабунного оленеводства и вовлечения в него части северных хантов (Квашнин, 2003). Все гаплотипы казымских хантов этой гаплогруппы значительно отличаются от гаплотипов тазовских ненцев.

Другие пять образцов казымских хантов относятся к гаплогруппе N1a2b1b1b1~ (B172, Z35108). К ней принадлежат все ранее обследованные мужчины ненцы из

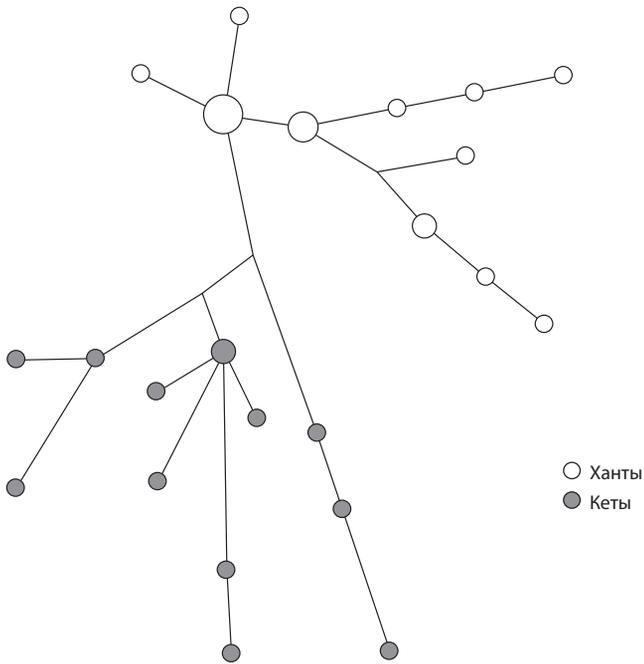


Рис. 3. Медианная сеть YSTR-гаплотипов гаплогруппы Q1b1a3b1a2~ у хантов и кетов.

фратрии Вануйто, относящиеся к родам Вануйто, Пуйко и Яунгат, и хантыйского по происхождению рода Пурунгуи. Четыре образца хантов отличаются по гаплотипам от ненцев, но один практически полностью совпадает с ними. Такое разделение на специфичные для хантов и близкие к ненцам гаплотипы совпадает с данными по гаплогруппе N1a1a1a1a2a1c1~. Очевидно, что в состав генофонда казымских хантов входят именно хантыйские по происхождению варианты этих гаплогрупп, но относительно недавно заключались браки и с лесными ненцами. Отсутствие этих гаплогрупп у русскиных хантов совпадает с данными распределения блоков IBD и коэффициента геномного инбридинга.

Распределение в исследованных популяциях различных гаплогрупп клады N Y-хромосомы согласуется с частотой угорского генетического компонента. Филогенетический анализ Y-хромосомных сублиний и гаплотипов различных гаплогрупп клады N показывает, что центр происхождения и расселения носителей угорского компонента в Южной, Западной Сибири и Восточной Европе – территория современного Алтая и Саян. Полученные результаты сопоставимы с данными этнологии, антропологии и лингвистики о вкладе уральского компонента в формирование различных народов Алтая и Саян и историческими ареалами угорских и других языков уральской языковой семьи.

Почти 40 % мужчин с. Казым принадлежит к гаплогруппе Q1b1a3b1a2~ (Z35974 xB32, B33, Z35993). Ближе всего к ней относится линия Q1b1a3b1a2~ (B33, Z35991), специфичная для популяции кетов. Этот вариант, кроме кетов, преобладает и у селькупов из Томской области и Красноярского края. Более удаленная линия Q1b1a3b1a~ (B30, YP1693 xZ35991) распространена в популяциях тувинцев с максимальной частотой в восточных горных районах Тывы (до 25 %). Образцы хантов демонстрируют

специфичный спектр гаплотипов с недавним эффектом основателя, который не наблюдается у кетов (рис. 3).

Распределение этих сублиний в популяциях хантов, кетов и тувинцев хорошо согласуется с долями совпадений по IBD-блокам между ними, графиком tSNE и по наличию угорского генетического компонента в этих популяциях по аутосомной части их генофонда. Присутствие этой линии у хантов связано не с недавним заимствованием у других аборигенных популяций (кетов и селькупов), а с тем, что она уже входила в состав расселявшихся предковых групп.

У трех мужчин д. Русскинская представлена совершенно другая гаплогруппа клады Q – Q1a2b~ (M25, L716, YP1674, YP1676). Это очень редкая гаплогруппа, не обнаруженная в других сибирских популяциях. С максимальной частотой она имеется у этнических туркмен Каракалпакстана, Ирана и Афганистана (Grugni et al., 2012; Схаляхо и др., 2016). У большинства других этносов ее частота очень низка. Гаплотипы хантов довольно сильно отличаются от других популяций. Скорее всего, присутствие у них этой линии – не следствие недавней метисации, а наследие угорских групп, переселившихся из Южной Сибири и Приуралья на север.

Последняя гаплогруппа, к которой относятся 16 мужчин хантов д. Русскинская, – R1a1a1b2-Y43850. Гаплотипы всех образцов достаточно близки, что говорит о недавнем эффекте основателя (рис. 4).

Специфичными терминальными SNP для хантов являются S7280, FGC687 и FGC38304. Наиболее близкие к этой линии варианты R1a1a1b2-Y43850 с большей частотой представлены у хакасов, шорцев и с меньшей – у тувинцев и северных алтайцев. По данным YFull, возраст этой гаплогруппы составляет примерно 3800 лет. Все эти образцы относятся к четырем различным линиям, которые разделились очень давно. Возраст кластера гаплотипов у хантов составил 933 года (SD = 336 лет), что примерно в полтора раза меньше возраста южносибирских линий. У хакасов сеока пилтир – 1469 лет (SD = 342 года) (Y39884 xY43109). Специфичная для бирюсинских хакасских сеоков туран, хызыл хая и шорских сеоков тарткын, шор-кызай и кара-шор линия этой гаплогруппы (Y62155.2) имеет примерно такой же возраст – 1315 лет (SD = 227 лет). Ветвь с более широким распространением в саяно-алтайских популяциях (Y43109) еще старше – 1566 лет (SD = 350 лет). Разница по SNP и STR между ее вариантами у хакасов, шорцев, тувинцев и северных алтайцев больше, чем с хантами.

Показана значительная неоднородность исследованных выборок хантов по составу и частотам гаплогрупп. Филогения разных линий двух гаплогрупп, N1a2b1b1 и R1a1a1b2-Y43850, указывает на их южносибирское происхождение в генофонде хантов. Территория Саян и Алтай была первичным очагом генерации разнообразия и экспансии численности предковых групп, носителей этих гаплогрупп на территории Сибири. Наиболее вероятно, что распространение большинства Y-хромосомных гаплогрупп у хантов происходило при начальном расселении угорских племен на север и запад.

Нужно учитывать тот факт, что ареал современных хантов находится севернее территории обитания их пред-

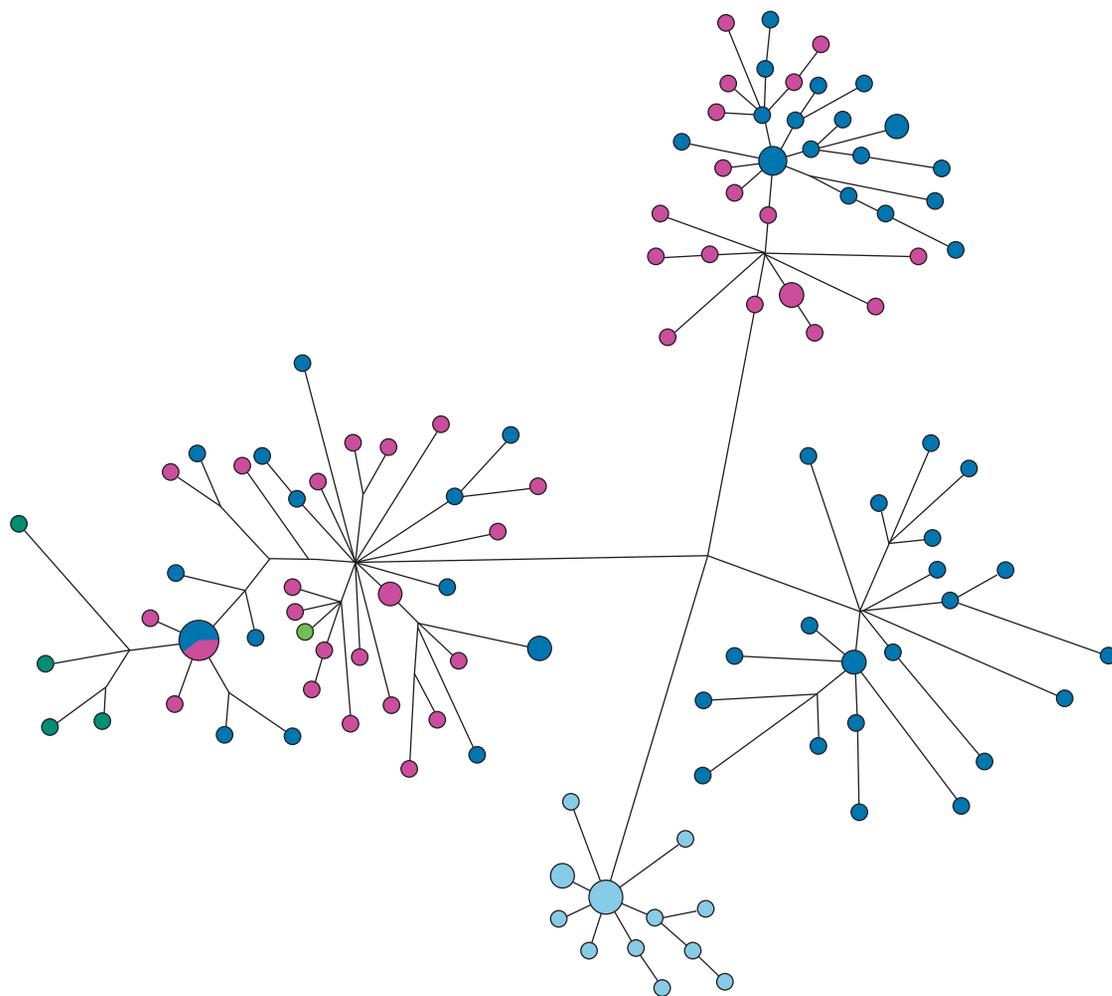


Рис. 4. Медианная сеть YSTR-гаплотипов гаплогруппы R1a1a1b2-Y43850 у хантов, хакасов, шорцев, тувинцев и алтайцев. Голубым цветом обозначены ханты, синим – хакасы, малиновым – шорцы, темно-зеленым – тувинцы, зеленым – алтайцы.

ков. Западно-Сибирский и Волго-Уральский регионы были местом вторичной генерации разнообразия, но не образования самой гаплогруппы N1a2. Окончательного мнения относительно места формирования этносов уральской языковой семьи на данный момент не существует, но многочисленные данные, в том числе и результаты исследований филогении и филогеографии гаплогрупп клады N, указывают на Южную Сибирь. Лингвистическая палеонтология описывает прауральский экологический ареал как территорию, ограниченную на западе Уральским хребтом, на севере – примерно Полярным кругом, на востоке – районом нижнего течения Ангары и Подкаменной Тунгуски и среднего течения Енисея, на юге – примерно современной южной границей западносибирской тайги от северных предгорий Саян и Алтая до нижнего течения Тобола и Среднего Урала включительно (Напольских, 2018).

Заключение

Таким образом, генофонд двух популяций хантов представляет собой разнородный набор по составу гаплогрупп Y-хромосомы, но очень близкий по аутосомным маркерам. Расширенный состав терминальных SNP для выявленных гаплогрупп позволил детально описать и уточнить разли-

чия в филогении и структуре отдельных этноспецифичных сублиний, определить их родство, следы экспансии численности в генофонде хантов. Результаты сравнительного анализа образцов мужчин свидетельствуют о близком генетическом родстве хантов с алтае-саянскими хакасами и тувинцами, а также с ненцами, коми, удмуртами и кетами. Специфичность гаплотипов, обнаружение различных терминальных SNP говорят о том, что ханты достаточно долго не входили в контакты с другими этносами. Исключение составляют лишь ненцы, в составе которых много хантыйских родов. Для северной популяции хантов с. Казым по Y-хромосомным линиям показан небольшой вклад лесных ненцев.

Полученные результаты не противоречат общепринятым версиям этногенеза хантов, но позволяют по-новому взглянуть на этот процесс. Основным фактором формирования генофонда хантов были их территориальная генетическая изоляция и более позднее смешение с пришлым самодийским населением, что при переходе на тундровое оленеводство привело к сильному демографическому росту их родов в составе ненцев. Относительно низкое генетическое разнообразие по аутосомным SNP и достаточно высокий уровень инбридинга у хантов это под-

тверждают. Новая информация о структуре генофонда хантов является важным дополнением к существующим антропологическим, археологическим, этнологическим и лингвистическим данным о формировании и родственных связях хантов с другими народами.

Список литературы / References

Брук С.И. Население мира. Этнодемографический справочник. М.: Наука, 1986.

[Brook S.I. The World Population. Ethnodemographic Reference Book. Moscow: Nauka Publ., 1986. (in Russian)]

Валихова Л.В., Харьков В.Н., Зарубин А.А., Колесников Н.А., Сваровская М.Г., Хитринская И.Ю., Штыгашева О.В., Волков В.Г., Степанов В.А. Генетическая взаимосвязь чулымских тюрков с хакасами и кетами по данным аутомомных SNP и гаплогруппам Y-хромосомы. *Генетика*. 2022;58(10):1177-1184. DOI 10.31857/S0016675822100113.

[Valikhova L.V., Kharkov V.N., Zarubin A.A., Kolesnikov N.A., Svarovskaya M.G., Khitrinskaya I.Yu., Shtygasheva O.V., Volkov V.G., Stepanov V.A. Genetic interrelation of the Chulym Turks with Khakass and Kets according to autosomal SNP data and Y-chromosome haplogroups. *Russ. J. Genet.* 2022;58(10):1228-1234. DOI 10.1134/S1022795422100118.]

Квашин Ю.Н. Гыданские ненцы: история формирования современной родовой структуры (XVIII–XX вв.). М.: ИЭА РАН, 2003.

[Kvashin Yu.N. Gydan Nenets: the History of the Formation of the Modern Genetic Structure (18-20th Centuries). Moscow: IEA RAS Publ., 2003. (in Russian)]

Напольских В.В. Очерки по этнической истории. Казань: Институт археологии им. А.Х. Халикова, 2018.

[Napolskikh V.V. Essays on Ethnic History. Kazan: Khalikov Institute of Archeology, 2018. (in Russian)]

Народы Западной Сибири: Ханты. Манси. Селькупы. Ненцы. Эңцы. Нганасаны. Кеты. М.: Наука, 2005.

[Peoples of West Siberia: Khanty, Mansi, Selkups, Nenets, Enets, Nganasans, Kets. Moscow: Nauka Publ., 2005. (in Russian)]

Народы России: Энциклопедия. М.: Большая Российская энциклопедия, 1994.

[The Peoples of Russia: Encyclopedia. Moscow: Bol'shaya Rossiyskaya Entsiklopediya Publ., 1994. (in Russian)]

Схалихо Р.А., Жабалин М.К., Юсупов Ю.М., Агджоян А.Т., Сабитов Ж.М., Гурьянов В.М., Балаганская О.А., Далимова Д.А., Давлетчурин Д.Х., Турдикулова Ш.У., Чухряева М.И., Асылгужин Р.Р., Акильжанова А.Р., Балановский О.П., Балановская Е.В. Генофонд туркмен Каракалпакстана в контексте популяций Центральной Азии (полиморфизм Y-хромосомы). *Вестн. Моск. ун-та. Серия 23. Антропология*. 2016;3:86-96.

[Skhalyakho R.A., Zhabagin M.K., Yusupov Yu.M., Agdzhojan A.T., Sabitov Zh.M., Gurianov V.M., Balaganskaya O.A., Dalimova D.A., Davletchurin D.Kh., Turdikulova Sh.U., Chukhryaeva M.I., Asilgujin R.R., Akilzhanova A.R., Balanovsky O.P., Balanovska E.V. Gene pool of Turkmens from Karakalpakstan in their Central Asian context (Y-chromosome polymorphism). *Vestnik Moskovskogo Universiteta. Seriya 23. Antropologiya = Moscow University Anthropology Bulletin*. 2016;3:86-96. (in Russian)]

Харьков В.Н., Валихова Л.В., Яковлева Е.Л., Сереброва В.Н., Колесников Н.А., Петелина Т.И., Хитринская И.Ю., Степанов В.А. Реконструкция происхождения гыданских ненцев на основе генетического анализа их родовой структуры с помощью нового

набора YSTR-маркеров. *Генетика*. 2021;57(12):1403-1414. DOI 10.31857/S0016675821120067.

[Kharkov V.N., Valikhova L.V., Yakovleva E.L., Serebrova V.N., Kolesnikov N.A., Petelina T.I., Khitrinskaya I.Yu., Stepanov V.A. Reconstruction of the origin of the Gydan Nenets based on genetic analysis of their tribal structure using a new set of YSTR markers. *Russ. J. Genet.* 2021;57(12):1414-1423. DOI 10.1134/S1022795421120061.]

Харьков В.Н., Новикова Л.М., Штыгашева О.В., Лузина Ф.А., Хитринская И.Ю., Волков В.Г., Степанов В.А. Генофонд хакасов и шорцев по маркерам Y-хромосомы: общие компоненты и генетическая структура родов. *Генетика*. 2020;56(7):826-833. DOI 10.31857/S0016675820070073.

[Kharkov V.N., Novikova L.M., Shtygasheva O.V., Luzina F.A., Khitrinskaya I.Y., Volkov V.G., Stepanov V.A. Gene pool of Khakass and Shors for Y chromosome markers: common components and tribal genetic structure. *Russ. J. Genet.* 2020;56(7):849-855. DOI 10.1134/S1022795420070073.]

Alexander D.H., Lange K. Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics*. 2011; 12:246. DOI 10.1186/1471-2105-12-246.

Alexander D.H., Novembre J., Lange K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res*. 2009;19(9):1655-1664. DOI 10.1101/gr.094052.109.

Bandelt H.J., Forster P., Röhl A. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* 1999;16(1):37-48. DOI 10.1093/oxfordjournals.molbev.a026036.

Browning B.L., Browning S.R. Improving the accuracy and efficiency of identity-by-descent detection in population data. *Genetics*. 2013; 194(2):459-471. DOI 10.1534/genetics.113.150029.

Browning S.R., Browning B.L. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am. J. Hum. Genet.* 2007;81(5):1084-1097. DOI 10.1086/521987.

Grugin V., Battaglia V., Hooshar Kashani B., Parolo S., Al-Zahery N., Achilli A., Olivieri A., Gandini F., Houshmand M., Sanati M.H., Torroni A., Semino O. Ancient migratory events in the Middle East: new clues from the Y-chromosome variation of modern Iranians. *PLoS One*. 2012;7(7):e41252. DOI 10.1371/journal.pone.0041252.

Guo Y., He J., Zhao S., Wu H., Zhong X., Sheng Q., Samuels D.C., Shyr Y., Long J. Illumina human exome genotyping array clustering and quality control. *Nat. Protoc.* 2014;9(11):2643-2662. DOI 10.1038/nprot.2014.174.

Gusev A., Palamara P.F., Aponte G., Zhuang Z., Darvasi A., Gregersen P., Pe'er I. The architecture of long-range haplotypes shared within and across populations. *Mol. Biol. Evol.* 2012;29(2):473-486. DOI 10.1093/molbev/msr133.

Pimenoff V.N., Comas D., Palo J.U., Vershubsky G., Kozlov A., Sajantila A. Northwestern Siberian Khanty and Mansi in the junction of West and East Eurasian gene pools as revealed by uniparental markers. *Eur. J. Hum. Genet.* 2008;16(10):1254-1264. DOI 10.1038/ejhg.2008.101.

Skotte L., Korneliusen T., Albrechtsen A. Estimating individual admixture proportions from next generation sequencing data. *Genetics*. 2013;195(3):693-702. DOI 10.1534/genetics.113.154138.

Zhivotovskiy L.A., Underhill P.A., Cinnioglu C., Kayser M., Morar B., Kivisild T., Scozzari R., Cruciani F., Destro-Bisol G., Spedini G., Chambers G.K., Herrera R.J., Yong K.K., Gresham D., Tournev I., Feldman M.W., Kalaydjieva L. The effective mutation rate at Y-chromosome STRs with application to human population divergence time. *Am. J. Hum. Genet.* 2004;74(1):50-61. DOI 10.1086/380911.

ORCID ID

V.N. Kharkov orcid.org/0000-0002-1679-2212

N.A. Kolesnikov orcid.org/0000-0001-8855-577X

V.A. Stepanov orcid.org/0000-0002-5166-331X

Благодарности. Исследование выполнено за счет гранта Российского научного фонда № 22-64-00060 (<https://rscf.ru/project/22-64-00060/>).

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию 14.10.2022. После доработки 23.11.2022. Принята к публикации 27.11.2022.

Перевод на английский язык <https://vavilov.elpub.ru/jour>

Идентичные по происхождению блоки в геномах коренного населения Сибири демонстрируют генетические связи между популяциями

Н.А. Колесников , В.Н. Харьков, К.В. Вагайцева, А.А. Зарубин, В.А. Степанов

Научно-исследовательский институт медицинской генетики, Томский национальный исследовательский медицинский центр Российской академии наук, Томск, Россия
 nikita.kolesnikov@medgenetics.ru

Аннотация. Генофонд коренного населения Сибири представляет собой уникальную систему с точки зрения исследования популяционно- и эволюционно-генетических процессов, анализа генетического разнообразия и реконструкции генетической истории популяций. Высокое этническое разнообразие является особенностью Сибири как одного из регионов периферийного расселения современного человека. Огромные пространства этого региона и малочисленность аборигенного населения способствовали формированию значительной территориальной и генетической подразделенности. На территории сибирской историко-этнографической провинции расселены около 40 коренных народностей. Проведено масштабное популяционное исследование генофонда коренных народов Сибири на уровне высокоплотного ДНК-микрочипа Infinium Multi-Ethnic Global-8, позволяющее заполнить существенный пробел в геногеографической картине населения Евразии. Для этого были отобраны и проанализированы фрагменты ДНК, унаследованные без рекомбинации каждой парой индивидов от их недавнего общего предка, т.е. сегменты (блоки), идентичные по происхождению (IBD). Распределение блоков IBD в популяциях Сибири хорошо согласуется с географической близостью популяций и их языковой принадлежностью. Чукчи, коряки и нивхи среди сибирских популяций формируют отдельный от основной группы Сибири кластер, причем чукчи и коряки являются более близкородственными. Образуются отдельные субкластеры эвенков и якутов, кетов и чулымцев, тувинцев и алтайцев внутри сибирского кластера. Анализ SNP, которые попадали в большее количество IBD-сегментов анализируемых популяций, позволил составить список из 5358 генов. По результатам расчета, обогащенные этими генами биологические процессы связаны с обнаружением химического раздражителя, участвующего в сенсорном восприятии запаха. Обогащенные найденными генами молекулярные пути связаны с метаболизмом линолевой, арахидоновой, тирозиновой кислот и путем обонятельной трансдукции. При этом анализ литературных данных показал, что некоторые из отобранных генов, которые встречались в большем количестве блоков IBD сразу в нескольких популяциях, могут играть роль в адаптации человека к факторам окружающей среды.
Ключевые слова: IBD; популяции человека; сибирские популяции.

Для цитирования: Колесников Н.А., Харьков В.Н., Вагайцева К.В., Зарубин А.А., Степанов В.А. Идентичные по происхождению блоки в геномах коренного населения Сибири демонстрируют генетические связи между популяциями. *Вавиловский журнал генетики и селекции*. 2023;27(1):55-62. DOI 10.18699/VJGB-23-08

Blocks identical by descent in the genomes of the indigenous population of Siberia demonstrate genetic links between populations

N.A. Kolesnikov , V.N. Kharkov, K.V. Vagaitseva, A.A. Zarubin, V.A. Stepanov

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

Abstract. The gene pool of the indigenous population of Siberia is a unique system for studying population and evolutionary genetic processes, analyzing genetic diversity, and reconstructing the genetic history of populations. High ethnic diversity is a feature of Siberia, as one of the regions of the peripheral settlement of modern human. The vast expanses of this region and the small number of aboriginal populations contributed to the formation of significant territorial and genetic subdivision. About 40 indigenous peoples are settled on the territory of the Siberian historical and ethnographic province. Within the framework of this work, a large-scale population study of the gene pool of the indigenous peoples of Siberia was carried out for the first time at the level of high-density biochips. This makes it possible to fill in a significant gap in the genogeographic picture of the Eurasian population. For this, DNA fragments were analyzed, which had been inherited without recombination by each pair of individuals from their recent common ancestor, that is, segments (blocks) identical by descent (IBD). The distribution of IBD blocks in the populations of Siberia is in good agreement with the geographical proximity of the populations and their linguistic affiliation. Among the Siberian populations, the Chukchi, Koryaks, and

Nivkhs form a separate cluster from the main Siberian group, with the Chukchi and Koryaks being more closely related. Separate subclusters of Evenks and Yakuts, Kets and Chulymy are formed within the Siberian cluster. Analysis of SNPs that fell into more IBD segments of the analyzed populations made it possible to compile a list of 5358 genes. According to the calculation results, biological processes enriched with these genes are associated with the detection of a chemical stimulus involved in the sensory perception of smell. Enriched for the genes found, molecular pathways are associated with the metabolism of linoleic, arachidonic, tyrosic acids and by olfactory transduction. At the same time, an analysis of the literature data showed that some of the selected genes, which were found in a larger number of IBD blocks in several populations at once, can play a role in genetic adaptation to environmental factors.

Key words: IBD; human populations; Siberian populations.

For citation: Kolesnikov N.A., Kharkov V.N., Vagaitseva K.V., Zarubin A.A., Stepanov V.A. Blocks identical by descent in the genomes of the indigenous population of Siberia demonstrate genetic links between populations. *Vavilovskii Zhurnal Genetiki i Seleksii = Vavilov Journal of Genetics and Breeding*. 2023;27(1):55-62. DOI 10.18699/VJGB-23-08

Введение

Генетико-демографические процессы в популяциях, колебания численности, метисационные события, миграции и естественный отбор оказывают влияние на структурированность генетического разнообразия в геномах отдельных индивидов и популяциях в целом. В частности, генетико-демографические процессы приводят к тому, что формируются общие по происхождению блоки сцепления (identity by descent, IBD). Сегментом, имеющим идентичные нуклеотидные последовательности, является IBD у двух или более лиц, если они унаследовали его от общего предка без рекомбинации, т. е. у этих людей сегмент имеет общее происхождение. Ожидаемая длина блока IBD зависит от количества поколений с момента появления последнего общего предка в локусе сегмента (Browning S.R., Browning B.L., 2010; Palamara et al., 2012).

Сегменты IBD можно использовать для оценки демографической истории популяций, включая эффект бутылочного горлышка и потоки генов в популяциях (Gusev et al., 2012). В недавних исследованиях были показаны различия в распределении IBD между африканскими, азиатскими и европейскими популяциями, а также сегменты IBD, общие с древними геномами, такими как геномы неандертальцев или денисовцев (Hochreiter, 2013).

Близкие биологические родственники имеют идентичные друг другу достаточно длинные фрагменты ДНК и, соответственно, в большинстве хромосом находятся идентичные по происхождению блоки значительной длины (> 66.7 сМ), в результате чего ожидаемая длина общих IBD составляет около 1700 сМ. Согласно расчетам, двоюродные и троюродные родственники, вероятно, будут иметь несколько регионов (более 2.5 сегментов) из-за наличия недавних предков, определяющих их родство. Предполагается, что для двоюродных родственников каждый IBD будет иметь общую длину больше 62 сМ и для троюродных – 25 сМ. Дальние родственники, относящиеся к четвероюродным или более дальним кузенам, с большой вероятностью несут один или несколько регионов от своего ближайшего общего предка. Такие пары включают в себя подавляющее большинство людей в конкретной популяции и обычно называются «неродственными», поскольку ожидается, что доля и число IBD по всему геному будут между ними относительно небольшими (Gusev et al., 2012).

Сегменты IBD могут помочь и в обнаружении сигналов естественного отбора в геноме человека. Поиск регионов с избытком сегментов IBD позволяет идентифицировать

регионы в геноме человека, которые находятся под очень недавним и сильным отбором, поскольку отбор обычно увеличивает число сегментов IBD среди людей в популяции (Albrechtsen et al., 2010; Han, Abney, 2011).

Относительно популяций коренных этносов Сибири было высказано предположение, что крупномасштабное рассредоточение и смешение популяций, вероятно, сыграли существенную роль в их предыстории и структуре, что может объяснить необычно высокую долю IBD между популяциями (Pugach et al., 2016).

Цель настоящего исследования – анализ структуры генофонда популяций коренного населения Сибири на основании выявления идентичных по происхождению блоков сцепления, их внутри- и межпопуляционного распределения.

Материалы и методы

Данные широкогеномного генотипирования получены с использованием микрочипов Infinium Multi-Ethnic Global-8 (Illumina), включающих свыше 1.7 млн маркеров. В анализе не было образцов, имеющих более 5 % пропущенных позиций, а также SNP с более чем 10 % отсутствующими генотипами. Данные были предварительно отфильтрованы по минимальной частоте редкого аллеля (MAF, minor allele frequency > 0.01). В итоге после фильтрации исходных данных в финальный массив данных вошло 886 889 аутомомных SNP.

Популяции коренного населения Сибири ($N = 477$) представлены: алтайцами (Б – с. Бешпельтир Чемальского муниципального района, $N = 24$ и К – с. Кулада Онгудайского района, $N = 25$), бурятами (А – пос. Агинское Агинского района, $N = 23$ и К – с. Курумкан Курумканского района, $N = 28$), калмыками ($N = 29$), кетами ($N = 15$), коряками ($N = 20$; материал коряков собран в Корякском автономном округе Камчатской области), чукчами ($N = 25$; популяционная выборка чукчей (образцы крови собраны в пос. Лорино, Сиреники, Янарыкот и Новое Чаплино Чукотского автономного округа) относится к береговой группе), нивхами ($N = 13$), татарами (Т – г. Томск, $N = 20$), тувинцами ($N = 28$), удэгейцами ($N = 15$), хантами (К – с. Казым Белоярского района, $N = 30$ и Р – д. Русскинская Сургутского района, $N = 26$), хакасами (Т – сагайцами Таштыпского района, $N = 29$ и Ш – качинцами Широинского района, $N = 26$), чулымцами ($N = 22$), эвенками (З – забайкальские (пос. Чара Каларского района, с. Моклакан и с. Тупик Тунгиро-Олёкминского района), $N = 25$ и Я – якутские, $N = 28$) и якутами ($N = 26$).

Таблица 1. Характеристики изученных популяционных выборок коренного населения Сибири

Популяция	Локализация	Объем выборки	Языковая принадлежность (семья/группа)	Антропологический тип
Алтайцы (Б)	Республика Алтай	24	Алтайская/тюркская	Монголоидный (центральноазиатский)
Алтайцы (К)		25		
Буряты (А)	Забайкальский край	23	Алтайская/монгольская	
Буряты (К)	Республика Бурятия	28		
Чукчи	Чукотский автономный округ	25	Чукотско-камчатские языки	Монголоидный (арктический)
Чулымцы	Томская область	22	Алтайская/тюркская	Уральский, монголоидный (южносибирский)
Эвенки (Я)	Якутия	28	Алтайская/тунгусо-маньчжурские языки	Монголоидный (байкальский)
Эвенки (З)	Забайкальский край	25		
Калмыки	Республика Калмыкия	29	Алтайская/монгольская	Монголоидный (центральноазиатский)
Кеты	Красноярский край	15	Енисейская	Енисейский
Хакасы (Ш)	Республика Хакасия	26	Алтайская/тюркская	Уральский, монголоидный (южносибирский)
Хакасы (Т)		29		
Ханты (К)	ХМАО – Югра	30	Уральская/обско-угорская	Уральский
Ханты (Р)		26		
Коряки	Камчатский край	20	Чукотско-камчатские языки	Монголоидный (арктический)
Нивхи	Хабаровский край	13	Палеоазиатский язык	Монголоидный (сахалиноамурский)
Татары (Т)	Томская область	20	Алтайская/тюркская	Уральский, монголоидный (южносибирский)
Тувинцы	Республика Тыва	28		Монголоидный (центральноазиатский)
Удэгейцы	Приморский край	15	Алтайская/тунгусо-маньчжурские языки	Монголоидный (байкальский)
Якуты	Якутия	26	Алтайская/тюркская	Монголоидный (центральноазиатский)

Материал депонирован в биоресурсной коллекции «Био-банк населения Северной Евразии». Характеристика изученных популяций представлена в табл. 1.

Фазирование генотипов проведено с помощью программного обеспечения Beagle 4.1 (Browning S.R., Browning B.L., 2007). Для анализа идентичных по происхождению блоков генома использовали алгоритм Refined IBD, версии refined-ibd.16May19.ad5.jar (Browning B.L., Browning S.R., 2013). Для сравнения популяций были получены суммы средних длин сегментов IBD между парами индивидов 1.5–1.999 cM, 2–3.999, 4–7.999, 8–15.999 и > 16 cM (далее в статье данные диапазоны будут записаны как 1.5–2 cM, 2–4, 4–8 и 8–16 cM). Тепловая карта с дендрограммой на основании логарифма от суммы средних длин сегментов IBD между парами индивидов для IBD построена с использованием пакета heatmap.2 в программной среде R.

Нами также были идентифицированы SNP, которые попадали в большее количество IBD сегментов анализируемых популяций (частота попадания SNP в IBD выше 99-го квантиля распределения частот), определена принадлежность этих SNP к генам и проведена оценка биологической значимости полученного списка генов. Для анализа использовали веб-ресурс WebGestalt (WEB-based Gene Set Analysis Toolkit), в частности анализ KEGG-путей и генных онтологий (Gene Ontology) методом ORA (over-representation analysis).

Результаты и обсуждение

Для более детального анализа генетического родства сибирских популяций и выяснения, в какой степени их генетическая структура может быть объяснена недавними локальными миграциями, мы выделили и проанализировали фрагменты ДНК, которые были унаследованы без рекомбинации каждой парой индивидов от их недавнего общего предка, т. е. сегменты (блоки), идентичные по происхождению (IBD).

Для популяций, населяющих территорию Сибири, характерна уникальная генетико-демографическая история, которая находит отражение в том числе и в распределении блоков IBD как внутри популяции, так и между ними. Нами рассчитаны суммы средних длин сегментов IBD между парами индивидов и на их результатах построена тепловая карта с дендрограммой на основании логарифма от суммы средних длин сегментов IBD (рис. 1).

Количество общих сегментов у представителей разных популяций согласуется с географией их проживания, так как на проживающие рядом народы могут влиять общие генетико-демографические процессы. Анализ тепловой карты демонстрирует кластеризацию популяций сибирской группы, связывающую народности по месту происхождения. Чукчи, коряки и нивхи формируют отдельный от основной группы сибирских популяций кластер, причём чукчи и коряки являются более близкородственными. Формируются отдельные субкластеры эвенков и якутов,

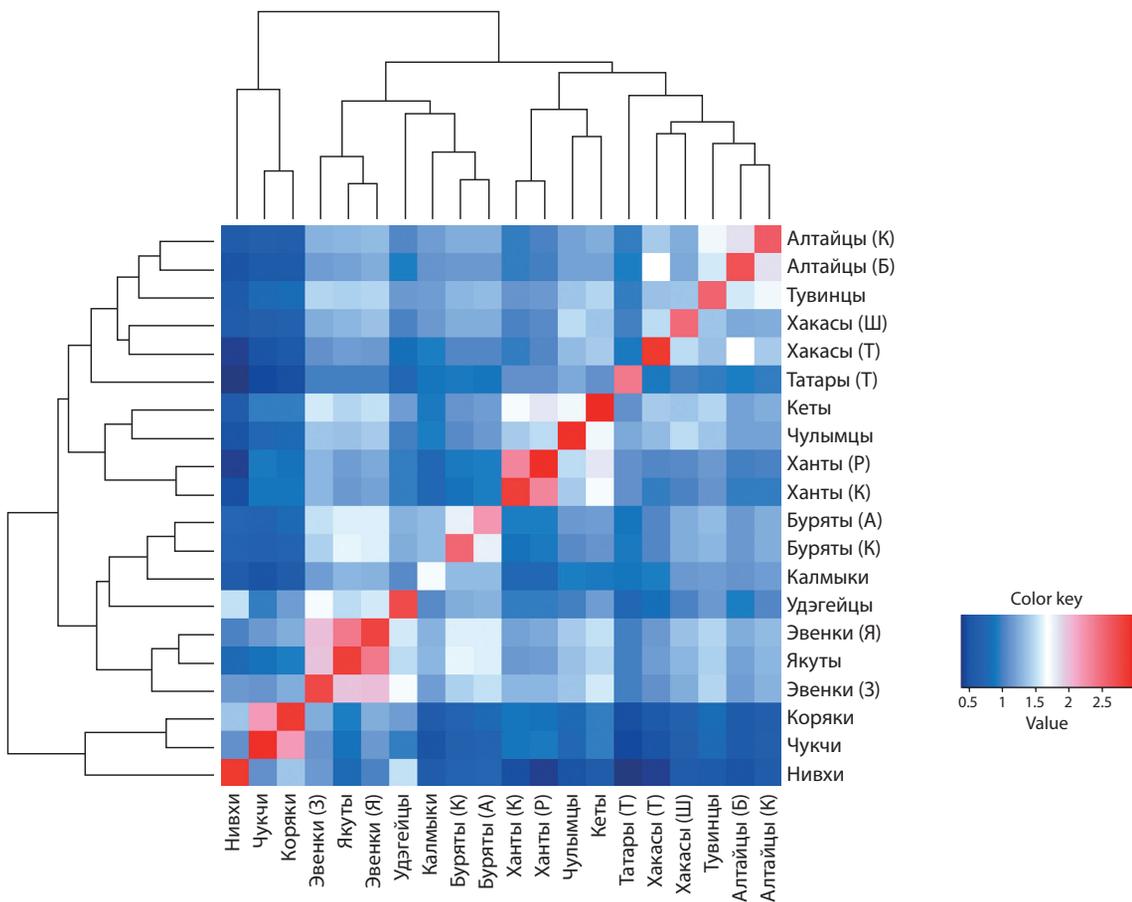


Рис. 1. Тепловая карта с дендрограммой на основании логорифма от суммы средних длин сегментов IBD (> 1.5 cM) между парами индивидов.

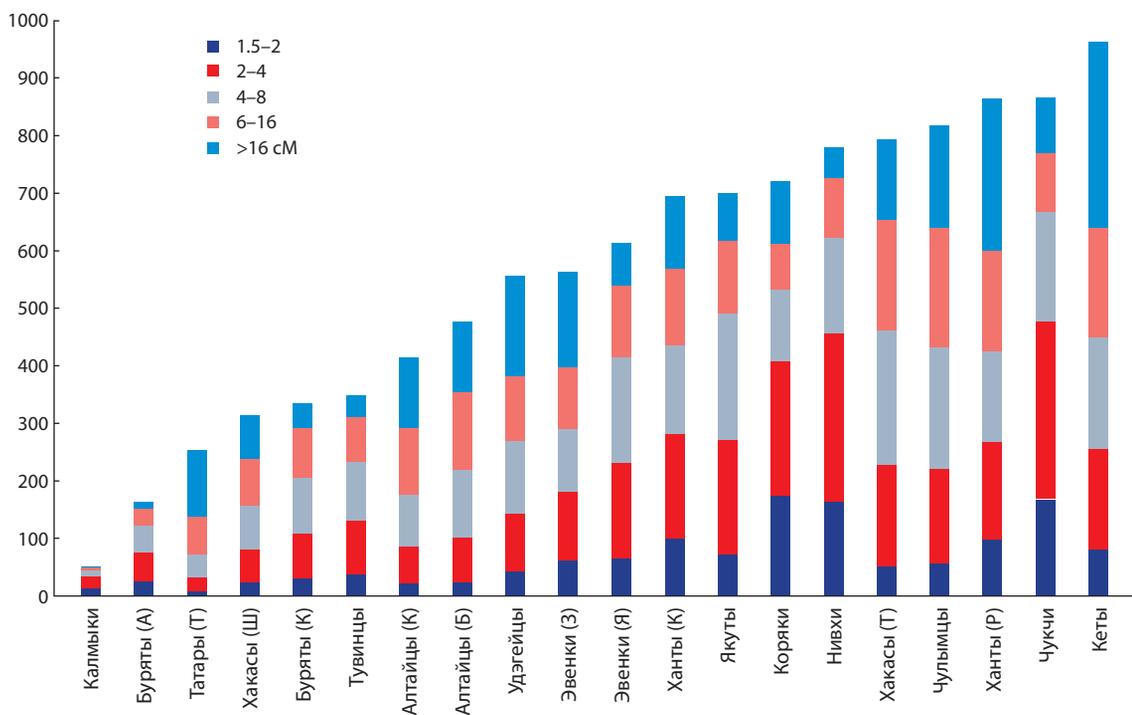


Рис. 2. Диаграмма суммы средних длин сегментов IBD между парами индивидов в исследуемых популяциях для IBD разных размеров (1.5–2, 2–4, 4–8, 8–16, >16 cM).

кетов и чулымцев, тувинцев и алтайцев внутри сибирского кластера.

При градации сегментов IBD по длине в целом тенденция сохраняется, но появляются некоторые отличия, точнее характеризующие недавние взаимоотношения народов. Для более длинных IBD кластеры больше соответствуют нынешнему географическому расположению популяций, отражая недавний обмен общими регионами. По трем диапазонам длин с размерами IBD (1.52–2, 2–4 и 4–8 сМ) популяции лучше разбиваются на близко географически расположенные пары: коряки-чукчи, якуты-эвенки. Кеты практически в равной степени имеют общие блоки IBD с чулымцами (18.7–27.2–7.7 сМ) и хантами (23.4–24.4–4.8 сМ для хантов (К) и 25.9–30.1–7.9 сМ для хантов (Р)). При этом среди хантов большее значение общих блоков IBD наблюдается у русскиных хантов, что соответствует их более близкому географическому расположению по сравнению с хантами Белоярского района.

Для более удаленных от кетов хакасов наблюдается меньшее значение общих блоков IBD практически в два-три раза по сравнению с хантами (9.8–10.4–2.5 сМ для хакасов (Ш) и 9.1–10.6–3.8 сМ для хакасов (Т)). Несмотря на то что эвенки, тувинцы и якуты еще более удалены, для тувинцев (12.3–12.8–2.1 сМ) и якутов (12.8–12.5–2.1 сМ) наблюдаются схожие значения общих блоков IBD. Для забайкальских эвенков (16.4–16.9–3.0 сМ) нужно отметить большие значения блоков IBD, чем с якутскими эвенками (14.7–14.0–2.6 сМ).

За некоторыми исключениями, общие сегменты IBD между сибирскими популяциями лучше объясняются географической близостью популяций, а не их языковой принадлежностью. Так, например, якутские эвенки, проживающие на территории Якутии, имеют больше общих IBD с якутами (252.7 сМ), чем с забайкальскими эвенками (102.5 сМ). При этом сумма средних длин сегментов IBD между парами индивидов между двумя популяциями эвенков сопоставима с суммой между забайкальскими эвенками и якутами (95.2 сМ).

С точки зрения анализа IBD внутри популяций в целом индивиды из популяций Крайнего Севера и Дальнего Востока (коряки, чукчи, нивхи) имеют больше общих IBD с образцами из той же группы, чем индивиды из южносибирских популяций, таких как алтайцы и тувинцы. При этом у чукчей, коряков и нивхов наибольший вклад вносят короткие IBD-фрагменты 1.5–4 сМ (55–57–59 %), что может говорить о бутылочном горлышке в прошлом во время миграций на север и северо-восток и/или сильной изоляции от остальных популяций, населяющих территорию Сибири. При этом у чулымцев, хантов (Р), населяющих центральную часть Сибири и имеющих наибольшую сумму средних длин сегментов IBD между парами индивидов, самый большой вклад вносят IBD длиной более 8 сМ (47–51 %), что говорит о недавнем значительном инбридинге внутри популяции.

Наиболее генетически гетерогенные сибирские популяции, имеющие минимальные значения по коэффициенту геномного инбридинга (F_{ROH}) (Колесников и др., 2021), также имеют минимальные значения по сумме средних

длин сегментов IBD. Особенно выражено это у популяций калмыков, агинских бурят и томских татар (рис. 2).

В популяциях бурятов нет значимых различий по средней суммарной длине между парами индивидов с другими популяциями, но есть существенная разница в распределении IBD внутри популяций. Так, у бурятов (К) значительно больше средняя суммарная длина между парами индивидов внутри популяции (335.4 сМ), по сравнению с бурятами (А) (1635 сМ), во многом за счет средних и длинных IBD. У агинских бурят гораздо больше доля коротких IBD (16–30–28–19–7 %), чем у курумканских (9–24–29–25–13 %). Несмотря на то что для бурятов, имеющих большую общую численность, отмечено практически удвоение численности с 237 тыс. человек в 1926 г. до 461 тыс. в 2010 г., данное различие между популяциями можно объяснить резким ростом численности населения пос. Агинское с 451 человека в 1908 г. до 4556 человек в 1939 г. и 15 тыс. – в 2010 г., и стабильностью населения пос. Курумкан с численностью 5617 человек в 1979 г. и 5465 – в 2010 г., находящегося в одном из отдаленных районов Бурятии.

Похожая динамика наблюдается у популяций чукчей (20–36–22–12–11 %), коряков (24–33–17–11–15 %), нивхов (21–38–21–13–7 %) и калмыков (29–39–22–7–3 %), за более резким уменьшением доли длинных IBD. Для чукчей, нивхов и коряков, имеющих небольшую численность (до 13 тыс.), характерно отсутствие резких колебаний численности популяций за последние 100 лет с ростом на 29–9–6 % соответственно, что также могло способствовать сокращению общих длинных IBD-сегментов при условии незначительного количества близкородственных браков внутри популяции.

Всего для 20 популяций получено 189314 SNP, попадающих в наибольшее количество IBD-сегментов анализируемых популяций (частота попадания SNP в IBD выше 99-го квантиля распределения частот). Из них 88530 SNP находятся в межгенных участках, остальные локализованы в области 5358 генов. В табл. 2 представлены гены, которые были показаны в четырех и более популяциях.

Из списка 5358 генов, наиболее часто попадающих в блоках IBD, 1694, по данным WebGestalt, оказались аннотированными по базе KEGG. В результате анализа с учетом поправки Бенджамини–Хохберга ($FDR = 0.05$) были выявлены обогащенные этими генами молекулярные KEGG-пути: путь метаболизма линолевой кислоты hsa00591 ($FDR = 0.0051$), включающий 17 генов (*CYP2C8*, *CYP2C9*, *PLA2G1B*, *PLB1*, *CYP1A2*, *CYP2C19*, *CYP3A4*, *PLA2G10*, *PLA2G2A*, *PLA2G2C*, *PLA2G2D*, *PLA2G2E*, *PLA2G2F*, *PLA2G4A*, *PLA2G4C*, *PLA2G5*, *PLA2G6*); путь метаболизма арахидоновой кислоты hsa00590 ($FDR = 0.0240$), включающий 27 генов (*CYP2C8*, *CYP2C9*, *PLA2G1B*, *PLB1*, *ALOX12*, *ALOX12B*, *ALOX15B*, *ALOX5*, *CYP2B6*, *CYP2C19*, *GPX1*, *GPX3*, *GPX7*, *PLA2G10*, *PLA2G2A*, *PLA2G2C*, *PLA2G2D*, *PLA2G2E*, *PLA2G2F*, *PLA2G4A*, *PLA2G4C*, *PLA2G5*, *PLA2G6*, *PTGIS*, *PTGS1*, *PTGS2*, *TBXAS1*); путь метаболизма тирозина hsa00350 ($FDR = 0.0240$), включающий 18 генов (*ADH1A*, *ADH1B*, *ADH1C*, *ADH4*, *ADH5*, *ADH6*, *ADH7*, *ALDH3B1*, *ALDH3B2*, *AOC2*, *AOC3*, *DDC*, *GOT1*, *HPD*, *IL4I1*, *PNMT*, *TYR*,

Таблица 2. Гены, попадающие в регионы с максимальным количеством блоков IBD более чем в четырех популяциях

Ген	Алтайцы (Б)	Алтайцы (К)	Буряты (А)	Буряты (К)	Чулымцы	Эвенки (Я)	Эвенки (З)	Калмыки	Кеты	Хакасы (Ш)	Хакасы (Т)	Ханты (К)	Ханты (Р)	Коряки	Нивхи	Татары (Т)	Тувинцы	Удгейцы	Якуты
AAGAB	+	+	+											+				+	
GSE1						+							+	+	+				+
IQCH	+	+	+											+				+	
IQCH-AS1	+	+	+											+				+	
SMAD3	+	+	+											+				+	
ANGPTL1	+						+					+	+						
RALGPS2	+						+					+	+						
ARFGAP1							+	+							+				+
C15orf61		+	+											+				+	
C16orf74						+			+				+		+				
CALML4		+	+											+				+	
CHRNA4							+	+							+				+
CLN6		+	+											+				+	
COL20A1							+	+							+				+
FEM1B		+	+											+				+	
FLJ16779							+	+							+				+
ITGA11		+	+											+				+	
LINC02206		+	+											+				+	
LOC100130587							+	+							+				+
LOC101929076		+	+											+				+	
LOC102723493		+	+											+				+	
MAP2K5		+	+											+				+	
MIR99AHG							+	+						+	+				
NKAIN4							+	+							+				+
PIAS1		+	+											+				+	
RALGPS2	+						+					+	+						
SKOR1		+	+											+				+	

TYRP1); путь обонятельной трансдукции hsa04740 (FDR = 4.55E-08), включающий 159 генов.

Метаболическое превращение полиненасыщенных жирных кислот (ПНЖК), таких как линолевая кислота, в биологически активные ПНЖК с длинной цепью (> 20 атомов углерода, ДЦ-ПНЖК), таких как арахидоновая кислота, необходимо для правильного обмена веществ. ДЦ-ПНЖК и их метаболиты являются важными структурными и сигнальными компонентами для многочисленных биологических систем, включая развитие и функционирование мозга, врожденный иммунитет и энергетический гомеостаз (Marszalek, Lodish, 2005; Calder, 2013). Существуют также пищевые источники предварительно сформированных ДЦ-ПНЖК в яйцах и некоторых видах мяса, содержащих арахидоновую кислоту (Horrocks, Yeo, 1999; Howe et al., 2006; Chilton et al., 2014). Найденные закономерности

в распределении IBD, содержащих гены, участвующие в метаболизме жирных кислот, могут указывать на недавний направленный отбор, связанный с адаптацией к особенностям питания в холодном климате, или отражать влияние западной диеты (Chilton et al., 2014).

Положительный отбор в генах, влияющих как на уровень ДЦ-ПНЖК, так и на метаболическую эффективность, посредством которой образуются ДЦ-ПНЖК в популяциях пигмеев на острове Флорес (Tucci et al., 2018), гренландских инуитов (Fumagalli et al., 2015) и коренных американцев (Amorim et al., 2017; Harris et al., 2019), предположительно также связан с особенностями питания в условиях холодного климата, хотя точное давление отбора неизвестно (Fumagalli et al., 2015). В качестве примера можно привести схожесть генотипов в генах, входящих в систему обоняния, которая может играть роль в фор-

мировании или поддержании социальных связей между индивидами внутри популяции (Christakis, Fowler, 2014). Для таких генотипов выявлены более высокие показатели положительного отбора (Fu et al., 2012).

Анализ геномных онтологий (3511 генов, по данным WebGestalt, оказались аннотированными по базе геномной онтологии) показал девять статистически значимых биологических процессов (с учетом поправки Бенджамини–Хохберга (FDR = 0.05)), связанных с обнаружением химического раздражителя, участвующего в сенсорном восприятии запаха (GO:0050911, GO:0007608, GO:0050907, GO:0050906, GO:0051606, GO:0009593, GO:0007600, GO:0007606, GO:0050877).

Анализ литературы также выявил, что ряд генов, попадающих в блоки IBD, может играть существенную роль в образовании онкологии и влиять на лечение, например ген *AAGAB*, попадающий в блоки IBD у пяти популяций. Ген *AAGAB* состоит из 10 экзонов, кодирующих белок из 315 аминокислот (aa), белок AAGAB (связывающий альфа- и гамма-адаптин). Он широко экспрессируется и взаимодействует с гамма-адаптинным и альфа-адаптинным адаптерными белковыми комплексами, AP1 и AP2, а также участвует в переносе мембран и играет роль в эндоцитозе и сортировке белков. Гетерозиготные мутации *AAGAB* вызывают точечную ладонно-подошвенную кератодермию типа 1 (PPKP1) – кожное заболевание, характеризующееся точечным гиперкератозом ладоней и подошв (Kiritsi et al., 2013). Ген *AAGAB* является многообещающим биомаркером ответа и исхода химиотерапии во время лечения рака молочной железы. Однако точная роль *AAGAB* в развитии рака молочной железы в настоящее время неясна и потенциально требует дальнейшего изучения (Bownes et al., 2019).

Другой ген, *GSE1*, который попадает у пяти популяций в блоки IBD, может функционировать как онкоген при раке молочной железы, желудка и предстательной железы, и также может быть важен при лечении пациентов с раком предстательной железы (Vamodu et al., 2021).

Ген *IQCH-AS1*, кодирующий антисмысловую РНК IQCH 1 (IQCH-AS1), коррелирует с выживаемостью и диагностикой онкологических больных, однако его роль в развитии рака щитовидной железы и химиочувствительности доксорубина остается неясной (Fei et al., 2022).

Участие гена *SMAD3* в регуляции генов, важных для развития клеток, таких как дифференцировка, рост и гибель, подразумевает, что изменение или подавление ее активности может привести к образованию или развитию рака.

Некоторые гены, представленные в табл. 1, могут играть роль в адаптации человека к факторам окружающей среды. Например, ген *IQCH* может быть регулирующим в сперматогенезе (Yin et al., 2005), а также ассоциирован у монголов с ростом взрослого человека (Kimura et al., 2008).

Ген *CHRNA4* кодирует субкомпонент никотиновых рецепторов $\alpha 4\beta 2$ в человеческом мозге. Показано, что люди, обладающие определенными генотипами гена *CHRNA4*, в задачах визуального поиска лучше справлялись с отслеживанием и идентификацией нескольких объектов (Espeseth et al., 2010). Полиморфизмы в гене *CHRNA4*, по-видимому, также вносят свой вклад в формирование

личности, влияя на степень чувствительности развития как к нормальным, так и неблагоприятным условиям окружающей среды (Grazioplene et al., 2013).

Ген *COL20A1*, кодирующий коллаген типа XX альфа 1, отмечен в ряде генов, имеющих несинонимичные изменения с высокой частотой у современного человека в сравнении с архаичными гоминидами, что, вероятно, способствовало развитию уникальных человеческих черт, и является интересным объектом для изучения (Kuhlwilm, Voesckx, 2019).

Закключение

Нами получена новая информация о структуре и составе генофондов коренных народов Сибири, их генетических взаимоотношений и генетико-демографических процессах на основании анализа распределения идентичных по происхождению блоков сцепления. Результаты исследования показывают кластеризацию сибирских популяций, связывающую народы по месту происхождения и отражающую общее происхождение и большую степень родства. Для популяций, населяющих территорию Сибири, характерна уникальная генетико-демографическая история, которая отражается в распределении блоков IBD как внутри популяции, так и между ними. Анализ блоков IBD значительно дополняет картину в изучении образования и взаимодействия этносов, но не дает однозначных ответов для популяций, развивающихся в условиях сложного этногенеза. За некоторыми исключениями, общие IBD в Сибири лучше объясняются географической близостью популяций, а не их языковой принадлежностью.

Анализ SNP, которые попадали в большее количество IBD-сегментов анализируемых популяций, позволил составить список из 5358 генов. По результатам расчетов, обогащенные этими генами биологические процессы связаны с обнаружением химического раздражителя, участвующего в сенсорном восприятии запаха. Обогащенные найденными генами молекулярные пути связаны с метаболизмом жирных кислот и путем обонятельной трансдукции. При этом анализ литературных данных показал, что некоторые из отобранных генов, встречающиеся в большем количестве блоков IBD сразу в нескольких популяциях, могут играть роль в адаптации человека к факторам окружающей среды и являются перспективными мишенями для дальнейшего изучения.

Список литературы / References

- Колесников Н.А., Харьков В.Н., Зарубин А.А., Раджабов М.О., Воевода М.И., Губина М.А., Хуснутдинова Э.К., Литвинов С.С., Екомасова Н.В., Штыгашева О.В., Максимова Н.Р., Сухомьясова А.Л., Степанов В.А. Особенности геномного распределения регионов высокой гомозиготности у коренного населения Северной Евразии на индивидуальном и популяционном уровнях на основе анализа SNP высокой плотности. *Генетика*. 2021; 57(11):1261-1275. DOI 10.31857/S0016675821110059.
- [Kolesnikov N.A., Kharkov V.N., Zarubin A.A., Radzhabov M.O., Voevoda M.I., Gubina M.A., Khusnutdinova E.K., Litvinov S.S., Ekomasova N.V., Shtygashева O.V., Maksimova N.R., Sukhomyasova A.L., Stepanov V.A. Features of the genomic distribution of runs of homozygosity in the indigenous population of Northern Eurasia at the individual and population levels based on high density SNP analysis. *Russ. J. Genet.* 2021;57(11):1271-1284. DOI 10.1134/S1022795421110053.]

- Albrechtsen A., Moltke I., Nielsen R. Natural selection and the distribution of identity-by-descent in the human genome. *Genetics*. 2010; 186(1):295-308. DOI 10.1534/genetics.110.113977.
- Amorim C.E.G., Nunes K., Meyer D., Comas D., Bortolini M.C., Salzano F.M., Hünemeier T. Genetic signature of natural selection in first Americans. *Proc. Natl. Acad. Sci. USA*. 2017;114(9):2195-2199. DOI 10.1073/pnas.1620541114.
- Bamodu O.A., Wang Y.H., Ho C.H., Hu S.W., Lin C.D., Tzou K.Y., Wu W.L., Chen K.C., Wu C.C. Genetic suppressor element 1 (GSE1) promotes the oncogenic and recurrent phenotypes of castration-resistant prostate cancer by targeting tumor-associated calcium signal transducer 2 (TACSTD2). *Cancers (Basel)*. 2021;13(16):3959. DOI 10.3390/cancers13163959.
- Bownes R.J., Turnbull A.K., Martinez-Perez C., Cameron D.A., Sims A.H., Oikonomidou O. On-treatment biomarkers can improve prediction of response to neoadjuvant chemotherapy in breast cancer. *Breast Cancer Res.* 2019;21(1):73. DOI 10.1186/s13058-019-1159-3.
- Browning B.L., Browning S.R. Improving the accuracy and efficiency of identity-by-descent detection in population data. *Genetics*. 2013; 194(2):459-471. DOI 10.1534/genetics.113.150029.
- Browning S.R., Browning B.L. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am. J. Hum. Genet.* 2007; 81(5):1084-1097. DOI 10.1086/521987.
- Browning S.R., Browning B.L. High-resolution detection of identity by descent in unrelated individuals. *Am. J. Hum. Genet.* 2010;86(4): 526-539. DOI 10.1016/j.ajhg.2010.02.021.
- Calder P.C. Long chain fatty acids and gene expression in inflammation and immunity. *Curr. Opin. Clin. Nutr. Metab. Care*. 2013;16(4): 425-433. DOI 10.1097/MCO.0b013e3283620616.
- Chilton F.H., Murphy R.C., Wilson B.A., Sergeant S., Ainsworth H., Seeds M.C., Mathias R.A. Diet-gene interactions and PUFA metabolism: A potential contributor to health disparities and human diseases. *Nutrients*. 2014;6(5):1993-2022. DOI 10.3390/nu6051993.
- Christakis N.A., Fowler J.H. Friendship and natural selection. *Proc. Natl. Acad. Sci. USA*. 2014;111(Suppl. 3):10796-10801. DOI 10.1073/pnas.1400825111.
- Espeseth T., Sneve M.H., Rootwelt H., Laeng B. Nicotinic receptor gene *CHRNA4* interacts with processing load in attention. *PLoS One*. 2010;5(12):e14407. DOI 10.1371/journal.pone.0014407.
- Fei Y., Li Y., Chen F. LncRNA-IQCH-AS1 sensitizes thyroid cancer cells to doxorubicin via modulating the miR-196a-5p/PPP2R1B signalling pathway. *J. Chemother.* 2022;1-9. DOI 10.1080/1120009X.2022.2082348.
- Fu F., Nowak M.A., Christakis N.A., Fowler J.H. The evolution of homophily. *Sci. Rep.* 2012;2:845. DOI 10.1038/srep00845.
- Fumagalli M., Moltke I., Grarup N., Racimo F., Bjerregaard P., Jørgensen M.E., Korneliusen T.S., Gerbault P., Skotte L., Linneberg A., Christensen C., Brandslund I., Jørgensen T., Huerta-Sánchez E., Schmidt E.B., Pedersen O., Hansen T., Albrechtsen A., Nielsen R. Greenlandic Inuit show genetic signatures of diet and climate adaptation. *Science*. 2015;349(6254):1343-1347. DOI 10.1126/science.aab2319.
- Grazioplene R.G., Deyoung C.G., Rogosch F.A., Cicchetti D. A novel differential susceptibility gene: *CHRNA4* and moderation of the effect of maltreatment on child personality. *J. Child Psychol. Psychiatry*. 2013;54(8):872-880. DOI 10.1111/jcpp.12031.
- Gusev A., Palamara P.F., Aponte G., Zhuang Z., Darvasi A., Gregersen P., Pe'er I. The architecture of long-range haplotypes shared within and across populations. *Mol. Biol. Evol.* 2012;29(2):473-486. DOI 10.1093/molbev/msr133.
- Han L., Abney M. Identity by descent estimation with dense genome-wide genotype data. *Genet. Epidemiol.* 2011;35(6):557-567. DOI 10.1002/gepi.20606.
- Harris D.N., Ruczinski I., Yanek L.R., Becker L.C., Becker D.M., Guio H., Cui T., Chilton F.H., Mathias R.A., O'Connor T.D. Evolution of hominin polyunsaturated fatty acid metabolism: from Africa to the New World. *Genome Biol. Evol.* 2019;11(5):1417-1430. DOI 10.1093/gbe/evz071.
- Hochreiter S. HapFABIA: Identification of very short segments of identity by descent characterized by rare variants in large sequencing data. *Nucleic Acids Res.* 2013;41(22):e202. DOI 10.1093/nar/gkt1013.
- Horrocks L.A., Yeo Y.K. Health benefits of docosahexaenoic acid (DHA). *Pharmacol. Res.* 1999;40(3):211-225. DOI 10.1006/phrs.1999.0495.
- Howe P., Meyer B., Record S., Baghurst K. Dietary intake of long-chain ω -3 polyunsaturated fatty acids: contribution of meat sources. *Nutrition*. 2006;22(1):47-53. DOI 10.1016/j.nut.2005.05.009.
- Kimura T., Kobayashi T., Munkhbat B., Oyungerel G., Bilegtsaikhan T., Anar D., Jambaldorj J., Munkhsaikhan S., Munkhtuvshin N., Hayaishi H., Oka A., Inoue I., Inoko H. Genome-wide association analysis with selective genotyping identifies candidate loci for adult height at 8q21.13 and 15q22.33-q23 in Mongolians. *Hum. Genet.* 2008; 123(6):655-660. DOI 10.1007/s00439-008-0512-x.
- Kiriti D., Chmel N., Arnold A.W., Jakob T., Bruckner-Tuderman L., Has C. Novel and recurrent *AAGAB* mutations: clinical variability and molecular consequences. *J. Invest. Dermatol.* 2013;133(10): 2483-2486. DOI 10.1038/jid.2013.171.
- Kuhlwil M., Boeckx C. A catalog of single nucleotide changes distinguishing modern humans from archaic hominins. *Sci. Rep.* 2019; 9(1):8463. DOI 10.1038/s41598-019-44877-x.
- Marszalek J.R., Lodish H.F. Docosahexaenoic acid, fatty acid-interacting proteins, and neuronal function: breastmilk and fish are good for you. *Annu. Rev. Cell Dev. Biol.* 2005;21:633-657. DOI 10.1146/annurev.cellbio.21.122303.120624.
- Palamara P.F., Lencz T., Darvasi A., Pe'er I. Length distributions of identity by descent reveal fine-scale demographic history. *Am. J. Hum. Genet.* 2012;91(5):809-822. DOI 10.1016/j.ajhg.2012.08.030.
- Pugach I., Matveev R., Spitsyn V., Makarov S., Novgorodov I., Osakovsky V., Stoneking M., Pakendorf B. The complex admixture history and recent southern origins of Siberian populations. *Mol. Biol. Evol.* 2016;33(7):1777-1795. DOI 10.1093/molbev/msw055.
- Tucci S., Vohr S.H., McCoy R.C., Vernot B., Robinson M.R., Barbieri C., Nelson B.J., Fu W., Purnomo G.A., Sudoyo H., Eichler E.E., Barbujani G., Visscher P.M., Akey J.M., Green R.E. Evolutionary history and adaptation of a human pygmy population of Flores Island, Indonesia. *Science*. 2018;361(6401):511-516. DOI 10.1126/science.aar8486.
- Yin L.L., Li J.M., Zhou Z.M., Sha J.H. Identification of a novel testis-specific gene and its potential roles in testis development/spermatogenesis. *Asian J. Androl.* 2005;7(2):127-137. DOI 10.1111/j.1745-7262.2005.00041.x.

ORCID ID

N.A. Kolesnikov orcid.org/0000-0001-8855-577X
V.N. Kharkov orcid.org/0000-0002-1679-2212
K.V. Vagaitseva orcid.org/0000-0003-4877-9749
A.A. Zarubin orcid.org/0000-0001-6568-6339
V.A. Stepanov orcid.org/0000-0002-5166-331X

Благодарности. Исследование выполнено за счет гранта Российского научного фонда № 22-64-00060 (<https://rscf.ru/project/22-64-00060/>).

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию 21.10.2022. После доработки 17.01.2023. Принята к публикации 24.01.2023.

Перевод на английский язык <https://vavilov.elpub.ru/jour>

Экспрессия генов *NUP153* и *YWHAB* с их канонических промоторов и альтернативных промоторов ретротранспозона LINE-1 в плаценте первого триместра беременности

В.В. Деменова¹✉, Е.Н. Толмачева¹, Т.В. Никитина¹, Е.А. Саженова¹, С.Ю. Юрьев², А.Ш. Махмутходжаев², А.С. Зуев¹, С.А. Филатова³, А.Е. Дмитриев³, Я.А. Даркова³, Л.П. Назаренко¹, И.Н. Лебедев^{1, 2}, С.А. Васильев^{1, 3}

¹ Научно-исследовательский институт медицинской генетики, Томский национальный исследовательский медицинский центр Российской академии наук, Томск, Россия

² Сибирский государственный медицинский университет Министерства здравоохранения Российской Федерации, Томск, Россия

³ Национальный исследовательский Томский государственный университет, Томск, Россия

✉ deviva2503@gmail.com

Аннотация. Для плаценты характерен уникальный гипометилированный геном. Благодаря этой особенности плаценты в первом триместре беременности наблюдается потенциальная возможность использования регуляторных элементов, полученных от ретровирусов и ретротранспозонов, которые во взрослом организме подавляются метилированием ДНК. Кроме того, у спонтанных абортусов и с нормальным кариотипом, и с анеуплоидиями отмечается аномальное повышение уровня метилирования ретротранспозона LINE-1 в трофобласте хориона, что может быть связано с нарушением транскрипции генов с использованием регуляторных элементов LINE-1. На сегодняшний день идентифицировано 988 генов, способных экспрессироваться с альтернативных промоторов LINE-1. Из них с помощью инструмента STRING были отобраны гены, продукты которых взаимодействуют с белками, экспрессирующимися на высоком уровне в плаценте и участвующими в дифференцировке трофобласта, *NUP153* и *YWHAB*. Целью настоящего исследования являлся анализ экспрессии генов *NUP153* и *YWHAB* с канонических и альтернативных промоторов ретротранспозона LINE-1 в зародышевой части плаценты спонтанных и медицинских абортусов первого триместра беременности. Определение уровня экспрессии генов проводили с помощью ПЦР в режиме реального времени в лимфоцитах взрослых индивидов ($n = 10$), в ворсинах хориона и экстраэмбриональной мезодерме медицинских абортусов ($n = 10$) и спонтанных абортусов с нормальным кариотипом ($n = 10$) и с наиболее частыми анеуплоидиями в I триместре беременности (трисомия 16 ($n = 8$) и моносомия X ($n = 6$)). Индекс метилирования LINE-1 оценивали в ворсинах хориона спонтанных абортусов с помощью таргетного бисульфитного массового параллельного секвенирования. Уровень экспрессии обоих генов с канонических промоторов был выше в лимфоцитах крови, чем в тканях плаценты ($p < 0.05$). Однако уровень экспрессии гена *NUP153* с альтернативного промотора LINE-1 был выше в 17 раз в ворсинах хориона и в 23 раза – в экстраэмбриональной мезодерме по сравнению с лимфоцитами ($p < 0.05$). Между группами спонтанных абортусов с моносомией X и остальными группами были выявлены статистически значимые различия. Уровень экспрессии *NUP153* и *YWHAB* с канонических промоторов был выше в группе спонтанных абортусов с моносомией X по сравнению со всеми другими группами ($p < 0.05$). Индекс метилирования LINE-1 отрицательно коррелировал с уровнем экспрессии генов как с канонических (*NUP153* – $R = -0.59$, *YWHAB* – $R = -0.52$, $p < 0.05$), так и с альтернативных промоторов LINE-1 (*NUP153* – $R = -0.46$, *YWHAB* – $R = -0.66$, $p < 0.05$). Таким образом, наблюдаемое нами повышение индекса метилирования LINE-1 в плаценте спонтанных абортусов связано с уровнем экспрессии генов *NUP153* и *YWHAB* не только с альтернативных, но и с канонических промоторов, что в дальнейшем может приводить к негативным последствиям для нормального эмбриогенеза.
Ключевые слова: невынашивание беременности; плацента; ретротранспозон LINE-1; метилирование ДНК; *NUP153*; *YWHAB*.

Для цитирования: Деменова В.В., Толмачева Е.Н., Никитина Т.В., Саженова Е.А., Юрьев С.Ю., Махмутходжаев А.Ш., Зуев А.С., Филатова С.А., Дмитриев А.Е., Даркова Я.А., Назаренко Л.П., Лебедев И.Н., Васильев С.А. Экспрессия генов *NUP153* и *YWHAB* с их канонических промоторов и альтернативных промоторов ретротранспозона LINE-1 в плаценте первого триместра беременности. *Вавиловский журнал генетики и селекции*. 2023;27(1):63-71. DOI 10.18699/VJGB-23-09

Expression of the *NUP153* and *YWHAB* genes from their canonical promoters and alternative promoters of the LINE-1 retrotransposon in the placenta of the first trimester of pregnancy

V.V. Demeneva¹✉, E.N. Tolmacheva¹, T.V. Nikitina¹, E.A. Sazhenova¹, S.Yu. Yuriev², A.Sh. Makhmutkhodzhaev², A.S. Zuev¹, S.A. Filatova³, A.E. Dmitriev³, Ya.A. Darkova³, L.P. Nazarenko¹, I.N. Lebedev^{1, 2}, S.A. Vasilyev^{1, 3}

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

² Siberian State Medical University, Tomsk, Russia

³ National Research Tomsk State University, Tomsk, Russia

✉ deviva2503@gmail.com

© Деменова В.В., Толмачева Е.Н., Никитина Т.В., Саженова Е.А., Юрьев С.Ю., Махмутходжаев А.Ш., Зуев А.С., Филатова С.А., Дмитриев А.Е., Даркова Я.А., Назаренко Л.П., Лебедев И.Н., Васильев С.А., 2023

Контент доступен под лицензией Creative Commons Attribution 4.0

Abstract. The placenta has a unique hypomethylated genome. Due to this feature of the placenta, there is a potential possibility of using regulatory elements derived from retroviruses and retrotransposons, which are suppressed by DNA methylation in the adult body. In addition, there is an abnormal increase in the level of methylation of the LINE-1 retrotransposon in the chorionic trophoblast in spontaneous abortions with both normal karyotype and aneuploidy on different chromosomes, which may be associated with impaired gene transcription using LINE-1 regulatory elements. To date, 988 genes that can be expressed from alternative LINE-1 promoters have been identified. Using the STRING tool, genes (*NUP153* and *YWHAB*) were selected, the products of which have significant functional relationships with proteins highly expressed in the placenta and involved in trophoblast differentiation. This study aimed to analyze the expression of the *NUP153* and *YWHAB* genes, highly active in the placenta, from canonical and alternative LINE-1 promoters in the germinal part of the placenta of spontaneous and induced abortions. Gene expression analysis was performed using real-time PCR in chorionic villi and extraembryonic mesoderm of induced abortions ($n = 10$), adult lymphocytes ($n = 10$), spontaneous abortions with normal karyotype ($n = 10$), and with the most frequent aneuploidies in the first trimester of pregnancy (trisomy 16 ($n = 8$) and monosomy X ($n = 6$)). The LINE-1 methylation index was assessed in the chorionic villi of spontaneous abortions using targeted bisulfite massive parallel sequencing. The level of expression of both genes from canonical promoters was higher in blood lymphocytes than in placental tissues ($p < 0.05$). However, the expression level of the *NUP153* gene from the alternative LINE-1 promoter was 17 times higher in chorionic villi and 23 times higher in extraembryonic mesoderm than in lymphocytes ($p < 0.05$). The expression level of *NUP153* and *YWHAB* from canonical promoters was higher in the group of spontaneous abortions with monosomy X compared to all other groups ($p < 0.05$). The LINE-1 methylation index negatively correlated with the level of gene expression from both canonical ($NUP153 - R = -0.59$, $YWHAB - R = -0.52$, $p < 0.05$) and alternative LINE-1 promoters ($NUP153 - R = -0.46$, $YWHAB - R = -0.66$, $p < 0.05$). Thus, the observed increase in the LINE-1 methylation index in the placenta of spontaneous abortions is associated with the level of expression of the *NUP153* and *YWHAB* genes not only from alternative but also from canonical promoters, which can subsequently lead to negative consequences for normal embryogenesis.

Key words: miscarriage; placenta; retrotransposon LINE-1; DNA methylation; *NUP153*; *YWHAB*.

For citation: Demeneva V.V., Tolmacheva E.N., Nikitina T.V., Sazhenova E.A., Yuriev S.Yu., Makhmutkhodzhaev A.Sh., Zuev A.S., Filatova S.A., Dmitriev A.E., Darkova Ya.A., Nazarenko L.P., Lebedev I.N., Vasilyev S.A. Expression of the *NUP153* and *YWHAB* genes from their canonical promoters and alternative promoters of the LINE-1 retrotransposon in the placenta of the first trimester of pregnancy. *Vavilovskii Zhurnal Genetiki i Seleksii = Vavilov Journal of Genetics and Breeding*. 2023;27(1):63-71. DOI 10.18699/VJGB-23-09

Введение

У человека репродуктивные потери наблюдаются чаще в первом триместре беременности, чем в другие периоды эмбриогенеза. Одной из наиболее частых причин ранней эмбриональной гибели являются нарушения числа хромосом (анеуплоидии), которые приводят к тяжелым аномалиям развития. Формирование мейотической и митотической форм анеуплоидий соответствует волнам эпигенетического репрограммирования, в частности этапу деметилирования генома в зиготе и на стадии дробления. На данном этапе ранняя бластоциста демонстрирует меньшее метилирование ДНК, чем клетки организма в любой другой момент онтогенеза (Smith et al., 2012). Для внутренней клеточной массы затем следует быстрая волна метилирования ДНК *de novo*, в то время как трофобластическая оболочка остается гипометилированной (Santos et al., 2010).

На протяжении всей беременности плацента обладает уникальным гипометилированным эпигенетическим ландшафтом по сравнению с другими экстраэмбриональными и эмбриональными тканями, что может говорить о ее особых функциях (Robinson, Price, 2015). Гипометилирование в плацентарной ДНК происходит в основном в «частично метилированных доменах» и по всему геному распределяется неравномерно. «Частично метилированными доменами» называют большие (> 100 т. п. н.) участки с пониженным метилированием ДНК, чередующиеся с областями с более высоким метилированием ДНК (Schroeder et al., 2013).

В плаценте также наблюдается пониженное метилирование ДНК некоторых типов повторяющихся элементов генома (Price et al., 2012). Известно, что ретротранспозон

LINE-1 (long interspersed nuclear element-1) представляет собой самый большой, занимающий приблизительно до 20 % генома, и наиболее эволюционно молодой класс ретротранспозонов у человека, сохраняющий способность к транспозиции (Ostertag et al., 2001). Транскрипционная активность LINE-1 подавляется метилированием ДНК в течение большинства периодов онтогенеза.

Важной особенностью LINE-1, требующей внимания, является высокий уровень его метилирования в лейкоцитах крови, независимо от возраста и пола, при этом в других тканях уровень метилирования LINE-1 имеет свои тканеспецифические различия (Chalitchagorn et al., 2004). Так, показано, что для плаценты как самостоятельного органа уровень метилирования ретротранспозонов не всегда совпадает с глобальным уровнем метилирования всего генома. В одном из исследований было показано значительное снижение уровня метилирования LINE-1 в тканях плаценты третьего триместра беременности по сравнению с первым. При этом не было обнаружено изменений уровня метилирования ДНК всего генома между первым и третьим триместрами (He et al., 2014).

Можно предположить о временной регуляции метилирования и активации LINE-1 в ходе нормального развития плаценты. Это поднимает вопрос о возможной функциональной роли последовательностей ретротранспозона LINE-1 в развитии плаценты. Ранее нами было обнаружено, что часть спонтанных абортусов с нормальным кариотипом характеризуются эпигенетическими нарушениями, сходными со спонтанными абортусами с анеуплоидией. Кроме того, для спонтанных абортусов с нормальным кариотипом был отмечен повышенный индекс метили-

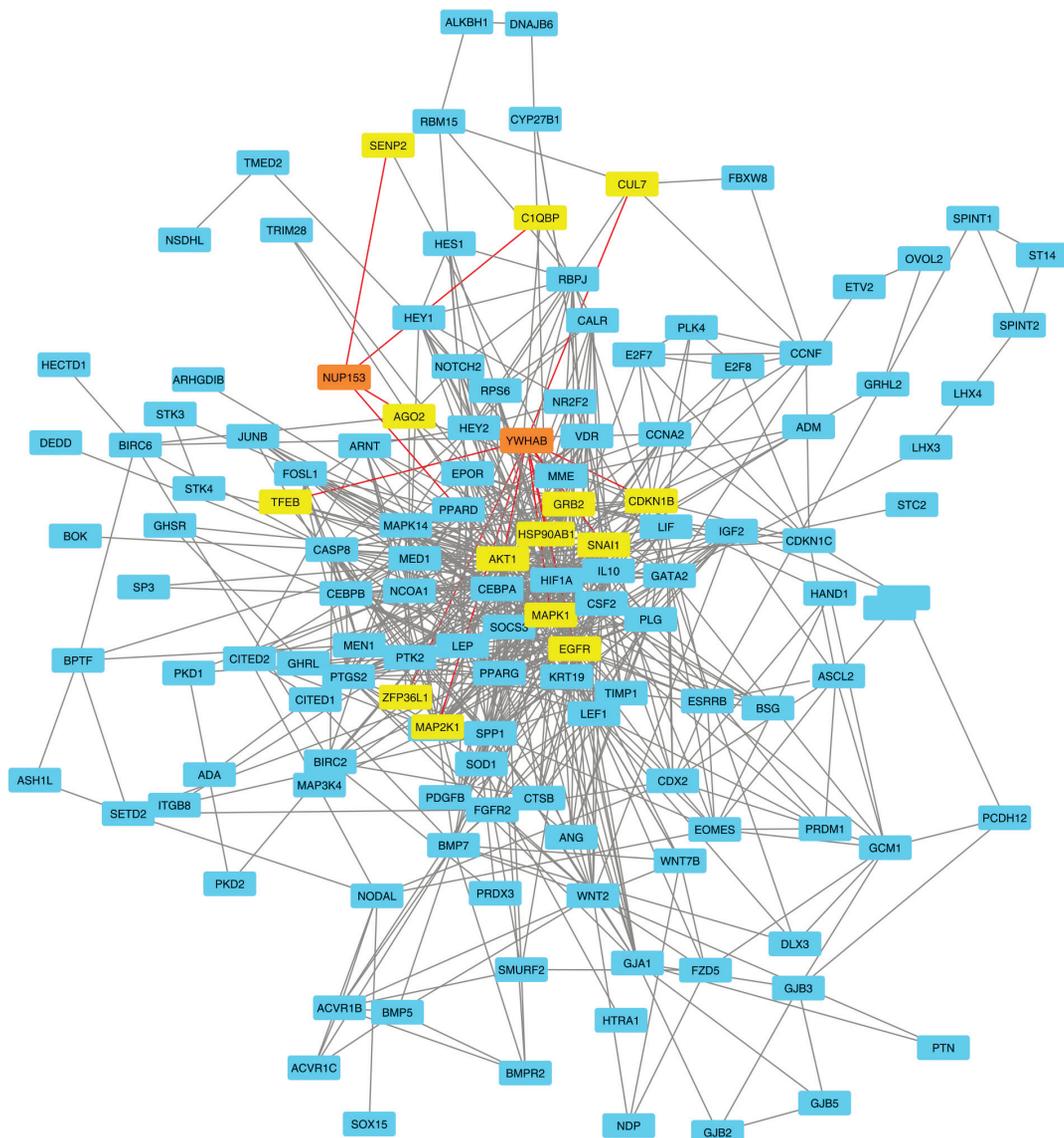


Рис. 1. Функционально значимые связи белков, участвующих в развитии плаценты (GO:0061450, GO:0097360, GO:0001890), с продуктами генов *NUP153* и *YWHAB* по данным STRING.

Желтым цветом показаны белки, имеющие функциональные связи (выделено красным) с продуктами генов *NUP153* и *YWHAB* (отмечены оранжевым цветом) (STRING score > 0.4).

рования в промоторе ретротранспозона LINE-1, наблюдаемый в группах спонтанных абортусов с трисомией по хромосоме 16 и моносомией X (Vasilyev et al., 2021b).

Одна из возможных ролей LINE-1 может состоять в использовании его регуляторных последовательностей для влияния на транскрипцию расположенных рядом генов. Такое влияние становится возможным благодаря тому, что LINE-1 включает в себя смысловой промотор, управляющий транскрипцией белков ORF1 и ORF2p, необходимых для ретротранспозиции, и антисмысловой промотор, управляющий транскрипцией химерных транскриптов – 5'-антисмысловых последовательностей LINE-1, сплайсированных с экзонами соседних генов (Denli et al., 2015). Антисмысловые транскрипты LINE-1 затрагивают до 4 % всех генов человека, а антисмысловые промоторы LINE-1 активно транскрибируются в различных типах клеток человека, в том числе и в эмбриональных тканях. Всего на

данный момент идентифицировано 988 генов, способных экспрессироваться с альтернативных промоторов LINE-1 (Criscione et al., 2016b). С учетом гипометилирования промоторов LINE-1 в плаценте возможно, что экспрессия множества генов в экстраэмбриональных тканях может происходить преимущественно с альтернативных промоторов LINE-1.

Среди генов, потенциально способных к экспрессии с альтернативных промоторов LINE-1, с помощью инструмента STRING было отобрано два, *NUP153* и *YWHAB*, продукты которых показали высокий уровень экспрессии в плаценте и функционально связаны с белками, участвующими в дифференцировке трофобласта (по данным Gene Ontology: GO:0061450, trophoblast cell migration; GO:0097360, chorionic trophoblast cell proliferation; GO:0001890, placenta development) (рис. 1). Ген *NUP153* функционирует как каркасный элемент в ядерной фазе

комплекса ядерных пор и необходим для нормального ядерно-цитоплазматического транспорта белков и мРНК как во время митоза соматических клеток (Bilig et al., 2019), так и в эмбриональных стволовых клетках мыши (Souquet et al., 2018). Ген *YWHAB* относится к группе генов, отвечающих за передачу сигнала путем связывания с фосфосерин-содержащими белками.

Белок, кодируемый геном, взаимодействует с фосфатазами *RAF1* и *CDC25* и может играть роль в митогенной передаче сигналов и механизмах клеточного цикла. В клеточной линии фибробластов, выделенной из эмбрионов мышей, было показано, что сверхэкспрессия *YWHAB* стимулирует и поддерживает независимый от прикрепления рост клеток (Sasaki et al., 2014).

Таким образом, целью настоящего исследования был анализ экспрессии генов *NUP153* и *YWHAB* с канонических и альтернативных промоторов LINE-1 в зародышевой части плаценты спонтанных и медицинских абортусов.

Материалы и методы

В качестве материала для исследования были отобраны образцы ворсин хориона и экстраэмбриональной мезодермы медицинских абортусов (МА) ($n = 10$, гестационный возраст 8.2 ± 2.3 нед), спонтанных абортусов (СА) с нормальным кариотипом ($n = 10$, гестационный возраст 7.2 ± 1.4 нед), трисомией по хромосоме 16 ($n = 8$, гестационный возраст 6.5 ± 0.8 нед) и моносомией X ($n = 6$, гестационный возраст 8.6 ± 0.7 нед). Образцы, взятые из биобанка «Биобанк населения Северной Азии» НИИ медицинской генетики Томского НИМЦ, получены в период с 2004 по 2021 г. и хранились в жидком азоте.

Исследование проводили с соблюдением этических норм в соответствии с Хельсинской декларацией Всемирной медицинской ассоциации. Для всех образцов из биобанка получено информированное согласие родителей на участие в эксперименте. Исследование было одобрено комитетом по биомедицинской этике НИИ медицинской генетики Томского НИМЦ (09.11.2020/№ 7).

Для определения кариотипа проведен стандартный цитогенетический анализ на прямых препаратах ворсин хориона и культурах фибробластов экстраэмбриональной мезодермы (Lebedev et al., 2004). Результаты кариотипирования для 14 образцов СА с трисомией и моносомией были подтверждены методом флуоресцентной гибридизации *in situ* (FISH). Уровень мозаицизма по анеуплоидии оценивали с нижним порогом 10 % и верхним порогом 90 %.

Для анализа моносомии X были использованы центромеро-специфичные ДНК-зонды на хромосомы X, а для анализа трисомии 16 – субтеломерные ДНК-зонды (16q и 16p). Анализ выполняли по описанной ранее методике (Vasilyev et al., 2010). Четыре образца имели мозаичный кариотип с уровнем трисомии от 10 до 90 %. Остальные 10 спонтанных абортусов с более высокой долей трисомии или моносомии были классифицированы как имеющие чистую анеуплоидию. В качестве группы сравнения использовали лимфоциты периферической крови родителей МА (5 семейных пар, возраст 30.8 ± 2.7), содержащиеся до начала эксперимента в реагенте «Лири» (Биолабмикс, Россия).

Выделение РНК проводили из ворсин хориона и экстраэмбриональной мезодермы фенол-хлороформным методом. С момента получения материала изучаемых образцов до начала выделения РНК все ткани хранились в жидком азоте. Разделение тканей предварительно осуществляли в RNeasy Lysis Buffer (Qiagen, США) для стабилизации РНК в образцах. Каждый образец гомогенизировали в ступке с жидким азотом, предварительно добавив 500 мкл реагента «Лири» (Биолабмикс). Лизат инкубировали сначала 5 мин при 55 °С, далее – 5 мин при комнатной температуре. Затем лизат центрифугировали при 12000 g в течение 10 мин для удаления нерастворенных фрагментов и отбирали супернатант в новую пробирку. На каждый 1 мл реагента «Лири» был добавлен 0.2 мл хлороформа, с последующим встряхиванием (вручную) в течение 15 с, инкубированием смеси в течение 10 мин при комнатной температуре и центрифугированием при 10000 g в течение 10 мин при 4 °С. Затем к водной фазе, содержащей РНК, добавляли 0.5 мл 100 % охлажденного изопропанола на каждый 1 мл реагента «Лири», инкубировали смесь при –20 °С в течение 10 мин и центрифугировали образец при 12000 g в течение 10 мин при 4 °С.

Осадок промывали дважды 80 % охлажденным этанолом при 10000 g в течение 5 мин при 4 °С. Сушили осадок 2 мин в концентрате (Eppendorf, США) (параметры: 45 °С, V-AL). Для растворения осадка добавляли 40 мкл ДЭПК-воды и 1 мкл RiboLock (Thermo, США) и ожидали 10 мин при комнатной температуре до полного растворения. Все образцы во время выделения на этапах инкубации при комнатной температуре находились на холодной подложке во избежание деградации РНК. После выделения все образцы хранились при –80 °С.

Перед проведением обратной транскрипции с получением кДНК осуществляли обработку РНК ДНКазой. Далее для обратной транскрипции использовали набор ОТ-М-MuLV-RN (Биолабмикс) со случайным гексапраймером. Смесь для проведения реакции обратной транскрипции включала 1.5 мкг РНК, 3 мкл гексапраймера, 4 мкл реакционного буфера KCl, 2 мкл 0.1 М ДДТ, 1 мкл 10 мМ смеси dNTP и 1 мкл ревертазы. Анализ экспрессии генов выполняли с помощью метода ПЦР в реальном времени. В качестве референсного гена был использован ген «домашнего хозяйства» *GAPDH*. Для генов *NUP153* и *YWHAB* были подобраны два вида праймеров: первый для длинных продуктов, которые экспрессируются только с канонических промоторов, а второй – для коротких продуктов, экспрессирующихся с альтернативных анти-смысловых промоторов LINE-1 (см. таблицу).

Ген *NUP153* включает 22 экзона, а короткий транскрипт с альтернативного промотора LINE-1 – только экзоны 21–22. Для детекции транскриптов гена *NUP153* с канонического промотора были подобраны праймеры на экзоны 16–17, а для детекции транскриптов с альтернативного промотора – на экзоны 21–22. С обычного промотора гена *YWHAB* транскрибируются два нормальных длинных транскрипта с экзонами 7 или 6, в связи с этим было решено использовать праймеры на каждый отдельный продукт. Для первого транскрипта с семью экзонами праймеры были подобраны в экзонах 1–2. Для второго транскрипта с шестью экзонами праймеры были

Последовательности олигонуклеотидных праймеров для оценки уровня экспрессии генов *NUP153*, *YWHAB* и *GAPDH* с помощью ПЦР в реальном времени

Ген	Транскрипт	Последовательность нуклеотидов
<i>NUP153</i>	Транскрипт с канонического промотора (NM_001278209.2, 22 экзона)	F 5'-TGTATGCTGAGAAACCAGGAAGTT-3' R 5'-GTAGAGTCTGCCTTATTCTGCACTA-3'
	Укороченный транскрипт с альтернативного промотора LINE-1 (2 экзона)	F 5'-CAGCATTACAGTGGGGTCAAAT-3' R 5'-CAACACCAATGTGACCTTTATTTCC-3'
<i>YWHAB</i>	Транскрипт с канонического промотора (NM_003404, 7 экзонов)	F 5'-GCTCGGAAGGGTCTTTGTTC-3' R 5'-TCTATCCACAGCCGAATGGG-3'
	Транскрипт с канонического промотора (NM_139323, 6 экзонов)	F 5'-GAGTAGTGGGCTTAGGAAGGAAGAG-3' R 5'-CTTTTATCCATTGTCATTCCCGTGG-3'
	Укороченный транскрипт с альтернативного промотора LINE-1 (4 экзона)	F 5'-CTGTAGCCTGGCAAAAACGG-3' R 5'-TCCGATGTCCACAGAGTGAGA-3'
<i>GAPDH</i>	Транскрипт с канонического промотора (NM_002046.7, 10 экзонов)	F 5'-GCCAGCCGAGCCACATC-3' R 5'-GGCAACAATATCCACTTTACCAGA-3'

Примечание. F – прямой праймер; R – обратный праймер.

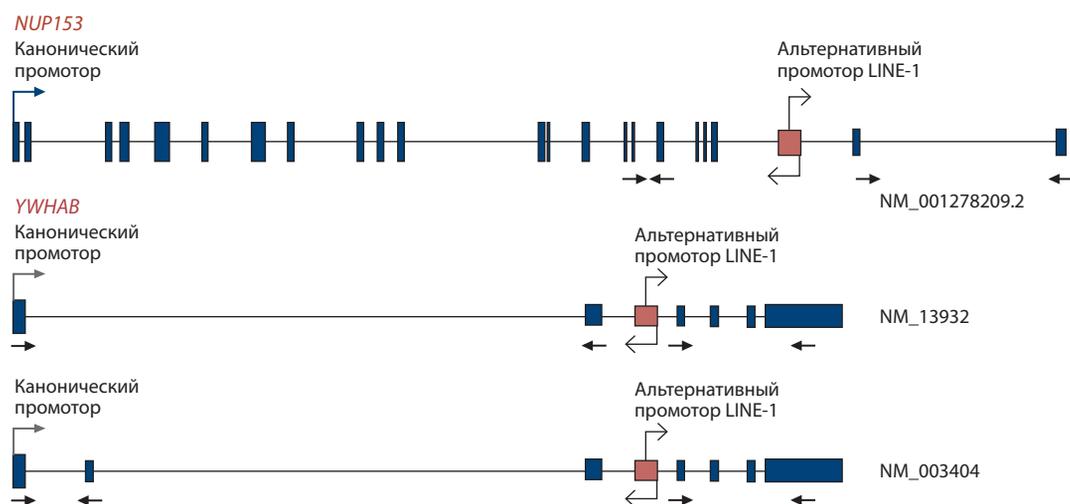


Рис. 2. Схема расположения альтернативных промоторов LINE-1 для генов *NUP153* и *YWHAB*.

Стрелками схематично показаны сайты гибридизации олигонуклеотидных праймеров. Стрелкой, начинающейся в начале элемента LINE-1, отмечено направление считывания с прямого промотора LINE-1, являющегося для него основным. Второй стрелкой, направленной в другую сторону, показано направление экспрессии с альтернативного промотора LINE-1, альтернативного и для изучаемых генов.

подобраны в экзонах 1–3. Праймеры короткого продукта с альтернативного промотора LINE-1 гена *YWHAB* были подобраны в экзонах 4–7 (рис. 2). За экспрессию с альтернативных промоторов генов принимали разность между уровнем экспрессии гена, оцененной с помощью праймеров, лежащих после альтернативного промотора, и уровнем экспрессии гена, оцененной с помощью праймеров, лежащих до него. И уже этот показатель был использован для анализа данных и отображен на графиках. Для гена *YWHAB* вычитали уровень экспрессии обоих транскриптов с канонического промотора.

Индекс метилирования оценивали в 19 CpG-сайтах промотора LINE-1 для ворсин хориона спонтанных абортусов с использованием таргетного бисульфитного

массового параллельного секвенирования. Приготовление библиотек и оценку проводили по ранее опубликованному протоколу (Vasilyev et al., 2021a). Статистический анализ данных выполняли с помощью программного обеспечения Statistica 10.0.

Результаты

Уровень экспрессии гена *NUP153* с канонического промотора был выше в лимфоцитах в 12.5 раза по сравнению с тканями плаценты ($p = 0.000001$). Уровень экспрессии гена *YWHAB* с канонического промотора был также в среднем выше в лимфоцитах крови, чем в тканях плаценты (в 4.6 раза) (транскрипт NM_13932 ($p = 0.00003$)). Уровень экспрессии транскрипта NM_003404 гена *YWHAB* имел

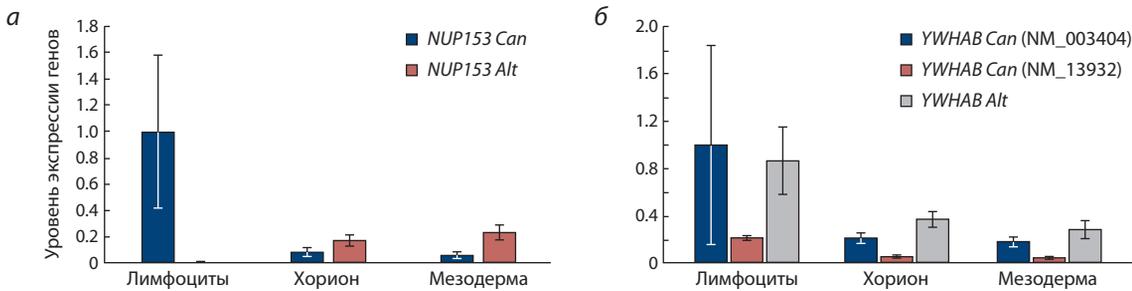


Рис. 3. Сравнение уровня экспрессии гена *NUP153* (а) и *YWHAB* (б) с канонических промоторов и альтернативных промоторов LINE-1 в лимфоцитах крови, хорионе и мезодерме плаценты.

Значения приведены в виде кратности различий относительно уровня экспрессии генов с канонического промотора в лимфоцитах взрослых индивидов. Для гена *YWHAB* представлены уровни экспрессии двух различных транскриптов с канонического промотора (NM_003404, NM_139323). *GAPDH* – референсный ген; *Can* – канонический промотор; *Alt* – альтернативный промотор LINE-1.

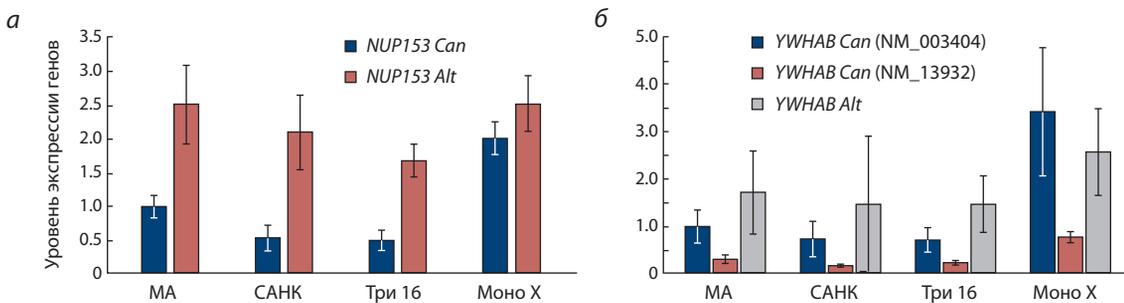


Рис. 4. Сравнение уровня экспрессии гена *NUP153* (а) и *YWHAB* (б) с канонических промоторов и альтернативных промоторов LINE-1 между группами спонтанных и медицинских абортусов.

Значения приведены как кратные различия относительно уровня экспрессии генов канонического промотора медицинских абортусов. САНК – спонтанные абортусы с нормальным кариотипом; Три 16 – спонтанные абортусы с трисомией по хромосоме 16; Моно X – спонтанные абортусы с моносомией X.

высокую вариабельность в лимфоцитах. Однако уровень экспрессии гена *NUP153* с альтернативных промоторов LINE-1 был статистически значимо выше в экстраэмбриональных тканях по сравнению с лимфоцитами взрослых индивидов (в ворсинах хориона – в 17 раз, в экстраэмбриональной мезодерме – в 23 раза, $p < 0.05$) (рис. 3). При сравнении групп спонтанных и медицинских абортусов между собой (ворсины хориона и экстраэмбриональная мезодерма) уровень экспрессии обоих генов с канонических промоторов был выше в группе СА с моносомией X, чем в группах СА с нормальным кариотипом (для гена *NUP153* – в 3.7 раза, для гена *YWHAB* – в 4.5 раза, $p < 0.05$), СА с трисомией 16 (для гена *NUP153* – в 3.9 раза, для гена *YWHAB* – в 4.3 раза, $p < 0.05$) и МА (для гена *YWHAB* – в 3.1 раза, $p < 0.05$) (рис. 4).

В ворсинах хориона спонтанных абортусов с различным кариотипом был оценен уровень метилирования в промоторе ретротранспозона LINE-1. Средний уровень метилирования LINE-1 в ворсинах хориона СА составил: с трисомией 16 – $41.9 \pm 5.8\%$, с моносомией X – $39.7 \pm 3.6\%$, с нормальным кариотипом – $38.4 \pm 3.9\%$. Индекс метилирования LINE-1 отрицательно коррелировал с уровнем экспрессии генов как с канонических (*NUP153* – $R = -0.59$, $p < 0.003$; *YWHAB* – $R = -0.52$, $p < 0.01$), так и с альтернативных промоторов LINE-1 (*NUP153* – $R = -0.46$, $p = 0.03$; *YWHAB* – $R = -0.66$, $p = 0.001$) (рис. 5).

Обсуждение

В настоящей работе обнаружено, что уровень экспрессии для генов *NUP153* и *YWHAB* в плаценте был ниже с канонических промоторов по сравнению с лимфоцитами крови взрослых индивидов, однако экспрессия гена *NUP153* с альтернативного промотора LINE-1 была выше в плаценте. Полученный результат свидетельствует в пользу гипотезы о том, что в плаценте вследствие гипометилированного эпигенетического ландшафта может активироваться экспрессия генов с альтернативных промоторов, приобретенных от ретровирусов и ретротранспозонов. Данное предположение поддерживается также и обогащением генов, тканеспецифично экспрессирующихся в плаценте, среди всех генов, способных к транскрипции с альтернативных промоторов LINE-1 (Ciscione et al., 2016a).

Нами не выявлено значимых отличий по уровню экспрессии генов *YWHAB* и *NUP153* с альтернативных промоторов между группами спонтанных абортусов с различным кариотипом и контрольной группой медицинских абортусов. При этом уровень экспрессии обоих генов с канонических промоторов был выше в группе спонтанных абортусов с моносомией X. Но обнаружено, что уровень экспрессии изученных генов изменяется у отдельных спонтанных абортусов в зависимости от изменения уровня метилирования в геноме. Полученные данные наглядно демонстрируют, что уровень экспрессии генов *NUP153*

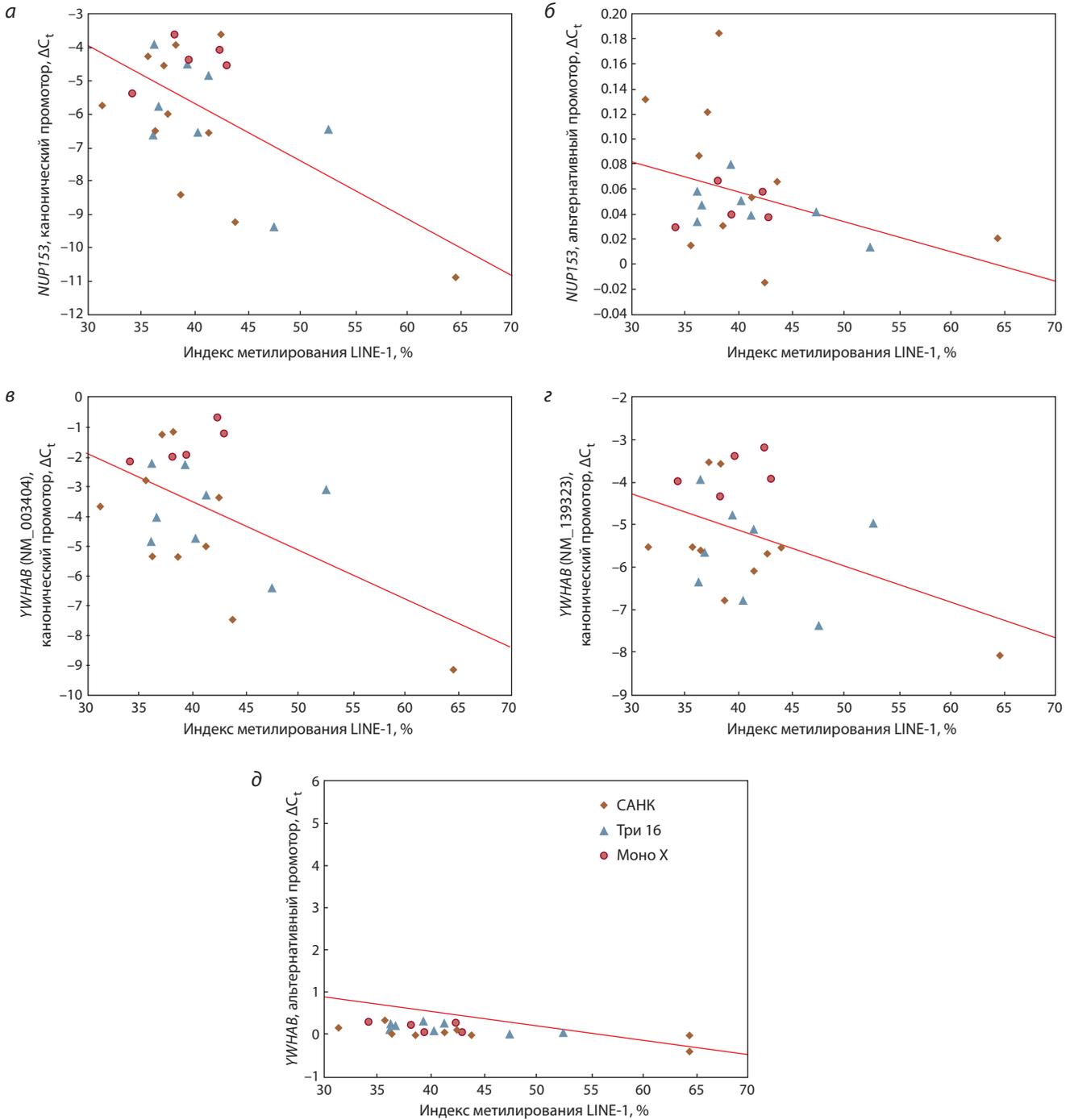


Рис. 5. Корреляция экспрессии генов *NUP153* и *YWHAВ* с индексом метилирования LINE-1 в трофобласте хориона спонтанных абортусов с нормальным кариотипом, трисомией 16 и моносомией X.

а – корреляция экспрессии гена *NUP153* с канонического промотора и индекса метилирования LINE-1; *б* – корреляция экспрессии гена *NUP153* с альтернативного промотора и индекса метилирования LINE-1; *в* – корреляция экспрессии гена *YWHAВ* (NM_003404) с канонического промотора и индекса метилирования LINE-1; *г* – корреляция экспрессии гена *YWHAВ* (NM_139323) с канонического промотора и индекса метилирования LINE-1; *д* – корреляция экспрессии гена *YWHAВ* с альтернативного промотора и индекса метилирования LINE-1. SANК – спонтанные абортусы с нормальным кариотипом; Три 16 – спонтанные абортусы с трисомией по хромосоме 16; Моно X – спонтанные абортусы с моносомией X.

и *YWHAВ* с канонических и альтернативных промоторов LINE-1 коррелирует с уровнем метилирования LINE-1: чем выше уровень метилирования LINE-1, тем ниже экспрессия.

Причин связи уровня метилирования LINE-1 с экспрессией исследуемых генов с обоих промоторов может быть

несколько. Во-первых, короткий транскрипт с альтернативного промотора может быть связан с активацией транскрипции гена с канонического промотора. Однако такой вариант представляется маловероятным, так как уровень экспрессии исследуемых генов с канонических промоторов на фоне гипометилирования генома в плаценте

был ниже, чем в лимфоцитах, для которых характерен уровень метилирования LINE-1 более 70 % (Rosser, An, 2012), хотя при верности высказанной гипотезы должно было быть наоборот. Во-вторых, уровень метилирования ретротранспозона LINE-1 может отражать глобальный уровень метилирования генома и уровень метилирования в каноническом промоторе исследуемых генов. Этот вариант представляется более вероятным, но также не снимает вопрос о сниженной экспрессии исследованных генов в плаценте на фоне гипометилированного эпигенетического ландшафта по сравнению с лимфоцитами взрослых индивидов. Очевидно, экспрессия исследуемых генов регулируется не только с помощью метилирования, но и тканеспецифичными факторами транскрипции.

Остается неясным, играет ли изменение экспрессии генов *NUP153* и *YWHAB* как с канонических, так и с альтернативных промоторов функциональную роль в жизнедеятельности плаценты или эти транскрипты – побочный продукт гипометилирования генома. Потенциально, нарушение работы гена *NUP153* может оказывать негативное влияние на ядерно-цитоплазматический транспорт белков и мРНК, а нарушение работы гена *YWHAB* может влиять на передачу клеточных сигналов.

Продукты генов *NUP153* и *YWHAB* имеют значимые функциональные связи с белками, участвующими в дифференцировке трофобласта (см. рис. 1). Для гена *NUP153* такими генами являются: *AGO2*, *SENP2*, *CIQBP*, *PPARD*. Для *YWHAB* список значимых связей шире – он взаимодействует с генами *TFEB*, *CUL7*, *ZFP36L*, *MAP2K1*, *AKT1*, *CDKN1B*, *SNAP1*, *MAPK1*, *EGFR*.

Нарушение работы каждого из этих генов оказывает негативное влияние на нормальное течение эмбриогенеза. Например, нормальная экспрессия *MAPK1* необходима для развития внеэмбриональной эктодермы во время плацентогенеза, а его отсутствие может приводить к гибели эмбриона из-за аномального развития и гиповаскуляризации плаценты (Bissonauth et al., 2006). Ген *CUL7* активно экспрессируется в клеточных линиях трофобласта. Дефицит белка Cul7 связан с задержкой внутриутробного развития из-за аномального развития плаценты, что может вызывать плохое снабжение плода кислородом (Fahlbusch et al., 2012). На более поздних сроках гестации дефицит может приводить к возникновению кожных или гиподермальных кровоизлияний, а также к развитию трофобласта с аномальной сосудистой структурой (Arai et al., 2003). Мутации *CUL7* в зародышевой линии связаны с синдромом 3-М и синдромом низкорослости якутов, оба из которых характеризуются пре- и постнатальной задержкой роста (Maksimova et al., 2007; Fu et al., 2010).

Ген *SENP2* принадлежит к семейству белков убиквитин-подобных и локализован в клетке в ядерных порах и цитоплазме (Talamillo et al., 2020). Мутации *SENP2* вызывают нарушение прогрессирования клеточного цикла во время развития трофобласта у мышей: удаление *SENP2* нарушает путь p53/Mdm2, влияя на клетки-предшественники трофобласта и на их созревание (Chiu et al., 2008). Во время дальнейшей дифференцировки *SENP2* оказывает влияние на нормальное развитие кардиомиоцитов. Сверхэкспрессия приводит к аномальной пролиферации кардиомиоцитов с нарушенной регуляцией циклина и

ингибиторов циклин-зависимой киназы, что приводит к врожденным аномалиям развития сердца (Kim et al., 2012). С другой стороны, делеции также вызывают дефекты развития миокарда из-за снижения пролиферации (Kang et al., 2010).

Логично предположить, что существующие функциональные связи генов *NUP153* и *YWHAB* с генами, участвующими в том числе и в дифференцировке трофобласта, могут идти как в негативном, так и в протективном направлении. Патологическое изменение экспрессии генов *NUP153* и *YWHAB* может потенциально приводить к нарушению работы других генов, формированию патологического фенотипа у пациентов или даже закончиться эмбриональной гибелью.

Заключение

Нами выявлено, что в тканях плаценты гены *NUP153* и *YWHAB* преимущественно экспрессируются с альтернативных промоторов LINE-1, находящихся в интронах. Несмотря на то что для всех групп (спонтанные и медицинские абортусы) экспрессия с альтернативных промоторов LINE-1 выше, чем с канонических, и не обнаружено значимых отличий по уровню экспрессии генов *YWHAB* и *NUP153* с альтернативных промоторов между группами, показана тенденция к общему снижению экспрессии для спонтанных абортусов по сравнению с медицинскими абортусами. Однако установлено, что уровень экспрессии изученных генов изменяется у отдельных спонтанных абортусов в зависимости от уровня метилирования в геноме. Полученные данные наглядно демонстрируют связь уровня экспрессии генов *NUP153* и *YWHAB* с канонических и альтернативных промоторов LINE-1 с уровнем метилирования LINE-1 в экстраэмбриональных тканях спонтанных абортусов: чем выше уровень метилирования LINE-1, тем ниже экспрессия генов.

Таким образом, повышение индекса метилирования LINE-1 в плаценте спонтанных абортусов может быть связано со снижением экспрессии генов не только с альтернативных, но и с канонических промоторов. Установленные особенности взаимосвязи уровня метилирования LINE-1 с экспрессией генов *NUP153* и *YWHAB* указывают на существующий механизм саморегуляции нормального эмбриогенеза, при нарушении которого последствием является эмбриональная гибель плода.

Список литературы / References

- Arai T., Kasper J.S., Skaar J.R., Ali S.H., Takahashi C., DeCaprio J.A. Targeted disruption of *p185/Cul7* gene results in abnormal vascular morphogenesis. *Proc. Natl. Acad. Sci. USA*. 2003;100(17):9855-9860. DOI 10.1073/pnas.1733908100.
- Bilir Ş., Kojidani T., Mori C., Osakada H., Kobayashi S., Koujin T., Hiraoka Y., Haraguchi T. Roles of Nup133, Nup153 and membrane fenestrations in assembly of the nuclear pore complex at the end of mitosis. *Genes Cells*. 2019;24(5):338-353. DOI 10.1111/gtc.12677.
- Bissonauth V., Roy S., Gravel M., Guillemette S., Charron J. Requirement for *Map2k1* (*Mek1*) in extra-embryonic ectoderm during placental development. *Development*. 2006;133(17):3429-3440. DOI 10.1242/dev.02526.
- Chalitchagorn K., Shuangshoti S., Hourpai N., Kongruttanachok N., Tangkijvanich P., Thong-ngam D., Voravud N., Sriuranpong V., Mutirangura A. Distinctive pattern of LINE-1 methylation level in

- normal tissues and the association with carcinogenesis. *Oncogene*. 2004;23(54):8841-8846. DOI 10.1038/sj.onc.1208137.
- Chiu S.Y., Asai N., Costantini F., Hsu W. SUMO-specific protease 2 is essential for modulating p53-mdm2 in development of trophoblast stem cell niches and lineages. *PLoS Biol*. 2008;6(12):e310. DOI 10.1371/journal.pbio.0060310.
- Criscione S.W., Teo Y.V., Neretti N. The chromatin landscape of cellular senescence. *Trends Genet*. 2016a;32(11):751-761. DOI 10.1016/j.tig.2016.09.005.
- Criscione S.W., Theodosakis N., Micevic G., Cornish T.B., Burns K.H., Neretti N., Rodić N. Genome-wide characterization of human L1 antisense promoter-driven transcripts. *BMC Genomics*. 2016b;17:463. DOI 10.1186/s12864-016-2800-5.
- Denli A.M., Narvaiza I., Kerman B.E., Pena M., Benner C., Marchetto M.C., Diedrich J.K., Aslanian A., Ma J., Moresco J.J., Moore L., Hunter T., Saghatelian A., Gage F.H. Primate-specific ORF0 contributes to retrotransposon-mediated diversity. *Cell*. 2015;163(3):583-593. DOI 10.1016/j.cell.2015.09.025.
- Fahlbusch F.B., Dawood Y., Hartner A., Menendez-Castro C., Nögel S.C., Tzschoppe A., Schneider H., Strissel P., Beckmann M.W., Schleussner E., Ruebner M., Dörr H.G., Schild R.L., Rascher W., Dötsch J. Cullin 7 and Fbxw 8 expression in trophoblastic cells is regulated via oxygen tension: implications for intrauterine growth restriction? *J. Matern. Fetal Neonatal Med*. 2012;25(11):2209-2215. DOI 10.3109/14767058.2012.684166.
- Fu J., Lv X., Lin H., Wu L., Wang R., Zhou Z., Zhang B., Wang Y.L., Tsang B.K., Zhu C., Wang H. Ubiquitin ligase cullin 7 induces epithelial-mesenchymal transition in human choriocarcinoma cells. *J. Biol. Chem*. 2010;285(14):10870-10879. DOI 10.1074/jbc.M109.004200.
- He Z.M., Li J., Hwa Y.L., Brost B., Fang Q., Jiang S.W. Transition of LINE-1 DNA methylation status and altered expression in first and third trimester placentas. *PLoS One*. 2014;9(5):96994. DOI 10.1371/journal.pone.0096994.
- Kang X., Qi Y., Zuo Y., Wang Q., Zou Y., Schwartz R.J., Cheng J., Yeh E.T.H. SUMO-specific protease 2 is essential for suppression of polycomb group protein-mediated gene silencing during embryonic development. *Mol. Cell*. 2010;38(2):191-201. DOI 10.1016/j.molcel.2010.03.005.
- Kim E.Y., Chen L., Ma Y., Yu W., Chang J., Moskowitz I.P., Wang J. Enhanced desumoylation in murine hearts by overexpressed SENP2 leads to congenital heart defects and cardiac dysfunction. *J. Mol. Cell. Cardiol*. 2012;52(3):638-649. DOI 10.1016/j.yjmcc.2011.11.011.
- 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.
- Maksimova N., Hara K., Miyashia A., Nikolaeva I., Shiga A., Nogovicina A., Sukhomyasova A., Argunov V., Shvedova A., Ikeuchi T., Nishizawa M., Kuwano R., Onodera O. Clinical, molecular and histopathological features of short stature syndrome with novel CUL7 mutation in Yakuts: new population isolate in Asia. *J. Med. Genet*. 2007;44(12):772-778. DOI 10.1136/jmg.2007.051979.
- Ostertag E.M., Kazazian H.H. Jr. Biology of mammalian L1 retrotransposons. *Annu. Rev. Genet*. 2001;35:501-538. DOI 10.1146/annurev.genet.35.102401.091032.
- Price E.M., Cotton A.M., Peñaherrera M.S., McFadden D.E., Kober M.S., Robinson W. Different measures of "genome-wide" DNA methylation exhibit unique properties in placental and somatic tissues. *Epigenetics*. 2012;7(6):652-663. DOI 10.4161/epi.20221.
- Robinson W.P., Price E.M. The human placental methylome. *Cold Spring Harb. Perspect. Med*. 2015;5(5):a023044. DOI 10.1101/cshperspect.a023044.
- Rosser J.M., An W. L1 expression and regulation in humans and rodents. *Front. Biosci. (Elite Ed)*. 2012;4(6):2203-2225. DOI 10.2741/537.
- Santos F., Hyslop L., Stojkovic P., Leary C., Murdoch A., Reik W., Stojkovic M., Herbert M., Dean W. Evaluation of epigenetic marks in human embryos derived from IVF and ICSI. *Hum. Reprod*. 2010;25(9):2387-2395. DOI 10.1093/humrep/deq151.
- Sasaki Y., Taya Y., Saito K., Fujita K., Aoba T., Fujiwara T. Molecular contribution to cleft palate production in cleft lip mice. *Congenit. Anom. (Kyoto)*. 2014;54(2):94-99. DOI 10.1111/cga.12038.
- Schroeder D.I., Blair J.D., Lott P., Yu H.O., Hong D., Cray F., Ashwood P., Walker C., Korf I., Robinson W.P., LaSalle J.M. The human placenta methylome. *Proc. Natl. Acad. Sci. USA*. 2013;110(15):6037-6042. DOI 10.1073/pnas.1215145110.
- Smith Z.D., Chan M.M., Mikkelsen T.S., Gu H., Gnirke A., Regev A., Meissner A. A unique regulatory phase of DNA methylation in the early mammalian embryo. *Nature*. 2012;484(7394):339-344. DOI 10.1038/nature10960.
- Souquet B., Freed E., Berto A., Andric V., Audugé N., Reina-San-Martin B., Lacy E., Doye V. Nup133 is required for proper nuclear pore basket assembly and dynamics in embryonic stem cells. *Cell Rep*. 2018;23(8):2443-2454. DOI 10.1016/j.celrep.2018.04.070.
- Talamillo O., Barroso-Gomila O., Giordano I., Ajuria L., Grillo M., Mayor U., Barrio R. The role of SUMOylation during development. *Biochem. Soc. Trans*. 2020;48(2):463-478. DOI 10.1042/BST20190390.
- Vasilyev S.A., Markov A.V., Vasilyeva O.Yu., Tolmacheva E.N., Zatulina L.A., Sharysh D.V., Zhigalina D.I., Demeneva V.V., Lebedev I.N. Method of targeted bisulfite massive parallel sequencing of the human LINE-1 retrotransposon promoter. *MethodsX*. 2021a;8:101445. DOI 10.1016/j.mex.2021.101445.
- Vasilyev S.A., Timoshevsky V.A., Lebedev I.N. Cytogenetic mechanisms of aneuploidy in somatic cells of chemonuclear industry 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., Zatulina 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*. 2021b;38(1):139-149. DOI 10.1007/s10815-020-02003-1.

ORCID ID

V.V. Demeneva orcid.org/0000-0002-5315-4914
E.N. Tolmacheva orcid.org/0000-0002-0716-4302
T.V. Nikitina orcid.org/0000-0002-4230-6855
E.A. Sazhenova orcid.org/0000-0003-3875-3932
S.Yu. Yuriev orcid.org/0000-0002-1343-5471
A.Sh. Makhmutkhodzhaev orcid.org/0000-0002-7541-0317

A.S. Zuev orcid.org/0000-0001-9474-9335
S.A. Filatova orcid.org/0000-0002-9344-0253
A.E. Dmitriev orcid.org/0000-0002-0070-4863
Ya.A. Darkova orcid.org/0000-0002-7117-9250
L.P. Nazarenko orcid.org/0000-0002-1861-433X
I.N. Lebedev orcid.org/0000-0002-0482-8046
S.A. Vasilyev orcid.org/0000-0002-5301-070X

Благодарности. Исследование проведено в рамках реализации темы государственного задания № 1022072600037-4.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию 13.10.2022. После доработки 26.12.2022. Принята к публикации 30.12.2022.

Перевод на английский язык <https://vavilov.elpub.ru/jour>

Изменение профиля метилирования ДНК в ткани печени при прогрессировании HCV-индуцированного фиброза до гепатоцеллюлярной карциномы

И.А. Гончарова , А.А. Зарубин, Н.П. Бабушкина, Ю.А. Королева, М.С. Назаренко

Научно-исследовательский институт медицинской генетики, Томский национальный исследовательский медицинский центр Российской академии наук, Томск, Россия
 irina.goncharova@medgenetics.ru

Аннотация. С использованием данных GSE73003 и GSE37988, представленных в базе данных GEODataSets (<https://www.ncbi.nlm.nih.gov/>), проведен сравнительный анализ уровня метилирования 27 578 CpG-сайтов между парными образцами опухолевой и окружающей опухоль тканями печени различной степени поражения (фиброз, цирроз) у больных HCV-индуцированной гепатоцеллюлярной карциномой (ГЦК), а также между опухолевой и нормальной тканью у больного ГЦК невирусной этиологии. Выявлено значительно меньшее число дифференциально метилированных сайтов между нормальной тканью печени и ГЦК невирусной этиологии, а также между ГЦК и фиброзом (32 и 40), чем между ГЦК и циррозом (2450 и 2304 соответственно по данным GSE73003 и GSE37988). По мере прогрессирования патологического изменения окружающей опухоль ткани уменьшается соотношение количества гипер-/гипометилированных дифференциально метилированных сайтов в опухоли. Так, в опухолевой ткани по сравнению с нормальной/фиброзом/циррозом печени гиперметилированы 75/62.5/47.7 % (GSE73003) и 16 % (GSE37988) CpG-сайтов соответственно. Стойкое гиперметилирование генов *ZNF154* и *ZNF540*, а также гипометилирование *CCL20* зарегистрировано в опухолевой ткани относительно как фиброза, так и цирроза печени. Белковые продукты генов *EDG4*, *CCL20*, *GPR109A* и *GRM8*, CpG-сайты которых характеризуются изменением уровня метилирования ДНК в опухоли на фоне цирроза и фиброза, принадлежат к категории «передачи сигналов рецепторов, связанных с G-белком». Однако изменение уровня метилирования «драйверных» для онкопатологии генов (*APC*, *CDKN2B*, *GSTP1*, *ELF4*, *TERT*, *WT1*) регистрируется в опухолевой ткани на фоне цирроза печени, но не фиброза. Среди гиперметилированных в опухолевой ткани генов на фоне цирроза печени наиболее представленными биологическими путями являются процессы развития, передачи межклеточных сигналов, регуляции транскрипции, связывания с белками Wnt-пути. Гены, гипометилированные в опухолевой ткани печени на фоне ее цирротического поражения, относятся к передаче обонятельных сигналов, нейроактивному взаимодействию лиганда с рецептором, кератинизации, иммунному ответу, ингибированию сериновых протеаз и метаболизму цинка. Гиперметилированные в опухоли гены локализуются в локусе 7p15.2 в регионе кластера *HOXA*, а гипометилированные CpG-сайты занимают протяженные области генома в кластерах генов обонятельных рецепторов (11p15.4), кератина и кератин-ассоциированных белков (12q13.13, 17q21.2 и 21q22.11), комплекса эпидермальной дифференцировки (1q21.3), а также функционирования иммунной системы – локусы 9p21.3 (кластер *IFNA*, *IFNB1*, *IFNW1*) и 19q13.41–19q13.42 (кластеры *KLK*, *SIGLEC*, *LILR*, *KIR*). Среди генов фиброгенеза или репарации ДНК *cg14143055* (*ADAMDEC1*) локализован в регионе связывания транскрипционных факторов семейства HOX, а *cg05921699* (*CD79A*), *cg06196379* (*TREM1*) и *cg10990993* (*MLH1*) расположены в области связывания транскрипционных факторов семейства белков ZNF. Таким образом, профиль метилирования ДНК в печени при HCV-индуцированной ГЦК является уникальным и различается в зависимости от степени поражения окружающей ткани – фиброз или цирроз.

Ключевые слова: метилирование ДНК; ХВГС; фиброз печени; цирроз печени; гепатоцеллюлярная карцинома.

Для цитирования: Гончарова И.А., Зарубин А.А., Бабушкина Н.П., Королева Ю.А., Назаренко М.С. Изменение профиля метилирования ДНК в ткани печени при прогрессировании HCV-индуцированного фиброза до гепатоцеллюлярной карциномы. *Вавиловский журнал генетики и селекции*. 2023;27(1):72-82. DOI 10.18699/VJGB-23-10

Changes in DNA methylation profile in liver tissue during progression of HCV-induced fibrosis to hepatocellular carcinoma

I.A. Goncharova , A.A. Zarubin, N.P. Babushkina, I.A. Koroleva, M.S. Nazarenko

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

Abstract. In this study we compared methylation levels of 27,578 CpG sites between paired samples of the tumor and surrounding liver tissues with various degrees of damage (fibrosis, cirrhosis) in HCV-induced hepatocellular carcinoma (HCC) patients, as well as between tumor and normal tissue in non-viral HCC patients, using GSE73003 and GSE37988 data from GEODataSets (<https://www.ncbi.nlm.nih.gov/>). A significantly lower number of differentially methylated sites (DMS) were found between HCC of non-viral etiology and normal liver tissue, as well as between

HCC and fibrosis (32 and 40), than between HCC and cirrhosis (2450 and 2304, respectively, according to GSE73003 and GSE37988 datasets). As the pathological changes in the tissue surrounding the tumor progress, the ratio of hyper-/hypomethylated DMSs in the tumor decreases. Thus, in tumor tissues compared with normal/fibrosis/cirrhosis of the liver, 75/62.5/47.7 % (GSE73003) and 16 % (GSE37988) of CpG sites are hypermethylated, respectively. Persistent hypermethylation of the *ZNF154* and *ZNF540* genes, as well as *CCL20* hypomethylation, were registered in tumor tissue in relation to both liver fibrosis and liver cirrhosis. Protein products of the *EDG4*, *CCL20*, *GPR109A*, and *GRM8* genes, whose CpG sites are characterized by changes in DNA methylation level in tumor tissue in the setting of cirrhosis and fibrosis, belong to "Signaling by G-protein-coupled receptors (GPCRs)" category. However, changes in the methylation level of the "driver" genes for oncopathology (*APC*, *CDKN2B*, *GSTP1*, *ELF4*, *TERT*, *WT1*) are registered in tumor tissue in the setting of liver cirrhosis but not fibrosis. Among the genes hypermethylated in tumor tissue in the setting of liver cirrhosis, the most represented biological pathways are developmental processes, cell-cell signaling, transcription regulation, Wnt-protein binding. Genes hypomethylated in liver tumor tissue in the setting of liver cirrhosis are related to olfactory signal transduction, neuroactive ligand-receptor interaction, keratinization, immune response, inhibition of serine proteases, and zinc metabolism. The genes hypermethylated in the tumor are located at the 7p15.2 locus in the *HOXA* cluster region, and the hypomethylated CpG sites occupy extended regions of the genome in the gene clusters of olfactory receptors (11p15.4), keratin and keratin-associated proteins (12q13.13, 17q21.2, and 21q22.11), epidermal differentiation complex (1q21.3), and immune system function loci 9p21.3 (*IFNA*, *IFNB1*, *IFNW1* cluster) and 19q13.41–19q13.42 (*KLK*, *SIGLEC*, *LILR*, *KIR* clusters). Among the genes of fibrogenesis or DNA repair, cg14143055 (*ADAMDEC1*) is located in the binding region of the HOX gene family transcription factors (TFs), while cg05921699 (*CD79A*), cg06196379 (*TREM1*) and cg10990993 (*MLH1*) are located in the binding region of the ZNF protein family transcription factor (TF). Thus, the DNA methylation profile in the liver in HCV-induced HCC is unique and differs depending on the degree of surrounding tissue lesion – liver fibrosis or liver cirrhosis.

Key words: DNA methylation; chronic hepatitis C; HCV; liver fibrosis; liver cirrhosis; hepatocellular carcinoma.

For citation: Goncharova I.A., Zarubin A.A., Babushkina N.P., Koroleva I.A., Nazarenko M.S. Changes in DNA methylation profile in liver tissue during progression of HCV-induced fibrosis to hepatocellular carcinoma. *Vavilovskii Zhurnal Genetiki i Selektii* = *Vavilov Journal of Genetics and Breeding*. 2023;27(1):72-82. DOI 10.18699/VJGB-23-10

Введение

Злокачественные новообразования печени характеризуются растущим уровнем заболеваемости во всем мире (Philips et al., 2021). Самая высокая заболеваемость и смертность наблюдаются в Восточной Азии и Африке, где ведущей причиной развития гепатоцеллюлярной карциномы (ГЦК) является хронический вирусный гепатит В и неалкогольная жировая болезнь печени (НЖБП). Вместе с тем в развитых странах одной из основных причин развития ГЦК считается хронический вирусный гепатит С (ХВГС), распространенность которого высока в Европе, а максимальные значения приходится на страны Восточной Европы, в том числе Россию (Goossens, Hoshida, 2015; Petruzzello et al., 2016).

Молекулярные механизмы развития ГЦК существенно различаются в зависимости от этиологии заболевания. Так, вирус гепатита В (HBV) может интегрироваться в геном гепатоцита хозяина, что приводит к прямому запуску канцерогенеза посредством активации протонкогенов и/или подавления активности генов-супрессоров опухолей (Levgero, Zucman-Rossi, 2016). В свою очередь вирус гепатита С (HCV), представляющий собой РНК-вирус, имеет ограниченные возможности интеграции в геном клетки печени хозяина и осуществляет свой канцерогенный потенциал путем включения многоэтапного процесса, приводящего через хроническое воспаление печени и прогрессирование фиброза к образованию и развитию опухолевых клонов. Риск развития ГЦК при хронической инфекции HCV напрямую связан с тяжестью фиброза печени, является редким событием при начальных стадиях фиброза и значительно выше у пациентов с циррозом (Khatun et al., 2021).

Среди различных факторов, определяющих подверженность HCV-инфекции и прогрессирование фиброза до ГЦК, немаловажную роль играет генетический и эпи-

генетический компонент. В частности, полногеномные ассоциативные исследования (GWAS) выявили около 140 локусов, из которых 84 отнесены к известным генам, белковые продукты которых вовлечены в ответ на HCV-инфекцию, противовирусную терапию, спонтанный клиренс вируса, развитие осложнений на интерферонотерапию (Kanz et al., 2005).

Идентифицированы гены, в том числе *EXO1*, *VCAN*, *KIT* и *MIR200C*, которые ассоциированы с развитием HCV-индуцированной ГЦК и рассматриваются в качестве потенциальных мишеней для фармакотерапии (Goossens, Hoshida, 2015; Schulze et al., 2015; Chen et al., 2021). Кроме того, показано прогностическое значение микроРНК, определяемых в ткани печени или сыворотке крови, при развитии HCV-индуцированной ГЦК (Aly et al., 2020; Yan et al., 2021).

Экспериментальные исследования aberrаций метилома ткани печени при ее патологии в зависимости от этиологических причин немногочисленны (Neumann et al., 2012; Hlady et al., 2014). Основными в отношении вирусной этиологии являются данные, полученные при сравнительном анализе парных опухолевых и непораженных участков печени у больных азиатского происхождения с ГЦК на платформе Illumina Infinium Human Methylation BeadChip 27k (Shen et al., 2012; Mah et al., 2014; Yamada et al., 2016). Ряд работ связан с реанализом имеющихся результатов по метилированию ДНК с привлечением дополнительных данных, в том числе полученных на микрочипе Illumina Human Methylation 450 BeadChip, из атласа генома рака TCGA (The Cancer Genome Atlas) (Fan et al., 2018; Meng et al., 2018; Wang Y. et al., 2019; Jiang et al., 2020; Zhao et al., 2021).

При сравнении списков дифференциально метилированных CpG-сайтов между анализируемыми участками ткани печени у пациентов с ГЦК в разных исследова-

ниях (Shen et al., 2012; Mah et al., 2014; Yamada et al., 2016) выявляется их существенное сходство. Например, список гиперметилированных в опухолевой ткани генов, представленных в работе (Yamada et al., 2016), на 93 % перекрывается с данными другого коллектива (Mah et al., 2014). Иная картина наблюдается при сравнении результатов реанализов. Так, в списках значимых для развития ГЦК генов, представленных в различных исследованиях, практически отсутствуют общие гены (Fan et al., 2018; Meng et al., 2018; Wang Y. et al., 2019; Jiang et al., 2020). Это объясняется разными критериями, выбранными для реанализа первичных данных, приведенных в репозитории GEO (Edgar et al., 2002; Barrett et al., 2013). Вместе с тем ни в одной из упомянутых работ не учитывалась этиология ГЦК, и в анализируемую группу включались как пациенты с носительством HBV или HCV, так и без носительства вирусов или их сочетаний.

Вклад метилирования ДНК в развитие HCV- и HBV-индуцированной ГЦК рассмотрен в метаанализах, включающих исследования таргетного метилирования генов-мишеней, ассоциированных с заболеваниями печени (Zhang et al., 2019, 2022). Выявлены гены, гиперметилированные в опухолевых тканях печени при ГЦК различной вирусной этиологии, которые, однако, в значительной степени являются общими, что не дает полного представления об особенностях профиля метилирования ДНК при воздействии вирусов гепатита В и С.

Нашим коллективом проводились работы в отношении генетических аспектов ХВГС. В результате установлены ассоциации полиморфизма генов фиброгенеза и репарации ДНК с патологией и патогенетически значимыми признаками, в том числе стадиями фиброза печени (Гончарова и др., 2020). Не исключено, что существуют особенности профиля метилирования ДНК в ткани печени на фоне фиброза и цирроза, индуцированные HCV и обуславливающие ГЦК.

Таким образом, целью настоящего исследования было выявление изменения профиля метилирования ДНК, в том числе в области генов, вовлеченных в фиброгенез или репарацию ДНК, в ткани печени при прогрессировании HCV-инфекции от фиброза печени до ГЦК с помощью реанализа первичных данных, находящихся в репозитории GEO.

Материалы и методы

В открытом доступе GEO имеются данные нескольких исследований, в которых анализировался профиль метилирования ДНК в печени у больных азиатского происхождения с ГЦК, обусловленной вирусными гепатитами В и С, на платформе Illumina Infinium Human Methylation BeadChip 27k (табл. 1). Для европеоидов данных по метилированию ДНК при ГЦК в репозитории GEO нет.

Из данных GSE73003 и GSE37988 для анализа выбраны пациенты, у которых диагностирован ХВГС по наличию суммарных антител к вирусу гепатита С (HCVab+) и отсутствию поверхностного антигена вирусного гепатита В (HBsAg-). Из набора GSE73003 в анализ вошли пациенты с HCV-индуцированной ГЦК, у которых неопухолевая ткань печени характеризовалась различными стадиями фибротического поражения: фиброз печени на фоне ХВГС ($n = 3$) и цирроз печени ($n = 8$). Кроме того, в работу включен один пациент с ГЦК неустановленной этиологии, но без HCVab и HBsAg, у которого окружающая ткань печени определена как нормальная (ГЦК_нормальная ткань/нормальная ткань).

Из массива GSE37988 в анализ включены пациенты с HCV-индуцированной ГЦК, у которых неопухолевая ткань печени находилась на стадии цирроза ($n = 6$). В настоящей работе мы не дифференцировали ткани и не использовали гистологические срезы, а опирались только на данные, представленные в GSE37988 и GSE73003 GEODataSets (<https://www.ncbi.nlm.nih.gov/>).

Таблица 1. Общая характеристика работ, связанных с анализом профиля метилирования ДНК в печени у больных ГЦК, обусловленной вирусными гепатитами В и С, с помощью метилочипа Illumina Infinium Human Methylation BeadChip 27k

Номер в GEO	Популяция	Количество пациентов с ГЦК, ткань печени	Результат	Литературный источник
GSE37988	Тайвань	$n = 62$, парные опухолевые/неопухолевые участки	684 CpG-сайта гиперметилированы и 1640 – гипометилированы в опухоли по сравнению с неопухолевыми тканями ($\Delta\beta \geq 0.20$, $FDR \leq 0.05$). В опухоли гиперметилирование подтверждено для генов <i>CDKL2</i> , <i>STEAP4</i> , <i>HIST1H3G</i> , <i>CDKN2A</i> и <i>ZNF154</i>	Shen et al., 2012
GSE57956	Сингапур	$n = 59$, парные опухолевые/неопухолевые участки	2037 CpG-сайтов гиперметилированы и 2379 – гипометилированы в опухоли по сравнению с неопухолевыми тканями ($\Delta\beta > 0.10$, $FDR < 0.05$). В опухоли гиперметилирование подтверждено для генов <i>SPDY1</i> , <i>TSPYL5</i> , <i>PKDREJ</i> , <i>ZNF154</i> , <i>TUBB6</i> , <i>CYB5R2</i> и <i>SH3YL1</i> , а гипометилирование – для <i>CYB11B1</i> и <i>SPRR3</i>	Mah et al., 2014
GSE73003	Япония	$n = 20$, парные опухолевые/неопухолевые участки	875 CpG-сайтов гиперметилированы и 1795 – гипометилированы в опухоли по сравнению с неопухолевыми тканями ($\Delta\beta > 0.15$, $FDR < 0.01$). Гиперметилирование в опухоли подтверждено для генов <i>APC</i> , <i>CDKN2A</i> , <i>GSTP1</i> , <i>AKR1B1</i> , <i>GRASP</i> , <i>MAP9</i> , <i>NXPЕ3</i> , <i>RSPH9</i> , <i>SPINT2</i> , <i>STEAP4</i> и <i>ZNF154</i>	Yamada et al., 2016

Поскольку в данных GSE57956 не приведена информация по этиологии патологии, в частности инфицированию вирусами гепатитов В и С, образцы тканей не вошли в настоящее исследование.

Кроме 27578 CpG-сайтов, представленных на метило-чипе Illumina Infinium Human Methylation BeadChip 27k, отдельно был оценен статус метилирования генов фиброгенеза и репарации ДНК, которые по результатам наших предыдущих исследований были ассоциированы с ХВГС, стадиями фиброза печени, скоростью прогрессирования фиброза до цирроза печени и коморбидными с ХВГС патологиями (Гончарова и др., 2020).

Статистический анализ данных осуществлен с помощью пакетов lumi, limma в программной среде R (Bioconductor). Коррекция на множественные сравнения выполнена по методу Benjamini–Hochberg (FDR).

В качестве параметра уровня метилирования ДНК использован показатель индекса метилирования β , который соответствует отношению флуоресцентных сигналов метилированных аллелей к сумме флуоресцентных сигналов метилированных и неметилированных аллелей. Индекс метилирования β варьирует от 0 (неметилированное состояние) до 1 (полное метилирование всех CpG-сайтов в данном положении). Дифференциально метилированными считались CpG-сайты с разницей среднего уровня метилирования между группами образцов $FDR < 0.05$ и $|\Delta\beta| \geq 0.2$, что превышает погрешность измерения на микрочипе и дополняет статистическую значимость различий биологически обоснованным критерием.

Функциональная аннотация белковых продуктов генов, включающих дифференциально метилированные CpG-сайты (ДМС), была выполнена с помощью программ Web-based GENE SeT AnaLysis Toolkit с редукцией категорий Weighted set cover (Liao et al., 2019) и Metascape (Zhou et al., 2019). Категории описываемых генов в терминах биологических процессов и молекулярных функций соответствуют классификатору базы данных Gene Ontology (GO), сигнальных и метаболических путей – KEGG и Reactome, лекарственных мишеней – DrugBank, хромосомной локализации – Chromosomal Location.

Дополнительно была проведена геномная аннотация ДМС генов фиброгенеза и репарации ДНК в клеточной линии гепатоцеллюлярной карциномы HepG2 с помощью ресурса UCSC (Kent et al., 2002), которая позволила охарактеризовать CpG-сайты, локализованные в промоторах генов, областях открытого хроматина, доступных для РНК-полимеразы II, сайтах связывания транскрипционных факторов (ТФ) и тем самым, возможно, влияющие на изменение экспрессии генов.

Результаты и обсуждение

Идентификация ДМС и их генов между опухолевой и неопухолевой тканями печени (норма без вирусов гепатита С и В, фиброз и цирроз на фоне ХВГС) у пациентов с ГЦК

При сравнительном анализе уровня метилирования 27578 CpG-сайтов между парными образцами, охарактеризованными как ГЦК, окруженная нормальной тканью, и нормальная ткань печени у пациента без вирусов гепа-

тита С и В (GSE73003), выявлено 32 ДМС, среди которых 24 CpG-сайта (21 ген) гиперметилированы и 8 CpG-сайтов (7 генов) гипометилированы в опухолевой относительно нормальной ткани (рис. 1, а). По два CpG-сайта выявлены в генах *RBM4*, *SOX9* и *SPAG8* (гиперметилированы в опухолевой ткани), а также в *ACTA2* (гипометилированы).

Двадцать CpG-сайтов с наибольшими различиями в уровне метилирования между опухолевой и нормальной тканями печени представлены в Прил. 1¹. Большая их часть расположена в области CpG-островков (16, или 80%). Среди них CpG-сайты генов *RBM4*, *TRIP12*, *BFSP1*, *FBP1*, *SGCE* и *PTPN4*, ранее показавших связь с развитием ГЦК (см. Прил. 1).

Между ГЦК, окруженной фиброзной тканью, и фиброзом при ХВГС (GSE73003) выявлено 40 дифференциально метилированных сайтов (см. рис. 1, б). В опухолевой ткани печени относительно фиброзной гиперметилированы 25 CpG-сайтов (24 гена) и гипометилированы 15 CpG-сайтов (15 генов). Значимые изменения уровня метилирования при онкотрансформации фиброзной ткани печени показаны для CpG-сайтов генов *ZNF154*, *DNM3*, *DLEC1*, *LYPD3*, *DDX49*, *NEFH*, *CCL20* и *NNMT*, которые были ранее связаны с развитием ГЦК (см. Прил. 1). Причем наиболее существенное гиперметилирование в опухоли по сравнению с фиброзом выявлено для двух CpG-сайтов, расположенных в области островка 1-го экзона гена *ZNF154* ($\Delta\beta = 0.593-0.596$, $FDR < 0.01$).

Из всех дифференциально метилированных генов (ДМГ) только белковый продукт *CCL20* является проангиогенным хемокином, индуцируемым HVC, который обеспечивает инвазию и миграцию эндотелиальных клеток при формировании ГЦК (Benkheil et al., 2018). Сайт cg21643045 гена *CCL20*, расположенный в 1-м экзоне, был гипометилирован в опухолевой ткани относительно фиброзной ($\Delta\beta = -0.382$, $FDR = 0.0235$).

При сравнении уровня метилирования ДНК между парными образцами тканей печени (опухолевой и пораженной циррозом) при ХВГС (GSE73003) выявлено 2450 ДМС (см. рис. 1, в). В ткани печени, пораженной опухолью, относительно неопухолевой ткани гиперметилированы 1168 CpG-сайтов (886 генов) и гипометилированы 1282 CpG-сайта (998 генов).

Из двадцати CpG-сайтов генов, показавших наиболее значимые изменения уровня метилирования при онкотрансформации ткани печени, пораженной циррозом, гены *GRM8*, *DNM3*, *DLEC1*, *ZNF154*, *WNK2*, *MFAP5*, *FOXD3*, *NEFH*, *MTNR1B*, *CCL20* и *RAB31* связаны с развитием ГЦК (см. Прил. 1). Более того, cg21790626 гена *ZNF154* и cg21643045 гена *CCL20* были гипер- и гипометилированы соответственно в опухолевой ткани относительно цирротической ($\Delta\beta = 0.598$, $FDR = 3.10 \times 10^{-7}$ и $\Delta\beta = -0.459$, $FDR = 1.43 \times 10^{-6}$).

Сравнительный анализ уровня метилирования 27578 CpG-сайтов между парными образцами опухолевой и неопухолевой ткани печени на фоне HCV-индуцированного цирроза печени (GSE37988) выявил 2304 ДМС (см. рис. 1, з). В опухолевой ткани печени относительно

¹ Приложения 1 и 2 см. по адресу:
<https://sites.icgbio.ru/vogis/download/pict-2023-27/appx1.pdf>

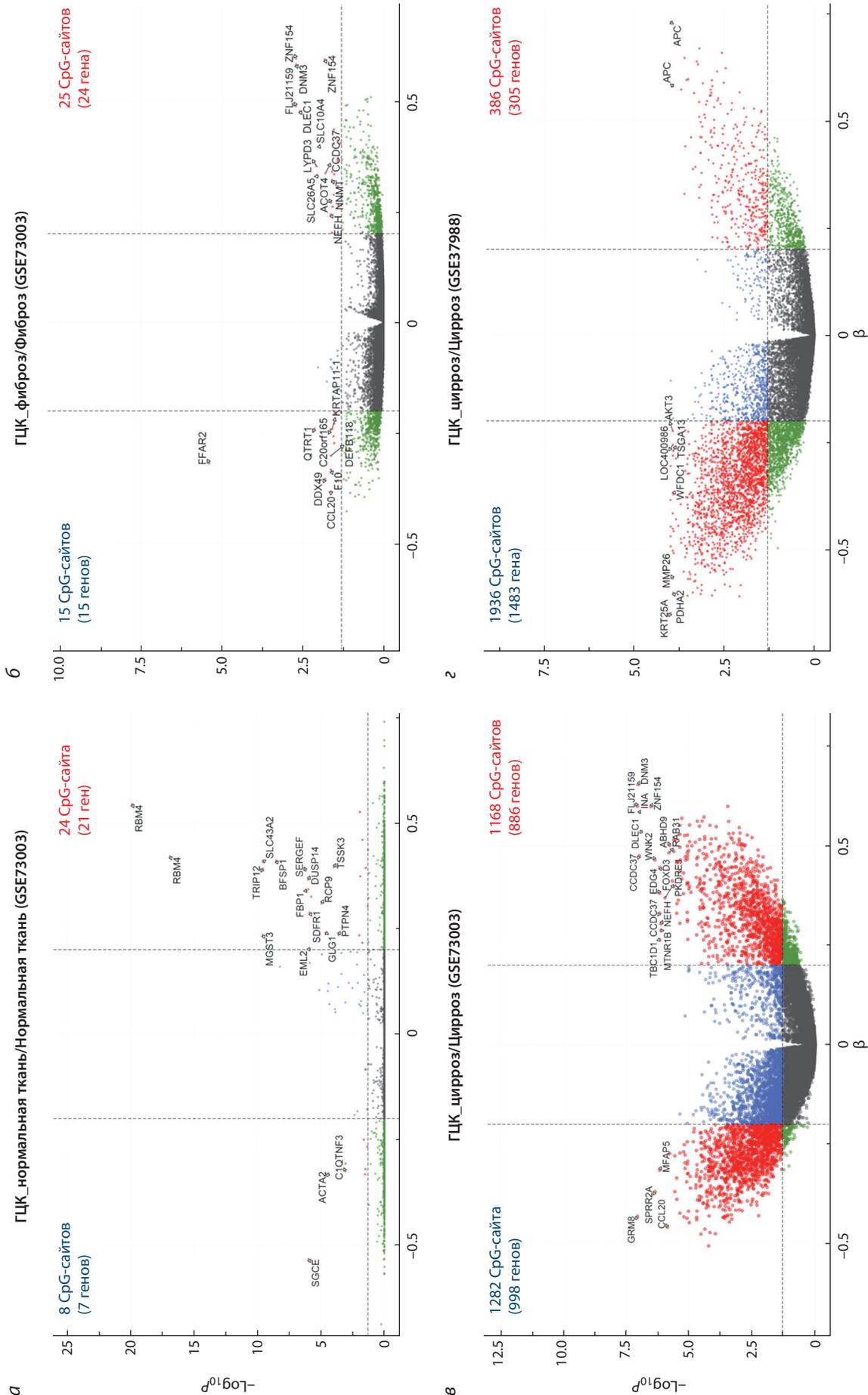


Рис. 1. Дифференциально метилированные CpG-сайты (гены) с различиями в уровне метилирования CpG-сайтов $|\Delta\beta| > 0.2$ и $FDR < 0.05$: а – между ГЦК с прилегающей нормальной тканью печени и нормальной тканью печени без вирусов гепатита С и В (GSE73003); б – между ГЦК с прилегающей фиброзной тканью печени и фиброзом печени при ХВГС (GSE73003); в – между ГЦК с прилегающей тканью печени, пораженной циррозом, и тканью с циррозом печени при ХВГС (GSE73003); г – между ГЦК с прилегающей тканью печени, пораженной циррозом, и тканью с циррозом печени при ХВГС (GSE37988).

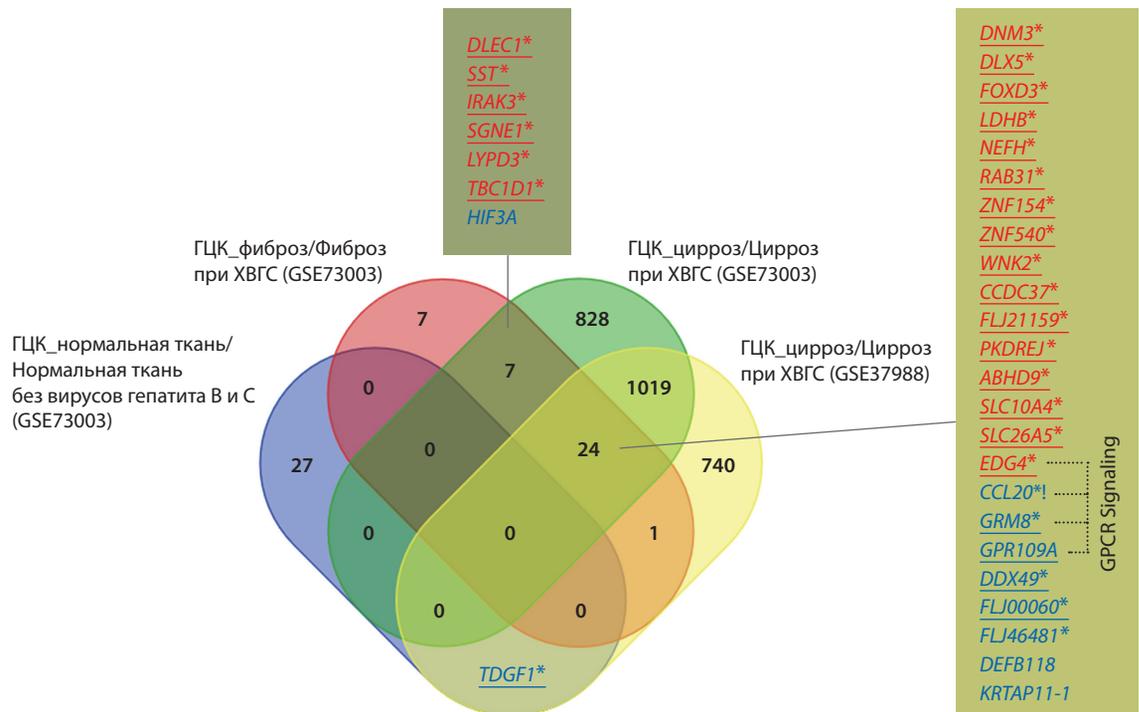


Рис. 2. Венн-диаграмма, отражающая количество общих ДМГ между опухолевой и прилегающей к ней тканями печени различной степени поражения (норма без вирусов С и В, фиброз и цирроз при ХВГС) у пациентов с ГЦК.

Синий/красный цвет – гипо-/гиперметилированные в опухолевой ткани относительно неопухолевой гены; подчеркивание – расположение ДМС в области CpG-островка; *!/ – вовлеченность гена в ГЦК/ГЦК при ХВГС.

пораженной циррозом гиперметилированы 386 CpG-сайтов (305 генов) и гипометилированы 1936 CpG-сайтов (1483 гена).

Гены и CpG-сайты, показавшие наиболее значимые изменения уровня метилирования при онкотрансформации ткани печени, пораженной циррозом по данным GSE37988, представлены в Прил. 1. Среди них гены *MAGEA3*, *APC*, *AKT3*, *MMP26* и *WFDC1* ассоциированы с ГЦК по результатам предыдущих исследований (см. Прил. 1). В отличие от данных GSE73003, меньшая часть CpG-сайтов (7 из 20, или 35 %) располагалась в области островков. Причем из них только два, *cg16970232* и *cg24332422* гена *APC*, были гиперметилированы в опухолевой ткани по сравнению с цирротической ($\Delta\beta = 0.730$, $FDR = 1.0 \times 10^{-4}$ и $\Delta\beta = 0.581$, $FDR = 1.2 \times 10^{-4}$).

Характеристика общих ДМГ между опухолевой и неопухолевой тканями печени (норма без вирусов гепатита С и В, фиброз и цирроз на фоне ХВГС) у пациентов с ГЦК

При сравнении списков генов, содержащих ДМС у пациентов с ГЦК между опухолевой и неопухолевой тканями в зависимости от степени поражения прилегающей к опухоли ткани печени, показано, что *ZNF154*, *DNM3*, *FLJ21159*, *DLEC1*, *CCDC37*, *NEFH*, *CCL20* и *KRTAP11-1* входят в состав топовых с максимальными различиями в уровнях метилирования CpG-сайтов между тканями (см. Прил. 1).

Дифференциально метилированные гены между опухолевыми тканями и фиброзом/циррозом печени (GSE73003) характеризуются наличием семи общих генов, шесть из

которых гиперметилированы в опухоли независимо от степени поражения окружающей ткани (рис. 2). Пять из семи ДМС общих генов локализованы в пределах CpG-островка. Для генов *DLEC1*, *SST*, *IRAK3*, *SGNE1*, *LYPD3* и *TBC1D1* ранее показана связь с развитием ГЦК, причем гены *DLEC1*, *IRAK3* и *SGNE1* были гиперметилированы в опухоли (Qiu et al., 2008; Kuo et al., 2015; Meng et al., 2018), что согласуется с результатами настоящего исследования.

Общими для опухоли при фиброзе и циррозе двух наборов данных (GSE73003 и GSE37988) являются 24 ДМГ, среди которых 16 (66.7 %) гиперметилированы и расположены в области CpG-островка (см. рис. 2). Ранее для 21 гена была показана связь с развитием ГЦК: *DNM3* характеризовался снижением экспрессии, а *FOXD3*, *LDHB*, *NEFH*, *ZNF154*, *FLJ21159*, *PKDREJ*, *ABHD9* и *WNK2* – гиперметилированием в тканях опухоли (Shen et al., 2012; Revill et al., 2013; Liu Z. et al., 2016; Meng et al., 2018; Miller et al., 2021). Гены *CCDC37*, *CCL20*, *DNM3*, *ZNF154* и *ZNF540* перекрываются со списком двадцати ДМГ при ГЦК без учета этиологии (Shen et al., 2012).

Среди восьми гипометилированных в опухоли генов для *CCL20*, *DDX49* и *GRM8* ранее было выявлено повышение экспрессии в сыворотке крови и/или опухолевой ткани у больных ГЦК, в том числе для *CCL20* при ГЦК на фоне ХВГС (Benkheil et al., 2018; Dai et al., 2021; Gao et al., 2022).

Функциональная аннотация общих 24 ДМГ между опухолевой и прилегающей к ней тканью печени различной степени поражения с использованием ресурса Metascape (см. рис. 2) показала связь гиперметилированного (*EDG4*)

и гипометилированных генов (*CCL20*, *GPR109A* и *GRM8*) с процессами, осуществляющими передачу сигналов от рецепторов, связанных с G-белком (Signaling by GPCR: R-HSA-372790). Более того, экспрессия гена *GRM8* в опухолевой ткани отрицательно коррелирует с выживаемостью пациентов с ГЦК, а уровень его метилирования входит в панель генов, важную для прогнозирования заболевания (Gao et al., 2022). Белкам GPCR отводят роль онкомодуляторов, aberrантная экспрессия которых влияет на различные сигнальные пути клеток, нарушая ангиогенез, инвазию, миграцию, метастазирование, иммунный ответ при инициации и прогрессировании ГЦК, что делает их привлекательными терапевтическими молекулярными мишенями (Peng et al., 2018).

В настоящем исследовании выявлено гиперметилирование CpG-сайтов генов *ZNF154* и *ZNF540*, кодирующих белки цинковых пальцев, в опухолевой ткани печени по сравнению с фиброзом и циррозом (см. рис. 2). Некоторые белки данной категории входят в состав сигнатуры прогностических маркеров выживаемости пациентов с HCV-индуцированной ГЦК и являются топовыми гиперметилированными генами при ГЦК различной этиологии (Shen et al., 2012; Wang X. et al., 2021). Анализ экспрессии данных генов в печени показал, что при ГЦК различной этиологии наблюдается репрессия транскрипции многих белков цинковых пальцев *ZNF* (Gonçalves et al., 2022). Вероятно, при ГЦК, индуцированной HCV, гены белков цинковых пальцев, в частности *ZNF154* и *ZNF540*, могут быть перспективными ранними маркерами онкотрансформации начиная с фиброза, а не только на фоне цирроза печени.

Ни один ген из списка 24 общих ДМГ между опухолевой и прилегающей к ней тканями печени при фиброзе и циррозе на фоне ХВГС двух наборов данных (GSE73003 и GSE37988) не входил в перечень известных молекулярных «драйверов» злокачественных новообразований, включая ГЦК (Hlady et al., 2014; Bailey et al., 2018; Cai et al., 2020; Molina-Sánchez et al., 2020; Zhang et al., 2022). Однако такие гены обнаружены среди ДМГ между опухолевой и цирротической тканями. В частности, в опухолевой ткани были гиперметилированы CpG-сайты, входящие в островки промоторов генов *APC*, *CDKN2B*, *GSTP1*, *ELF4* и *TERT*, а разные CpG-сайты гена *WT1* характеризовались разнонаправленным изменением их уровня метилирования.

Функциональная аннотация ДМГ между опухолевой и неопухолевой тканями печени (норма без вирусов гепатита С и В, фиброз и цирроз на фоне ХВГС) у пациентов с ГЦК

По наиболее представленным биологическим путям и основным молекулярным функциям гены гипо- и гиперметилированных CpG-сайтов в опухолевой ткани по сравнению с цирротическими тканями у пациентов с ГЦК при ХВГС сходны между данными GSE73003 и GSE37988 (Прил. 2). Так, для генов, CpG-сайты которых гиперметилированы в опухолевой ткани, наиболее представленными являются биологические процессы, относящиеся к процессам развития (GO:0007399, GO:0009790, GO:0048468, FDR < 2.2×10^{-16}) и передаче межклеточных сигналов (GO:0007267, FDR < 2.2×10^{-16} , см. Прил. 2).

Эти результаты частично согласуются с данными (Shen et al., 2012), где среди наиболее значимых при ГЦК различной этиологии выделяются процессы развития. Гены, содержащие гиперметилированные CpG-сайты при HCV-индуцированной ГЦК, по молекулярным функциям сходны с выявленными при ГЦК различной этиологии (Shen et al., 2012) и включают регуляцию транскрипции и связывание с ДНК (GO:0003700; GO:0140110, GO:0003677, FDR < 0.0002), а также связывание белков Wnt-пути (GO:0017147, FDR = 1.3×10^{-4}).

Гиперметилированные гены локализованы на хромосоме 7 (7p15.2) в регионе кластера *HOXA* (FDR = 2.3×10^{-5} , см. Прил. 2). Ранее, при выявлении сигнатуры метилирования ДНК в ткани печени при ГЦК, было показано, что 39 из 214 CpG-сайтов ассоциированы с изменением экспрессии генов, в том числе локализованных в регионе chr7:27144326–27145664 в непосредственной близости к гомеобоксным факторам транскрипции (*HOXA6*, *HOXA3*, *HOXA5*, *HOXA7* и *HOXA4*), которые участвуют в онкогенезе, пролиферации и миграции клеток (Gonçalves et al., 2022).

Гипометилированные гены при HCV-индуцированной ГЦК относятся в основном к следующим биологическим процессам: иммунный и защитный ответы (GO:0006955, GO:0006952, FDR < 2.2×10^{-16}); сигнальный путь рецепторов, связанных с G-белком (GO:0007186, FDR < 6.0×10^{-10}); дифференцировка эпителиальных клеток (GO:0030855, FDR < 2.2×10^{-16} , см. Прил. 2), что частично согласуется с данными, полученными для ГЦК различной этиологии (Shen et al., 2012).

По молекулярным функциям гипометилированных генов при HCV-индуцированной ГЦК различной этиологии (Shen et al., 2012), с одной стороны, выявляется сходство в отношении ряда категорий, таких как связывание с рецепторами различных антигенов, а с другой – активность ингибиторов пептидазы, в том числе серинового типа, только при HCV-индуцированной карциноме (см. Прил. 2). В настоящее время известно, что ингибитор сериновой протеазы, секретируемый опухолевыми клетками печени (SPINK1 или LC-SPINK), представляет собой белок, содержание которого значительно повышается в сыворотке крови индивидов при ГЦК вирусной этиологии (Lu et al., 2020).

Согласно базам данных KEGG и Reactome, наиболее значимыми молекулярными путями для гипометилированных в опухоли генов были передача обонятельных сигналов (hsa04740, FDR < 2.2×10^{-16}), взаимодействие между цитокиновыми рецепторами (hsa04060, FDR < 7.8×10^{-7}) и нейроактивное взаимодействие лиганда с рецептором (hsa04080, FDR < 0.0007), передача сигналов рецепторов, связанных с G-белком (R-HSA-372790, FDR < 3.5×10^{-6}), кератинизация (R-HSA-6805567, FDR < 2.2×10^{-16}) и иммунная система (R-HSA-168256, FDR < 3.4×10^{-6} , см. Прил. 2). По-видимому, это обусловлено тем, что в опухоли гипометилирование ДНК захватывает протяженные области генома в кластерах генов обонятельных рецепторов (11p15.4), кератина и кератин-ассоциированных белков (12q13.13, 17q21.2 и 21q22.11), комплекса эпидермальной дифференцировки (1q21.3), а также функционирования иммунной системы – локусы 9p21.3 (кластер *IFNA*, *IFNB1*,

IFNW1) и 19q13.41–19q13.42 (кластеры *LILR*, *KIR*, *KLK*, *ZNF*, *SIGLEC*, см. Прил. 2).

Нарушение эпигенетической регуляции иммунной системы является общей чертой при раке различной локализации (Berglund et al., 2021). Передача обонятельных сигналов и нейрорактивное взаимодействие лиганда с рецептором входят в путь передачи сигналов рецепторов, связанных с G-белком, обогащение которого тоже часто встречается при злокачественных новообразованиях (Wei et al., 2012). Эктопическая экспрессия генов обонятельных рецепторов, связанная в том числе с эпигенетическими механизмами, по-видимому, обеспечивает инвазивность и метастазирование опухолевых клеток на поздних стадиях процесса (Fessahaye et al., 2021). Нарушение процесса кератинизации – реже регистрируемое событие в опухоли, для которого показана связь с уровнем гидроксиметилирования ДНК при раке головы и шеи в зависимости от носительства вируса папилломы человека (Liu S. et al., 2020), а также обогащение гипометилированных генов при раке молочной железы (Holm et al., 2016).

База данных DrugBank указывает на то, что гипометилированные при HCV-индуцированной ГЦК гены участвуют в метаболизме цинка (DB01593), добавление которого в пищу может быть рекомендовано для снижения риска ГЦК после эрадикации HCV противовирусными препаратами прямого действия (Hosui et al., 2021). Не исключено, что существует связь между дефицитом цинка и гипометилированием ДНК в отдельных генах (Azimi et al., 2022).

Профиль метилирования генов фиброгенеза и репарации ДНК

Среди генов, белковые продукты которых участвуют в процессах фиброгенеза или репарации ДНК, из разряда ранее показавших ассоциации с заболеваниями печени в исследованиях нашей научной группы выявлено десять дифференциально метилированных CpG-сайтов (табл. 2). Сайты cg03876618 гена *IGFBP7* и cg14323109 гена *KDR*,

расположенные в области CpG-островков, были гиперметилированы в опухоли по сравнению с окружающей цирротически измененной тканью. CpG-сайты генов *ADAMDEC1*, *CD79A*, *MMP3* и *TREM1* дифференциально метилированы по данным GSE37988 и GSE73003 (см. табл. 2).

Проведение геномной аннотации с помощью ресурса UCSC показало, что в клеточной линии гепатоцеллюлярной карциномы HepG2 в состав активного промотора входят cg01053621 (*APOA2*) и cg10990993 (*MLH1*), а в регионах связывания с субъединицей А РНК-полимеразы II локализованы cg03876618 (*IGFBP7*), cg14323109 (*KDR*), cg16466334 (*MMP3*), cg06196379 (*TREM1*) и cg01053621 (*APOA2*).

Сайт cg14143055 гена *ADAMDEC1*, гипометилированный в опухолевой ткани, локализован в регионе связывания транскрипционных факторов семейства HOX, которые играют немаловажную роль в онкогенезе различных опухолей, в том числе ГЦК (Gonçalves et al., 2022). Вместе с тем HCV инфекция и экспрессия корового белка вируса запускают активацию генов HOX (Kasai et al., 2021), что может быть одним из факторов развития HCV-индуцированной ГЦК.

Гипометилированные в опухолевой ткани при HCV-индуцированной ГЦК cg05921699 (*CD79A*), cg06196379 (*TREM1*), cg10990993 (*MLH1*) локализованы в области связывания TF семейства белков цинковых пальцев ZNF, которые кроме регуляции транскрипции осуществляют индукцию белок-белковых взаимодействий, посттранскрипционную регуляцию, липидный обмен, иммунные реакции и влияют на развитие многих форм рака, включая ГЦК (Li et al., 2022).

В заключение следует отметить, что проведенное исследование имеет ограничение в связи с малочисленностью выборок больных с нормальной и фибротической окружающей опухоль тканями при ХВГС, так как в большинстве случаев ГЦК развивается на фоне пораженной циррозом

Таблица 2. ДМС генов, участвующих в фиброгенезе и репарации ДНК, между опухолевой и тканью печени, пораженной циррозом у пациентов с ГЦК

CpG-сайт	Ген	Расстояние до TSS/локализация в CpG-островке	ГЦК/Цирроз			
			$\beta \pm SD$ (GSE73003)	FDR	$\beta \pm SD$ (GSE37988)	FDR
cg14143055	<i>ADAMDEC1</i>	1374/нет	0.46 ± 0.15/0.70 ± 0.04	0.0106	0.34 ± 0.16/0.73 ± 0.05	0.0038
cg05921699	<i>CD79A</i>	477/нет	0.52 ± 0.17/0.74 ± 0.03	0.0136	0.45 ± 0.20/0.74 ± 0.05	0.0469
cg16466334	<i>MMP3</i>	16/нет	0.35 ± 0.16/0.66 ± 0.02	0.0029	0.37 ± 0.22/0.71 ± 0.05	0.0300
cg06196379	<i>TREM1</i>	428/нет	0.21 ± 0.07/0.44 ± 0.02	0.0004	0.13 ± 0.12/0.36 ± 0.05	0.0139
cg03876618	<i>IGFBP7</i>	505/да	0.55 ± 0.22/0.19 ± 0.02	0.0011	–	–
cg14323109	<i>KDR</i>	181/да	0.34 ± 0.20/0.08 ± 0.02	0.0124	–	–
cg10990993	<i>MLH1</i>	1347/нет	0.19 ± 0.04/0.42 ± 0.09	0.0036	–	–
cg01053621	<i>APOA2</i>	573/нет	–	–	0.18 ± 0.12/0.47 ± 0.09	0.0119
cg06531741	<i>HTR3B</i>	139/нет	–	–	0.45 ± 0.24/0.79 ± 0.03	0.0437
cg03017475	<i>TAS2R38</i>	852/нет	–	–	0.25 ± 0.11/0.55 ± 0.07	0.0001

Примечание. TSS – точка инициации транскрипции; β – уровень метилирования; SD – стандартное отклонение. Полу жирным шрифтом выделены ДМС/ДМГ, гиперметилированные в опухолевой ткани по сравнению с пораженной циррозом тканью печени. Прочерк – CpG-сайты не являются дифференциально метилированными.

ткани и другие случаи наблюдаются значительно реже. В работе не была учтена внутриопухолевая и клеточная гетерогенность тканей, которая тесно связана с профилем метилирования ДНК (Hlady et al., 2017). Принимая во внимание тот факт, что фокус работы заключался в анализе профиля метилирования ДНК печени при HCV-индуцированной ГЦК, сложно однозначно выделить CpG-сайты, специфичные для данной патологии. Методическим ограничением является невозможность разграничения метилирования ДНК и гидроксиметилирования ДНК, поскольку она подвергается бисульфитной модификации перед гибридизацией на метилочипе.

Заключение

Сравнительный анализ профиля метилирования ДНК в печени у пациентов с ГЦК между опухолью и неопухолевыми тканями с различной степенью поражения (нормальная ткань, HCV-индуцированный фиброз, HCV-индуцированный цирроз) показал значительно меньшее число ДМС между ГЦК и нормальной тканью без вирусов гепатита С и В/фиброзом печени при ХВГС (32 и 40), чем между ГЦК и циррозом печени при ХВГС в данных GSE73003 и GSE37988 (2450 и 2304 соответственно).

Исходя из того, что уровень фиброза коррелирует с функцией печени, а цирроз представляет собой основной фактор риска развития ГЦК (Roehlen et al., 2020), можно было бы ожидать, что нормальная и фибротическая ткани печени будут максимально удалены по эпигенетическому профилю от ГЦК, и по мере прогрессирования фиброза до цирроза печени количество ДМС между опухолевыми и окружающей тканями будет сокращаться. Вместе с тем наблюдается обратная картина: чем значительнее поражение окружающей опухоль ткани печени, тем большие различия в уровне метилирования ДНК наблюдаются между ними. Возможно, нормальная ткань печени или ткань с минимальным фибротическим поражением способствует сдерживанию функциональной разбалансировки генома опухоли, приводя к минимальным отличиям профиля метилирования ДНК между данными тканями. Это предположение косвенно подтверждается тем, что изменение уровня метилирования «драйверных» для ГЦК генов регистрируется на фоне цирроза, но не фиброза.

По мере прогрессирования патологического изменения окружающей опухоль ткани печени уменьшается соотношение количества гипер-/гипометилированных ДМС в опухоли. Так, у пациентов с ГЦК в опухолевой ткани по сравнению с нормальной гиперметилированы 24 CpG-сайта, или 75 % от всех ДМС. По сравнению с тканью печени, пораженной фиброзом при ХВГС, в опухолевой ткани гиперметилированы 25 из 40 ДМС, что составляет 62.5 %. В том случае, когда окружающая опухоль ткань представляет собой цирроз печени, количество гиперметилированных CpG-сайтов в опухолевой ткани относительно группы сравнения составляет 47.7 и 16 % (GSE73003 и GSE37988 соответственно). В предыдущих исследованиях также было выявлено преобладание количества гипометилированных CpG-сайтов протяженных участков генома, в том числе в области генов и межгенных регионов, в опухоли при ГЦК относительно окружающей цирротической ткани печени (Shen et al., 2012; Hlady et al., 2014; Yamada

et al., 2016; Yan et al., 2021). В настоящей работе впервые показано, что у пациентов с ГЦК опухоль на фоне неизменной ткани печени и при фиброзе печени при ХВГС характеризуется большей долей гиперметилированных CpG-сайтов, а количество гипометилированных сайтов возрастает в опухолевой ткани при циррозе.

В исследованиях, посвященных изучению профиля метилирования генов в печени при ГЦК, основное внимание уделяется гиперметилированным генам, включая гены семейств ZNF и HOX, среди которых проводится поиск значимых для развития заболевания маркеров. Вместе с тем сравнительный анализ показал, что при HCV-индуцированной ГЦК в опухолевой ткани наблюдается большее число гиперметилированных CpG-сайтов только по сравнению с окружающей тканью с признаками фиброза. В случае же, когда окружающая опухоль ткань представляет собой цирроз печени, большинство локусов в опухолевой ткани гипометилированы, что, по-видимому, является поздним событием, происходящим при переходе от фибротического поражения тканей печени к злокачественной трансформации.

В связи с этим при HCV-индуцированной ГЦК необходимо уделять также внимание гипометилированным локусам, которые, как показали результаты данного исследования, относятся к белкам GPCR (*CCL20*, *GPR109A* и *GRM8*), локализованы в сайтах связывания ТФ HOX (*ADAMDECI*), ZNF (*CD79A*, *MLH1*) или в области генов ингибиторов сериновых протеаз, одна из которых – SPINK1 – в настоящее время рассматривается как маркер, способный выявить ГЦК вирусной этиологии на ранней стадии. Кроме того, в нашей работе гипометилированные ДМС были локализованы в генах, связанных с метаболизмом цинка, который, как известно, играет роль в патогенезе многих заболеваний, включая ГЦК.

Таким образом, при проведении исследований по оценке профиля метилирования ДНК в печени при ГЦК необходимо учитывать функциональное состояние и степень поражения окружающей опухоль ткани, так как спектр ДМГ значительно различается между парами опухоль/неопухоль ткани в зависимости от того, является ли она условно нормальной, с признаками фиброза или циррозом. Для выявления прогностических маркеров ГЦК, в том числе при жидкостной биопсии, следует учитывать этиологию заболевания, поскольку спектр ДМС и ДМГ HCV-индуцированной ГЦК лишь частично перекрывается с таковыми, выявленными при анализе данной патологии другой природы.

Список литературы / References

- Гончарова И.А., Назаренко М.С., Белобородова Е.В., Марков А.В., Пузырев В.П. Генетическая структура предрасположенности прогрессирования фиброза печени различной этиологии. *Мед. генетика*. 2020;19(8):83-84. DOI 10.25557/2073-7998.2020.08.83-84. [Goncharova I.A., Nazarenko M.S., Beloborodova E.V., Markov A.V., Puzyrev V.P. Genetic structure of a predisposition to the progression of liver fibrosis of various etiologies. *Meditinskaya Genetika = Medical Genetics*. 2020;19(8):83-84. DOI 10.25557/2073-7998.2020.08.83-84. (in Russian)]
- Aly D.M., Gohar N.A.-H., El-Hady A., Khairy M., Abdullatif M.M. Serum microRNA let-7a-1/let-7d/let-7f and miRNA 143/145 gene expression profiles as potential biomarkers in HCV induced hepato-

- cellular carcinoma. *Asian Pac. J. Cancer Prev.* 2020;21(2):555-562. DOI 10.31557/APJCP.2020.21.2.555.
- Azimi Z., Isa M.R., Khan J., Wang S.M., Ismail Z. Association of zinc level with DNA methylation and its consequences: a systematic review. *Heliyon.* 2022;8(10):e10815. DOI 10.1016/j.heliyon.2022.e10815.
- Bailey M.H., Tokheim C., Porta-Pardo E., Sengupta S., Bertrand D., Weerasinghe A., Colaprico A., Wendl M.C., Kim J., Reardon B., Ng P.K., Jeong K.J., Cao S., Wang Z., Gao J., Gao Q., Wang F., Liu E.M., Mularoni L., Rubio-Perez C., Nagarajan N., Cortés-Ciriano I., Zhou D.C., Liang W.W., Hess J.M., Yellapantula V.D., Tamborero D., Gonzalez-Perez A., Suphavilai C., Ko J.Y., Khurana E., Park P.J., Van Allen E.M., Liang H., MC3 Working Group, Cancer Genome Atlas Research Network, Lawrence M.S., Godzik A., Lopez-Bigas N., Stuart J., Wheeler D., Getz G., Chen K., Lazar A.J., Mills G.B., Karchin R., Ding L. Comprehensive characterization of cancer driver genes and mutations. *Cell.* 2018;173(2):371-385.e18. DOI 10.1016/j.cell.2018.02.060.
- Barrett T., Wilhite S.E., Ledoux P., Evangelista C., Kim I.F., Tomashevsky M., Marshall K.A., Phillippy K.H., Sherman P.M., Holtko M., Yefanov A., Lee H., Zhang N., Robertson C.L., Serova N., Davis S., Soboleva A. NCBI GEO: archive for functional genomics data sets – update. *Nucleic Acids Res.* 2013;41(D1):D991-D995. DOI 10.1093/nar/gks1193.
- Benkheil M., Van Haele M., Roskams T., Laporte M., Noppen S., Abbasi K., Delang L., Neyts J., Liekens S. CCL20, a direct-acting pro-angiogenic chemokine induced by hepatitis C virus (HCV): potential role in HCV-related liver cancer. *Exp. Cell Res.* 2018;372(2):168-177. DOI 10.1016/j.yexcr.2018.09.023.
- Berglund A., Putney R.M., Hamaidi I., Kim S. Epigenetic dysregulation of immune-related pathways in cancer: bioinformatics tools and visualization. *Exp. Mol. Med.* 2021;53(5):761-771. DOI 10.1038/s12276-021-00612-z.
- Cai Y., Tian Y., Wang J., Wei W., Tang Q., Lu L., Luo Z., Li W., Lu Y., Pu J., Yang Z. Identification of driver genes regulating the T-cell – infiltrating levels in hepatocellular carcinoma. *Front. Genet.* 2020;11:560546. DOI 10.3389/fgene.2020.560546.
- Chen G., Zhang W., Ben Y. Identification of key regulators of hepatitis C virus-induced hepatocellular carcinoma by integrating whole-genome and transcriptome sequencing data. *Front. Genet.* 2021;12:741608. DOI 10.3389/fgene.2021.741608.
- Dai H., Feng J., Nan Z., Wei L., Lin F., Jin R., Zhang S., Wang X., Pan L. Morphine may act via DDX49 to inhibit hepatocellular carcinoma cell growth. *Aging (Albany NY).* 2021;13(9):12766-12779. DOI 10.18632/aging.202946.
- Edgar R., Domrachev M., Lash A.E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* 2002;30(1):207-210. DOI 10.1093/nar/30.1.207.
- Fan G., Tu Y., Chen C., Sun H., Wan C., Cai X. DNA methylation biomarkers for hepatocellular carcinoma. *Cancer Cell Int.* 2018;18:140. DOI 10.1186/s12935-018-0629-5.
- Fessahaye G., Khalid R., Mohamed K.-H., Alsidqi M., Khalid A., Mohamed H., Ibrahim M. Breast cancer and smell: hints from epigenetic and functional alteration of the olfaction. (Preprint). 2021. DOI 10.21203/rs.3.rs-910345/v1.
- Gao Y., Liu J., Zhao D., Diao G. A novel prognostic model for identifying the risk of hepatocellular carcinoma based on angiogenesis factors. *Front. Genet.* 2022;13:857215. DOI 10.3389/fgene.2022.857215.
- Gonçalves E., Gonçalves-Reis M., Pereira-Leal J.B., Cardoso J. DNA methylation fingerprint of hepatocellular carcinoma from tissue and liquid biopsies. *Sci. Rep.* 2022;12(1):11512. DOI 10.1038/s41598-022-15058-0.
- Goossens N., Hoshida Y. Hepatitis C virus-induced hepatocellular carcinoma. *Clin. Mol. Hepatol.* 2015;21(2):105-114. DOI 10.3350/cmh.2015.21.2.105.
- Hlady R.A., Tiedemann R.L., Puszyk W., Zendejas I., Roberts L.R., Choi J.H., Liu C., Robertson K.D. Epigenetic signatures of alcohol abuse and hepatitis infection during human hepatocarcinogenesis. *Oncotarget.* 2014;5(19):9425-9443. DOI 10.18632/oncotarget.2444.
- Hlady R.A., Zhou D., Puszyk W., Roberts L.R., Liu C., Robertson K.D. Initiation of aberrant DNA methylation patterns and heterogeneity in precancerous lesions of human hepatocellular cancer. *Epigenetics.* 2017;12(3):215-225. DOI 10.1080/15592294.2016.1277297.
- Holm K., Staaf J., Lauss M., Aine M., Lindgren D., Bendahl P.-O., Vallon-Christersson J., Barkardottir R.B., Höglund M., Borg Å., Jönsson G., Ringnér M. An integrated genomics analysis of epigenetic subtypes in human breast tumors links DNA methylation patterns to chromatin states in normal mammary cells. *Breast Cancer Res.* 2016;18(1):27. DOI 10.1186/s13058-016-0685-5.
- Hosui A., Tanimoto T., Okahara T., Ashida M., Ohnishi K., Wakahara Y., Kusumoto Y., Yamaguchi T., Sueyoshi Y., Hirao M., Yamada T., Hiramatsu N. Oral zinc supplementation decreases the risk of HCC development in patients with HCV eradicated by DAA. *Hepatol. Commun.* 2021;5(12):2001-2008. DOI 10.1002/hep4.1782.
- Jiang H.Y., Ning G., Wang Y.S., Lv W.-B. 14-CpG-based signature improves the prognosis prediction of hepatocellular carcinoma patients. *BioMed Res. Int.* 2020;2020:9762067. DOI 10.1155/2020/9762067.
- Kanz C., Aldebert P., Althorpe N., Baker W., Baldwin A., Bates K., Browne P., van den Broek A., Castro M., Cochrane G., Duggan K., Eberhardt R., Faruque N., Gamble J., Diez F.G., Harte N., Kulikova T., Lin Q., Lombard V., Lopez R., Mancuso R., McHale M., Nardone F., Silventoinen V., Sobhany S., Stoehr P., Tuli M.A., Tzouvara K., Vaughan R., Wu D., Zhu W., Apweiler R. The EMBL nucleotide sequence database. *Nucleic Acids Res.* 2005;33:D29-D33. DOI 10.1093/nar/gki098.
- Kasai H., Mochizuki K., Tanaka T., Yamashita A., Matsuura Y., Morishi K. Induction of HOX genes by hepatitis C virus infection via impairment of histone H2A monoubiquitination. *J. Virol.* 2021;95(6):e01784-20. DOI 10.1128/JVI.01784-20.
- Kent W.J., Sugnet C.W., Furey T.S., Roskin K.M., Pringle T.H., Zahler A.M., Haussler D. The human genome browser at UCSC. *Genome Res.* 2002;12(6):996-1006. DOI 10.1101/gr.229102.
- Khatun M., Ray R., Ray R.B. Hepatitis C virus associated hepatocellular carcinoma. *Adv. Cancer Res.* 2021;149:103-142. DOI 10.1016/bs.acr.2020.10.003.
- Kuo C.C., Shih Y.L., Su H.Y., Yan M.D., Hsieh C.B., Liu C.Y., Huang W.T., Yu M.H., Lin Y.W. Methylation of *IRAK3* is a novel prognostic marker in hepatocellular carcinoma. *World J. Gastroenterol.* 2015;21(13):3960-3969. DOI 10.3748/wjg.v21.i13.3960.
- Levrero M., Zucman-Rossi J. Mechanisms of HBV-induced hepatocellular carcinoma. *J. Hepatol.* 2016;64(1 Suppl.):S84-S101. DOI 10.1016/j.jhep.2016.02.021.
- Li X., Han M., Zhang H., Liu F., Pan Y., Zhu J., Liao Z., Chen X., Zhang B. Structures and biological functions of zinc finger proteins and their roles in hepatocellular carcinoma. *Biomark. Res.* 2022;10(1):2. DOI 10.1186/s40364-021-00345-1.
- Liao Y., Wang J., Jaehnig E.J., Shi Z., Zhang B. WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Res.* 2019;47:W199-W205. DOI 10.1093/nar/gkz401.
- Liu S., de Medeiros M.C., Fernandez E.M., Zarins K.R., Cavalcante R.G., Qin T., Wolf G.T., Figueroa M.E., D’Silva N.J., Rozek L.S., Sartor M.A. 5-Hydroxymethylation highlights the heterogeneity in keratinization and cell junctions in head and neck cancers. *Clin. Epigenetics.* 2020;12(1):175. DOI 10.1186/s13148-020-00965-8.
- Liu Z., Yan H., Zhang J. Blood DNA methylation markers in potentially identified Chinese patients with hepatocellular carcinoma. *Pak. J. Pharm. Sci.* 2016;29(4 Suppl.):1451-1456.
- Lu F., Shah P.A., Rao A., Gifford-Hollingsworth C., Chen A., Trey G., Soryal M., Talat A., Aslam A., Nasir B., Choudhry S., Ishtiaq R., Sanoff H., Conteh L.F., Noonan A., Hu K.Q., Schmidt C., Fu M., Civan J., Xiao G., Lau D.T., Lu X. Liver cancer-specific serine protease inhibitor kazal is a potentially novel biomarker for the early detection of hepatocellular carcinoma. *Clin. Transl. Gastroenterol.* 2020;11(12):e00271. DOI 10.14309/ctg.0000000000000271.

- Mah W.C., Thurnherr T., Chow P.K., Chung A.Y., Ooi L.L., Toh H.C., Teh B.T., Sauntharajah Y., Lee C.G. Methylation profiles reveal distinct subgroup of hepatocellular carcinoma patients with poor prognosis. *PLoS One*. 2014;9(8):e104158. DOI 10.1371/journal.pone.0104158.
- Meng C., Shen X., Jiang W. Potential biomarkers of HCC based on gene expression and DNA methylation profiles. *Oncol. Lett.* 2018; 16(3):3183-3192. DOI 10.3892/ol.2018.9020.
- Miller B.F., Petrykowska H.M., Elnitski L. Assessing *ZNF154* methylation in patient plasma as a multicancer marker in liquid biopsies from colon, liver, ovarian and pancreatic cancer patients. *Sci. Rep.* 2021;11(1):221. DOI 10.1038/s41598-020-80345-7.
- Molina-Sánchez P., Ruiz de Galarreta M., Yao M.A., Lindblad K.E., Bresnahan E., Bitterman E., Martin T.C., Rubenstein T., Nie K., Golas J., Choudhary S., Bárcena-Varela M., Elmas A., Miguela V., Ding Y., Kan Z., Grinspan L.T., Huang K.L., Parsons R.E., Shields D.J., Rollins R.A., Lujambio A. Cooperation between distinct cancer driver genes underlies intertumor heterogeneity in hepatocellular carcinoma. *Gastroenterology*. 2020;159(6):2203-2220.e14. DOI 10.1053/j.gastro.2020.08.015.
- Neumann O., Kesselmeier M., Geffers R., Pellegrino R., Radlwimmer B., Hoffmann K., Ehemann V., Schemmer P., Schirmacher P., Bermejo J.L., Longerich T. Methylome analysis and integrative profiling of human HCCs identify novel protumorigenic factors. *Hepatology*. 2012;56(5):1817-1827. DOI 10.1002/hep.25870.
- Peng W.T., Sun W.Y., Li X.R., Sun J.C., Du J.J., Wei W. Emerging roles of G protein-coupled receptors in hepatocellular carcinoma. *Int. J. Mol. Sci.* 2018;19(5):1366. DOI 10.3390/ijms19051366.
- Petruzzello A., Marigliano S., Loquercio G., Cozzolino A., Cacciapuoti C. Global epidemiology of hepatitis C virus infection: an up-date of the distribution and circulation of hepatitis C virus genotypes. *World J. Gastroenterol.* 2016;22(34):7824-7840. DOI 10.3748/wjg.v22.i34.7824.
- Philips C.A., Rajesh S., Nair D.C., Ahamed R., Abduljaleel J.K., Augustine P. Hepatocellular carcinoma in 2021: an exhaustive update. *Cureus*. 2021;13(11):e19274. DOI 10.7759/cureus.19274.
- Qiu G.H., Salto-Tellez M., Ross J.A., Yeo W., Cui Y., Wheelhouse N., Chen G.G., Harrison D., Lai P., Tao Q., Hooi S.C. The tumor suppressor gene *DLEC1* is frequently silenced by DNA methylation in hepatocellular carcinoma and induces G1 arrest in cell cycle. *J. Hepatol.* 2008;48(3):433-441. DOI 10.1016/j.jhep.2007.11.015.
- Revell K., Wang T., Lachenmayer A., Kojima K., Harrington A., Li J., Hoshida Y., Llovet J.M., Powers S. Genome-wide methylation analysis and epigenetic unmasking identify tumor suppressor genes in hepatocellular carcinoma. *Gastroenterology*. 2013;145(6):1424-1435.e25. DOI 10.1053/j.gastro.2013.08.055.
- Roehlen N., Crouch E., Baumert T.F. Liver fibrosis: mechanistic concepts and therapeutic perspectives. *Cells*. 2020;9(4):875. DOI 10.3390/cells9040875.
- Schulze K., Imbeaud S., Letouzé E., Alexandrov L.B., Calderaro J., Rebouissou S., Couchy G., Meiller C., Shinde J., Soysouvanh F., Calatayud A.L., Pinyol R., Pelletier L., Balabaud C., Laurent A., Blanc J.F., Mazzaferro V., Calvo F., Villanueva A., Nault J.C., Bioulac-Sage P., Stratton M.R., Llovet J.M., Zucman-Rossi J. Exome sequencing of hepatocellular carcinomas identifies new mutational signatures and potential therapeutic targets. *Nat. Genet.* 2015;47(5):505-511. DOI 10.1038/ng.3252.
- Shen J., Wang S., Zhang Y.J., Kappil M., Wu H.C., Kibriya M.G., Wang Q., Jasmine F., Ahsan H., Lee P.H., Yu M.W., Chen C.J., Santella R.M. Genome-wide DNA methylation profiles in hepatocellular carcinoma. *Hepatology*. 2012;55(6):1799-1808. DOI 10.1002/hep.25569.
- Wang X., Xing Z., Xu H., Yang H., Xing T. Development and validation of epithelial mesenchymal transition-related prognostic model for hepatocellular carcinoma. *Aging (Albany NY)*. 2021;13(10):13822-13845. DOI 10.18632/aging.202976.
- Wang Y., Ruan Z., Yu S., Tian T., Liang X., Jing L., Li W., Wang X., Xiang L., Claret F.X., Nan K., Guo H. A four-methylated mRNA signature-based risk score system predicts survival in patients with hepatocellular carcinoma. *Aging (Albany NY)*. 2019;11(1):160-173. DOI 10.18632/aging.101738.
- Wei P., Tang H., Li D. Insights into pancreatic cancer etiology from pathway analysis of genome-wide association study data. *PLoS One*. 2012;7(10):e46887. DOI 10.1371/journal.pone.0046887.
- Yamada N., Yasui K., Dohi O., Gen Y., Tomie A., Kitaichi T., Iwai N., Mitsuyoshi H., Sumida Y., Moriguchi M., Yamaguchi K., Nishikawa T., Umemura A., Naito Y., Tanaka S., Arai S., Itoh Y. Genome-wide DNA methylation analysis in hepatocellular carcinoma. *Oncol. Rep.* 2016;35(4):2228-2236. DOI 10.3892/or.2016.4619.
- Yan P., Pang P., Hu X., Wang A., Zhang H., Ma Y., Zhang K., Ye Y., Zhou B., Mao J. Specific miRNAs in naïve T cells associated with hepatitis C virus-induced hepatocellular carcinoma. *J. Cancer*. 2021;12(1):1-9. DOI 10.7150/jca.49594.
- Zhang C., Huang C., Sui X., Zhong X., Yang W., Hu X., Li Y. Association between gene methylation and HBV infection in hepatocellular carcinoma: a meta-analysis. *J. Cancer*. 2019;10(25):6457-6465. DOI 10.7150/jca.33005.
- Zhang C., Zhang W., Yuan Z., Yang W., Hu X., Duan S., Wei Q. Contribution of DNA methylation to the risk of hepatitis C virus-associated hepatocellular carcinoma: a meta-analysis. *Pathol. Res. Pract.* 2022;238:154136. DOI 10.1016/j.prp.2022.154136.
- Zhao P., Malik S., Xing S. Epigenetic mechanisms involved in HCV-induced hepatocellular carcinoma (HCC). *Front. Oncol.* 2021;11:677926. DOI 10.3389/fonc.2021.677926.
- Zhou Y., Zhou B., Pache L., Chang M., Khodabakhshi A.H., Tanaseichuk O., Benner C., Chanda S.K. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* 2019;10:1523. DOI 10.1038/s41467-019-09234-6.

ORCID ID

I.A. Goncharova orcid.org/0000-0002-9527-7015
A.A. Zarubin orcid.org/0000-0001-6568-6339
N.P. Babushkina orcid.org/0000-0001-6133-8986
I.A. Koroleva orcid.org/0000-0003-1498-6934
M.S. Nazarenko orcid.org/0000-0002-0673-4094

Благодарности. Работа выполнена при финансировании госзадания Министерства науки и высшего образования, № 122020300041-7.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию 18.10.2022. После доработки 08.12.2022. Принята к публикации 15.12.2022.

Перевод на английский язык <https://vavilov.elpub.ru/jour>

Влияние предварительной обработки образцов периферической крови человека на качество Hi-C библиотек

М.М. Гридина¹✉, Э. Весна¹, М.Е. Миньженкова², Н.В. Шилова², О.П. Рыжкова², Л.П. Назаренко³, Е.О. Беляева³, И.Н. Лебедев³, В.С. Фишман¹

¹ Федеральное исследовательское учреждение Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия

² Медико-генетический научный центр им. академика Н.П. Бочкова, Москва, Россия

³ Томский национальный исследовательский медицинский центр Российской академии наук, Томск, Россия

✉ gridinam@gmail.com

Аннотация. Метод захвата конформации хроматина в его полногеномном варианте (Hi-C) – мощный инструмент не только для выявления закономерностей пространственной организации генома, но и для понимания влияния их нарушения на развитие заболеваний. Кроме того, метод может быть использован для детекции хромосомных перестроек, в том числе сбалансированных транслокаций и инверсий. Применение метода Hi-C для поиска хромосомных перестроек получает все более широкое распространение. Это связано с тем, что современные высокопроизводительные методы анализа генома позволяют эффективно детектировать точечные мутации и несбалансированные хромосомные перестройки. Однако чувствительность этих методов для определения сбалансированных транслокаций и инверсий остается достаточно низкой. Хранение образцов цельной крови может влиять на количество и целостность выделяемой из них геномной ДНК, а кроме того, приводит к искажению результатов последующих анализов в том случае, если хранение осуществлялось в ненадлежащих условиях. Метод Hi-C крайне требователен к исходному материалу, так как необходимым условием для его успешного применения и получения качественных данных является сохранение пространственной укладки хроматина внутри ядра. Цель нашего исследования состояла в том, чтобы определить оптимальные условия хранения крови для проведения последующего анализа Hi-C. Были выбраны 10 различных условий хранения образцов крови и пробоподготовки. Для каждого условия приготовлены Hi-C библиотеки и отсекарованы, после чего оценивалось качество полученных библиотек. В результате сформулированы требования к хранению и подготовке образцов, необходимые для получения качественных Hi-C данных. Нами установлен минимальный объем образца крови, достаточный для проведения Hi-C анализа. Помимо этого, мы определили способы выделения ядерных элементов крови и их долгосрочного хранения, наиболее подходящие для последующего проведения Hi-C анализа. Основное требование, сформулированное нами, – не замораживать цельную кровь.

Ключевые слова: Hi-C; периферическая кровь человека; хранение образцов крови.

Для цитирования: Гридина М.М., Весна Э., Миньженкова М.Е., Шилова Н.В., Рыжкова О.П., Назаренко Л.П., Беляева Е.О., Лебедев И.Н., Фишман В.С. Влияние предварительной обработки образцов периферической крови человека на качество Hi-C библиотек. *Вавиловский журнал генетики и селекции*. 2023;27(1):83-87. DOI 10.18699/VJGB-23-11

Influence of human peripheral blood samples preprocessing on the quality of Hi-C libraries

M.M. Gridina¹✉, E. Vesna¹, M.E. Minzhenkova², N.V. Shilova², O.P. Ryzhkova², L.P. Nazarenko³, E.O. Belyaeva³, I.N. Lebedev³, V.S. Fishman¹

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

² Research Centre for Medical Genetics, Moscow, Russia

³ Tomsk National Research Medical Center of the Russian Academy of Sciences, Tomsk, Russia

✉ gridinam@gmail.com

Abstract. The genome-wide variant of the chromatin conformation capture technique (Hi-C) is a powerful tool for revealing patterns of genome spatial organization, as well as for understanding the effects of their disturbance on disease development. In addition, Hi-C can be used to detect chromosomal rearrangements, including balanced translocations and inversions. The use of the Hi-C method for the detection of chromosomal rearrangements is becoming more widespread. Modern high-throughput methods of genome analysis can effectively reveal point mutations and unbalanced chromosomal rearrangements. However, their sensitivity for determining translocations and inversions remains rather low. The storage of whole blood samples can affect the amount and integrity of genomic DNA, and it can distort the results of subsequent analyses if the storage was not under proper conditions. The Hi-C method is extremely demanding on the input material. The necessary condition for successfully applying Hi-C and obtaining high-quality data is the preservation of the spatial chromatin organization within the nucleus. The purpose of this study was to determine

the optimal storage conditions of blood samples for subsequent Hi-C analysis. We selected 10 different conditions for blood storage and sample processing. For each condition, we prepared and sequenced Hi-C libraries. The quality of the obtained data was compared. As a result of the work, we formulated the requirements for the storage and processing of samples to obtain high-quality Hi-C data. We have established the minimum volume of blood sufficient for conducting Hi-C analysis. In addition, we have identified the most suitable methods for isolation of peripheral blood mononuclear cells and their long-term storage. The main requirement we have formulated is not to freeze whole blood.

Key words: Hi-C; human peripheral blood; blood samples storage.

For citation: Gridina M.M., Vesna E., Minzhenkova M.E., Shilova N.V., Ryzhkova O.P., Nazarenko L.P., Belyaeva E.O., Lebedev I.N., Fishman V.S. Influence of human peripheral blood samples preprocessing on the quality of Hi-C libraries. *Vavilovskii Zhurnal Genetiki i Selekcii* = *Vavilov Journal of Genetics and Breeding*. 2023;27(1):83-87. DOI 10.18699/VJGB-23-11

Введение

Сочетание метода захвата конформации хроматина с полногеномным секвенированием привело к разработке простого и эффективного протокола Hi-C, который позволяет изучать архитектуру хроматина в масштабах всего генома (Lieberman-Aiden et al., 2009; Rao et al., 2014). Наряду с многочисленными сведениями об организации и динамике хроматина в ядре, результаты Hi-C показали, что зависимость между трехмерным расстоянием в пространстве ядра и «нуклеотидным» расстоянием в геномных координатах можно описать степенной функцией во всех изученных типах клеток. Это означает, что хромосомные перестройки не только оказывают эффект на частоту контактов районов, непосредственно расположенных в точках хромосомных разрывов, но и изменяют паттерн трехмерных контактов широкой области вокруг границы перестройки (Mozheiko, Fishman, 2019). Исходя из этой закономерности недавно были предложены методы детекции хромосомных перестроек на основе анализа изменений трехмерной организации ядра (Harewood et al., 2017; Chakraborty, Ay, 2018; Díaz et al., 2018; Fishman et al., 2018; Melo et al., 2020). Данные методы позволяют в том числе детектировать сбалансированные хромосомные перестройки, что до сих пор остается трудно разрешимой задачей (Hakim et al., 2012; Dong et al., 2017). Кроме того, из данных Hi-C можно извлечь информацию об однонуклеотидных вариациях (Mozheiko, Fishman, 2019), важную для генетической диагностики.

Кровь – один из удобных материалов для проведения генетической диагностики. Правильное обращение с образцами крови перед выделением нуклеиновых кислот является критическим фактором для полногеномных исследований. Продолжительное хранение и неадекватные условия приводят к уменьшению количества выделяемой ДНК (Nederhand et al., 2003; Malentacchi et al., 2015; Schröder, Steimer, 2018) и к ее деградации (Ross et al., 1990; Permenter et al., 2015). Высокая степень деградации ДНК представляет серьезную проблему для последующих молекулярно-биологических анализов (Palmirota et al., 2011; Malentacchi et al., 2015). Например, увеличение времени хранения образца крови приводит к завышенной оценке уровня метилирования ДНК, что может быть связано с разной стабильностью метилированной и неметилированной ДНК (Schröder, Steimer, 2018).

Ключевыми этапами протокола Hi-C являются фрагментация и лигирование хроматина. Для получения качественных данных необходимо, чтобы оба этих этапа проходили *in nucleus*, т. е. в условиях максимального со-

хранения целостности ядра. Таким образом, в отличие от методов анализа последовательности ДНК, метод Hi-C предъявляет дополнительные требования к качеству исходного материала. В связи с этим нам кажется актуальным определить подходящие условия хранения образцов крови, предназначенных для проведения Hi-C анализа.

Материалы и методы

Кровь для исследования брали из локтевой вены в вакуумные пробирки с ЭДТА. Условия хранения и подготовки образцов для фиксации указаны в таблице и на рис. 1.

Выделение ядерных элементов из 3 мл цельной крови проводили одним из следующих способов:

- лизис эритроцитов при помощи буфера RBCL (BioLegend) согласно инструкциям производителя. После лизиса клеточный осадок однократно промывали фосфатно-солевым раствором (PBS);
- центрифугирование 300 г в течение 10 мин. Сыворотку крови, включая интерфазу, переносили в PBS, повторно центрифугировали 300 г 10 мин;
- центрифугирование в градиенте Histopaque-1077 Hybri-max (Sigma, Великобритания) в соответствии с протоколом производителя.

Криоконсервацию выделенных ядерных элементов крови проводили в среде для заморозки клеток: 10 % DMSO, 90 % KSR (Thermo Fisher Scientific). Клетки замораживали при -80°C и хранили в жидком азоте. После разморозки клетки однократно промывали PBS.

Описание сравниваемых условий

№ п/п	Краткое описание условия хранения и подготовки образцов для фиксации (время указано от момента забора крови)
1	Менее 4 ч; RBCLB
2	Менее 4 ч; RBCLB; заморозка KSR+DMSO
3	Менее 4 ч; центрифугирование
4	Менее 4 ч; центрифугирование; заморозка KSR+DMSO
5	24 ч RT; RBCLB
6	-20°C в течение 4 дней; RBCLB
7	2 дня при $+4^{\circ}\text{C}$; RBCLB
8	4 дня при $+4^{\circ}\text{C}$; RBCLB
9	7 дней при $+4^{\circ}\text{C}$; RBCLB
10	Менее 4 ч; Histopaque-1077 Hybri-max; заморозка KSR+DMSO



Рис. 1. Схема обработки образца крови перед фиксацией.

Клетки подсчитывали и ресуспендировали в PBS в концентрации 1 млн клеток/мл. Фиксацию клеток, приготовление Hi-C библиотек и анализ данных выполняли, как описано в (Gridina et al., 2021), используя для фрагментации хроматина DNКазу I (Thermo Fisher Scientific) или S1 нуклеазу (Thermo Fisher Scientific). Для приготовления библиотек для секвенирования использовали HAPA Hyper prep и QIAseq® FX DNA Library Kit (Qiagen), согласно инструкциям производителей. Концентрацию библиотек определяли флуориметром Qubit 3.0 (Thermo Fisher Scientific). Библиотеки секвенировали на HiSeq XTen (Illumina) парными прочтениями 150 п. н.

Результаты и обсуждение

Первым шагом протокола Hi-C, необходимым для того, чтобы сохранить пространственную организацию хроматина внутри ядра такой, какая она есть в клетке, является фиксация материала. Мы столкнулись с тем, что не всегда есть возможность доставить образец в лабораторию для фиксации в день забора крови, и решили систематически оценить влияние условий хранения крови перед проведением Hi-C анализа на качество получаемых данных. Были выбраны 10 условий, которые включали в себя: разные способы выделения ядерных элементов из цельной крови, разные время и температуру хранения образца, возможность заморозки ядерных элементов крови перед фиксацией для длительного хранения (см. рис. 1 и таблицу).

Несмотря на то что забор крови считается малоинвазивной процедурой, очевидно, что существуют определенные ограничения на объем крови, который можно получить от пациента. Особенно если пациент – маленький ребенок, или если он имеет определенные проблемы системы свертываемости крови. Для проведения Hi-C анализа необходимо 1.5–2.5 млн клеток. В норме в 1 мл крови содержится $(4–11) \times 10^6$ клеток. Для тестирования каждого условия было взято по 3 мл цельной крови в двух репликах. После удаления из образцов эритроцитов, перед фиксацией проводили подсчет ядерных элементов крови (рис. 2). В случае условия № 3 (выделение ядерных элементов без обработки RBCL) было обнаружено существенно большее число клеток. При подсчете перед фиксацией мы не определяли долю живых клеток, поэтому возможно, что в условиях № 3 сохраняются погибающие клетки, которые в других случаях подвергаются лизису при использовании RBCL буфера (Brown et al., 2016) или заморозке. Аргументом в поддержку этого предположения может служить значительно более низкое число клеток в образцах № 4 и 10, которые не обрабатывались RBCL, но подвергались замораживанию, а не в условиях № 2.

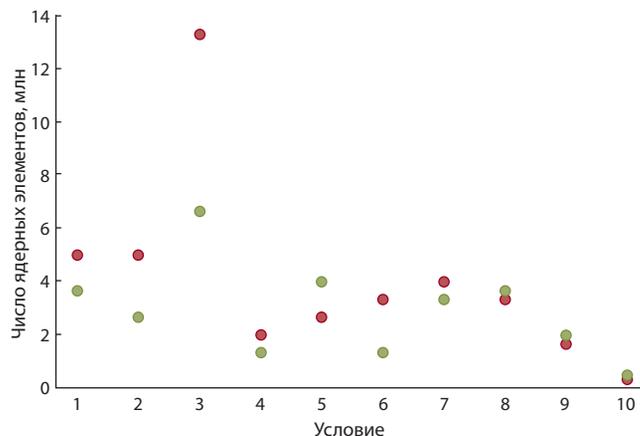


Рис. 2. Число ядерных элементов в 1 мл крови.

Эксперименты выполнены в двух независимых репликах, обозначенных на графике разным цветом. Номера по горизонтальной оси соответствуют разным условиям хранения и подготовки образцов, описанным в таблице.

Стоит отметить, что во всех образцах, кроме № 6, не было признаков выраженного гемолиза перед началом выделения ядерных элементов крови; для № 6 не было возможности оценить этот параметр. Гемолиза следует избегать, так как это один из основных факторов, негативно влияющих на количество ДНК при выделении (Sabouh et al., 2012), что может быть связано с деградацией ДНК под действием нуклеаз, высвобождающихся из разрушающихся клеток.

В процессе лизиса эритроцитов и последующих отмывок в некоторых образцах формировались конгломераты клеток, которые было невозможно диссоциировать. Конгломераты были в обеих репликах в условиях № 6, 8, 9 и 10. Следовательно, для этих образцов невозможно сказать, насколько точно было подсчитано число клеток и насколько равномерно клетки были распределены по аликвотам.

Для приготовления Hi-C библиотек брали по 2.5 млн фиксированных клеток. Чтобы оценить качество Hi-C библиотек, делают следующие контроли: геномная ДНК, ДНК после фрагментации хроматина и после лигирования хроматина (Belaghzal et al., 2017). Фореграмма контролей для всех условий давала похожую картину (рис. 3).

Полученные библиотеки секвенировали, прочтения были выровнены на геном человека сборки hg19 (GRCh_37) и проведена оценка качества библиотек. Все библиотеки имели высокую долю выровненных прочтений (рис. 4, а).

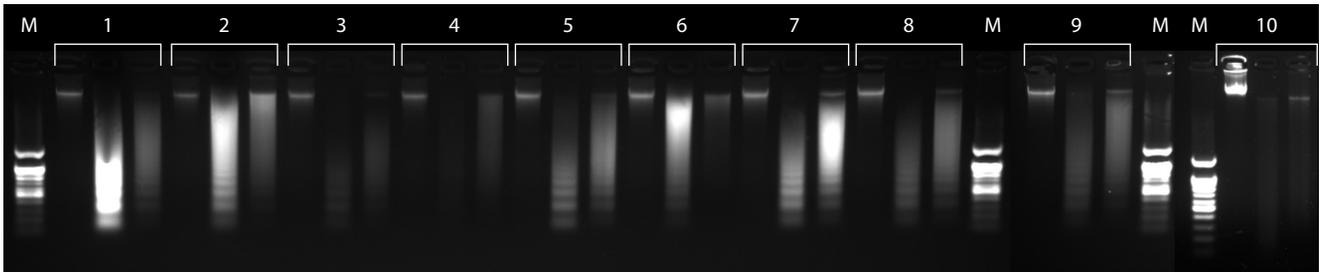


Рис. 3. Контроли фрагментации и лигирования хроматина в Hi-C экспериментах.

Номера соответствуют разным условиям хранения и подготовки образцов. Порядок нанесения: геномная ДНК, ДНК после фрагментации, ДНК после лигирования. М – ДНК-маркер 100 п. н.

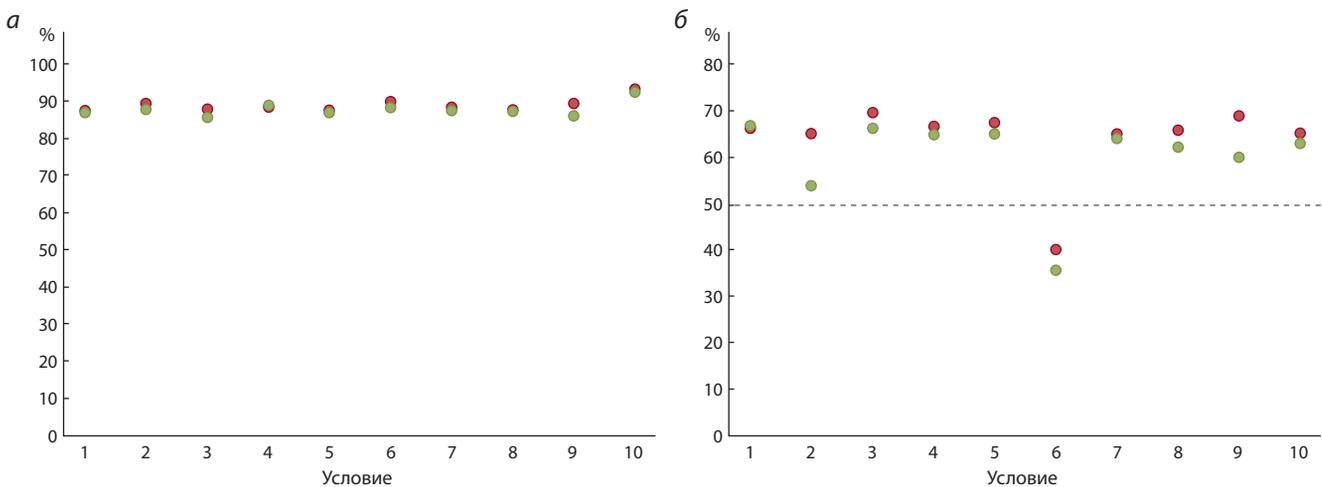


Рис. 4. Показатели качества Hi-C библиотек: а – доля выровненных прочтений; б – доля *cis* контактов среди Hi-C контактов в библиотеках.

Эксперименты выполнены в двух независимых репликах, обозначенных на графиках разным цветом. Номера по горизонтальной оси соответствуют разным условиям хранения и подготовки образцов, описанным в таблице.

Ранее нами было показано (Gridina et al., 2021), что наиболее важной для оценки качества библиотек является доля *cis* контактов (соотношение *cis/all* (FF and RR orient)) (см. рис. 4, б). Этот показатель отражает долю Hi-C прочтений, попадающих на одну хромосому, среди всех Hi-C прочтений. Он был сопоставим для всех библиотек, за исключением № 6, где составил 40.3 и 35.7 %. Это означает, что полученные Hi-C данные будут неинформативны из-за того, что лигирование большого числа фрагментов происходило случайным образом. В данном случае мы замораживали образцы крови без добавления каких-либо криопротекторов и хранили в течение 4 дней при -20°C . Столь низкая доля *cis* контактов может быть следствием возникновения случайных разрывов нитей ДНК, которые возникают при замораживании клеток без криопротекторов (Narayanan et al., 2001; Peng et al., 2008; Al-Salmani et al., 2011). С другой стороны, такой способ замораживания приводит к образованию кристаллов льда внутри клетки и в результате – к разрывам клеточных и ядерных структур (Mazur, 1984). В обоих случаях может происходить выход фрагментов ДНК из ядра и их лигирование в растворе, что приводит к образованию неинформативных фрагментов ДНК.

Заключение

В данной работе мы систематически оценили различные условия хранения и предфиксационной обработки образцов крови. В результате исследования были сформулированы следующие рекомендации по хранению и предпроцессингу образцов крови для дальнейшего проведения Hi-C анализа:

- если нет возможности доставить образец в день забора крови, его можно хранить при $+4^{\circ}\text{C}$ в течение как минимум 7 дней;
- для криоконсервации лучше предварительно провести обработку буфером RBCL;
- достаточно 1–2 мл цельной крови (человека без признаков лейкопении), однако если предполагается хранение образца более 48 ч, этот объем следует увеличить до 4–6 мл;
- категорически нельзя замораживать цельную кровь.

Список литературы / References

Al-Salmani K., Abbas H.H., Schulpen S., Karbaschi M., Abdalla I., Bowman K.J., So K.K., Evans M.D., Jones G.D., Godschalk R.W., Cooke M.S. Simplified method for the collection, storage, and comet assay analysis of DNA damage in whole blood. *Free Radic. Biol. Med.* 2011;51(3):719-725. DOI 10.1016/j.freeradbiomed.2011.05.020.

- Belaghzal H., Dekker J., Gibcus J.H. Hi-C 2.0: an optimized Hi-C procedure for high-resolution genome-wide mapping of chromosome conformation. *Methods*. 2017;123:56-65. DOI 10.1016/j.ymeth.2017.04.004.
- Brown W.E., Hu J.C., Athanasiou K.A. Ammonium-chloride-potassium lysing buffer treatment of fully differentiated cells increases cell purity and resulting neotissue functional properties. *Tissue Eng. Part C Methods*. 2016;22(9):895-903. DOI 10.1089/ten.tec.2016.0184.
- Caboux E., Lallemand C., Ferro G., Hémon B., Mendy M., Biessy C., Sims M., Wareham N., Britten A., Boland A., Hutchinson A., Siddiq A., Vineis P., Riboli E., Romieu I., Rinaldi S., Gunter M.J., Peeters P.H.M., van der Schouw Y.T., Travis R., Bueno-de-Mesquita H.B., Canzian F., Sánchez M.-J., Skeie G., Olsen K.S., Lund E., Bilbao R., Sala N., Barricarte A., Palli D., Navarro C., Panico S., Redondo M.L., Polidoro S., Dossus L., Boutron-Ruault M.C., Clavel-Chapelon F., Trichopoulou A., Trichopoulos D., Lagiou P., Boeing H., Fisher E., Tumino R., Agnoli C., Hainaut P. Sources of pre-analytical variations in yield of DNA extracted from blood samples: analysis of 50,000 DNA samples in EPIC. *PLoS One*. 2012;7(7):e39821. DOI 10.1371/journal.pone.0039821.
- Chakraborty A., Ay F. Identification of copy number variations and translocations in cancer cells from Hi-C data. *Bioinformatics*. 2018; 34(2):338-345. DOI 10.1093/bioinformatics/btx664.
- Díaz N., Kruse K., Erdmann T., Staiger A.M., Ott G., Lenz G., Vaquerizas J.M. Chromatin conformation analysis of primary patient tissue using a low input Hi-C method. *Nat. Commun.* 2018;9(1):4938. DOI 10.1038/s41467-018-06961-0.
- Dong Z., Wang H., Chen H., Jiang H., Yuan J., Yang Z., Wang W.-J., Xu F., Guo X., Cao Y., Zhu Z., Geng C., Cheung W.C., Kwok Y.K., Yang H., Leung T.Y., Morton C.C., Cheung S.W., Choy K.W. Identification of balanced chromosomal rearrangements previously unknown among participants in the 1000 Genomes Project: implications for interpretation of structural variation in genomes and the future of clinical cytogenetics. *Genet. Med.* 2018;20(7):697-707. DOI 10.1038/gim.2017.170. Epub 2017. Nov. 2.
- Fishman V.S., Salnikov P.A., Battulin N.R. Interpreting chromosomal rearrangements in the context of 3-dimensional genome organization: a practical guide for medical genetics. *Biochemistry (Mosc.)*. 2018;83(4):393-401. DOI 10.1134/S0006297918040107.
- Gridina M., Mozheiko E., Valeev E., Nazarenko L.P., Lopatkina M.E., Markova Z.G., Yablonskaya M.I., Voinova V.Y., Shilova N.V., Lebedev I.N., Fishman V. A cookbook for DNase Hi-C. *Epigenetics Chromatin*. 2021;14(1):15. DOI 10.1186/s13072-021-00389-5.
- Hakim O., Resch W., Yamane A., Klein I., Kieffer-Kwon K.-R., Jankovic M., Oliveira T., Bothmer A., Voss T.C., Ansarah-Sobrinho C., Mathe E., Liang G., Cobell J., Nakahashi H., Robbani D.F., Nussenzweig A., Hager G.L., Nussenzweig M.C., Casellas R. DNA damage defines sites of recurrent chromosomal translocations in B lymphocytes. *Nature*. 2012;484(7392):69-74. DOI 10.1038/nature10909.
- Harewood L., Kishore K., Eldridge M.D., Wingett S., Pearson D., Schoenfelder S., Collins V.P., Fraser P. Hi-C as a tool for precise detection and characterisation of chromosomal rearrangements and copy number variation in human tumours. *Genome Biol.* 2017; 18(1):125. DOI 10.1186/s13059-017-1253-8.
- Lieberman-Aiden E., van Berkum N.L., Williams L., Imakaev M., Ragozcy T., Telling A., Amit I., Lajoie B.R., Sabo P.J., Dorschner M.O., Sandstrom R., Bernstein B., Bender M.A., Groudine M., Gnirke A., Stamatoyannopoulos J., Mirny L.A., Lander E.S., Dekker J. Comprehensive mapping of long range interactions reveals folding principles of the human genome. *Science*. 2009;326(5950):289-293. DOI 10.1126/science.1181369.
- Malentacchi F., Ciniselli C.M., Pazzagli M., Verderio P., Barraud L., Hartmann C.C., Pizzamiglio S., Weisbuch S., Wyrich R., Gelmini S. Influence of pre-analytical procedures on genomic DNA integrity in blood samples: the SPIDIA experience. *Clin. Chim. Acta*. 2015;440: 205-210. DOI 10.1016/j.cca.2014.12.004.
- Mazur P. Freezing of living cells: mechanisms and implications. *Am. J. Physiol.* 1984;247(3):C125-C142. DOI 10.1152/ajpcell.1984.247.3.C125.
- Melo U.S., Schöpflin R., Acuna-Hidalgo R., Mensah M.A., Fischer-Zirmsak B., Holtgrewe M., Klever M.-K., Türkmen S., Heinrich V., Pluym I.D., Matoso E., Bernardo de Sousa S., Louro P., Hülsemann W., Cohen M., Dufke A., Latos-Bieleńska A., Vingron M., Kalscheuer V., Quintero-Rivera F., Spielmann M., Mundlos S. Hi-C identifies complex genomic rearrangements and TAD-shuffling in developmental diseases. *Am. J. Hum. Genet.* 2020;106(6):872-884. DOI 10.1016/j.ajhg.2020.04.016.
- Mozheiko E.A., Fishman V.S. Detection of point mutations and chromosomal translocations based on massive parallel sequencing of enriched 3C libraries. *Russ. J. Genet.* 2019;55(10):1273-1281. DOI 10.1134/S1022795419100089.
- Narayanan S., O'Donovan M.R., Duthie S.J. Lysis of whole blood *in vitro* causes DNA strand breaks in human lymphocytes. *Mutagenesis*. 2001;16(6):455-459. DOI 10.1093/mutage/16.6.455.
- Nederhand R.J., Droog S., Klufft C., Simoons M.L., De Maat M.P.M., Investigators of the EUROPA trial. Logistics and quality control for DNA sampling in large multicenter studies. *J. Thromb. Haemost.* 2003;1(5):987-991. DOI 10.1046/j.1538-7836.2003.00216.x.
- Palmirotta R., Ludovici G., De Marchis M.L., Savonarola A., Leone B., Spila A., De Angelis F., Della Morte D., Ferroni P., Guadagni F. Pre-analytical procedures for DNA studies: the experience of the inter-institutional multidisciplinary BioBank (BioBIM). *Biopreserv. Biobank*. 2011;9(1):35-45. DOI 10.1089/bio.2010.0027.
- Peng L., Wang S., Yin S., Li C., Li Z., Wang S., Liu Q. Autophosphorylation of H2AX in a cell-specific frozen dependent way. *Cryobiology*. 2008;57(2):175-177. DOI 10.1016/j.cryobiol.2008.06.005.
- Permenter J., Ishwar A., Rounsavall A., Smith M., Faske J., Sailey C.J., Alfaro M.P. Quantitative analysis of genomic DNA degradation in whole blood under various storage conditions for molecular diagnostic testing. *Mol. Cell. Probes*. 2015;29(6):449-453. DOI 10.1016/j.mcp.2015.07.002.
- Rao S.S.P., Huntley M.H., Durand N.C., Stamenova E.K., Bochkov I.D., Robinson J.T., Sanborn A.L., Machol I., Omer A.D., Lander E.S., Aiden E.L. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*. 2014;159(7):1665-1680. DOI 10.1016/j.cell.2014.11.021.
- Ross K.S., Haites N.E., Kelly K.F. Repeated freezing and thawing of peripheral blood and DNA in suspension: effects on DNA yield and integrity. *J. Med. Genet.* 1990;27(9):569-570. DOI 10.1136/jmg.27.9.569.
- Schröder C., Steimer W. gDNA extraction yield and methylation status of blood samples are affected by long-term storage conditions. *PLoS One*. 2018;13(2):e0192414. DOI 10.1371/journal.pone.0192414.

ORCID ID

M.M. Gridina orcid.org/0000-0002-7972-5949
E. Vesna orcid.org/0000-0003-3480-3963
M.E. Minzhenkova orcid.org/0000-0001-5458-0408

N.V. Shilova orcid.org/0000-0002-0641-1084
I.N. Lebedev orcid.org/0000-0002-0482-8046
V.S. Fishman orcid.org/0000-0002-5573-3100

Благодарности. Исследование выполнено при финансовой поддержке РНФ в рамках проекта № 22-24-00190. В исследовании использованы образцы биоколлекции «Биобанк населения Северной Евразии» НИИ медицинской генетики Томского НИМЦ РАН. Авторы выражают благодарность Александре Ян и Артему Нурисламову.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию 08.11.2022. После доработки 29.12.2022. Принята к публикации 30.12.2022.

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”.

Регистрационное свидетельство ПИ № ФС77-45870 выдано Федеральной службой по надзору в сфере связи, информационных технологий и массовых коммуникаций 20 июля 2011 г.

«Вавиловский журнал генетики и селекции» включен 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, Российский индекс научного цитирования, ВИНТИ, базы данных Emerging Sources Citation Index (Web of Science), Zoological Record (Web of Science), Scopus, PubMed Central, Ebsco, DOAJ, Ulrich's Periodicals Directory, Google Scholar, Russian Science Citation Index на платформе Web of Science, каталог научных ресурсов открытого доступа ROAD.

Открытый доступ к полным текстам:

русскоязычная версия – на сайте ИЦиГ СО РАН, <https://sites.icgbio.ru/vogis/>
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 на журнал обязательна.

✉ e-mail: 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.

Подписано в печать 20.02.2023. Выход в свет 28.02.2023. Формат 60 × 84 1/8. Усл. печ. л. 10.23.

Уч.-изд. л. 12.2. Тираж 150 экз. (1-й завод 1–45 экз.) Заказ № 27. Цена свободная.

Отпечатано в Сибирском отделении РАН, Морской проспект, 2, Новосибирск, 630090.

