

ISSN 1999-9127

Государственное научное учреждение
**«ИНСТИТУТ ГЕНЕТИКИ И ЦИТОЛОГИИ
НАЦИОНАЛЬНОЙ АКАДЕМИИ НАУК БЕЛАРУСИ»**

МОЛЕКУЛЯРНАЯ И ПРИКЛАДНАЯ ГЕНЕТИКА

**СБОРНИК НАУЧНЫХ ТРУДОВ
ТОМ 39**

Издается с 2005 года
Выходит два раза в год

Минск
2025

УДК [577.21+575] (082)

Молекулярная и прикладная генетика: сб. науч. тр. / Ин-т генетики и цитологии НАН Беларуси; редкол.: А. В. Кильчевский (гл. ред.) [и др.]. – Минск: Ин-т генетики и цитологии НАН Беларуси, 2025. – Т. 39. – 162 с. – ISSN 1999-9127.

В сборнике научных трудов публикуются обзорные и экспериментальные статьи в области молекулярной и прикладной генетики растений, микроорганизмов, животных, человека, отражающие исследования генетических процессов на молекулярном, клеточном, организменном и популяционном уровнях. Особое внимание уделяется наиболее актуальным проблемам геномики, генетической и клеточной инженерии. Публикуются результаты изучения генетических основ селекции растений, животных и микроорганизмов, разработки эффективных биотехнологий для сельского хозяйства, здравоохранения, охраны окружающей среды, биобезопасности.

Сборник предназначен для специалистов, работающих в области генетики, преподавателей, аспирантов и студентов ВУЗов биологического, сельскохозяйственного и медицинского профиля.

Редакционная коллегия:

А. В. Кильчевский — главный редактор,
Л. Э. Макарина-Кибак — заместитель главного редактора,
О. Ю. Баранов, К. У. Вильчук, С. И. Гриб, Е. В. Гузенко, Н. И. Дубовец,
А. Н. Евтушенков, А. П. Ермишин, А. И. Ковалевич, Ж. В. Колядич, С. А. Красный,
В. А. Лемеш, Н. П. Максимова, М. Е. Михайлова, П. М. Морозик, И. Б. Моссэ,
М. Е. Никифоров, В. Е. Падутов, Ф. И. Привалов, В. Н. Решетников, В. В. Титок,
И. П. Шейко, Р. И. Шейко — члены редколлегии;
Е. М. Гришкова — ответственный секретарь.

УДК [577.21+575] (082)

ISSN 1999-9127

СОДЕРЖАНИЕ

<i>О. Г. Бабак, Е. В. Дрозд, Н. А. Мартиновская, Н. В. Анисимова, К. К. Яцевич, И. Е. Баева, Т. В. Никонович, И. Г. Пугачева, М. М. Добродькин, Х. Р. Родригес, Т. Э. Рамос, Е. В. Гонсалес, Ю. К. Эррера, А. В. Кильчевский</i> Применение методов молекулярного маркирования и фитопатологической оценки при создании нового селекционного материала томата с комплексной устойчивостью к болезням		7
<i>А. Н. Верчук, А. В. Кильчевский, С. В. Кубрак, Н. В. Савина, Л. В. Милько, С. С. Савчук, В. Н. Лебедько, Д. В. Дубовик, Е. В. Никитина</i> Генотипирование лекарственных растений рода <i>Thymus</i> L., произрастающих на территории Беларуси и Узбекистана.....		21
<i>А. Н. Заинчковская, Е. А. Фомина, О. Ю. Урбанович</i> Исследование уровня экспрессии генов <i>CBF12</i> и <i>CBF14</i> у сортов озимой пшеницы в моделируемых условиях холодового стресса		33
<i>Е. В. Гузенко, А. И. Царь</i> Генетическая структура и вариабельность популяций пчел <i>Apis mellifera mellifera</i> в Республике Беларусь		41
<i>М. Е. Михайлова, Е. В. Снытков, Н. И. Тиханович, Н. А. Камыш, Р. И. Шейко, А. А. Сермягин, Н. А. Зиновьева</i> Сравнительный анализ генетической структуры красного белорусского и голштинского скота отечественной селекции по микросателлитным локусам		54
<i>И. И. Саванович, И. А. Малеваная, Л. Э. Макарина-Кибак, М. Д. Амелянович, П. М. Морозик</i> Генетические маркеры в персонализации приема кофеина для повышения скорости реакции спортсменов		65
<i>Н. В. Никитченко, И. Ю. Бакутенко, А. М. Чичко, Е. В. Сечко, Н. И. Рябоконт</i> Микросателлитные локусы генов <i>PSMA6</i> , <i>FAM177A1</i> и <i>PRORP</i> как генетические маркеры предрасположенности к развитию воспалительного процесса различной этиологии в суставах у детей		74
<i>Е. В. Волочник, Д. Р. Капуза, И. В. Пахомова, О. С. Вишкова, Т. В. Шман</i> Изменения кариотипа лейкоэмических клеток при рецидиве острого лимфобластного лейкоза из В-клеток предшественников у пациентов детского и молодого взрослого возраста.....		84
<i>Е. В. Снытков, М. Д. Амелянович</i> Поиск полиморфных локусов в дифференциально экспрессируемых генах, ассоциированных с развитием туберкулеза		100
<i>Е. В. Гузова, М. А. Лиховец, А. А. Гавричков, И. С. Трусов, И. Н. Ильющёнок, Т. В. Романовская, В. В. Гринев</i> Идентификация неаннотированных белков в клетках острого миелоидного лейкоза человека с помощью мульти-OMICs подхода		112

<i>И. С. Казловский, Н. В. Поклонская, И. В. Бельская, Т. В. Амвросьева, А. И. Зинченко</i> Метод создания генетических конструкций, содержащих гены полноразмерных капсидных белков норовирусов человека	126
<i>О. Д. Левданский, А. М. Молчун</i> Сравнительный анализ программного обеспечения для сборки малых геномов <i>de novo</i>	138
<i>Я. И. Шейко, А. М. Слуквин, С. С. Ниязова, Р. И. Шейко</i> Изучение клеточных параметров с помощью системы нейросетевого анализа изображений клеточных культур	145
К юбилею академика Кильчевского Александра Владимировича, главного редактора сборника научных трудов «Молекулярная и прикладная генетика»	156

CONTENTS

<i>O. G. Babak, E. V. Drozd, N. A. Martinovskaya, N. V. Anisimova, K. K. Yatsevich, I. E. Bayeva, T. V. Nikonovich, I. G. Puhachova, M. M. Dobrodkin, G. R. Rodríguez, T. H. Ramos, E. B. Gonzalez, Y. C. Herrera, A. V. Kilchevsky</i> Employing molecular marking and phytopathological evaluation methods in the development of new tomato breeding material with complex disease resistance.....	7
<i>A. N. Verchuk, A. V. Kilchevsky, S. V. Kubrak, N. V. Savina, L. V. Milko, S. S. Savchuk, V. N. Lebedko, D. V. Dubovik, E. V. Nikitina</i> Genotyping of medicinal plants of the genus <i>Thymus</i> L. growing in the territory of Belarus and Uzbekistan	21
<i>A. N. Zainchkovskaya, A. A. Famina, O. Yu. Urbanovich</i> Study of the expression level of the <i>CBF12</i> and <i>CBF14</i> genes in winter wheat varieties under simulated cold stress conditions	33
<i>E. V. Guzenko, A. I. Tsar</i> Genetic structure and variability of <i>Apis mellifera mellifera</i> populations in the Republic of Belarus.....	41
<i>M. E. Mikhailova, E. V. Snytkov, N. I. Tikhanovich, N. A. Kamysh, R. I. Sheiko, A. A. Sermyagin, N. A. Zinovyeva</i> Comparative analysis of the genetic structure of the red Belarusian and Holstein cattle of domestic breeding based on microsatellite loci.....	54
<i>I. I. Savanovich, I. A. Malyovanaya, L. E. Makaryna-Kibak, M. D. Ameliyanovich, P. M. Marozik</i> Genetic markers in personalizing caffeine intake to boost reaction time in athletes.....	65
<i>N. V. Nikitchenko, I. Y. Bakutenko, A. M. Tchichko, E. V. Sechko, N. I. Ryabokon</i> Microsatellite loci of the <i>PSMA6</i> , <i>FAM177A1</i> and <i>PRORP</i> genes as genetic predisposition markers to the development of the inflammatory process of various etiology in children's joints	74
<i>A. V. Valochnik, D. R. Kapuza, I. V. Pakhomava, V. S. Vshyukova, T. V. Shman</i> Karyotype changes of leukemic cells in children and young adults with relapsed B-cell precursor acute lymphoblastic leukemia.	84
<i>E. V. Snytkov, M. D. Ameliyanovich</i> Search for polymorphic loci in differentially expressed genes associated with tuberculosis development.....	100
<i>K. V. Huzava, M. A. Likhavets, A. A. Gavrichkov, I. S. Trusau, I. M. Ilyushonak, T. V. Ramanouskaya, V. V. Grinev</i> Identification of unannotated proteins in human acute myeloid leukemia cells using a multi-omics approach	112
<i>I. Kazlouski, N. Paklonskaya, I. Belskaya, T. Amvrosieva, A. Zinchenko</i> Method for designing genetic constructs containing the genes of full-length capsid proteins of human noroviruses	126

<i>A. D. Liaudanski, A. M. Molchun</i>	
Comparative analysis of <i>de novo</i> small genome assembly software.....	138
<i>Ya. I. Sheiko, A. M. Slukvin, S. S. Niyazova, R. I. Sheiko</i>	
Study of cellular parameters using neural network analysis of cell culture images.....	145
To mark the anniversary of Academician Alexander Kilchevsky, editor-in-chief of <i>The Molecular & Applied Genetics</i> proceedings.....	156

О. Г. Бабак¹, Е. В. Дрозд¹, Н. А. Мартиновская¹, Н. В. Анисимова¹, К. К. Яцевич¹, И. Е. Баева²,
Т. В. Никонович², И. Г. Пугачева², М. М. Добродькин², Х. Р. Родригес³, Т. Э. Рамос³,
Е. В. Гонсалес³, Ю. К. Эррера³, А. В. Кильчевский¹

ПРИМЕНЕНИЕ МЕТОДОВ МОЛЕКУЛЯРНОГО МАРКИРОВАНИЯ И ФИТОПАТОЛОГИЧЕСКОЙ ОЦЕНКИ ПРИ СОЗДАНИИ НОВОГО СЕЛЕКЦИОННОГО МАТЕРИАЛА ТОМАТА С КОМПЛЕКСНОЙ УСТОЙЧИВОСТЬЮ К БОЛЕЗНЯМ

¹Государственное научное учреждение
«Институт генетики и цитологии Национальной академии наук Беларуси»
Республика Беларусь, 220072, г. Минск, ул. Академическая, 27
e-mail: O.Babak@igc.by

²Учреждение образования «Белорусская государственная орденов Октябрьской Революции
и Трудового Красного Знамени сельскохозяйственная академия»
Республика Беларусь, 213410, г. Горки, ул. Мичурина, 5

³Научно-исследовательский институт им. Лилианы Димитровой
Республика Куба, 33500, Маябеке, ш. Бехукаль-Кивикан, км 33,5

По результатам секвенирования R-генов *Ph2(Ch)* (LOC101246197) и *Ph3* (KJ563933) у образцов томата белорусской и кубинской селекции выявлены структурные особенности аллелей, связанные с контролем резистентности образцов к *Phytophthora infestans*. Разработаны CAPS-маркер Ph-2(R)_HindIII и SCAR-маркер Ph3 del22/11, позволяющие типировать аллели, связанные с устойчивостью / восприимчивостью к данному патогену. ДНК-типирование образцов белорусско-кубинской коллекции на наличие аллелей, связанных с устойчивостью к болезням, позволило выявить формы, сочетающие в себе 7–9 аллелей устойчивости к болезням: Lyto (7 R), LRs-30(7 R), Pablo (8 R), Alty (8 R), Sile (9 R), LR-13 (9 R). Белорусский сорт Ирма показал максимальную устойчивость к болезням и продуктивность в условиях тропического климата Республики Куба. Создан новый селекционный материал с повышенной устойчивостью к фитофторе в условиях Республики Беларусь.

Ключевые слова: томат, гены устойчивости к фитофторозу, ДНК-типирование, маркеры генов устойчивости к болезням, селекционный материал.

Для цитирования: Применение методов молекулярного маркирования и фитопатологической оценки при создании нового селекционного материала томата с комплексной устойчивостью к болезням / О. Г. Бабак, Е. В. Дрозд, Н. А. Мартиновская [и др.] // Молекулярная и прикладная генетика: сб. науч. тр. / Ин-т генетики и цитологии НАН Беларуси; редкол.: А. В. Кильчевский (гл. ред.) [и др.]. – Минск, 2025. – Т. 39. – С. 7–20.

Введение

Резистентность растений томата к болезням, вызываемым возбудителями различной природы (грибные, бактериальные, вирусные инфекции), является как отдельным направлением селекции, так и обязательным условием при создании новых высокопродуктивных сортов и гибридов с улучшенным качеством плодов. При этом климатические факторы окружающей среды обеспечивают условия для доминирующего проявления тех или иных патогенов. Так, в условиях открытого грунта

на территории Беларуси самым вредоносным возбудителем болезней томата является оомицет *Phytophthora infestans* (Mont.) de Bary. Фитофтороз (*late blight*), вызываемый данным возбудителем, может привести к 100% потере урожая. При этом в условиях жаркого климата с высокими температурами и недостатком влаги первое место в заражении растений занимают вирусные инфекции. Принимая во внимание то, что приспособление растений к определенным условиям среды способствует закреплению в генотипе уникальных алле-

лей, обеспечивающих развитие механизмов защиты от наиболее распространенных инфекций, вовлечение в селекционный процесс образцов, приспособленных к контрастным условиям среды, имеет важное значение при создании форм с комплексной устойчивостью к болезням.

В наших исследованиях по селекции томата широко используются методы молекулярного маркирования генетических детерминант устойчивости к возбудителям болезней наряду с другими целевыми хозяйственно-ценными признаками. В рамках ранее проведенных исследований в Институте генетики и цитологии НАН Беларуси апробированы и адаптированы методики ДНК-типирования аллелей к ряду заболеваний, совместно с Белорусской государственной сельскохозяйственной академией выполнены молекулярная (с помощью известных маркеров устойчивости к болезням) и полевая оценка материала белорусской селекции [1, 2]. Особое внимание уделяется изучению генетики устойчивости томата к фитофторозу с использованием маркеров к генам *Ph2* (ХМ_004249610.4) и *Ph3* (KJ563933), предложенных американскими и нидерландскими исследователями [3, 4, 5]. В 2021 г. китайскими учеными [6] было показано, что гену *Ph2* соответствует последовательность Solyc10g085460, близко расположенная к последовательности ХМ_004249610.4, в связи с чем представляет интерес изучение полиморфизма данного гена на новом селекционном материале. К настоящему времени в Институте генетики и цитологии разработана «Технология маркер-сопутствующего отбора форм томата с высокими биохимическими и технологическими свойствами плодов» (Методические указания), которая включает маркеры к 22 аллелям, связанным с устойчивостью к грибным, бактериальным, вирусным болезням [7, 8].

Целью данного этапа исследований было изучение новых генетических источников устойчивости к болезням на основе секвенирования R-генов и разработанных маркеров, а также фитопатологической оценки материала географически разнообразного происхождения — белорусской и кубинской селекции, и отбор нового селекционного материала с комплексной устойчивостью к болезням. Для достижения данной цели последовательно

решались следующие задачи: секвенирование и анализ генов устойчивости к фитофторозу у образцов белорусско-кубинской коллекции; разработка и апробация новых молекулярных маркеров на основе выявленного полиморфизма аллелей (отличий в нуклеотидной последовательности) у изучаемых форм; оценка популяций F_5 белорусской селекции с помощью вновь разработанных маркеров устойчивости к болезням; ДНК-типирование образцов белорусско-кубинской коллекции с помощью системы ДНК-маркеров к аллелям, связанным с устойчивостью к болезням; полевая оценка устойчивости образцов в условиях Кубы и Беларуси; отбор нового селекционного материала с комплексной устойчивостью к болезням.

Материалы и методы

В качестве исследуемого материала использовались образцы томата открытого и защищенного грунта коллекций Института генетики и цитологии НАН Беларуси и Белорусской государственной сельскохозяйственной академии (БГСХА):

сорта Редгрант, Зорка, Гарант, Чирок, Беркут, Тамара, Ирма, Дивиденд, Спажытак, Источник, Звездопад;

образцы семи популяций F_5 гибридов томата для открытого грунта, у большинства из которых материнской формой был сорт Ирма с повышенной полевой устойчивостью к фитофторозу, а отцовскими формами были образцы с различными аллелями генов качества плодов (*b*, *og^c*, *t*, *rin*, *alc*): Ирма × Линия 19-612, Ирма × Линия 19-645, Ирма × Линия 217, Ирма × Линия 16-57, Ирма × Линия 16-8, Ирма × Дивиденд, Ирма × Линия 221;

образцы трех популяций F_6 из гибрида С 9464 × 19/15-4.

Также использовались сорта и селекционные формы коллекции Института плодовоовощеводства им. Лилианы Димитровой (Республика Куба): LD56, Pablo, LR-13, L22/11, L29/6, L33/1, Alty, L-43, Lyto, Sile, LRs-30, Daniel, L41/3.

Секвенирование последовательностей генов *Ph2(Ch)* (LOC101246197) и *Ph3* проводили с подобранными геноспецифическими праймерами на основе последовательности участка хромосомы 10 томата (NC_015447.3), где локализован ген *Ph2*, а также известных

последовательностей гена *Ph3* (KJ563933, NC_090808.1), локализованного на хромосоме 9 (табл. 1). Использовали набор реагентов Big Dye® Terminator v 3.1 Cycle Sequencing Kit

согласно методике производителя. Для дизайна праймеров использовали программу Primer-BLAST на сайте Центра биотехнологической информации NCBI [9].

Таблица 1

Праймеры для секвенирования генов *Ph2* и *Ph3* у томата

Ген, локус, белок	Наименование праймеров	Последовательность праймера, 5' → 3'	Температура отжига, °C	Ожидаемый размер ПЦР-продукта, н. п.
<i>Ph2</i> , LOC101246197 <i>Solanum lycopersicum</i> putative disease resistance protein RGA3 [9]	Ph-2_01-F/R	GTGAAGCTAAGGAGCAAACGC	60	854
		CACGCCACAAATCATCCAGG		
	Ph-2_02-F/R	GACAGGGCAGAAACTTGAGC	60	974
		TGATGCCTGATGCGGACAAA		
	Ph-2_03-F/R	AAGATAAGCGAGCCAGGACG	60	970
		TGCGCAACTTCTCAAGCCTA		
	Ph-2_04-F/R	TCGACATCTTGAGCTGGTAGG	60	1040
		GCCGTTTGCACCTCTTTAGC		
	Ph-2_05-F/R	ACGGACATTTGCACTGGGAT	60	796
		CACTCACCGGCTCATCTCTG		
	Ph-2_06-F/R	CACCTCTCCCAACCTGTACG	60	995
		CTGGGCGTTACCATTGGACT		
<i>Ph3</i> KJ563933, Ph-3 resistance protein [9]	Ph-3_01-F/R	TGGTCGTTGACTCCAACACT	60	764
		TTGGCTGCTGCGAGACATAG		
	Ph-3_02-F/R	AATAGTTGGCATGCCCGGTT	60	990
		ACCCTAGAAACATTGCCGGG		
	Ph-3_03 F/R	ACTTCTTTACACCGAGCACA	60	745
		TTCAAATGCACGTTTCCGCTC		
	Ph-3_04 F/R	TTCGTGTC AACCGAGTGAGC	60	808
		TTTCGGCTAGTGACACGCTG		

Молекулярно-генетические методы выделения ДНК, ПЦР-анализ, рестрикция, оценка продуктов амплификации и рестрикции в агарозном или полиакриламидном гелях проводились согласно разработанным в Институте генетики и цитологии НАН Беларуси Методическим указаниям [7]. В таблице 2 представлены маркеры, используемые для ДНК-типирования аллелей, связанных с устойчивостью/восприимчивостью к патогенам (R/S-аллелей) у образцов томата белорусско-кубинской коллекции, и ссылки на литературные источники, приведенные в Методических указаниях и ранее опубликованных работах [1, 7, 8].

Фитопатологическую оценку в научно-исследовательском институте плодоовощеводства им. Лилианы Димитровой проводили в полевых и лабораторных условиях. Наличие болезней и степень вредоносности оценивались по методологии, установленной для каждого заболевания: фитофтороза, альтернариоза, кладоспориоза, фузариоза, вертициллеза, поражения вирусом желтой курчавости листьев (*TYLCV*) [10]. Для идентификации грибных патогенов образцы зараженной растительной ткани отбирали из надземной части, преимущественно с листьев, помещали во влажные камеры для создания оптималь-

Таблица 2
Наименование и характеристика маркеров, используемых для ДНК-типирования образцов *S. lycopersicum*

Ген	Тип и название маркера	Праймеры для ПЦР	Т отжига, °C	Размер ПЦР-продукта, п. н.	Литературный источник
<i>Mi-1</i>	SCAR, Mi-1.2	Mi23F: TGGAAAAATGTTGAATTTCTTTTG Mi23R: GCATAATATATGGCTTGTTCACCC	55	R – 380; S – 430	Seah S. et al., 2007
<i>Mi-9</i>	SCAR, C8B-Mi9	F: ACCCACGCCCCATCAATG R: TGCAAGAGGGTGAATATGAGTGC	59	R – 400; S – 360	El-Sappah A. H., 2019
<i>Ve</i>	CAPS, Ve (<i>Xba</i> I)	F: CGAACTTGAC TACATTGACC CTG R: CAGTCTTGAAAGGTTGCTCAGCC	59	R – 410, 332; S – 410, 310, 22	Je Min Lee et al., 2015
<i>I-2</i>	SCAR, I-2	I-2/5F: CAAGGAACTGCGTCTGTCTG I-2/5R: ATGAGCAATTTGTGGCCAGT	55	R(I-2) – 633; R(I-2C) – 566; S – 693, 760	Davis R. M. et al., 2010
<i>I-3</i>	SCAR, P7-43D	F: GGTAAAGAGATGCGATGATTATGTGGAG R: GTCTTTACCACAGGAACCTTTATCACC	60	R – 875; S – 650	Je Min Lee. et al., 2015
<i>I-7</i>	CAPS, I-7 (<i>Age</i> I)	F: AAGAAGTCCCTTCTTCCCTTA R: GGAATAACCAAGGGGTGTT	56	R – 608, 200; S – 808	Je Min Lee. et al., 2015
<i>Cf2, Cf5</i>	SCAR, 2-5Cf	F: GCTATCTTTGGGTATCAACTT R: AGATGACATCGACAAAATGTG	58	R(Cf-2) – 1600; R(Cf-5) – 1163, 880; S – нет продукта	Kanwar J. S. et al., 1980
<i>Cf5</i>	SCAR, CF5	F: GTAATATCAGTGACCTTCACA R: ATTTTCCAAACTGAAAAG	55	R – 1600; S – нет продукта	Dixon M. S. et al., 1998
<i>Cf4</i>	SCAR, Cf4/4A	F: TTACGACAGAAAGAACTCTTTCTTGG R: AGTCCGCTTACGTTGGATGG	55	R(Cf-4) – 816; R(Cf-4A) – 965; S – нет продукта	Parniske M. et al., 1997 Takken F. L. et al., 1998
<i>Cf9</i>	SCAR, Cf9/9DC	CS5: TTTCCAACTTACAAATCCCTTC DS1: GAGAGCTCAACCTTTACGAA CS1: GCCGTTCAAAGTTGGGTGTT	55	R(Cf-9) – 378; R(9DC) – 507; S – нет продукта	Van der Hoorn R. A. et al., 2001
<i>Ph-2</i>	CAPS, dTG63 (<i>Hinf</i> I)	F: CTACTCTTTCTATGCAATTGAATTG R: AATAAATTTCAACCATAGAAATGATT	55	R – 246; S – 220	Panthee D. R. et al., 2012
<i>Ph-3</i>	SCAR, NC-LB-9-6678	F: CCTTAATGCAATAGGCAAT R: ATTTGAATGTTCTGGATTGG	52	R – 600; S – 900	Panthee D. R. et al., 2012
<i>Ph-3</i>	SCAR	F: CTACTCGTGCAAGAAGGTAC R: TCCACATCACCTGCCAGTTG	55	R – 176; S – 154	Kim H. J., 2015

Окончание таблицы 2

Ген	Тип и название маркера	Праймеры для ПЦР	Т отжига, °C	Размер ПЦР-продукта, п. н.	Литературный источник
<i>Pto</i>	CAPS, <i>Pto</i> _RsaI	F: ATCTACCCACAATGAGCATGAGCTC R: GTGCATACTCCAGTTTCCAC	56	R – 552; S – 113, 439	Yang W. et al., 2005, Je Min Lee. et al., 2015
<i>Rx4</i>	SCAR, <i>pcc12</i> InDel	F: TCCACATCAAAATGCGTTTCT R: TTCCAATCCTTTCCATTTCG	60	R – 113; S – 119	Yang W. et al., 2005
<i>Tm1</i>	SCAR, <i>Tm1</i> :8por	F: CCACTGTATGATTCTGCTAGTGAA R1: AGCTTTAACAATAATAAGAAATAAGAC R2: GCAAGCTAAGGTTTACATATATGCC	56	R – 131; S – 212	Je Min Lee. et al., 2015
<i>Tm-2</i>	CAPS, <i>Tm2RS-f3/r3</i> <i>HpaI</i>	F: TGGAGGGGAATATTTGTGGA R: ACTTCAGACAACCCATTCCG	61	R(<i>Tm2</i>) – 458, 245; S – 703	Shi A. et al., 2011
<i>Tm-2²</i>	CAPS, <i>Tm2RS-f3/r3</i> <i>BsiHKAI</i>	F: TGGAGGGGAATATTTGTGGA R: ACTTCAGACAACCCATTCCG	61	R(<i>Tm2²</i>) – 358, 353; R(<i>Tm2</i>) – 703	Shi A. et al., 2011
<i>Sw-5</i>	SCAR, <i>Sw5-2</i>	F: CGGAACCTGTAACTTGACTG R: GAGCTCTCATCCATTTTCCG	60	R(<i>Sw-5b</i>) – 541; S – нег продукта	Shi A. et al., 2011
<i>Ty2</i>	SCAR, <i>Ty-2</i>	F: TGGCTCATCCTGAAAGCTGATAGCGC R: AGTGATACATCCTTGCCATTGACT	50	R – 900; S – 790	Je Min Lee. et al., 2015
<i>Ty3</i>	SCAR, <i>Ty-3</i>	F: GGTAGTGGAATGATGCTGCTC R: GCTCTGCCATTGTCCCCATATAAACC	50	R(<i>Ty-3</i>) – 450; R(<i>Ty-3a</i>) – 630; S – 320	Je Min Lee. et al., 2015

ных условий влажности и температуры для развития гриба, стимулирования роста мицелия и образования структур бесполого размножения. Затем морфология репродуктивных структур патогена исследовалась в лаборатории с помощью микроскопа [11–13]. Оценка в полевых условиях проводилась на естественном фоне. Исследование было проведено с 11 белорусскими сортами и четырьмя кубинскими контрольными сортами и линиями в оптимальный сезон для выращивания томатов на Кубе (выращивание в грунте с декабря по март). Наряду с фитопатологической оценкой проводили оценку урожайности белорусских форм в сравнении с контрольными кубинскими формами Daniel, Pablo, L-43, Lyto.

Оценку зараженности растений основными болезнями в условиях Беларуси проводили на естественном инфекционном фоне на опытном поле Белорусской государственной сельскохозяйственной академии. Определяли распространенность фитофтороза, альтернариоза, а также вирусных и бактериальных инфекций в целом. Оценку распространенности болезней определяли в процентном отношении растений с признаками заболевания к общему числу испытываемых растений, независимо от степени поражения. Степень поражения растений болезнями определяли по 4-бальной шкале поражений, предложенной А. В. Алпатьевым [14]. Для диагностики применяли визуальные методы оценки симптомов: наличие темно-бурых пятен при фитофторозе; бурых, коричневых или черных пятен с кон-

центрическими кольцами при альтернариозе; характерных признаков вирусных инфекций и бактериальных поражений. Оценка каждой формы выполнялась на 10–15 растениях.

Результаты и обсуждение

Секвенирование и анализ генов устойчивости к фитофторозу, разработка маркеров для типирования R/S-аллелей

В связи с появлением данных в литературе о новых источниках устойчивости к фитофторозу выполнено секвенирование гена *Ph2(Ch)* (LOC101246197) у образцов, демонстрирующих различную степень устойчивости к данному заболеванию: Дивиденд, Ирма, Pablo, Alty, Sile. Полевые испытания образцов в климатических условиях Беларуси и Кубы подтверждают максимальную устойчивость сорта Ирма. Секвенирование последовательностей гена *Ph-2* выявило отличия по 32 SNP между R- и S-аллелями. На основе SNP G/A в позиции 2925 разработан и успешно апробирован CAPS-маркер Ph-2 (R) *Hind*III (F: TCGACATCTTGAGCTGGTAGG, R: GCCGTTTGCACCTCTTTAGC) для типирования R/S-аллелей данного гена (рис. 1). У сорта Ирма выявлен аллель, полностью идентичный таковому у сорта Moboline LA3152, который обуславливает устойчивость к возбудителю фитофтороза у томата [6]. По результатам ДНК-типирования коллекции образцов R-аллель также выявлен у 4-х образцов: LR-13, L43, L41/3, Lyto.

Наряду с образцами, включенными в состав



Рис. 1. Полиморфный участок гена *Ph2(Ch)* с выделенным сайтом рестрикции для CAPS-маркера Ph2(R) *Hind*III на геномной ДНК томата

белорусско-кубинской коллекции, выполнено ДНК-типирование популяций F_5 – F_6 томата для открытого грунта селекции БГСХА и Института генетики и цитологии НАН Беларуси. На рисунке 2 представлен пример результатов ДНК-типирования образцов популяций F_5 – F_6 .

Согласно результатам ДНК-маркирования, отобраны формы с R-аллелем гена *Ph2(Ch)* из популяций F_6 С9464 × 15/15-4, из F_5 Ирма × Линия 217, F_5 Ирма × Линия 16-57, F_5 Ирма × Линия 16-8, F_5 Ирма × Дивиденд, F_5 Ирма × Линия 19-612.

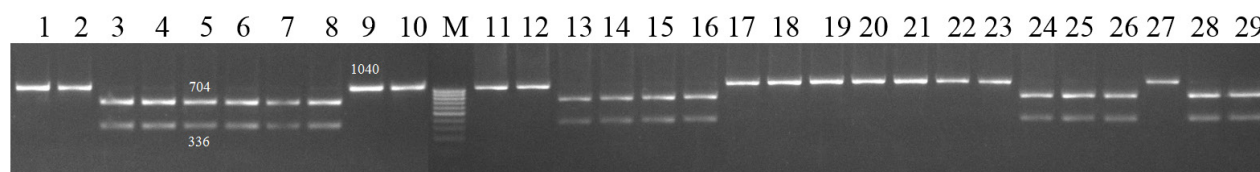


Рис. 2. Результаты электрофореза продуктов рестрикции ампликонов, синтезируемых в ходе ПЦР с маркером *Ph-2(Ch)*_HindIII у образцов популяций F_5 – F_6 : 1–2 — № 104 из F_6 С 9464 × 19/15-4; 3–4 — № 106 из F_6 С 9464 × 19/15-4; 5–8 — № 119 из F_5 Ирма × Линия 221; 9–12 — № 117 из F_5 Ирма × Линия 217; 13–17 — № 115 из F_5 Ирма × Линия 16-57; 18–21 — № 113 из F_5 Ирма × Линия 16-8; 22–23 — № 111 из F_5 Ирма × Дивиденд; 24–26 — № 109 из F_5 Ирма × Линия 19-645; 27–29 — № 107 из F_5 Ирма × Линия 19-612; М — маркер молекулярного веса ДНК (100 bp)

На основании данных о происхождении устойчивого аллеля *Ph3* из *S. pimpinellifolium* и результатов выполненного ранее молекулярного маркирования образцов с помощью маркера NC-LB-9-6678 аллелей гена *Ph3* [5], а также фенотипической оценки проявления признаков на данном этапе работы для секвенирования отобраны образцы Ирма, Alty, Sile, LR-13, F_2 Дзивосны (*S. lycopersicum* K[−] — *ph3*), *S. pimpinellifolium* (K⁺ — *Ph3*), Блэк Бриллиант (*S. lycopersicum* K⁺ — *Ph3*).

Согласно результатам выравнивания полученных сиквенсов, у сортов с аллелем *ph3*, типизируемым с помощью SCAR маркера NC-LB-9-6678, разница представляет собой вставку фрагмента LTR ретротранспозона семейства *Ty1-soria*, размер которого составляет 306 п. н. [15]. Встроенный ретротранспозон является целевым различием, выявляемым с помощью большинства разработанных маркеров. Однако сравнительный анализ полученных нами последовательностей показал, что у аллеля *ph-3* из *S. lycopersicum* (NC_090808.1) в начале гена присутствуют две делеции размерами по 11 п. н. на расстоянии 107–118, 177–188 п. н. от старт-кодона. Данные делеции ведут к сдвигу рамки считывания и появлению раннего стоп-кодона (рис. 3), что, в свою очередь, ведет к синтезу короткой белковой последовательности размером 59 аминокислот, в сравнении с 851-й аминокислотой в последовательности, синтезируемой на основе и-РНК аллеля *Ph3*

из *S. pimpinellifolium* (KJ563933). Полученные данные говорят о том, что встроенный ретротранспозон не может оказывать влияния на функциональность белка, т. к. расположен после сформированного раннего стоп-кодона. Разработан новый SCAR-маркер *ph3del22/11*, а также подтверждена эффективность ранее разработанного маркера [16], позволяющего типировать наличие указанных делеций, связанных с восприимчивостью к фитофторозу.

Сопоставление результатов разделения форм с помощью маркеров, разработанных на основе двух вышеуказанных полиморфизмов, показывает, что у большинства исследуемых форм оба маркера позволяют разделять R/S аллели. При этом у образца LR-13 без вставки ретротранспозона, который по результатам ДНК-типирования с помощью маркера NC-LB-9-6678 является устойчивым, выявлено наличие одной делеции (11 п. н.), которая также ведет к формированию раннего стоп-кодона, нефункционального белка и восприимчивости к фитофторозу (рис. 3), что указывает на неполное соответствие результатов отбора устойчивых и восприимчивых образцов данными маркерами.

Наряду с указанными полиморфизмами у образцов Блэк Бриллиант и Sile выявлен ряд SNP-замен, отличающих их от других образцов с аллелем *Ph-3* и между собой. Образец Sile также содержит делецию 9 п. н., приводящую к выпадению трех аминокислот

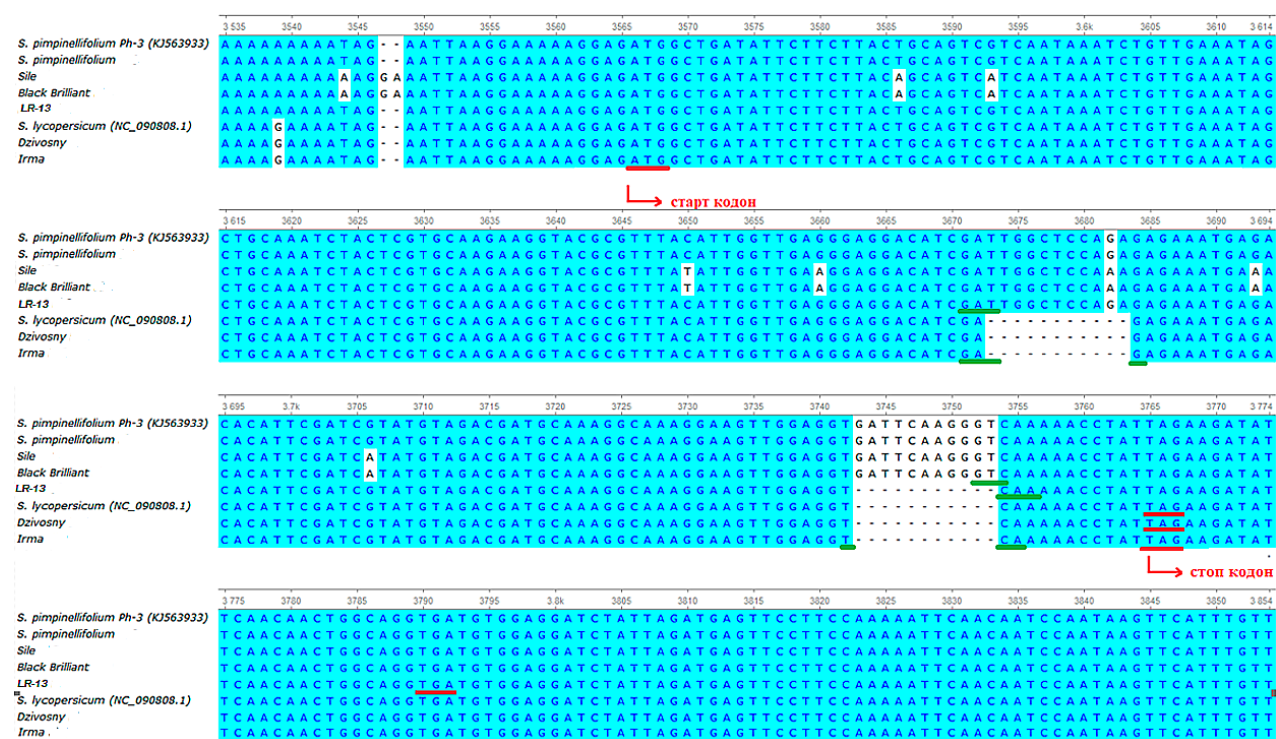


Рис. 3. Участок блока выравнивания гена *Ph-3* у образцов томата со структурными различиями, приводящими к формированию укороченной аминокислотной последовательности

в пептиде (рис. 3). Полученные результаты показывают высокий генетический полиморфизм аллелей и подтверждают важность его детального анализа при разработке маркеров и изучении их связи с устойчивостью форм при фитопатологической оценке.

На рисунке 4 представлены результаты апробации разработанного *ph3del22/11* SCAR-маркера. При этом маркерный ампликон размером 153 п. н. соответствует ампликону 907 п. н., выявляющему неустойчивый аллель *ph3* маркером NC-LB-9-6678, а ампликон 175 п. н. соответствует размеру 601 п. н. у образцов

с функциональным белком устойчивости. Новый SCAR-маркер позволяет выявить полиморфизм, связанный с формированием нефункционального белка у форм с одной 11-нуклеотидной делецией, приводящей к формированию нефункционального белка.

Типирование образцов белорусско-кубинской коллекции на наличие маркеров аллелей, связанных с устойчивостью к болезням, и сопоставление с результатами фитопатологического анализа

Поиск новых источников комплексной устойчивости осуществлен путем ДНК-ти-

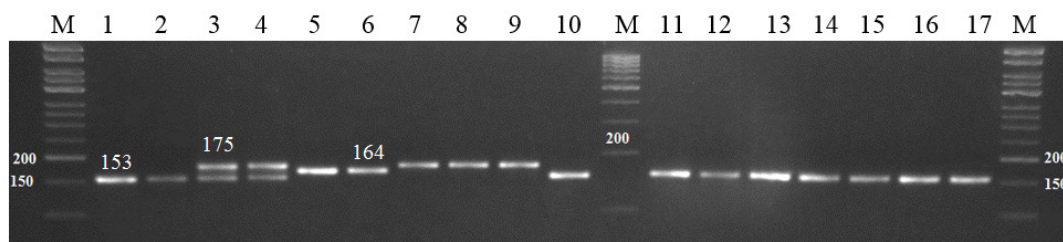


Рис. 4. Результаты ДНК-типирования образцов томата с SCAR-маркером *Ph3 del22/11*: 1–2 — д. 66-1 (F_2 Дзивосны *ph3*); 3–4 — д. 66-7 (F_2 Дзивосны *ph3/Ph3*); 5–6 — LR-13; 7–8 — Sile; 9 — Блэк бриллиант; 10–11 — ИСИ63; 12–13 — LD-56; 14–15 — Alty; 16–17 — L-43 (153 и 164 п. н. — фрагменты, выявляющие S-аллели, 175 п. н. — фрагмент, выявляющий R-аллель)

пирования изучаемого материала томата на наличие маркеров устойчивости к мелойдогенозу (*Mi-1.2*, *Mi-9*), фитофторозу (*Ph3*, *Ph2*, *Ph2(Ch)*), кладоспориозу (*Cf2*, *Cf4*, *Cf5*, *Cf9*), фузариозу (*I-2*, *I-2C*, *I-3*, *I-7*), вертициллезу (*Ve*), бактериальной пятнистости (*Rx-4*, *Pto*) и крапчатости (*Pto*), к вирусным болезням: желтой курчавости листьев (*Ty2*, *Ty3*), мозаике томата (*Tm-1*, *Tm-2*) и бронзовости плодов (*Sw-5b*).

ДНК-типирование образцов томата белорусской селекции для открытого грунта выявило наличие аллеля *Ph2(Ch)* у образца Ирма. Аллель устойчивости к фузариозу *I-2C* был идентифицирован в генотипах сортов Источник, Чирок; аллель *I-3* устойчивости к фузариозу — в генотипах сортов Зорка, Гарант, Чирок, Беркут, Тамара, Ирма, Дивиденд, Спажытак, Источник, Звездопад; аллель *Mi-9* устойчивости к мелойдогенозу (нематоды) — у всех образцов коллекции. Наличие известных аллелей устойчивости к бактериальным пятнистостям, кладоспориозу у изучаемых форм

не выявлено. Результаты ДНК-типирования образцов кубинской коллекции представлены в таблице 3. Молекулярный анализ коллекции позволил выделить формы, сочетающие в генотипе 7–11 аллелей, связанных с устойчивостью к болезням: Pablo, Alty, Sile, L 41/3, L 22/11, LRs-30, LR-13.

Результаты оценки устойчивости образцов белорусско-кубинской коллекции к болезням в полевых условиях представлены в таблице 4. Учет развития фитофтороза проводился в два этапа, остальных болезней — однократно. В процессе наблюдений у исследуемых образцов были обнаружены симптомы фитофтороза, альтернариоза, вирусных болезней и бактериальной крапчатости. Наиболее восприимчивыми к листовым пятнистостям были сорта Дивиденд, L-43(1), Спажытак, Зорка, Тамара, Звездопад, которые характеризовались уровнем пораженности более 50% листовой поверхности. В целом образцы кубинской коллекции оказались более устойчивыми к альтернариозу, чем белорусские. Плоды

Таблица 3

Результаты ДНК-типирования образцов кубинской коллекции на наличие маркеров, связанных с устойчивостью к болезням

Наименование образца	Выявленные аллели, связанные с устойчивостью к болезням
LD56	<i>Mi-9</i> , <i>Ph2</i> , <i>I-3/i-3</i> , <i>Cf5</i> , <i>Sw-5b</i>
Pablo	<i>Mi-9</i> , <i>Ph2</i> , <i>I-2/I-2C</i> , <i>I-3/i-3</i> , <i>Pto</i> , <i>Cf5</i> , <i>Rx-4</i>
Lyto	<i>Mi-9</i> , <i>Ph2</i> , <i>Ph2(Ch)</i> , <i>I-3/i-3</i> , <i>Ty2</i>
Alty	<i>Mi-9</i> , <i>Ph2</i> , <i>I-2/i2</i> , <i>I-3/i-3</i> , <i>I-7/i-7</i> , <i>Cf4/Cf4A</i> , <i>Ve</i>
Sile	<i>Mi-9</i> , <i>Ph2</i> , <i>Ph3</i> , <i>I-3/i-3</i> , <i>I-7</i> , <i>Cf4/Cf4A</i> , <i>Cf5</i> , <i>Ty3a</i> , <i>Ve</i>
Daniel	<i>Mi-9</i> , <i>I-3/i-3</i> , <i>Cf5</i> , <i>Ve</i>
L 29/6	<i>Mi-1.2</i> , <i>Mi-9</i> , <i>I-2/i2</i> , <i>I-3/i-3</i> , <i>I-7</i> , <i>Cf4/Cf4A</i> , <i>Cf5</i> , <i>Cf9</i> , <i>Pto</i> , <i>Rx-4</i>
L 33/1	<i>Mi-1.2</i> , <i>Mi-9</i> , <i>I-2/i2</i> , <i>I-3/i-3</i> , <i>Cf4/Cf4A</i> , <i>Cf5</i> , <i>Cf9</i> , <i>Pto</i> , <i>Rx-4</i>
L-43	<i>Mi-1.2 (mi-1.2)</i> , <i>Mi-9</i> , <i>Ph2(Ch)</i> , <i>I-3/i-3</i> , <i>I-7</i> , <i>Cf5</i>
LRs-30	<i>Mi-9</i> , <i>Ph2</i> , <i>Ph2(Ch)/ph2(Ch)</i> , <i>I-3/i-3</i> , <i>I-7/i7</i> , <i>Cf4/Cf4A</i> , <i>Cf5</i> , <i>Cf9</i> , <i>Pto/pto</i>
LR-13	<i>Mi-1.2</i> , <i>Mi-9</i> , <i>Ph2(Ch)</i> , <i>Ph3</i> , <i>I-2C</i> , <i>I-3/i-3</i> , <i>Cf4/Cf4A</i> , <i>Cf5</i> , <i>Pto</i> , <i>Rx-4</i> , <i>Sw-5b</i>
L 22/11	<i>Mi-1.2/mi-1.2</i> , <i>Mi-9</i> , <i>Ph2</i> , <i>I-2/I-2C</i> , <i>I-3/i-3</i> , <i>Cf4/Cf4A</i> , <i>Cf5</i> , <i>Cf9</i> , <i>Rx-4</i> , <i>Pto</i> , <i>Ve/ve</i>
L 41/3	<i>Mi-1.2</i> , <i>Mi-9</i> , <i>Ph2</i> , <i>Ph2(Ch)</i> , <i>Ph3</i> , <i>I-2</i> , <i>I-3/i-3</i> , <i>Cf5</i> , <i>Cf9</i> , <i>I-7</i> , <i>Ve</i>

сортов Ирма, L 22/11(2), L 41/3, LRs-30, Sile, L-43(3) характеризовались отсутствием или единичными пятнами поражения фитофторозом. У образцов L-43(2) и Daniel отмечались слабые симптомы вирусного поражения (12,5 и 7,1% соответственно). У образца Lyto, кроме фитофтороза, наблюдалось распространение бактериальной пятнистости у 23% растений.

Сопоставление результатов развития фитофтороза в полевых условиях с результатами молекулярного маркирования показало, что наименьшая степень развития болезни с полным отсутствием признаков на плодах на момент

учета была у образцов Ирма, L 41/3, LRs-30, каждый из которых содержал в генотипе аллель устойчивости к фитофторозу: *Ph2(Ch)*. Образец L 41/3 характеризовался наличием трех маркеров, связанных с устойчивостью к фитофторозу. Сравнение распространенности болезни у образцов с аллелями *Ph2* (LOC101246981) и *Ph2(Ch)* (LOC101246197) показывает достаточно высокий уровень резистентности у форм с аллелем *Ph2(Ch)* и низкий — с аллелем *Ph2*. Показана наибольшая эффективность разделения материала по R/S-аллелям гена *Ph3* с помощью нового

Таблица 4

Результаты полевой оценки устойчивости к болезням образцов томата (2025, Беларусь)

Сорт	Фитофтороз				Альтернариоз		Вирусные болезни	Бактериальные болезни
	развитие, балл	распространенность, % 19.08/3.09.2025			развитие, балл	распространенность, % 19.08.2025		
	листья	листья	стебли	плоды	листья	листья		
Дивиденд	3,0/3,2	100	100/100	80/80	2,8	100	—	—
Ирма	2,4/2,4	100	—/18	—/—	2,5	100	—	—
Alty	1,4/1,9	100	40/40	—/80	0,1	80	—	—
L 33/1	1,6/1,6	100	30/30	—/31	1,0	100	—	—
L 29/6	1,2/1,6	100	23/100	8/85	0,1	100	—	—
L 22/11	1,4/2,3	100	33/100	17/100	0,1	83,3	—	—
L 22/11	1,2/1,2	100	27/27	—/—	0,1	100	—	—
LRs-30	1,0/1,0	100	—/—	—/—	0,1	100	—	—
LR-13	1,2/1,7	100	—/—	—/62	0,1	30	—	—
Sile	1,4/1,6	100	14/86	—/14	1,0	100	—	—
Lyto	1,2/2,3	100	92/92	23/31	0,1	100	—	23
LD56	1,7/2,2	100	70/100	—/60	1,0	100	—	—
Pablo	1,1/1,7	100	11/100	6/66	0,1	100	—	—
Daniel	2,4/2,8	100	100/100	—/57	1,0	100	7,1	—
L 41/3	1,4/2,0	100	—/20	—/15	0,1	100	—	—
L-43(1)	1,8/1,8	100	15/15	—/15	0,1	100	—	—
L-43(2)	1,4/1,8	100	50/88	25/63	1,0	100	12,5	—
L-43(3)	2,2/3,8	100	91,7/100	83/100	1,0	100	—	—
Спажытак	1,8/3,4	100	30/100	60/100	0,1	100	—	—
Источник	1,2/1,7	100	15,3/100	15/100	0,2	100	—	—

Окончание таблицы 4

Сорт	Фитофтороз				Альтернариоз		Вирусные болезни	Бактериальные болезни
	развитие, балл	распространенность, % 19.08/3.09.2025			развитие, балл	распространенность, % 19.08.2025		
		листья	листья	стебли		листья	листья	
Зорка	2,4/3,6	100	100/100	45/100	2,8	100	—	—
Гарант	2,6/3,6	100	100/100	55/100	2,8	100	—	—
Чирок	2,4/3,2	100	80/100	40/95	2,4	100	—	—
Беркут	2,4/3,4	100	80/100	50/90	2,4	100	—	—
Тамара	2,8/3,8	100	100/100	60/100	2,6	100	—	—
Ред Гранд	2,2/3,2	100	75/100	30/88	2,4	100	—	—
Звездапад	2,6/3,4	100	100/100	40/100	2,8	100	—	—

Примечание. «—» симптомы заражения не обнаружены

SCAR-маркера ph3del22/11, а также ранее разработанного Ph3-SCAR маркера [16].

Результаты оценки устойчивости белорусских образцов к фитофторозу, альтернариозу, кладоспориозу, а также вирусу желтой курчавости листьев при их испытании в условиях Кубы представлены на рисунке 5. Согласно полученным данным, наибольшую степень

поражения грибными заболеваниями имел сорт Редгрант (3–4 балла). Сорт Беркут показал степень поражения фитофторозом и кладоспориозом на уровне трех баллов. При этом у большинства образцов признаков поражения указанными патогенами не наблюдалось. Низкое поражение такими грибными инфекциями, как *Phytophthora infestans*

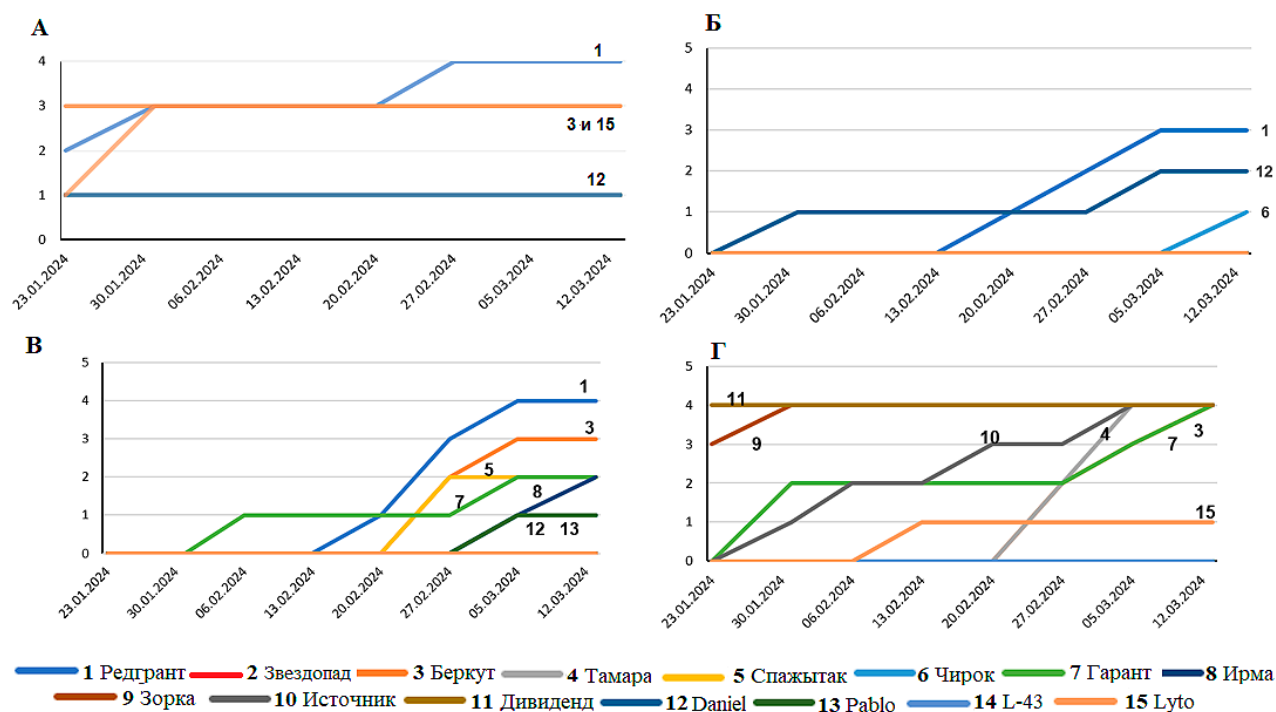


Рис. 5. Характер развития болезней у образцов томата: А — *Phytophthora infestans*, Б — *Alternaria alternata*, В — *Cladosporium fulvum*, Г — *Tomato Yellow Leaf Curl Virus*

и *Alternaria alternata*, связано прежде всего с сухими и жаркими условиями выращивания растений. Наиболее вредоносным в условиях Кубы для томата является вирус желтого скручивания листьев (*Tomato yellow leaf curl virus*, *TYLCV*). Поражение на уровне 4 балла наблюдалось у шести из одиннадцати испытываемых белорусских сортов: Дивиденд, Беркут, Гарант, Источник, Зорка, Тамара, при этом из четырех контрольных кубинских образцов поражение на уровне 1 балл имел сорт Lyto. Визуальных симптомов поражения *Fusarium oxysporum* (фузариозом) и *Verticillium dahliae*

(вертициллезом) не наблюдалось. Наилучшим образом в жарких и сухих условиях проявил себя сорт Ирма, на котором не было существенных проявлений поражения изучаемыми патогенами. Также результаты оценки признаков урожайности (табл. 5) показали его значительное превосходство над остальными образцами.

Согласно представленным данным, сорт Ирма превзошел по продуктивности и урожайности лучшие кубинские образцы Lyto и L-43 более чем в 2 раза. Существенное превышение лучших кубинских форм по продуктивности

Таблица 5

Показатели урожайности сортов белорусской селекции в условиях Кубы (Куба, 2024)

Наименование образцов	Средняя масса плода, г ± SX	Продуктивность, кг/растение	Урожайность, т/га
Редгрант	43,0 ± 10,2	1,29 ± 0,25	28,43 ± 1,63
Беркут	89,8 ± 12,4	2,16 ± 0,32	47,42 ± 2,26
Тамара	53,0 ± 8,6	1,59 ± 0,28	34,88 ± 1,03
Спажытак	40,5 ± 6,3	0,40 ± 0,18	8,90 ± 0,68
Чирок	48,0 ± 9,6	2,85 ± 0,34	62,68 ± 12,12
Гарант	34,5 ± 5,6	1,49 ± 0,26	32,71 ± 1,03
Ирма	33,0 ± 7,1	3,99 ± 0,32	87,74 ± 3,48
Звездопад	22,6 ± 6,3	0,40 ± 0,12	8,82 ± 1,54
Зорка	39,5 ± 8,6	1,43 ± 0,46	31,61 ± 1,73
Источник	27,1 ± 5,2	0,06 ± 0,02	1,40 ± 0,26
Дивиденд	125,4 ± 18,1	0,84 ± 0,12	18,38 ± 1,03
Daniel	102,0 ± 12,4	0,60 ± 0,14	13,25 ± 0,89
Pablo	90,7 ± 12,6	0,31 ± 0,12	6,86 ± 0,42
L-43	93,0 ± 10,2	1,90 ± 0,24	41,88 ± 1,26
Lyto	78,5 ± 6,4	1,79 ± 0,16	39,42 ± 2,03
Среднее	60,8	1,40	30,91
НСР₀₅	15,6	0,28	16,25

также наблюдалось у сортов Беркут и Чирок. Таким образом, по результатам оценки в двух экологических зонах сорт Ирма, содержащий аллель *Ph2(Ch)*, продемонстрировал максимальную устойчивость к болезням.

Заключение

На основании сопоставления результатов молекулярного анализа образцов и полевой

фитопатологической оценки показана высокая эффективность выявления форм с повышенной резистентностью с помощью разработанных ДНК-маркеров.

По результатам секвенирования гена *Ph-3* у образцов томата выявлены полиморфизмы (две делеции размером по 11 п. п.), приводящие к формированию нефункционального белка, что связано с усилением восприимчиво-

сти форм томата к возбудителю фитофтороза. Выявлен ранее неописанный аллель у формы LR-13, формирующий нефункциональный белок. Разработан и успешно апробирован SCAR-маркер *ph3 del22/11*, позволяющий типировать данный аллель наряду с известными ранее R/S-аллелями гена *Ph-3*. Показано, что типирование образцов с помощью маркеров, выявляющих наличие/отсутствие фрагмента LTR ретротранспозона семейства *Ty1- copia*, не связано напрямую с устойчивостью к фитофторозу.

Максимальная эффективность отбора форм с повышенной устойчивостью к фитофторозу связана с отбором форм с R-аллелем гена *Ph2(Ch)* (LOC101246197). Разработан CAPS-маркер *Ph-2(R)* *HindIII* *CC-Ase* для типирования аллелей R/S данного гена. Сорт Ирма, содержащий аллель *Ph2(Ch)*, продемонстрировал максимальную устойчивость к болезням при испытании в двух экологических зонах: умеренного и тропического климата.

В результате молекулярно-генетических исследований отобраны лучшие образцы томата из белорусско-кубинской коллекции для дальнейшего использования в селекционной работе, содержащие в генотипе 7–11 аллелей, связанных с устойчивостью к болезням.

Список использованных источников

1. Генетические основы селекции растений.
2. Частная генетика растений / А. В. Кильчевский, Л. В. Хотылева, М. М. Добродькин [и др.] // Нац. акад. наук Беларуси, Институт генетики и цитологии. – Минск : Беларуская навука. – 2010. – Гл. 9. Томат. – С. 422–508.
3. Оценка исходного материала томата методами классической и маркер-сопутствующей селекции для создания устойчивых к фитопатогенам гетерозисных гибридов / И. Е. Зайцева, И. Г. Пугачева, М. М. Добродькин [и др.] // Вестник БГСХА. – 2018. – № 2. – С. 135–139.
4. Panthee, D. R. A reexamination of molecular markers for use in marker-assisted breeding in tomato / D. R. Panthee, M. R. Foolad // *Euphytica*. – 2012. – Vol. 184, № 2. – P. 165–179.
5. Molecular markers associated with *Ph-3* gene conferring late blight resistance in tomato / D. R. Panthee, G. G. Randy, I. Ragy, A. Candice // *American Journal of Plant Sciences*. – 2015. – Vol. 6, № 13. – P. 2144–2150.
6. Оценка и применение молекулярных маркеров в селекции на устойчивость томата (*Solanum Lycopersicum* L.) к фитофторе (*Phytophthora Infestans*) / О. Г. Бабак, Е. В. Дрозд, Н. А. Некрашевич [и др.] // Молекулярная и прикладная генетика. – 2021. – Т. 31. – С. 22–30.
7. Fine mapping of the *Ph-2* gene conferring resistance to late blight (*Phytophthora infestans*) in tomato X. Zhi, J. Shu, Z. Zheng [et al.] / *Plant Disease*. – 2021. – Vol. 105, № 4. – P. 851–858.
8. Технология маркер-сопутствующего отбора форм томата с высокими биохимическими и технологическими свойствами плодов: Методические рекомендации / Министерство сельского хозяйства и продовольствия Республики Беларусь; Национальная академия наук Беларуси; Институт генетики и цитологии Национальной академии наук Беларуси; сост.: О. Г. Бабак, Н. А. Некрашевич, Н. В. Анисимова [и др.]. – Минск : Право и экономика, 2023. – 74 с.
9. Использование молекулярных маркеров, связанных с устойчивостью к биотическим и абиотическим факторам среды, при создании селекционного материала томата и перца в Беларуси / О. Г. Бабак, Е. В. Дрозд, Н. А. Некрашевич [и др.] // Овощи России. – 2025. – № 1. – С. 5–13.
10. GenBank : [сайт]. – 2024. – URL: <https://www.ncbi.nlm.nih.gov/genbank> (дата обращения: 10.09.2025).
11. Agrios, G. N. Plant Pathology / G. N. Agrios. – Amsterdam : Elsevier Academic Press, 2005. – P. 398–401.
12. Simmons, E. G. Alternaria: an identification manual / E. G. Simmons. – Netherlands : CBS Centre for Fungal Biodiversity, 2007. – 775 p.
13. Gallegly, M. E. Phytophthora: identifying species by morphology and DNA fingerprints / M. E. Gallegly, C. Hong. – Minnesota : American Phytopathological Society Press, 2008. – 158 p.
14. Ristaino, J. B. Key to the Identification of Common Phytophthora Species / J. B. Ristaino // American Phytopathological Society // J. B. Ristaino. – 2011. – 1 CD-ROM.
15. Методические указания по селекции и семеноводству овощных культур, возделываемых в защищённом грунте (томаты, перец) / А. В. Алпатьев, П. Ф. Сокол, В. В. Хренова [и др.]. – М. : ВАСХНИЛ, 1976. – 85 с.
16. Мартынов, В. В. Особенности первич-

ной структуры гена *Ph-3*, выявленные при создании нового маркера устойчивости томата к фитофторозу / В. В. Мартынов, Е. Г. Козарь, И. А. Енгальчева // Сельскохозяйственная биология. – 2022. – Т. 57, № 153. – С. 954–964.

16. Gene-based molecular marker system for

multiple disease resistances in tomato against Tomato yellow leaf curl virus, late blight, and verticillium wilt / H. J. Kim, J. M. Lee, C. S. Oh [et al.] // Euphytica. – 2015. – Vol. 205, № 2. – P. 599–613.

O. G. Babak¹, E. V. Drozd¹, N. A. Martinovskaya¹, N. V. Anisimova¹, K. K. Yatsevich¹, I. E. Bayeva², T. V. Nikonovich², I. G. Puhachova², M. M. Dobrodin², G. R. Rodríguez³, T. H. Ramos³, E. B. Gonzalez³, Y. C. Herrera³, A. V. Kilchevsky¹

EMPLOYING MOLECULAR MARKING AND PHYTOPATHOLOGICAL EVALUATION METHODS IN THE DEVELOPMENT OF NEW TOMATO BREEDING MATERIAL WITH COMPLEX DISEASE RESISTANCE

¹State Scientific Institution

“Institute of Genetics and Cytology of the National Academy of Sciences of Belarus”

27 Akademicheskaya St., 220072 Minsk, the Republic of Belarus

e-mail: O.Babak@igc.by

²Educational Institution

“Belarusian State Agricultural Academy”

5 Michurin St., 213410 Gorki, the Republic of Belarus

³“Liliana Dimitrova” Horticultural Research Institute

Bejucal-Quivicán highway, Km 33 ½, Quivicán Municipality, Mayabeque Province, the Republic of Cuba

Based on the sequencing results of the R-genes *Ph2(Ch)* (LOC101246197) and *Ph3* (KJ563933) of the tomato samples of Belarusian and Cuban selection, structural features of alleles associated with the control of samples resistance to *Phytophthora infestans* were identified. The CAPS marker Ph-2(R)_HindIII and the SCAR marker Ph3 del22/11, allowing the typing of alleles associated with resistance/susceptibility to this pathogen, have been developed. DNA typing of the Belarusian–Cuban collection samples for the presence of disease-resistant alleles made it possible to identify the forms combining 7–9 disease resistant alleles: Lyto (7 R), LRs-30 (7 R), Pablo (8 R), Alty (8 R), Sile (9 R) and LR-13 (9 R). The Belarusian variety Irma demonstrated the highest disease resistance and productivity in the tropical climate of the Republic of Cuba. New breeding material with enhanced resistance to late blight in the conditions of the Republic of Belarus was produced.

Keywords: tomato, late blight resistance genes, DNA typing, markers of disease resistance genes, breeding material.

Дата поступления в редакцию: 17 сентября 2025 г.

А. Н. Верчук¹, А. В. Кильчевский¹, С. В. Кубрак¹, Н. В. Савина¹, Л. В. Милько¹, С. С. Савчук²,
В. Н. Лебедько², Д. В. Дубовик², Е. В. Никитина³

ГЕНОТИПИРОВАНИЕ ЛЕКАРСТВЕННЫХ РАСТЕНИЙ РОДА *THYMUS* L., ПРОИЗРАСТАЮЩИХ НА ТЕРРИТОРИИ БЕЛАРУСИ И УЗБЕКИСТАНА

¹Государственное научное учреждение

«Институт генетики и цитологии Национальной академии наук Беларуси»

Республика Беларусь, 220072, г. Минск, ул. Академическая, 27

e-mail: a.n.verchuk@mail.ru

²Государственное научное учреждение

«Институт экспериментальной ботаники им. В.Ф. Купревича Национальной академии наук Беларуси»

Республика Беларусь, 220072, г. Минск, ул. Академическая, 27

³Институт ботаники Академии наук Республики Узбекистан

Республика Узбекистан, 100128, г. Ташкент, ул. Дурмон йули, 32

В работе представлены результаты изучения последовательностей трех ДНК-штрихкодов — *matK*, *psbA-trnH* и *ITS1-4* — на видах рода *Thymus*, произрастающих на территории Беларуси и Узбекистана, с целью идентификации спорных таксонов и подбора эффективных ДНК-маркеров. Все ДНК-штрихкоды выявили относительно высокую внутривидовую и низкую межвидовую генетическую изменчивость в пределах рода *Thymus*. Наименьшую дистанцию между видами показали маркеры *matK* и *ITS1-4*, что указывает на их недостаточную информативность при ДНК-штрихкодировании данной группы растений. Наибольшую информационную ценность представляет маркер *psbA-trnH*, который можно рекомендовать в качестве компонента двух-трех-локусной системы ДНК-штрихкодов для идентификации таких таксономически сложных групп, как род *Thymus*.

Ключевые слова: ДНК-штрихкодирование, *ITS1-4*, *matK*, *psbA-trnH*, *Thymus*.

Для цитирования: Генотипирование лекарственных растений рода *Thymus* L., произрастающих на территории Беларуси и Узбекистана / А. Н. Верчук, А. В. Кильчевский, С. В. Кубрак [и др.] // Молекулярная и прикладная генетика: сб. науч. тр. / Ин-т генетики и цитологии НАН Беларуси; редкол.: А. В. Кильчевский (гл. ред.) [и др.]. – Минск, 2025. – Т. 39. – С. 21–32.

Введение

Род *Thymus* L. (Тимьян) — крупный, полиморфный и таксономически сложный род семейства *Lamiaceae* (Яснотковые). В настоящее время род насчитывает около 350 видов, относится к трибе *Mentheae* семейства яснотковых (*Lamiaceae*), подсемейству *Nepetoideae* [1–3]. Представители рода — ароматные, тонковетвистые полукустарники. Отдельные виды *Thymus* используются как ароматические и пряные растения, а их фармакологические свойства известны с древности, благодаря чему они находят широкое применение в традиционной медицине. В связи с этим тимьяны представляют значительный интерес для фармацевтической, пищевой и парфюмерной промышленности [4–6]. Из растительного

сырья тимьяна получают препараты, обладающие антисептическим, дезинфицирующим, обезболивающим действием; сырье используют как в традиционной, так и в народной медицине. Виды тимьяна способны расти в различных экологических условиях (лесные поляны и опушки, боровые пески, песчаные и глинистые степи, каменистые склоны, скалы, карбонатные обнажения, горные тундры) и распространены по всей Евразии, в Северной Африке и на юге Гренландии [1, 2].

На территории Беларуси встречаются четыре близких вида рода *Thymus*: тимьян ползучий (*Thymus serpyllum* L.), тимьян блошиный (*T. pulegioides* L.), тимьян Маршалла (*T. marschallianus* Willd.) и тимьян обыкновенный (*T. vulgaris* L.).

Первые два считаются аборигенными и наиболее распространенными, тимьян Маршалла встречается редко, тимьян обыкновенный — средиземноморский вид, культивируемый в Беларуси [7]. Виды тимьян ползучий (*T. serpyllum*) и тимьян обыкновенный (*T. vulgaris*) являются разрешенными к медицинскому применению в Беларуси [8].

В пределах Узбекистана достоверно отмечены четыре вида рода *Thymus*: *T. seravschanicus* Klokov, *T. incertus* Klokov, *T. dmitrievae* Gamajun. и *T. subnervosus* Vved., Nabiev & Tulyag. (эндемик Нуратау). При этом 10-томное издание «Флора Узбекистана» (1961) [9] указывало для территории страны лишь два вида — *Th. seravschanicus* и *Th. incertus*. В кадастрах региональных флор, для флоры Самаркандской и Навоийской областей приводятся *T. seravschanicus* и *T. subnervosus*; для флоры Ташкентской области — *T. seravschanicus*, *T. incertus*. Вид *T. dmitrievae* по имеющимся сборам известен исключительно из Ташкентской области [10]. Ряд названий, ранее трактованных как самостоятельные (*T. bucharicus* Klokov, *T. aschurbajevii* Klokov, *T. cuneatus* Klokov; «Флора СССР» (1954)) [11], в современных сводках рассматриваются как синонимы *T. seravschanicus*, что влияет на итоговое число видов в пределах страны.

Поскольку виды рода *Thymus* относятся к категории лекарственных и ароматических растений, традиционно научный интерес направлен на выяснение биохимического состава. Основными компонентами эфирного масла растений рода *Thymus* являются фенолы тимол и карвакрол, содержание которых может достигать до 30%. Различные виды тимьянов отличаются как по общему содержанию эфирного масла, так и по его количественному и качественному составу, при этом полиморфизм компонентного состава эфирного масла наблюдается как на межвидовом, так и на внутривидовом уровне [12].

Растения рода *Thymus* идентичны по внешнему виду, что затрудняет их идентификацию, поэтому при заготовке аптечного сырья не всегда возможно различить виды тимьянов, часть из которых не разрешена к официальному применению. Кроме того, род *Thymus* является примером сложной группы, охватывающей множество таксонов, связанных

между собой гибридизацией, перекрывающейся фенотипической изменчивостью и другими признаками сетчатой эволюции, что усложняет задачу таксономистов. Некоторые представители тимусов до сих пор имеют неопределенный статус [13].

Разрешение проблем идентификации и уточнения таксономического статуса требует сочетания классических методов систематики с молекулярно-генетическими подходами. Применение штрихкодирования к растениям таксономически сложных групп сталкивается с рядом трудностей. Выбор подходящего маркера, способного различить близкородственные и морфологически схожие виды, зависит от многих факторов, среди которых полиплоидия, степень родства интересующих таксонов, поток генов, гибридизация [14]. В связи с этим цель данного исследования состояла в идентификации отдельных видов рода *Thymus* L., произрастающих на территории Беларуси и Узбекистана, с помощью трех маркеров: *ITS1-4*, *matK*, *psbA-trnH*. Ранее генотипирование этих видов белорусской флоры с помощью метода ДНК-штрихкодирования не проводилось.

Материалы и методы

В данное исследование включено пять видов семейства *Thymus* белорусской флоры: *Thymus marschallianus* Willd., *Thymus vulgaris* L., *Thymus serpyllum* L., *Thymus pulegioides* L., *Thymus oblongifolius* (Hybrid *Thymus ovatus* Mill. × *Thymus serpyllum* L.) и два вида флоры Узбекистана: *Thymus incertus* Klok., *Thymus seravschanicus* Klok.

Тотальную ДНК выделяли из высушенной в силикагеле растительной ткани с использованием коммерческого набора GeneJET Plant Genomic DNA Purification Kit (Thermo Fisher Scientific, США) согласно протоколу производителя. Амплификацию ДНК-штрихкодов проводили в финальном объеме 8–10 мкл; основные компоненты — готовая смесь для проведения ПЦР Quick-load Taq 2X Master Mix (Праймтех, Беларусь), праймеры в конечной концентрации 0,3 пмоль/мкл и матрица ДНК в конечной концентрации 10 нг/мкл.

В качестве генетических маркеров использовали последовательности универсальных штрихкодов для растений: *psbA-trnH*, *matK*

и *ITS1-4*. Последовательности праймеров для амплификации целевых продуктов и размеры ожидаемых фрагментов приведены в таблице 1.

Качество продуктов амплификации проверяли в 2%-ом агарозном геле. Полученные ПЦР-продукты очищали набором для ферментной очистки (Exo I и rSAP). Образцы секвенировали в прямом и обратном направлении с использованием набора реагентов Brilliant Dye Terminator v3.1 Cycle Sequencing Kit (Nimagen, Нидерланды) с последующим анализом продуктов на автоматическом анализаторе ABI 3500 DNA Analyzer (Applied Biosystems, США). Хроматограммы, полученные в результате секвенирования, были проанализированы в Sequence Scanner 1.0. Полученные последовательности сравнивали с последовательностями в международных базах данных NCBI BLAST (Basic Local Alignment Search Tool) [15]. Множественное

выравнивание последовательностей и построение дендрограмм проводили с помощью программы MEGA X v.10.2.4, дендрограммы строили на основе матрицы попарных генетических расстояний методом ближайшего соседа (Neighbour-Joining).

Для проведения филогенетического анализа исследуемая выборка была расширена нуклеотидными последовательностями представителей рода *Thymus*, дополнительно взятыми из базы данных Национального центра биотехнологической информации NCBI (National Center for Biotechnology Information).

Статистический анализ был проведен в программе MEGA X v.10.2.4 на основании двухпараметрической модели Кимуры (K2P, Kimura 2-parameter) с построением политомической филограммы, используя дискретный метод наибольшего правдоподобия (Test Maximum likelihood method).

Таблица 1

Праймеры для амплификации маркерных последовательностей ДНК-штрихкодов

ДНК-штрихкод	Последовательность праймеров (прямой / обратный)	Ожидаемый размер ПЦР-продукта, п. о.
<i>psbA-trnH</i>	<i>psbA3_F</i> 5'-GTTATGCATGAACGTAATGCTC-3' <i>trnHf_05</i> 5'-CGCGCATGGTGGATTCAATCC-3'	~350–400
<i>ITS1-4</i>	<i>ITS1_18S</i> 5'-TCCGTAGGTGAACCTGCGG-3' <i>ITS4_26S</i> 5'-TCCTCCGCTTATTGATATGC-3'	~600
<i>matK</i>	<i>matK-390f</i> 5'-CGATCTATTCAATTCATTTTC-3' <i>matK-1326r</i> 5'-TCTAGCACACGAAAGAAGT-3'	~740–800

Результаты и обсуждение

Род *Thymus* L. является одним из самых многочисленных в семействе Яснотковые (*Lamiaceae*), представлен низкорослыми кустарничками и полукустарничками, а также травянистыми растениями. Род характеризуется космополитичностью, но принято считать, что центр разнообразия находится в Средиземноморье. При удалении от Средиземноморского бассейна разнообразие видов уменьшается. Представители рода распространены почти по всей Евразии (за исключением тропиков), в Северной Африке и Гренландии [16]. Род *Thymus* представляет собой так называемую taxonomically complex groups (TCGs) — одну из сложных в систематическом отношении

групп [13, 17]. Характеристика видов осложняется из-за высокой изменчивости морфологических признаков у представителей рода. Существует большое количество морфологически трудноразличимых форм. Отдельные виды тимьянов легко образуют гибриды, при этом появляется большое количество переходных форм. При перекрывании ареалов отдельных видов тимьянов образуются устойчивые формы, имеющие гибридное происхождение и статус вида. В целях усовершенствования систематики вида одни авторы выделяют множество мелких видов, в то время как другие предлагают определить их в подвиды и гибридные формы [12, 18].

По таксономической системе классификации

цветковых растений, учитывающей результаты молекулярного анализа ДНК [19], род насчитывает 341 подтвержденный вид и 182 таксона, ожидающих подтверждения. Род относится к трибе *Mentheae*, подсемейству *Nepetoideae* и разделен на восемь секций: *Micantes*, *Mastichina*, *Piperella*, *Thymus*, *Teucrioides*, *Pseudothymra*, *Hyphodromi* и *Serpyllum*. Многие виды, включая широко распространенный тимьян ползучий, отличаются большим разнообразием трудно дифференцированных форм, что приводит к противоречивым данным по видовой численности в роду. Сложная систематика *Thymus* в настоящее время продолжает уточняться с помощью методов молекулярной биологии [13, 20].

Согласно данным литературы, наиболее используемыми для генетической идентификации видовой принадлежности растений рода тимус являются ДНК-штрихкоды ядерной

последовательности *ITS1-4* и два пластидных маркера *psbA-trnH* и *matK*. Однако результаты применения этих последовательностей для различения видов в пределах *Thymus* spp. в работах ряда авторов неоднозначны [13, 17, 20].

В нашем исследовании длина видовых консенсусных маркерных последовательностей составила в среднем 746 п. н. для *matK*, 338 п. н. для *psbA-trnH* и 536 п. н. для *ITS1-4*. Наибольшую вариабельность и наибольшее число парсимонически значимых сайтов проявил маркер *psbA-trnH* — 28,57%. Это согласуется с работой Furan М. А. (2023), в которой именно *psbA-trnH* показал наибольшее количество мутаций и самую низкую долю консервативных участков по сравнению с другими маркерами. Для двух других маркеров — *matK* и *ITS1-4* — число парсимонически значимых сайтов в нашем исследовании составило менее 1% (табл. 2).

Таблица 2

Вариабельность ДНК-штрихкодов

ДНК-маркер	Длина последовательности, п. н.	Количество вариабельных сайтов	Парсимонически значимые сайты	
			количество	%
<i>matK</i>	746	8	1	0,13
<i>psbA-trnH</i>	338	49	14	28,57
<i>ITS1-4</i>	536	7	2	0,37

Для расширенного филогенетического анализа исследуемая выборка видов *Thymus* была дополнена аналогичными последовательностями из NCBI. Для каждого маркера на основании его вариабельности были рассчитаны значения внутривидовых и межвидовых дистанций, построены политомические филограммы на основе дискретного метода наибольшего правдоподобия и двухпараметрической модели Кимуры.

Согласно полученным данным, наибольшие внутривидовые различия по маркеру *ITS1-4* были отмечены у *T. marschallianus* (0,00604 или 0,6%). Виды *T. vulgaris*, *T. serpyllum*, *T. pulegioides* характеризовались относительно одинаковым значением внутривидового полиморфизма, у *T. seravshanicus* два проанализированных сиквенса были идентичны (табл. 3).

Межвидовой полиморфизм по ядерному маркеру *ITS1-4* был наиболее выражен меж-

Таблица 3

Значения межвидовых и внутривидовых дистанций представителей рода *Thymus*, рассчитанные по вариабельности маркера *ITS1-4*

Виды	<i>Thymus vulgaris</i>	<i>Thymus serpyllum</i>	<i>Thymus seravshanicus</i>	<i>Thymus pulegioides</i>	<i>Thymus oblongifolius</i>	<i>Thymus marschallianus</i>
Внутривидовая дистанция	0,00420 ± 0,00160	0,00442 ± 0,00170	0,0 ± 0,0	0,00452 ± 0,00222	n/c	0,00604 ± 0,00292

Окончание таблицы 3

Виды	<i>Thymus vulgaris</i>	<i>Thymus serpyllum</i>	<i>Thymus seravshanicus</i>	<i>Thymus pulegioides</i>	<i>Thymus oblongifolius</i>	<i>Thymus marschallianus</i>
<i>T. vulgaris</i> (12)		0,00333*	0,00395*	0,00436*	0,00088*	0,00360*
<i>T. serpyllum</i> (12)	0,00926		0,00195*	0,00287*	0,00324*	0,00183*
<i>T. seravshanicus</i> (2)	0,00907	0,00358		0,00387*	0,00388*	0,00206*
<i>T. pulegioides</i> (4)	0,01250	0,00641	0,00793		0,00431*	0,00344*
<i>T. oblongifolius</i> (1)	0,00227	0,00697	0,00678	0,01020		0,00352*
<i>T. marschallianus</i> (3)	0,00984	0,00497	0,00376	0,00869	0,00754	

Примечание. В первом столбце в скобках указано количество проанализированных сиквенсов; под диагональю — значения межвидовых дистанций, над диагональю — значения соответствующих стандартных ошибок; п/с — рассчитать значение внутривидовой дистанции невозможно, вид представлен одной последовательностью; * — значения стандартных ошибок

ду *T. vulgaris* и видами: *T. pulegioides* (1,25%), *T. serpyllum* (0,93%), *T. seravshanicus* (0,91%), *T. marschallianus* (0,98%) и между видами *T. oblongifolius* и *T. pulegioides* (1,02%). Тем не менее, разрешающая способность маркера оказалась низкой. Среднее значение межвидовой дистанции для маркера *ITS1-4* составило 0,73%. Соотношение межвидовой и внутривидовой изменчивости для маркера *ITS1-4* на анализируемой выборке было низким — 1,91, что свидетельствует об отсутствии четкого разрыва и, соответственно, ограничивает индивидуальное применение данного маркера.

На филограмме, построенной на основании межвидовых дистанций видов *Thymus* по маркеру *ITS1-4* (рис. 1), выделилось два крупных кластера, один из которых почти полностью образован представителями вида *Thymus vulgaris*. В пределах этого кластера, в свою очередь, выделяются несколько групп ветвей, образованных представителями вида *T. vulgaris*, имеющих наибольшее внутривидовое сходство. Таким образом, образцы *T. vulgaris* сформировали собственную кладу с уникальными гаплотипами. Однако основание клады, образованное большинством образцов, включает в себя дополнительно последовательность *Thymus oblongifolius*, что может затруднить видовую идентификацию этих двух видов.

Второй крупный кластер разделен на два подкластера, один из которых образован преимущественно образцами *Thymus serpyllum*. Во второй подкластер включены все осталь-

ные виды, которые не образуют отдельных ветвей и практически не имеют дистанции друг с другом, что затрудняет их точную дифференциацию.

Филогенетический анализ, проведенный для собственных (полученных авторами) ядерных последовательностей, позволил разделить исследуемые виды только на две группы, соответствующие существующим секциям рода *Thymus*: секция 1 — *Thymus serpyllum* L.; секция 2 — остальные образцы, не дифференцированные данным маркером.

По хлоропластному маркеру *matK* наибольший внутривидовой полиморфизм отмечен у *T. serpyllum* (0,63%). Два образца *T. marschallianus* были идентичны, семь образцов *T. vulgaris* и пять образцов *T. pulegioides* характеризовались относительно невысоким значением внутривидового полиморфизма (табл. 4).

Межвидовой полиморфизм по маркеру *matK* был наиболее выражен между видом *T. serpyllum* и видами *T. pulegioides* (0,53%), *T. marschallianus* (0,53%). Наиболее близкими по *matK* оказались виды *T. seravshanicus* и *T. oblongifolius*, *T. seravshanicus* и *T. marschallianus*. Среднее значение межвидовой дистанции для маркера *matK* было 0,35%. Соотношение межвидовой и внутривидовой изменчивости для *matK* в нашем исследовании составило 1,05, что свидетельствует об отсутствии четкого разрыва между видами по данному маркеру и ограничивает его индивидуальное применение.

На построенной по маркеру *matK* филограм-

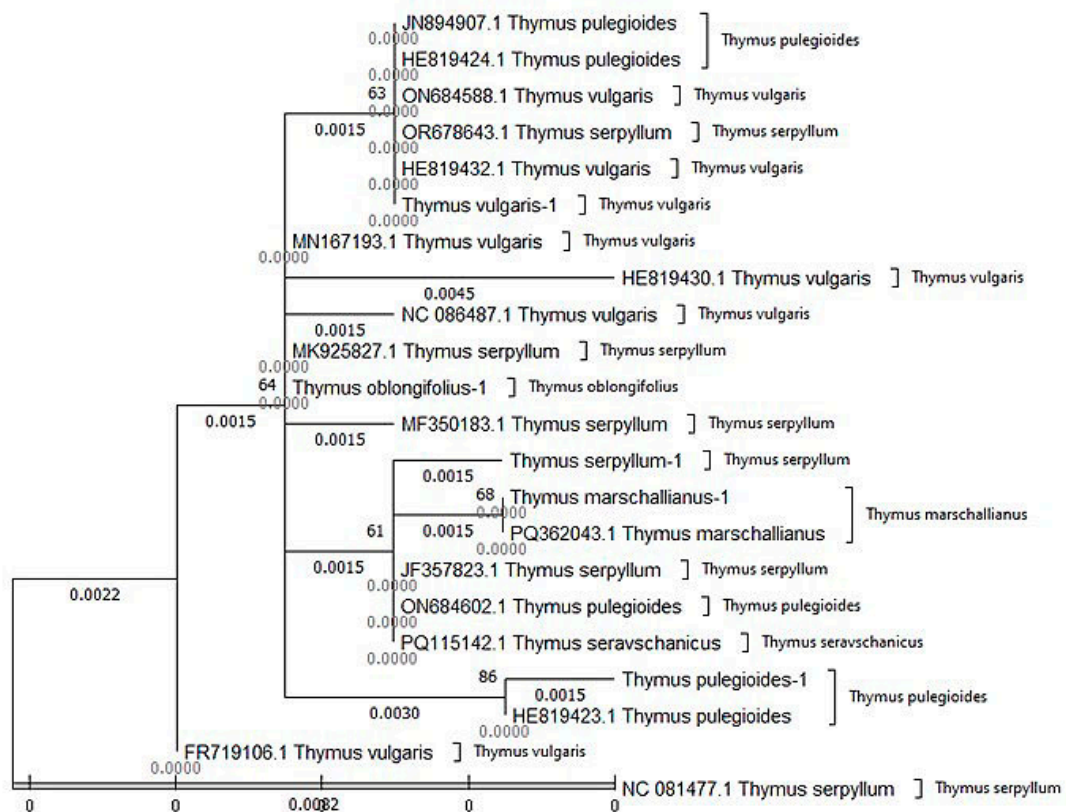


Рис. 1. Политомическая филограмма, построенная дискретным методом наибольшего правдоподобия по маркеру *ITS1-4*

Таблица 4

Значения межвидовых и внутривидовых дистанций представителей рода *Thymus*, рассчитанные по вариабельности маркера *matK*

Виды	<i>Thymus vulgaris</i>	<i>Thymus serpyllum</i>	<i>Thymus seravschanicus</i>	<i>Thymus pulegioides</i>	<i>Thymus oblongifolius</i>	<i>Thymus marschallianus</i>
Внутривидовая дистанция	0,003 ± 0,00127	0,00634 ± 0,0019	n/c	0,00391 ± 0,00174	n/c	0,0 ± 0,0
<i>T. vulgaris</i> (12)		0,00127*	0,00163*	0,00132*	0,00082*	0,00232*
<i>T. serpyllum</i> (12)	0,00470		0,00126*	0,00148*	0,00095*	0,00204*
<i>T. seravschanicus</i> (2)	0,00322	0,00377		0,00157*	0,00145*	0,00153*
<i>T. pulegioides</i> (4)	0,00361	0,00528	0,00331		0,00115*	0,00225*
<i>T. oblongifolius</i> (1)	0,00172	0,00327	0,00150	0,00240		0,00218*
<i>T. marschallianus</i> (3)	0,00472	0,00528	0,00150	0,00481	0,00300	

Примечание. В первом столбце в скобках указано количество проанализированных сиквенсов; под диагональю — значения межвидовых дистанций, над диагональю — значения соответствующих стандартных ошибок; n/c — рассчитать значение внутривидовой дистанции невозможно, вид представлен одной последовательностью; * — значение стандартных ошибок

ме (рис. 2) один из образцов *T. serpyllum* образовал внешнюю ветвь по отношению ко всем остальным образцам, аналогичным образом выделился один из образцов *T. vulgaris*.

Остальные образцы сформировали общий кластер, внутри которого можно выделить четыре неоднородных подкластера. Образцы *T. vulgaris* распределились между двумя подкластерами, в пределах которых они слабо отличались от образцов *T. serpyllum*, *T. pulegioides* и *T. oblongifolius*. Шесть исследуемых образцов *Thymus serpyllum* уверенно разделяются только в половине случаев, остальные последовательности имеют нулевую дистанцию с последовательностями, происходящими от иных видов. Секвенированная последовательность *T. seravshanicus*, хоть и находится в одном подкластере с вида-

ми *T. serpyllum* и *T. pulegioides*, но в отличие от них, имеет шестинуклеотидную вставку (TGG AAT). Два образца *T. pulegioides* формируют значительно удаленную от остальных образцов ветвь, достаточную для однозначного определения видовой принадлежности, однако остальные три образца попали в ветви, образованные последовательностями иных исследуемых видов. Единственный образец *T. oblongifolius* не образует собственной ветви, однако на конце имеет однонуклеотидную вставку (G), которая позволяет идентифицировать последовательность данного вида. Обе последовательности *T. marschallianus* сформировали отдельную ветвь.

По хлоропластному маркеру *psbA-trnH* наибольшие внутривидовые различия наблюдались у *T. serpyllum* (табл. 5).

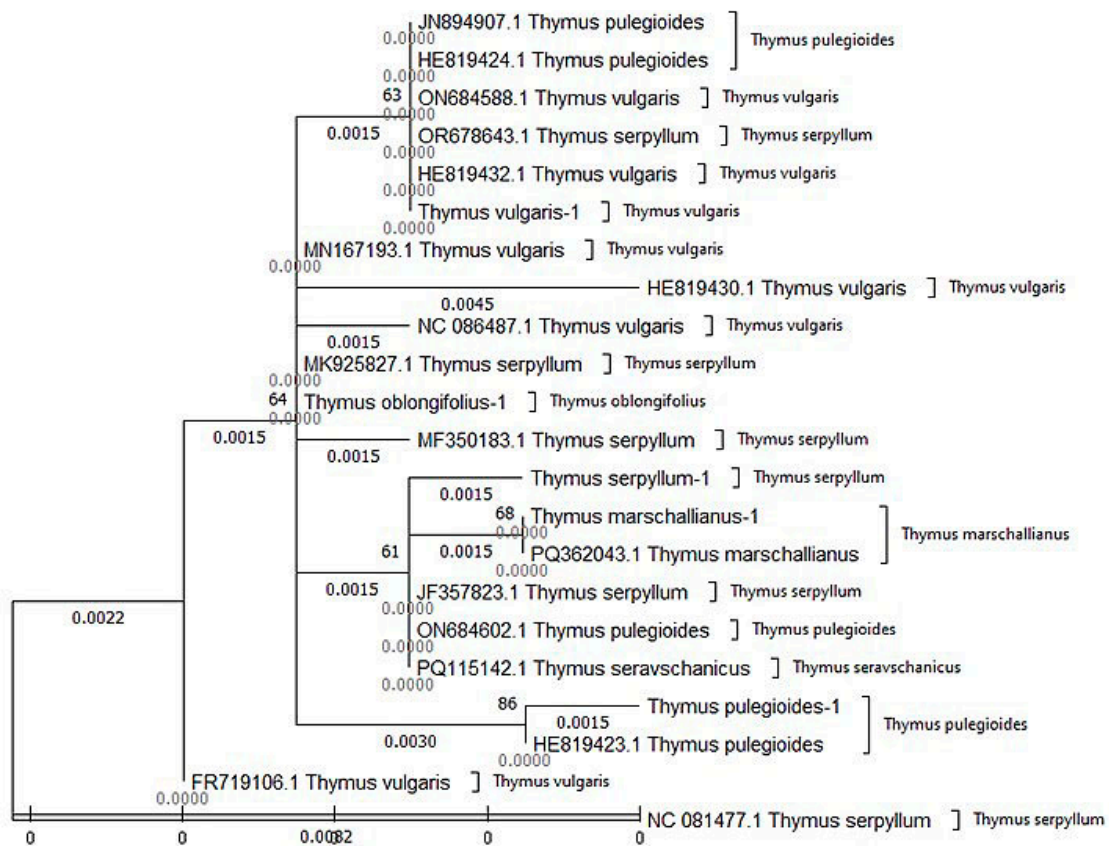


Рис. 2. Политомическая филограмма, построенная дискретным методом наибольшего правдоподобия по маркеру *matK*

Межвидовой полиморфизм был наиболее выражен между *T. marschallianus* и всеми остальными видами: *T. oblongifolius* (6,53%), *T. vulgaris* (7,03%), *T. pulegioides* (7,05%),

T. seravshanicus (7,94%), *T. serpyllum* (8,25%), *T. incertus* (8,97%). Межвидовой полиморфизм между *T. incertus* и остальными видами был немного ниже — в среднем 6,11%. Среднее

Таблица 5

Значения межвидовых и внутривидовых дистанций представителей рода *Thymus*, рассчитанные по варибельности *psbA-trnH*

Виды	<i>Thymus vulgaris</i>	<i>Thymus serpyllum</i>	<i>Thymus seravshanicus</i>	<i>Thymus pulegioides</i>	<i>Thymus oblongifolius</i>	<i>Thymus marschallianus</i>	<i>Thymus incertus</i>
Внутривидовая дистанция	0,01686 ± 0,00566	0,03302 ± 0,00840	0,00369 ± 0,00543	0,016 ± 0,00543	n/c	n/c	n/c
<i>T. vulgaris</i> (4)		0,0058*	0,0063*	0,0042*	0,0081*	0,0165*	0,0129*
<i>T. serpyllum</i> (4)	0,0252		0,0045*	0,0056*	0,0092*	0,0170*	0,0132*
<i>T. seravshanicus</i> (2)	0,0178	0,0180		0,0050*	0,0094*	0,0180*	0,0154*
<i>T. pulegioides</i> (7)	0,0148	0,0231	0,0136		0,0083*	0,0164*	0,0129*
<i>T. oblongifolius</i> (1)	0,0235	0,0341	0,0243	0,0246		0,0163*	0,0165*
<i>T. marschallianus</i> (1)	0,0703	0,0825	0,0794	0,0705	0,0653		0,0193*
<i>T. incertus</i> (1)	0,0476	0,0555	0,0594	0,0485	0,0656	0,0897	

Примечание. В первом столбце в скобках указано количество проанализированных сиквенсов; под диагональю — значения межвидовых дистанций, над диагональю — значения соответствующих стандартных ошибок; n/c — рассчитать значение внутривидовой дистанции невозможно, вид представлен одной последовательностью; * — значение стандартных ошибок

значение межвидовой дистанции в целом для маркера *psbA-trnH* было 4,54%. Соотношение межвидовой и внутривидовой изменчивости для *psbA-trnH* составило 2,61, что считается небольшим значением для видовой идентификации, однако в нашем исследовании этот показатель оказался наиболее высоким среди всех исследуемых маркеров.

На филограмме (рис. 3), построенной по маркеру *psbA-trnH*, внешнюю ветвь образовал вид *T. marschallianus*. Остальные образцы сгруппировались в два кластера. Четыре исследуемых образца *T. vulgaris* вошли в первый кластер, при этом два из них обладают уникальными для данного вида гаплотипами и могут безошибочно определяться.

Оставшиеся два образца не разделяются и максимально схожи с образцами *T. pulegioides*. Вид *T. pulegioides*, который представлен семью образцами, имеет только два видоспецифичных гаплотипа, вошедших обособленной группой

во второй кластер. Остальные нуклеотидные последовательности *T. pulegioides* не имели удаления от видов *T. vulgaris* и *T. seravshanicus*. Вид *T. serpyllum*, представленный в исследовании четырьмя последовательностями, обладает уникальными гаплотипами, из которых одна образует отдельную ветвь с *T. incertus*, а три близки по данному маркеру *T. seravshanicus*. В свою очередь, обе последовательности *T. seravshanicus* имели минимальные дистанции с последовательностями видов *T. serpyllum* и *T. pulegioides*. Виды, представленные по одному образцу — *T. oblongifolius*, *T. marschallianus*, *T. incertus* — имели уникальные для данного исследования гаплотипы и успешно отделялись от других видов.

При филогенетическом анализе полученных авторами последовательностей *psbA-trnH* без учета депонированных из NCBI, получено аналогичное разделение исследуемых видов на группы (рис. 4).

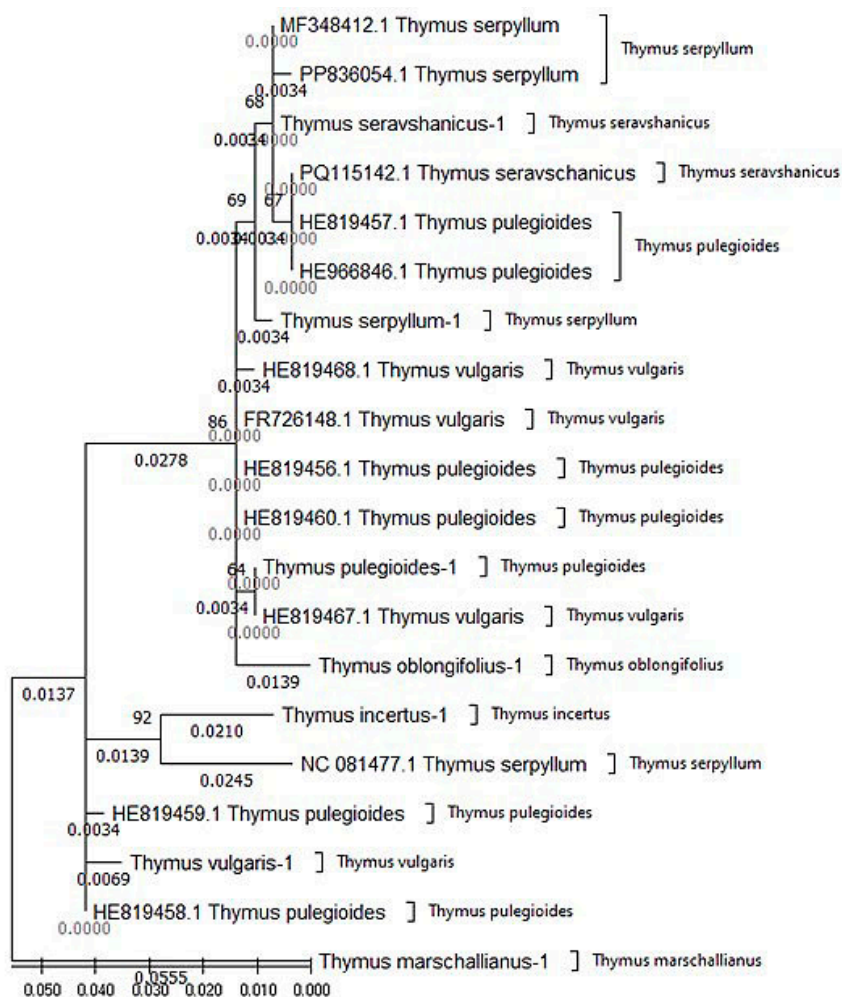


Рис. 3. Политомическая филограмма, построенная дискретным методом наибольшего правдоподобия по маркеру *psbA-trnH*

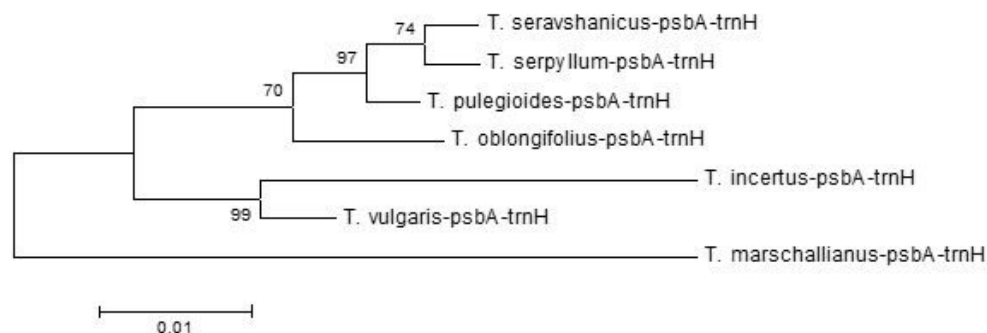


Рис. 4. Дендрограмма, построенная на основе сравнения последовательностей региона *psbA-trnH* у семи представителей рода *Thymus*

Заключение

Таким образом, генетические расстояния между видами рода *Thymus* различались в зависимости от используемых ДНК-штрихко-

дов. Маркеры *ITS1-4* и *matK*, рекомендуемые в литературе для идентификации тимьянов, показали самый низкий уровень генетической дифференциации на анализируемых нами ви-

дах. Только с помощью маркера *psbA-trnH* выявлен «пробел штрихкодирования» (barcoding gap), позволяющий дифференцировать таксоны на уровне видов. Также *psbA-trnH* продемонстрировал наибольшую межвидовую дистанцию, однако его внутривидовая изменчивость сопоставима с межвидовой, что может снижать разрешающую способность ДНК-маркера. Наименьшую межвидовую дистанцию выявил *matK*, что ставит под сомнение его эффективность для разграничения видов в исследуемых таксонах. *ITS1-4* показал неожиданно низкую межвидовую дистанцию, что также указывает на его недостаточную информативность для ДНК-штрихкодирования в данной группе растений. Тем не менее, межвидовая дистанция по всем трем маркерам выше нуля, что говорит о наличии полиморфизма в нуклеотидном составе каждого вида.

Несмотря на эффективность *psbA-trnH* в данном исследовании, в результате филогенетического анализа по этому маркеру морфологически различные виды, принадлежащие к разным секциям, были сгруппированы вместе, а близкородственные виды оказались генетически различными. Подобные результаты типичны для таксономически сложных групп, к которым и относят род *Thymus*, и все еще требуют всестороннего исследования. Высокая внутривидовая и относительно низкая межвидовая генетическая изменчивость не позволяет выделить дискретные видоспецифичные ДНК-штрихкоды. Поэтому для идентификации сложных таксонов рекомендуется подбирать комбинацию из нескольких ДНК-штрихкодов. Одним из них можно считать маркер *psbA-trnH*, который согласно нашим результатам, имеет информационную ценность для таксономической классификации тимьянов.

Работа выполнена при финансовой поддержке Белорусского Республиканского Фонда Фундаментальных Исследований (договор с БРФФИ № Б23УЗБ-030 от 20.11.2023 г.).

Список использованных источников

1. Bartolucci, F. Typification of names and taxonomic notes within the genus *Thymus* L. (*Lamiaceae*) / F. Bartolucci, L. Peruzzi, N. Pas-salacqua // *Taxon*. – 2013. – Vol. 62, № 6. – P. 1168–1174.
2. Characterization of the complete chloroplast genome of the important herbal medicinal plant, *Thymus quinquecostatus* (*Lamiaceae*) / J.-H. Song, S. Yang, G. Choi, I. Park // *Journal of Asia-Pacific Biodiversity*. – 2023. – Vol. 16, № 1. – P. 140–143.
3. The Families and Genera of Vascular Plants / R. Harley, S. Atkins, A. Budantsev, P. D. Cantino. – Berlin: Springer, 2004. – P. 167–275.
4. Bower, A. The health benefits of selected culinary herbs and spices found in the traditional Mediterranean diet / A. Bower, S. Marquez, E. Gonzalez de Mejia // *Critical Reviews in Food Science and Nutrition*. – 2016. – Vol. 56, № 16. – P. 2728–2746.
5. Nieto, G. A review on applications and uses of *Thymus* in the food industry / G. Nieto // *Plants*. – 2020. – Vol. 9, № 8. – P. 961. – DOI: 10.3390/plants9080961.
6. Phytochemical Fingerprinting and *In Vitro* Antimicrobial and Antioxidant Activity of the Aerial Parts of *Thymus marschallianus* Willd. and *Thymus seravschanicus* Klovov Growing Widely in Southern Kazakhstan. / B. S. Zhumakanova, I. Korona-Główniak, K. Skalicka-Woźniak [et al.] // *Molecules*. – 2021. – Vol. 26. – P. 3193. – DOI: 10.3390/molecules26113193.
7. Бузук, А. Г. Изменчивость химического состава эфирного масла *Thymus pulegioides* L. / А. Г. Бузук, Р. А. Юрченко, Г. Н. Бузук // *Вестник фармации*. – 2012. – Т. 55, № 1. – С. 19–25.
8. Савчук, С. С. Видовое разнообразие лекарственных растений, используемых белорусами / С. С. Савчук. – Минск : Право и экономика, 2023. – 84 с.
9. Флора Узбекистана: в 6 т. / гл. ред. К. З. Закиров. – Ташкент : Изд-во Акад. наук УзССР, 1961. – Т. 5: Loganiaceae – Rubiaceae / ред. тома А. И. Введенский. – 668 с.
10. Плантиум. Растения и лишайники России и сопредельных стран: открытый онлайн атлас и определитель растений: [сайт]. – 2007–2025. – URL: <https://www.plantarium.ru/> (дата обращения: 11.09.2025).
11. Флора СССР : в 30 т. / гл. ред. В. Л. Комаров. – Москва–Ленинград : Изд-во Академии наук СССР, 1954. – Т. 21 : Семейство Яснотковые (*Labiatae*) / Б. К. Шишкин. – 520 с.
12. Винокурова, О. А. Сравнительная характеристика различных видов тимьяна: состав,

- свойства, применение / О. А. Винокурова, О. В. Тринеева, А. И. Сливкин // Разработка и регистрация лекарственных средств. – 2016. – № 4. – С. 134–150.
13. DNA Barcoding Study of Representative Thymus Species in Bulgaria. / I. Aneva, P. Zhelev, G. Bonchev [et al.] // Plants. – 2022. – Vol. 11. – P. 270. – DOI:10.3390/plants11030270.
14. Hollingsworth, P. M. Choosing and using a plant DNA barcode / P. M. Hollingsworth, S. W. Graham, D. P. Little // PLoS ONE. – 2011. – Vol. 6. – DOI:10.1371/journal.pone.0019254.
15. BLAST: Basic Local Alignment Search Tool.: [database]. – URL: <https://blast.ncbi.nlm.nih.gov/Blast.cgi> (accessed on 11 Sep 2025).
16. Аль-Карави, Х. А. Х. Особенности роста, развития и биохимического состава хозяйственно значимых представителей рода *Thymus* L. в условиях нечерноземной зоны Российской Федерации: дис. ... канд. биол. наук : 03.02.01 / Ханан Ахмед Хади Аль-Карави; Российск. гос. аграрный ун-т. – Москва, 2018. – 158 с.
17. DNA barcoding to analyse taxonomically complex groups in plants: The case of *Thymus* (Lamiaceae) / S. Federici, A. Galimberti, F. Bartolucci [et al.] // Bot. J. Linn. Soc. – 2013. – Vol. 171. – P. 687–699.
18. Старчак, Ю. А. Фармакогностическое изучение растений рода тимьян (*Thymus* L.) как перспективного источника получения фитопрепаратов: автореф. дис. ... д-ра фарм. наук : 14.04.02 / Старчак Юлия Анатольевна; Сам. гос. мед. ун-т. – Самара, 2016. – 46 с.
19. Euro+Med 2006+ [continuously updated]: Euro+Med PlantBase — the information resource for Euro-Mediterranean plant diversity. Published at <http://www.europlusmed.org> [date of access: 14.12.2023].
20. Furan, M. A. Elevating thyme species identification: exploiting key chloroplast genes (matK, rbcL, and psbA-trnH) through DNA barcoding and phylogenetic analysis / M. A. Furan // BSI Eng Sci. – 2023. – Vol. 6, № 4. – P. 572–582.

A. N. Verchuk¹, A. V. Kilchevsky¹, S. V. Kubrak¹, N. V. Savina¹, L. V. Milko¹, S. S. Savchuk²,
V. N. Lebedko², D. V. Dubovik², E. V. Nikitina³

GENOTYPING OF MEDICINAL PLANTS OF THE GENUS *THYMUS* L. GROWING IN THE TERRITORY OF BELARUS AND UZBEKISTAN

¹State Scientific Institution

“Institute of Genetics and Cytology of the National Academy of Sciences of Belarus”

27 Akademicheskaya St., 220072 Minsk, the Republic of Belarus

e-mail: a.n.verchuk@mail.ru

²State Scientific Institution

“V. F. Kuprevich Institute of Experimental Botany of the National Academy of Sciences of Belarus”

27 Akademicheskaya St., 220072 Minsk, the Republic of Belarus

³Institute of Botany of the Academy of Sciences of the Republic of Uzbekistan

32 Durmon Yuli St., 100125 Tashkent, the Republic of Uzbekistan

The paper presents research results on the sequences of three DNA barcodes — *matK*, *psbA-trnH* and *ITS1-4* — in relation to the species of the genus *Thymus* growing in Belarus and Uzbekistan to identify controversial taxa and to select effective DNA markers. All DNA barcodes revealed relatively high intraspecific and low interspecific genetic variability within the genus *Thymus*. The smallest distance between the species was demonstrated by *matK* and *ITS1-4* markers, which indicates their insufficient informational value in the DNA barcoding of this group of plants. The *psbA-trnH* marker is of the greatest informational value and can be recommended as a component of a two- or three-locus DNA barcode system for identifying such taxonomically complex groups as the genus *Thymus*.

Keywords: DNA barcoding, *ITS1-4*, *matK*, *psbA-trnH*, *Thymus*.

Дата поступления в редакцию: 11 сентября 2025 г.

А. Н. Заинчковская, Е. А. Фомина, О. Ю. Урбанович

ИССЛЕДОВАНИЕ УРОВНЯ ЭКСПРЕССИИ ГЕНОВ *CBF12* И *CBF14* У СОРТОВ ОЗИМОЙ ПШЕНИЦЫ В МОДЕЛИРУЕМЫХ УСЛОВИЯХ ХОЛОДОВОГО СТРЕССА

Государственное научное учреждение
«Институт генетики и цитологии Национальной академии наук Беларуси»
Республика Беларусь, 220141, г. Минск, ул. Ф. Скорины, 34
e-mail: O.Urbnovich@igc.by

Проведено исследование уровня экспрессии генов *CBF12* и *CBF14* методом ПЦР в режиме реального времени в условиях моделирования холодового стресса у сортов озимой пшеницы Иркутская, Гирлянда и Кубус, обладающих различным уровнем морозостойкости. Как для гена *CBF12*, так и для гена *CBF14* наиболее высокий уровень экспрессии наблюдался после снижения температуры до -2 и -4 °C. При этом уровень экспрессии гена *CBF12* у исследуемых сортов был значительно выше по сравнению с геном *CBF14*. Изучение уровня экспрессии гена *CBF12* показало наибольшее увеличение уровня экспрессии данного гена у сортов Иркутская и Кубус в 1122,04 и 1645,28 раза соответственно через 24 ч после наступления условий -4 °C, и у сорта Гирлянда в 2190,41 раза через 2 ч после наступления условий -4 °C. Максимальный показатель уровня экспрессии гена *CBF14* у сорта Гирлянда в 73,55 раза был показан при -4 °C через 2 ч, а у сортов Иркутская и Кубус — в 22,90 и 69,04 раза соответственно при -2 °C через 2 ч. В целом, полученные результаты свидетельствуют о том, что уровень экспрессии генов *CBF12* и *CBF14* значительно изменяется в ответ на воздействие холода в моделируемых условиях холодового стресса на этапах снижения температурного режима до -2 и -4 °C, что может быть использовано в селекционном процессе, направленном на создание устойчивых сортов пшеницы.

Ключевые слова: озимая пшеница, холодовой стресс, морозоустойчивость, транскрипционные факторы, ген *CBF12*, ген *CBF14*.

Для цитирования: Заинчковская, А. Н. Исследование уровня экспрессии генов *CBF12* и *CBF14* у сортов озимой пшеницы в моделируемых условиях холодового стресса / А. Н. Заинчковская, Е. А. Фомина, О. Ю. Урбанович // Молекулярная и прикладная генетика: сб. науч. тр. / Ин-т генетики и цитологии НАН Беларуси; редкол.: А. В. Кильчевский (гл. ред.) [и др.]. – Минск, 2025. – Т. 39. – С. 33–40.

Введение

Пшеница является одной из основных культур, важных для глобальной продовольственной безопасности. По данным Продовольственной и сельскохозяйственной организации Объединенных Наций (ФАО), в 2023–2024 гг. общее производство пшеницы в мире составило около 789 млн т [1]. В Республике Беларусь пшенице также уделяется большое значение. По данным Национального статистического комитета (на июль 2025 г.), под посевные площади пшеницы в нашей стране было отведено 755 тыс. га, большую часть из которых (около 619 тыс. га, или 82%) занимают сорта озимой пшеницы. В целом доля посевных площадей, занятых посевами пшеницы в Республике Бе-

ларусь, составляет 30,4% от всех площадей, занятых зерновыми и зернобобовыми культурами, и 13,2% от всех посевных площадей в стране [2].

Главная причина доминирования озимых сортов заключается в том, что в основных районах возделывания они дают большие урожаи, чем яровые.

В развитии озимой пшеницы можно выделить два периода. Первый проходит осенью до заморозков, второй начинается весной и заканчивается плодоношением и отмиранием растений. В период роста озимая пшеница часто подвергается различным неблагоприятным внешним воздействиям, включающим отрицательные температуры, которые могут

приводить к изреживанию или полной гибели посевов. В Республике Беларусь погодные условия в зимне-весенний период могут быть достаточно суровыми. Это и низкие температуры зимой, отсутствие снежного покрова в отдельные периоды, оттепели и возвратные заморозки и др.

Важная роль в адаптации растений к действию температурного стресса принадлежит генам *Cor/Lea* семейства [3]. Кодируемые данным семейством генов белки представляют собой надсемейство гидрофильных внутриклеточных белков, накопление которых (до 1–2% от общего количества клеточных белков) в растительной клетке происходит при клеточной дегидратации. Они способствуют защите целостности клеточных структур при замораживании. *Cor/Lea* гены локализованы на разных хромосомах пшеницы и экспрессируются при низкой температуре [3, 4]. У многих растений, в том числе и у пшеницы, экспрессия генов *Cor/Lea* регулируется семейством факторов транскрипции, называемым CRT-связывающие факторы (C-Repeat Binding Factors, CBF), путем связывания элементов CRT [4].

CBF-факторы представляют собой довольно крупное подсемейство факторов транскрипции, которые принадлежат к семейству белков фактора связывания элемента, чувствительно к этилену (APETALA2/ethylene-responsive element binding factor, AP2/ERF), и активируются в ответ на осмотические стрессы, такие как холод или засуха. Домен AP2/ERF связывается с CRT/DRE элементами в промоторной области различных генов, участвующих в ответе на абиотический стресс, также известных как «регулон CBF». Данные гены принимают участие в защите от неблагоприятных последствий потери воды, вызванной заморозками и засухой, с помощью биосинтеза осмопротекторных белков, активности, связанной с метаболизмом углеводов, и транспорта сахаров [5–7].

Для гексаплоидной пшеницы описано 65 *TaCBF* генов, 27 из которых являются паралогами с 1–3 гомеологичными копиями, расположенными в субгеномах A, B и D [8].

Первые работы по изучению *CBF* генов / локуса *FR-2* у пшеницы были проведены при картировании популяций диплоидной пшеницы однозернянки (*Triticum monosocum* L.)

[9–15]. Была показана значительная связь между уровнем экспрессии генов *TmCBF12*, *TmCBF14*, *TmCBF15* и *TmCBF16* и устойчивостью к заморозкам, измеренной как способность к отрастанию побегов растения после стресса. Более того, в результате картирования с высокой плотностью также было установлено, что гены *TmCBF12*, *TmCBF14* и *TmCBF15* являются кандидатами на роль вносящих большой вклад в уровень морозоустойчивости у линий G3116 и DV92 [12].

Благодаря исследованиям, проведенным на *T. monosocum*, в мягкой пшенице также были идентифицированы *CBF* гены. Для диплоидной и полиплоидной пшеницы было выявлено меньшее количество копий гена *CBF14* в субгеноме B по сравнению с субгеномами A и D [16]. Делеции одновременно 6, 9 и 11 генов *TaCBF* были обнаружены S. Pearce с соавторами в локусе *FR-B2* тетраплоидной и гексаплоидной пшениц. Указанные делеции, охватывающие центральную часть локуса с расположенными в ней генами *TaCBF-B12*, *TaCBFB14* и *TaCBF-B15*, были выявлены в основном у образцов культивируемой яровой пшеницы и ассоциированы с более низким уровнем морозостойкости по сравнению с локусом *FR-B2* дикого типа без делеций [17]. T. Würschum с соавторами провели оценку фенотипического разнообразия по признаку зимостойкости среди 407 образцов озимой пшеницы селекции различных европейских стран, в ходе которой также было подтверждено влияние числа копий гена *TaCBF-A14* на фенотипическое проявление указанного признака [18].

В связи с показанной выше важностью генов *CBF12* и *CBF14* в ответной реакции пшеницы на воздействие холодового стресса, целью нашего исследования было изучение влияния воздействия низких температур на изменения уровня экспрессии указанных генов у сортов озимой пшеницы различного генетического происхождения в моделируемых условиях для поиска критериев отбора устойчивых генотипов.

Материалы и методы

Растительный материал и условия выращивания растений

Для исследования были выбраны сорта ози-

мой пшеницы Иркутская, Гирлянда и Кубус, отличающиеся по уровню морозостойкости. Сорта Гирлянда и Кубус были предоставлены отделом зерновых колосовых культур РУП «Научно-практический центр НАН Беларуси по земледелию» (г. Жодино), сорт Иркутская — Сибирским институтом физиологии и биохимии растений Сибирского отделения Российской академии наук. Сорт российской селекции Иркутская способен перезимовывать в условиях Восточной Сибири. Сорт белорусской селекции Гирлянда является районированным и обладает по результатам полевых испытаний высоким уровнем морозоустойчивости в климатических условиях Беларуси. В отличие от него, сорт немецкой селекции Кубус в климатических условиях Беларуси обладает низким уровнем морозоустойчивости.

Условия выращивания

Растения были выращены при определенном режиме (**1 стадия:** 25 дней при +16 / +10 °С (12 ч день / 12 ч ночь); **2 стадия:** 5 дней при +10 / +4 °С (12 ч день / 12 ч ночь); **3 стадия:** +4 °С (48 ч ночь); **4 стадия:** +2 °С (60 ч ночь), **5 стадия:** –2 °С (24 ч ночь); **6 стадия:** 5 дней при +2 °С (12 ч ночь) / –2 °С (12 ч ночь); **7 стадия:** –4 °С (24 ч ночь)) в климатической камере Pol-Eko-Aparatura KK1200Fit при относительной влажности 80% и интенсивности освещения 15 000 лк. Контрольную группу растений выращивали при +16 / +10 °С (12 ч день / 12 ч ночь) на протяжении всего эксперимента. Данный режим подобран в соответствии с условиями, максимально приближенными к тем, которым подвергаются растения в осенне-зимний период в Республике Беларусь.

Уровень экспрессии указанных генов измеряли на стадиях 3 (+4 °С), 5 (–2 °С), 6 (+2 °С / –2 °С), 7 (–4 °С) через 2 и 24 часа после наступления каждой из них.

Получение препаратов РНК, синтез кДНК, проведение ПЦР в режиме реального времени, а также статистическая обработка данных проводились по методикам, представленным в исследованиях А. Н. Заинчковской и др. [19–21].

Последовательности праймеров для генов *CBF12*, *CBF14* и референсного гена, кодирующего первичную последовательность актина, приведены в таблице 1.

Результаты и обсуждение

Как упоминалось ранее, гены *CBF12* и *CBF14* вносят значительный вклад в формирование холодоустойчивости у озимой пшеницы. Помимо указанных выше исследований, J. Zhu с соавторами показали различие в количестве копий гена *TaCBFA14* как среди чувствительных и устойчивых к холоду гаплотипов, так и среди гаплотипов с яровым и озимым типом развития. Более того, повышенное число копий гена *CBF14* среди образцов с озимым типом развития было связано с более высоким уровнем экспрессии указанного гена у толерантных гаплотипов. Данными авторами также было установлено различие в количестве копий и для генов *TaCBF-A12* и *TaCBF-A15* и показано наличие в них двух инсерций/делеций и десяти SNP [25]. Вклад количества копий гена в вариабельность *FR-A2* локуса был подтвержден А.-N. Sieber с соавторами для твердой пшеницы, где приблизительно 90% генотипических различий данного локуса объяснялось различным числом копий гена *CBF-A14*.

Таблица 1

Нуклеотидные последовательности праймеров, использованных для анализа экспрессии генов *CBF12* и *CBF14*

Праймер	Последовательность 5' → 3'	Температура отжига (T_m), °С	Литературный источник
TaCBF12	F AAATGGACGCGGGCACGTACTA R TCATCAGTGGTTCCATAGCGCC	60	[22]
TaCBF14	F CCACCAAATATGGGAGGAAA R CTTTCACAATGAACGAGCA	60	[23]
TaACT	F ACCTTCAGTTGCCCAGCAAT R TCGACCGCTGGCATACAAG	56–64	[24]

[26]. В реальных условиях растения подвергаются воздействию холода в течение более длительного промежутка времени. В связи с этим для нас представлял интерес изучение уровня экспрессии генов *CBF12* и *CBF14* в моделируемых условиях, приближенных к реальным.

В нашем исследовании было изучено изменение показателей уровня экспрессии указанных генов методом ПЦР в режиме реального времени в моделируемых условиях холодового стресса у сортов озимой пшеницы Иркутская, Гирлянда и Кубус, обладающих различным уровнем морозостойкости.

Исследование уровня экспрессии гена *CBF12* (рис. 1) показало наибольшее увеличение уровня экспрессии данного гена у сорта Гирлянда в 2190,41 раза через 2 ч после наступления условий -4°C , а также у сортов Кубус и Иркутская — в 1645,28 и 1122,04 раза соответственно через 24 ч после наступления условий -4°C по сравнению с растениями из контрольной группы. Следует отметить, что у исследованных сортов наблюдалось увеличение уровня экспрессии данного гена начиная с начального этапа моделируемых условий холодового стресса, т. е. при $+4^{\circ}\text{C}$ до наступления условий -2°C через 2 ч после наступления указанных условий для сорта Кубус, а также через 24 ч для сортов Иркутская и Гирлянда. Далее было установлено снижение уровня экспрессии гена

CBF12 в исследованных сортах до наступления условий $+2^{\circ}\text{C} / -2^{\circ}\text{C}$ через 24 ч и последующее увеличение уровня его экспрессии. Следует отметить, что уровень экспрессии указанного гена для сорта белорусской селекции Гирлянда практически на всех этапах моделируемых условий холодового стресса был выше, чем у сортов российской селекции Иркутская и немецкой селекции Кубус. В целом высокие значения и различия в уровне экспрессии гена *CBF12* у исследованных сортов свидетельствуют о его значительном вкладе в ответ на воздействие моделируемых условий холодового стресса. Различия в уровне экспрессии указанного гена были также установлены F. Sutton с соавторами для морозоустойчивых и морозочувствительных линий, полученных путем азидного мутагенеза сорта озимой пшеницы «Winoka» и прошедших холодовую акклиматизацию (яровизацию) [22]. В исследованиях Y. Motomura с соавторами устойчивый к заморозкам сорт Мироновская 808 демонстрировал более высокий уровень экспрессии гена *CBF12* по сравнению с чувствительным к заморозкам сортом Chinese Spring под воздействием низких температур [27].

Анализ изменения уровня экспрессии гена *CBF14* (рис. 2) показал значительное увеличение уровня его экспрессии в 22,90, 60,29 и 69,04 раза через 2 ч после наступления этапа

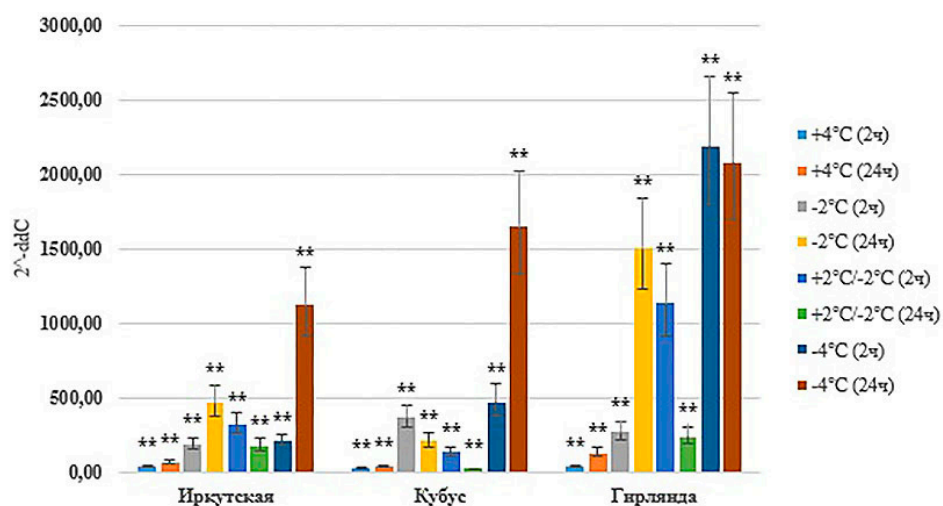


Рис. 1. Относительный уровень экспрессии гена *CBF12* у сортов мягкой озимой пшеницы с различным уровнем холодоустойчивости в моделируемых условиях холодового стресса. На диаграмме представлены средние значения и стандартные отклонения. Уровень экспрессии гена *CBF12* у растений, не подвергающихся холодовому стрессу, равен 0. $P < 0,01$ (**) по сравнению с контролем

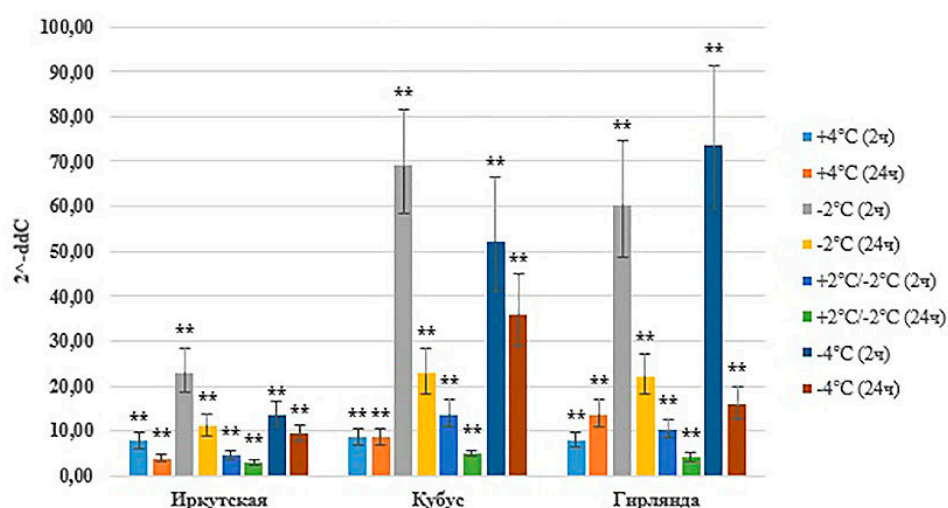


Рис. 2. Относительный уровень экспрессии гена *CBF14* у сортов мягкой озимой пшеницы с различным уровнем холодоустойчивости в моделируемых условиях холодного стресса. На диаграмме представлены средние значения и стандартные отклонения. Уровень экспрессии гена *CBF14* у растений, не подвергающихся холодному стрессу, равен 1. $P < 0,01$ (**) по сравнению с контролем

моделируемых условий -2°C у сортов Иркутская, Гирлянда и Кубус соответственно. Еще один пик экспрессии указанного гена наблюдался у всех исследуемых сортов через 2 ч после наступления условий -4°C . Наибольший показатель уровня экспрессии гена *CBF14* у сорта Гирлянда (в 73,55 раза) был показан при -4°C через 2 ч, а у сортов Иркутская и Кубус — в 22,90 и 69,04 раза соответственно при -2°C через 2 ч. В целом, для сортов Гирлянда и Кубус уровень экспрессии гена *CBF14* был выше по сравнению с сортом Иркутская. Уменьшение уровня экспрессии данного гена в морозочувствительных линиях и его увеличение в морозоустойчивых линиях было показано в упомянутом выше исследовании F. Sutton с соавторами [22]. Различные уровни экспрессии были выявлены M. Ahres с соавторами для двух сортов мягкой пшеницы и двух линий, полученных в результате скрещивания незимостойкого ярового сорта Manitou и очень морозостойкого озимого сорта Norstar [28].

Следует отметить, что как для гена *CBF12*, так и для гена *CBF14* наиболее высокий уровень экспрессии наблюдался при наступлении условий моделируемого эксперимента -2 и -4°C . При этом уровень экспрессии гена *CBF12* в исследуемых сортах был значительно выше по сравнению с геном *CBF14*. Среди сортов самая высокая экспрессия гена *CBF12*

была установлена для сорта Гирлянда, а гена *CBF14* — для сортов Кубус и Гирлянда. Как ранее упоминалось, сорт Иркутская способен перезимовывать в условиях Восточной Сибири, где температуры значительно ниже тех, при которых проводилось наше исследование. Возможно, при более низких температурах происходит новый пик экспрессии указанных генов. Высокие показатели экспрессии генов *CBF12* и *CBF14* для сорта Гирлянда можно объяснить тем, что моделируемые условия были приближены к климатическим условиям, характерным для Республики Беларусь. Именно для условий нашей страны данный сорт и был районирован. Не стоит также забывать о том, что морозостойкость является полигенным признаком, и на ее проявление, помимо исследованных, оказывают влияние еще и другие *CBF* гены. В связи с этим представляет интерес изучение уровня экспрессии других *CBF* генов, оказывающих влияние на холодоустойчивость.

Таким образом, изучение уровня экспрессии генов *CBF12* и *CBF14* методом ПЦР в реальном времени в сортах озимой пшеницы Иркутская, Гирлянда и Кубус, обладающих различным уровнем холодоустойчивости, показало, что данные гены вносят вклад в ответ на воздействие холода в моделируемых условиях холодного стресса и могут служить

потенциальными маркерами выживаемости растений при низких температурах.

Заключение

В результате исследования было продемонстрировано влияние холода на значения уровня экспрессии генов *CBF12* и *CBF14* в сортах озимой пшеницы Иркутская, Гирлянда и Кубус, обладающих различным уровнем морозостойкости, в моделируемых условиях, приближенных к условиям осенне-зимнего периода в Республике Беларусь.

Анализ уровня экспрессии гена *CBF12* показал, что наибольшее увеличение уровня экспрессии данного гена у сортов Иркутская и Кубус наблюдалось через 24 ч после наступления условий -4°C , а у сорта Гирлянда — через 2 ч после наступления условий -4°C . Для сорта белорусской селекции Гирлянда уровень экспрессии указанного гена практически на всех этапах моделируемых условий холодостресса был выше, чем у сортов российской селекции Иркутская и немецкой селекции Кубус. Профили экспрессии гена *CBF14* были схожи для сортов Кубус и Гирлянда практически на всех этапах моделируемого эксперимента, и для данных сортов уровень его экспрессии был выше по сравнению с сортом Иркутская. Следует отметить, что показатели экспрессии гена *CBF14* на протяжении всего эксперимента были значительно ниже по сравнению с геном *CBF12*. Максимальное значение уровня его экспрессии было выявлено также у сорта Гирлянда через 2 ч после наступления этапа -4°C . При этом значения данного показателя для гена *CBF14* были почти в 30 раз ниже, чем для гена *CBF12*. Оба исследованных гена демонстрировали наиболее высокий уровень экспрессии при наступлении условий моделируемого эксперимента -2 и -4°C .

Если рассматривать особенности показателей экспрессии по сортам, то наибольшую реакцию с самыми выраженными показателями уровня экспрессии гена *CBF12* демонстрирует сорт Гирлянда, гена *CBF14* — сорта Кубус и Гирлянда. В целом устойчивый к холоду сорт белорусской селекции Гирлянда отвечает на воздействие холодом более высоким уровнем экспрессии генов *CBF12* и *CBF14*.

Таким образом, изучение уровня экспрессии генов *CBF12* и *CBF14* методом ПЦР в реаль-

ном времени в сортах озимой пшеницы Иркутская, Гирлянда и Кубус показало, что они вносят вклад в ответ на воздействие холода в моделируемых условиях холодостресса. Полученные результаты и разработанные методы определения уровня экспрессии указанных генов могут быть использованы для оценки потенциала сортов к холодостойкости в селекционных программах по созданию устойчивых к морозам растений в Республике Беларусь и странах со схожими с ней климатическими условиями.

Работа выполнена при финансовой поддержке гранта БРФФИ Б23УЗБ-027.

Список использованных источников

1. Продовольственная и сельскохозяйственная организация Объединенных Наций : [сайт]. — Рим, 2025. — URL: <https://www.fao.org/documents/card/en/c/CC3939EN> (дата обращения: 06.08.2025).
2. Сельское хозяйство Республики Беларусь: Статистический буклет / Национальный статистический комитет Республики Беларусь, Государственный комитет по имуществу Республики Беларусь ; редкол.: И. В. Медведева [и др.]. — Минск : Информационно-вычислительный центр Национального статистического комитета республики Беларусь. — 2025. — 36 с.
3. COR/LEA proteins as indicators of frost tolerance in *Triticeae*: a comparison of controlled versus field conditions / K. Kosová, M. Klíma, I. T. Prášil [et al.] // *Plants*. — 2021. — Vol. 10, № 4. — P. 1–13.
4. Kobayashi, F. Transcriptional activation of *Cor/Lea* genes and increase in abiotic stress tolerance through expression of a wheat *DREB2* homolog in transgenic tobacco / F. Kobayashi, M. Ishibashi, S. Takumi // *Transgenic Res.* — 2008. — Vol. 17, № 5. — P. 755–767.
5. Regulation of the *Arabidopsis* CBF regulon by a complex low-temperature regulatory network / S. Park, C. Lee, C. J. Doherty [et al.] // *Plant J.* — 2015. — Vol. 82, № 2. — P. 193–207.
6. DREB1/CBF transcription factors: their structure, function and role in abiotic stress tolerance in plants / M. Akhtar, A. Jaiswal, G. Taj [et al.] // *J. Genet.* — 2012. — Vol. 91, № 3. — P. 385–395.
7. Heidarvand, L. What happens in plant molecular responses to cold stress? / L. Heidarvand,

- R. Maali Amiri // *Acta Physiol. Plant* – 2010. – Vol. 32, № 3. – P. 419–431.
8. Wheat *CBF* gene family: identification of polymorphisms in the *CBF* coding sequence / S. Mohseni, H. Che, Z. Djillali [et al.] // *Genome*. – 2012. – Vol. 55, № 12. – P. 865–881.
9. The expression of several *Cbf* genes at the *Fr-A2* locus is linked to frost resistance in wheat / A. Vagujfalvi, A. Aprile, A. Miller [et al.] // *Mol. Genet. Genomics*. – 2005. – Vol. 274, № 5. – P. 506–514.
10. The cold-regulated transcriptional activator *Cbf3* is linked to the frost-tolerance locus *Fr-A2* on wheat chromosome 5A / A. Vágújfalvi, G. Galiba, L. Cattivelli [et al.] // *Mol. Gen. Genom.* – 2003. – Vol. 269, № 1. – P. 60–67.
11. Miller, A. K. A cluster of 11 CBF transcription factors is located at the frost tolerance locus *Fr-A^m2* in *Triticum monococcum* / A. K. Miller, G. Galiba, J. Dubcovsky // *Mol. Genet. Genomics*. – 2006. – Vol. 275, № 2. – P. 193–203.
12. Identification of candidate *CBF* genes for the frost tolerance locus *Fr-A(m)2* in *Triticum monococcum* / A. K. Knox, C. X. Li, A. Vagujfalvi [et al.] // *Plant Mol Biol.* – 2008. – Vol. 67, № 3. – P. 257–270.
13. Mapping genes affecting flowering time and frost resistance on chromosome 5B of wheat / B. Toth, G. Galiba, E. Feher [et al.] // *Theor. Appl. Genet.* – 2003. – Vol. 107, № 3. – P. 509–514.
14. Identification of quantitative trait loci and associated candidate genes for low-temperature tolerance in cold-hardy winter wheat / M. Baga, S. V. Chodaparambil, A. E. Limin [et al.] // *Funct Integr Genomics*. – 2007. – Vol. 7, № 1. – P. 53–68.
15. Integration of genetic and genomics resources in einkorn wheat enables precision mapping of important traits / G. Saripalli, L. Adhikari, C. Amos [et al.] // *Commun. Biol.* – 2023. – Vol. 6, № 1. – P. 1–14.
16. Dhillon, T. *Cbf14* copy number variation in the A, B, and D genomes of diploid and polyploid wheat / T. Dhillon, E. J. Stockinger // *Theor. Appl. Genet* – 2013. – Vol. 126, № 11. – P. 2777–2789.
17. Large deletions in the *CBF* gene cluster at the *Fr-B2* locus are associated with reduced frost tolerance in wheat / S. Pearce, J. Zhu, Á. Boldizsár [et al.] // *Theor. Appl. Genet.* – 2013. – Vol. 126, № 11. – P. 2683–2697.
18. Copy number variations of *CBF* genes at the *Fr-A2* locus are essential components of winter hardiness in wheat / T. Würschum, H. F. C. Longin, V. Hahn [et al.] // *Plant J.* – 2017. – Vol. 89, № 4. – P. 764–773.
19. Экспрессия генов *TaDHN9-A*, *TaDHN13* и *TaDHN20* у сортов яровой пшеницы в ответ на засуху / А. Н. Заинчковская, Е. А. Фомина, Ф. Н. Кушанов [и др.] // Молекулярная и прикладная генетика: сб. науч. тр. / Институт генетики и цитологии НАН Беларуси; редкол.: А.В. Кильчевский (гл. ред.) [и др.]. – 2023. – Т. 35. – С. 27–34.
20. Заинчковская, А. Н. Изучение уровня экспрессии генов *HvGT-8*, *HvGT-10* и *HvGT-42* у сортов озимого ячменя в моделируемых условиях холодного стресса / А. Н. Заинчковская, Е. А. Фомина, О. Ю. Урбанович // Молекулярная и прикладная генетика: сб. науч. тр. / Институт генетики и цитологии НАН Беларуси; редкол.: А.В. Кильчевский (гл. ред.) [и др.]. – 2024. – Т. 36. – С. 43–51.
21. Изучение влияния холодного стресса на уровень экспрессии генов *HvCBF14* и *HvCOR14b* у сортообразцов озимого ячменя в моделируемых условиях / А. Н. Заинчковская, Е. А. Фомина, А. В. Ярошевич [и др.] // Молекулярная и прикладная генетика : сб. науч. тр. / Институт генетики и цитологии НАН Беларуси; редкол.: А.В. Кильчевский (гл. ред.) [и др.]. – 2024. – Т. 37. – С. 26–35.
22. *Cbf* genes of the *Fr-A2* allele are differentially regulated between long-term cold acclimated crown tissue of freeze-resistant and - susceptible, winter wheat mutant lines / F. Sutton, D.-G. Chen, X. Ge [et al.] // *BMC Plant Biology*. – 2009. – Vol. 9, № 34. – P. 1–9.
23. Light-quality and temperature-dependent *CBF14* gene expression modulates freezing tolerance in cereals / A. Novák, Á. Boldizsár, É. Ádám [et al.] // *J. Exp. Bot.* – 2016. – Vol. 67, № 5. – P. 1285–1295.
24. Assessment of candidate reference genes for the expression studies with brassinosteroids in *Lolium perenne* and *Triticum aestivum* / B. Jurczyk, E. Pocięcha, A. Janeczko [et al.] // *Journal of Plant Physiology*. – 2014. – Vol. 171, № 16. – P. 1541–1544.
25. Copy number and haplotype variation at the *VRN-A1* and central *FR-A2* loci are associated with frost tolerance in hexaploid wheat / J. Zhu, S. Pearce, A. Burke [et al.] // *Theor. Appl. Genet.* – 2014. – Vol. 127, № 5. – P. 1183–1197.

26. Copy number variation of *CBF-A14* at the *Fr-A2* locus determines frost tolerance in winter durum wheat / A.-N. Sieber, C. F. H. Longin, W. L. Leiser, [et al.] // *Theor. Appl. Genet.* – 2016. – Vol. 129, № 6. – P. 1087–1097.
27. A major quantitative trait locus for cold-responsive gene expression is linked to frost-resistance gene *Fr-A2* in common wheat / Y. Motomura, F. Kobayashi, J. C. M. Iehisa [et al.] // *Breed Sci.* – 2013. – Vol. 63, № 1. – P. 58–67.
28. Investigating the impact of spring (*Vrn-A1*) and winter (*vrn-A1*) vernalization alleles on frost tolerance induced by light spectrum and low temperatures in different wheat backgrounds / M. Ahres, T. Palmai, Z. Farkas [et al.] // *Environmental and Experimental Botany.* – 2025. – Vol. 229. – P. 1–15.

A. N. Zainchkovskaya, A. A. Famina, O. Yu. Urbanovich

STUDY OF THE EXPRESSION LEVEL OF THE *CBF12* AND *CBF14* GENES IN WINTER WHEAT VARIETIES UNDER SIMULATED COLD STRESS CONDITIONS

State Scientific Institution

“Institute of Genetics and Cytology of the National Academy of Sciences of Belarus”

34 F. Skorina St., 220141 Minsk, the Republic of Belarus

e-mail: O.Urbanovich@igc.by

The expression level of the *CBF12* and *CBF14* genes was studied using real-time PCR under conditions of simulated cold stress in winter wheat varieties Irkutskaya, Girlyanda, and Cubus, which have different levels of frost resistance. For both the *CBF12* and *CBF14* genes, the highest expression level was observed after a temperature decrease to -2°C and -4°C . Moreover, the expression level of the *CBF12* gene in the studied varieties was significantly higher compared to the *CBF14* gene. The study of the *CBF12* gene expression level showed the greatest increase in the expression level of this gene in the Irkutskaya and Cubus varieties by 1122.04 and 1645.28 times, respectively, in 24 hours after the onset of -4°C conditions; and in the Girlyanda variety by 2190.41 times in two hours after the onset of -4°C conditions. The maximum expression level of the *CBF14* gene in the Girlyanda variety was 73.55 times higher at -4°C after 2 hours, while in the Irkutskaya and Cubus varieties, it was 22.90 and 69.04 times higher, respectively, at -2°C after 2 hours. Overall, these results indicate that the expression level of the *CBF12* and *CBF14* genes changes significantly in response to cold exposure under simulated cold stress conditions at the temperature reduction stages to -2°C and -4°C , which can be used in the breeding process aimed at creating resistant wheat varieties.

Keywords: winter wheat, cold stress, frost resistance, transcription factors, *CBF12* gene, *CBF14* gene.

Дата поступления в редакцию: 26 августа 2025 г.

Е. В. Гузенко, А. И. Царь

ГЕНЕТИЧЕСКАЯ СТРУКТУРА И ВАРИАБЕЛЬНОСТЬ ПОПУЛЯЦИЙ ПЧЕЛ *APIS MELLIFERA MELLIFERA* В РЕСПУБЛИКЕ БЕЛАРУСЬ

Государственное научное учреждение
«Институт генетики и цитологии Национальной академии наук Беларуси»
Республика Беларусь, 220072, г. Минск, ул. Академическая, 27
e-mail: e.guzenko@igc.by

Темная лесная пчела *A. m. mellifera* — единственный из известных подвидов медоносных пчел, идеально адаптировавшийся к жизни в холодном климате Северной Европы. Для Беларуси темная лесная пчела является аборигенным подвидом, который считался исчезнувшим в результате гибридизации. До недавнего времени отсутствовали достоверные данные о существовании темной лесной пчелы на территории Беларуси. Объектом данного исследования являлись пасечные и бортевые медоносные пчелы, собранные на территории республики. С помощью комплекса молекулярно-генетических маркеров идентифицирован аборигенный подвид *A. m. mellifera*. Анализ STR-локусов показал высокую генетическую вариабельность образцов в целом и особенно собранных в Гомельской и Витебской областях. Анализ генетических дистанций установил существование структурных группировок. Индивидуальный молекулярно-генетический анализ выявил высокую степень метизации в пчелосемьях. Обнаружены генетически однородные пчелосемьи подвида *A. m. mellifera*, которые могут стать основой микропопуляции для восстановления и сохранения аборигенных медоносных пчел в Беларуси.

Ключевые слова: медоносная пчела, *Apis mellifera mellifera*, локус мтДНК COI-COII, STR-маркеры, популяционно-генетическая структура, метизация, генетическая однородность.

Для цитирования: Гузенко, Е. В. Генетическая структура и вариабельность популяций пчел *Apis mellifera mellifera* в Республике Беларусь / Е. В. Гузенко, А. И. Царь // Молекулярная и прикладная генетика: сб. науч. тр. / Ин-т генетики и цитологии НАН Беларуси; редкол.: А. В. Кильчевский (гл. ред.) [и др.]. — Минск, 2025. — Т. 39. — С. 41–53.

Введение

В последние десятилетия мировая популяция пчел значительно сократилась и продолжает сокращаться из-за потери среды обитания, интенсивных методов ведения сельского хозяйства, изменения климата и чрезмерного использования таких агрохимикатов, как пестициды [1]. По данным Межправительственной научно-политической платформы по биоразнообразию и экосистемным услугам (IPBES), западная медоносная пчела является самым распространенным управляемым опылителем в мире. Более 80 млн ульев ежегодно производят около 1,6 млн т меда [2]. Опылители вносят непосредственный вклад в обеспечение продовольственной безопасности. По данным специалистов по вопросам пчеловодства Продовольственной и сельскохозяйствен-

ной организации (ФАО) ООН, треть мирового производства продуктов питания зависит от пчел [1]. Отмечается также, что генетическое разнообразие местных подвидов — одно из самых важных наследий, которое мы можем оставить будущим поколениям пчеловодов и фермеров.

Медоносные пчелы демонстрируют значительную географическую вариативность, что связано с адаптацией к изменяющимся климатическим условиям и разнообразию растений, а также к преобладающим вредителям и патогенам. Потребность в высоких продуктивных показателях пчелиных колоний в сочетании с желательными поведенческими характеристиками привела к значительным изменениям, вызванным научно необоснованной селекцией пчел. Таким образом, исторически сложивша-

яся схема географического распределения подвидов разрушается из-за массового импорта и растущей мировой практики торговли пчеломатками и пчелопакетами.

Указанная деятельность ставит под угрозу местные аборигенные расы и экотипы, способствуя гибридизации [3, 4], а в некоторых регионах — полной замене аборигенных подвидов неместными пчелами с высокой коммерческой доходностью. Например, замена *A. m. mellifera* в Северной и Центральной Европе на *A. m. carnica* или *A. m. ligustica* [5]. Обратной стороной этих экономически обусловленных процессов является растущая тенденция к единообразию популяций медоносных пчел, что приводит к потере как генетического разнообразия, так и специфических адаптаций к местным условиям [3, 4]. Сокращение генетического разнообразия влияет и на способность популяций медоносных пчел адаптироваться к новым негативным факторам, например, поражению клещами рода *Varroa* и *Tropilaelaps*. Аборигенные подвиды медоносных пчел являются важными резервуарами региональной изменчивости и местных адаптаций; их исчезновение означает утрату уникальных сочетаний признаков, сформировавшихся в результате естественного отбора в течение длительного времени, что весьма важно для более устойчивого пчеловодства [6].

Мировым научным сообществом признается необходимость поощрения региональных усилий по сохранению и селекции аборигенных подвидов пчел и поддержанию их на изолированных пасеках. Для достижения этой цели необходимо иметь базу данных эталонных генотипов, которая бы учитывала естественную генетическую вариабельность медоносных пчел.

Для Беларуси темная лесная пчела является аборигенным подвидом, который до недавнего времени считался исчезнувшим в результате гибридизации [7]. В настоящее время в нашей стране аборигенную медоносную пчелу с большей вероятностью можно встретить в регионах, где преобладает бортовое пчеловодство и где никогда не было пасечного разведения пчел. Таким географическим районом является Припятское Полесье (Гомельская область). Также вероятность обнаружить аборигенную пчелу выше в северных

районах республики (Витебская область, приграничные территории с Литвой). Результаты последних исследований показали, что на территории Литвы 52% пасечных пчел относится к подвиду *A. m. mellifera*. При анализе образцов обнаружено 20 уникальных межгенных спейсерных митотипов COI-COII, 11 из которых были найдены впервые [8].

Цель настоящего исследования — с помощью комплекса молекулярно-генетических маркеров идентифицировать аборигенный подвид *A. m. mellifera* на территории Беларуси, провести популяционно-генетические исследования и установить степень гибридности пчелосемей. Для этого необходимо установить таксономическую принадлежность пчел по анализу митохондриальной ДНК; изучить полиморфизм STR-локусов ядерной ДНК пчел, принадлежащих эволюционной линии М; провести популяционно-генетический анализ и установить филогенетические связи, генетическую однородность и гибридность исследуемых популяций.

Материалы и методы

Материалом для исследований служила коллেকция медоносных пчел, которая была собрана в географических районах: Припятское Полесье (Гомельская область), Витебская область, Минская область, Брестская область, Гродненская область. Рабочих пчел отбирали из бортей, ульев, ловушек, расположенных на окраинах деревень. Сбор материала проходил в течение июля-августа 2021–2022 гг. самостоятельно владельцами пасек или бортниками. Для сравнения в исследование также взяты образцы из Российской Федерации (Чувашия, Башкирия), предположительно принадлежащие к подвиду темная лесная пчела. Всего проанализировано 159 пчелосемей.

Пчел фиксировали в 96%-ом этаноле и хранили до выделения ДНК при -10°C . Выделение ДНК проводили из торакса (груди) пчелы. Материал измельчали, ресуспендировали в 200 мкл ТЕ буфера (10 mM Tris-HCl, pH 7,5; 1 mM EDTA). Для выделения ДНК использовали Genomic DNA Purification Kit (#K0512) (Thermo Fisher Scientific) в соответствии с инструкцией производителя. Анализ полиморфизма локуса COI-COII в мтДНК проводили с использованием ПЦР со специфическими

праймерами (GGCAGAATAAGTGCATTG; CAATATCATTTGATGACC). Реакционная смесь включала 20 нг геномной ДНК, по 5 пмоль прямого и обратного праймера, 2 мМ смеси дНТП (dNTPs), 50 мМ $MgCl_2$ и 5 ед./мкл Taq-полимеразы. ПЦР проводили в термоциклере MyCycler™ (BioRad, США) в следующих условиях: 92 °С в течение 3 мин, 30 циклов с параметрами: денатурация при температуре 92 °С в течение 30 сек, отжиг праймеров при температуре 50 °С в течение 1 мин 30 сек, элонгация при 63 °С в течение 2 мин. Конечная элонгация при 72 °С 10 мин. Качество прохождения реакции оценивалось по гель-электрофорезу. При наличии продукта реакции проводили рестрикционный анализ (DraI-тест) с использованием эндонуклеазы DraI (SibEnzyme®) в соответствии с инструкцией производителя. Продукты рестрикции оценивали в 1,5%-ом агарозном геле.

Анализ полиморфизма микросателлитных повторов ДНК проводили с использованием ПЦР с флуоресцентно-мечеными праймерами (A24, A88, A113, AP043, HB-THE-03, HB-C16-01, HB-C16-05, A28a). Реакционная смесь включала 20 нг геномной ДНК, по 5 пмоль прямого и обратного праймера, 2 мМ смеси dNTPs, 50 мМ $MgCl_2$ и 5 ед./мкл Taq-полимеразы. ПЦР проводили в термоциклере Thermal Cycler C1000 (BioRad, США) в следующих условиях: 94 °С в течение 3 мин, 30 циклов с параметрами: денатурация при температуре 94 °С в течение 30 сек, отжиг праймеров в течение 30 сек (температура отжига подбиралась в зависимости от праймера), элонгация при 72 °С в течение 30 сек. Конечная элонгация при 72 °С 10 мин.

Филогенетические отношения изучали путем расчета парных генетических расстояний Nei [9] с использованием GenAlEx 6.41 [10]. Кластерный анализ был проведен с помощью программы STRUCTURE 2.3.4. [11]. Для визуализации результатов и их математического подтверждения методами Evanno [12] была использована веб-программа STRUCTURE Harvester [13]. Данная веб-программа была использована для детекции подходящего количества кластеров путем их индивидуального перебора. Статистическую обработку данных проводили с использованием программы STRUCTURE v.2.3.4, GenAlEx v.6.5. Были рас-

считаны следующие показатели: среднее число аллелей на локус (N_a), эффективное число аллелей (N_e), уровни ожидаемой (H_e) и наблюдаемой (H_o) гетерозиготности, значение информационного индекса Шеннона (I), коэффициенты F_{IS} и F_{ST} .

Результаты и обсуждение

Генетическое разнообразие локуса COI-COII мтДНК

Классическим методом для идентификации пород (подвидов) пчел до последнего времени оставался морфометрический анализ, предложенный Кожевниковым Г. А. еще в 1900 г. Однако морфометрические методы в условиях повсеместной гибридизации пчел из-за опосредованной человеком интрогрессии (активный экспорт пчелопакетов) не всегда достоверны.

Подвиды медоносных пчел можно различать по генетическим особенностям мтДНК. Количество митохондрий в яйцеклетке намного больше, чем в сперматозоиде, и даже если отдельные отцовские митохондрии попадают в яйцеклетку при оплодотворении, то они блокируются, и зигота получает только материнский набор митохондрий [14]. Это приводит к отсутствию рекомбинаций в мтДНК, и мутационный процесс оказывается единственным путем формирования ее полиморфизма.

Анализ митохондриального генома *A. mellifera* показал, что спейсерный участок, локализованный между генами COI и COII, образован последовательностью гена tRNKLeu и сложными повторами P и Q элементов [15]. Пчелы разных эволюционных ветвей характеризуются разными комбинациями элементов P и Q. Наиболее коротким фрагментом локуса COI-COII мтДНК, содержащим единственный элемент Q (350 п. н.), характеризуются подвиды пчел эволюционной ветви C [16]. Подвиды пчел, относящиеся к эволюционным ветвям A, M, O, Y и Z, характеризуются более длинным фрагментом локуса COI-COII мтДНК и содержат разные варианты элемента P и от 1 до 5 копий элемента Q [17]. У среднерусской породы (темная лесная пчела), поддерживаемой в России, регистрируются несколько вариантов локуса COI-COII: PQQ, PQQQ, PQQQQ (от 500 до 1500 п. н.).

С учетом особенностей организации пчелосемей для изучения полиморфизма мтДНК

можно использовать одну рабочую пчелу (или одного трутня), личинку, особь при выходе из ячейки расплода или нелетную пчелу от одной семьи. Так как образцы для исследования доставляли пчеловоды, и мы не могли контролировать процесс сбора материала, нами случайным образом отбирались по 10 особей из 50 пчел каждой семьи. Такой подход позволяет проводить более точную оценку генетическо-

го разнообразия пчел от каждой пчелосемьи.

Из 159 проанализированных пчелосемей 28 показали принадлежность к эволюционной линии М, к которой относится темная лесная пчела, из них 20 семей обитали на территории Беларуси, 8 — на территории России. Обнаружено три варианта комбинаций Р и Q элементов в локусе COI-COII: PQ, PQQ, PQQQ (табл. 1).

Таблица 1

Обнаруженные варианты локуса COI-COII мтДНК у изучаемых образцов эволюционной линии М, собранных на территории Беларуси и России

№	Шифр образца	Расположение пасеки, борты, ловушки	Вариант локуса COI-COII
1	16	Минская область, Минский район, д. Раубичи. Местная пчела	PQ
2	17	Минская область, Минский район, д. Раубичи. Местная пчела	PQ PQQ
3	18	Витебская область, Глубокский район, д. Сазановщина. Местная пчела	PQQ
4	19	Витебская область, Глубокский район, д. Шуневцы. Местная пчела	PQQQ
5	20	Витебская область, Глубокский район, д. Морги. Местная пчела	PQ
6	37	Гомельская область, Лельчицкий район, д. Дзержинск. Местная пчела	PQQ
7	38	Гомельская область, Лельчицкий район, д. Дзержинск. Бортевая пчела	PQQ
8	39	Гомельская область, Лельчицкий район, д. Дзержинск. Бортевая пчела	PQQ
9	40	Гомельская область, Лельчицкий район, д. Дзержинск. Бортевая пчела	PQQ
10	41	Гомельская область, Лельчицкий район, д. Дзержинск. Бортевая пчела	PQQ
11	42	Гомельская область, Лельчицкий район, д. Дзержинск. Бортевая пчела	PQQ
12	43	Гомельская область, Лельчицкий район, д. Дзержинск. Бортевая пчела	PQQ
13	45	Витебская область, г. Глубокое. Местные пчелы	PQQQ
14	46	Витебская область, г. Глубокое. Местные пчелы	PQQ
15	65	Витебская область, Лепельский район, д. Суша	PQ
16	87	Брестская область, Пружанский район. Рой в ловушке	PQ
17	101	Гродненская область, Островецкий район, д. Глинище. Рой в ловушке	PQQ
18	103	Гродненская область, Щучинский район, д. Орля	PQ PQQ
19	104	Гродненская область, Щучинский район, д. Ошурки. Рой в ловушке.	PQQ
20	121	Гродненская область, Гродненский район, д. Романово. Рой в ловушке	PQQ
21	66	Чувашия, Красноармейский район	PQQ
22	69	Чувашия, Красноармейский район	PQ
23	71	Чувашия, Батыревский район	PQQ
24	74	Чувашия, Аликовский район	PQQQ
25	77	Чувашия, Моргоушский район	PQQQ
26	78	Чувашия, Моргоушский район	PQQQ
27	82	Башкирия, Буззянский район	PQQQ
28	85	Башкирия, Бурзянский район	PQ

При индивидуальном анализе пчел нами обнаружен внутрисемейный полиморфизм по локусу COI-COII в образцах № 17, 103. Рабочие особи данной семьи имели разные гаплотипы: у образца № 17 у десяти особей амплифицировался фрагмент размером 600 п. н., у девяти особей — 800 п. н.; в семье № 103 у четырех особей амплифицировался фрагмент размером 600 п. н., у шести особей — 800 п. н. (рис. 1).

У медоносных пчел возможно наследование мтДНК как по материнской, так и по отцовской линиям. Данное явление называется гетероплазмией и до конца не изучено. Наличие разных типов мтДНК может влиять на метаболизм клетки и жизнеспособность пчелы. Обнаруженный внутрисемейный полиморфизм также может быть следствием замены пчеломатки при пчеловодстве, а также некачественным сбором биологического материала.

При анализе остальных пчелосемей внутрисемейный полиморфизм по локусу COI-COII выявлен не был.

Генетическая структура анализируемых популяций темной лесной пчелы отличается в зависимости от территории обитания: для образцов Гомельской области характерна ком-

бинация PQQ (800 п. н.), для образцов, собранных на территории других областей, типичных комбинаций не обнаружено.

Особи с сочетанием элементов PQ, PQQ, PQQQ имеют западно-европейское происхождение и могут принадлежать подвидам *A. m. mellifera* (темная лесная пчела), *A. m. iberensis* и *A. m. sinisxinyuan*. Мы считаем, что анализируемые образцы относятся к аборигенному для Беларуси подвиду медоносных пчел *A. m. mellifera*, поскольку *A. m. iberiensis* (Engel, 1999) встречается на Пиренейском полуострове (Испания и Португалия), некоторых близлежащих островах и имеет гаплотипы и А-линии, и М-линии [18]. Подвид *A. m. sinisxinyuan* [19] был обнаружен в 2016 г. исследователями Китая в префектуре Синьюань, расположенной в Синьцзян-Уйгурском автономном районе. Эволюционная история данной группы окончательно не сформулирована. Завоз пчел данных подвидов для коммерческого разведения в Беларуси маловероятен.

Небольшой размер и высокое содержание в клетках мтДНК значительно облегчают исследование ее полиморфизма и позволяют эффективно изучать разнообразие митохон-

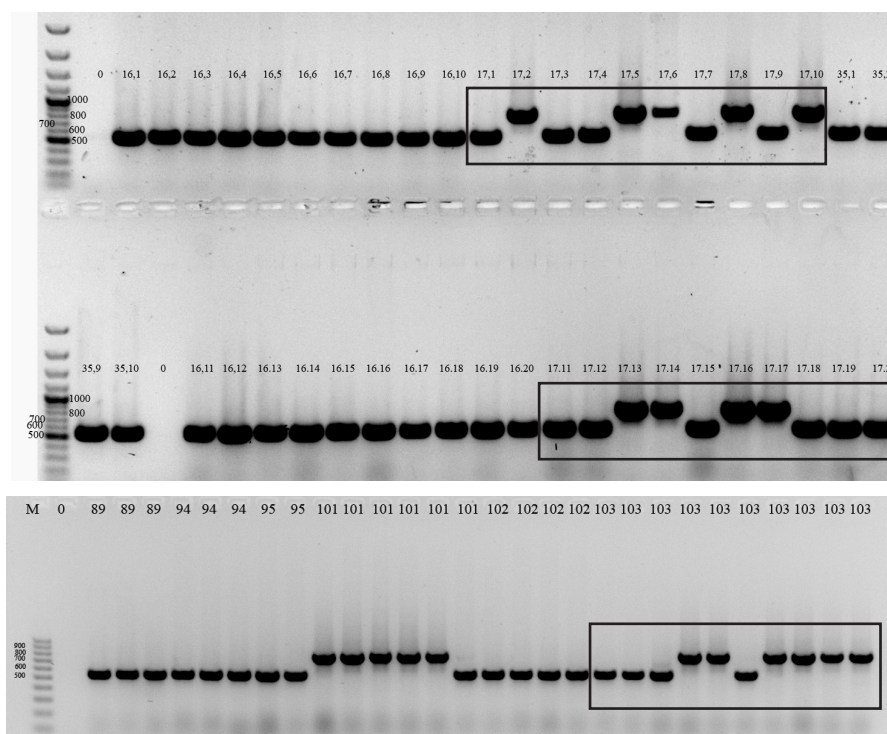


Рис. 1. Электрофореграмма продуктов амплификации локуса COI-COII рабочих пчел семей 16 (16.1–16.10), 17 (17.1–17.10; 17.11–17.20), 103.

дриального генофонда. Однако специфичность локуса COI-COII мтДНК может установить происхождение медоносной пчелы только по материнской линии.

Генетическое разнообразие STR-локусов ядерной ДНК

По результатам предыдущих исследований определен набор STR-маркеров, позволяющих достоверно устанавливать внутрисемейный и межсемейный полиморфизм, чистоту и гибридность пчелосемей, обитающих на территории Беларуси [20].

Индивидуальный анализ 166 рабочих пчел, принадлежащих 20 пчелосемьям, выявил 179 аллелей размером от 65 до 340 п. н. (табл. 2).

Количество выявленных аллелей в среднем на локус составило 22,375, что является высоким значением; количество эффективных аллелей колебалось от 11,492 до 3,404 в зависимости от локуса и в среднем составило 6,829. Среднее значение индекса Шеннона — 2,230, что указывает на достаточно высокое генетическое разнообразие изучаемых образцов *A. m. mellifera*.

Для описываемой популяции среднее значение H_0 было незначительно ниже H_e , что говорит о некотором дефиците гетерозигот (табл. 3). Установленное значение коэффициента инбридинга было достаточно низким — 2,4%.

Таблица 2

Характеристика проанализированных STR-локусов

STR-локус	Номер хромосомы	Количество выявленных аллелей	Размер аллеля, п. н.	
			минимальный	максимальный
A24	7	19	76	216
A88	8	19	99	255
A113	6	29	133	230
AP043	3	24	119	241
HB-THE-03	1	18	135	255
HB-C16-01	16	40	181	340
HB-C16-05	1	19	65	110
A28a	14	11	100	149

Таблица 3

Популяционно-генетические параметры исследуемых пчелосемей по восьми STR-локусам

Локус	N	N _a	N _e	I	H ₀	H _e	uHe	F
HB-C16-05	162	19	8,454	2,432	0,846	0,882	0,884	0,041
HB-C16-01	163	40	11,492	2,991	0,840	0,913	0,916	0,079
HB-THE-03	165	18	6,995	2,224	0,752	0,857	0,860	0,123
A88	163	19	3,404	1,754	0,693	0,706	0,708	0,018
A113	165	29	8,394	2,578	0,915	0,881	0,884	-0,039
AP043	166	24	7,381	2,333	0,849	0,865	0,867	0,017
A24	165	19	4,641	1,896	0,788	0,785	0,787	-0,004
A28a	164	11	3,876	1,633	0,774	0,742	0,744	-0,044
Mean	164,125	22,375	6,829	2,230	0,807	0,829	0,831	0,024
SE	0,479	3,105	0,967	0,161	0,024	0,026	0,027	0,020

На рисунке 2 представлена гистограмма с результатами, иллюстрирующими популяционно-генетическую структуру совокупности пчелосемей. Показано, что количество выявленных аллелей практически равно количеству уникальных аллелей.

Частота встречаемости редких аллелей

микросателлитных локусов варьировала от 0,003 во всех STR-локусах до 0,5 в локусе A88 (146 п. н.). На рисунке 3 приведена дендрограмма, отражающая частоту встречаемости 40 аллелей наиболее полиморфного STR-локуса HB-C16-01 в генотипах *A. m. mellifera*. Большинство аллелей являются редкими

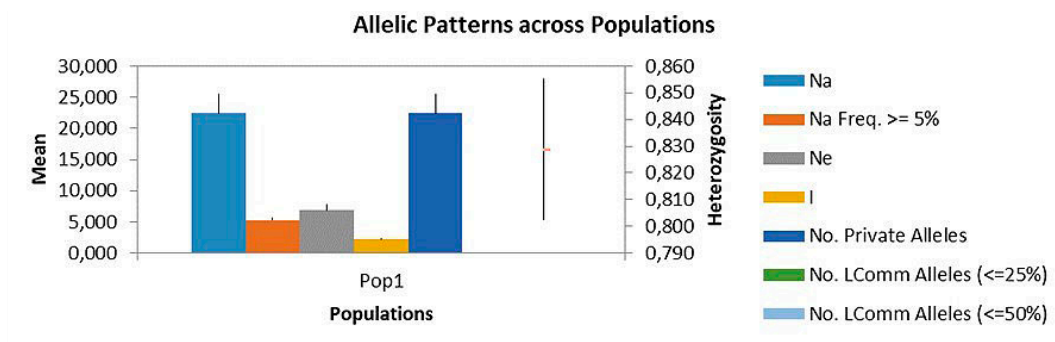


Рис. 2. Популяционно-генетическая характеристика исследуемых пчелосемей

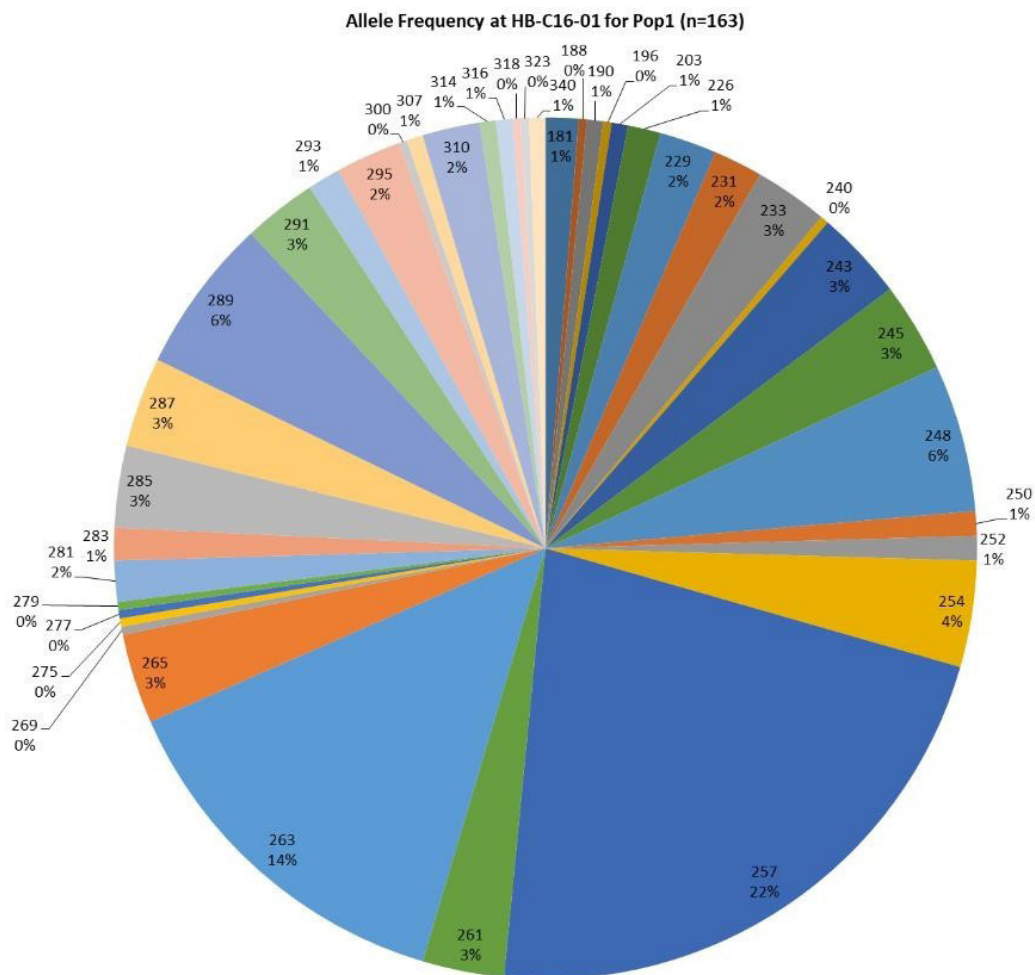


Рис. 3. Дендрограмма частоты встречаемости аллелей локуса HB-C16-01 в генотипах *A. m. mellifera* белорусской популяции

и встречаются с частотой менее 2%. На рисунке 4 приведена дендрограмма, отражающая частоту встречаемости 11 аллелей наименее полиморфного STR-локуса A28a.

Количество выявленных аллелей варьировало от 7,3 до 16,75 и было наибольшим в популяции Гомельской области. Количество эффективных аллелей также было наибольшим у образцов Гомельской популяции — 6,711. Среднее значение индекса Шеннона в Гомельской популяции равнялось 2,216, что указывает на достаточно высокое генетическое разнообразие данной популяции. Второе место по генетическому разнообразию занимала популяция Витебской области — 1,905; далее следовала Минская

популяция и Гродненская. Доля гетерозигот в популяциях Витебской и Гродненской областей была выше, чем в Минской и Гомельской. В последних зафиксирован незначительный дефицит гетерозигот. На рисунке 5 представлена обобщающая информация по исследованным популяциям медоносных пчел, из которой следует, что популяции Гомельской области (регион 2) и Витебской области (регион 4) обладали большим генетическим разнообразием и большим количеством уникальных аллелей по сравнению с популяциями остальных областей.

В сравнительный анализ включены образцы эволюционной линии М из регионов Башкирии и Чувашии Российской Федерации (рис. 6). У данных образцов (номера 69–77) по исследованным STR-локусам полиморфизм и количество уникальных аллелей были на уровне изучаемых белорусских образцов.

Панели различных микросателлитных маркеров активно используются для изучения интрогрессии, метизации, чистопородности пчел во всем мире. Так, популяции аборигенного для Саудовской Аравии подвида *A. m. jemenitica* [21] изучались с использованием микросателлитных маркеров A7, A24, A28, A88, A113, B124, Ap43 и Ap81 с целью выяснения степени интрогрессии и гибридизации с пчелами, которые активно импортировались в страну из других регионов. С использованием анализа восьми микросателлитных маркеров был изучен полиморфизм трех популяций иранской медоносной пчелы *A. m. meda* на северо-западе Ирана [22]. Генетическое разнообразие медоносных пчел островных по-

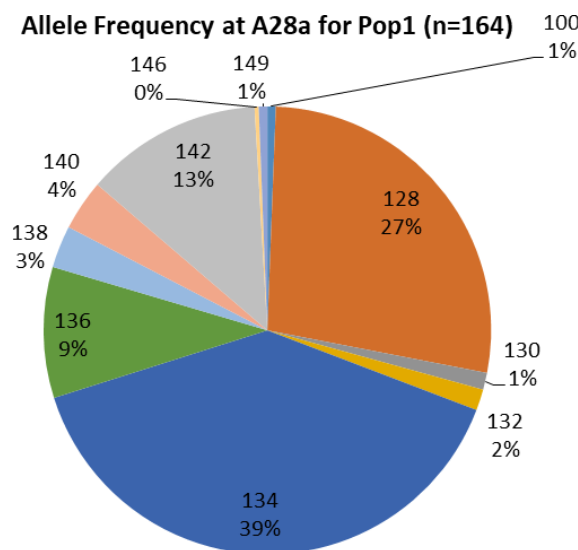


Рис. 4. Дендрограмма частоты встречаемости аллелей локуса A28a в генотипах *A. m. mellifera* белорусской популяции

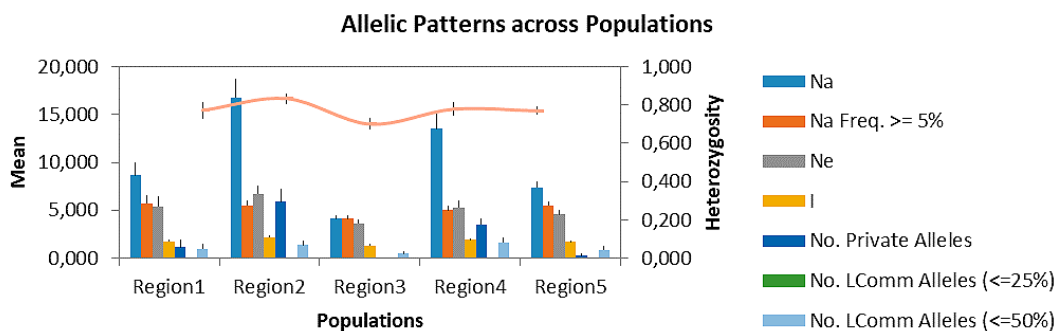


Рис. 5. Популяционно-генетическая структура пчелосемей эволюционной линии М в зависимости от региона Беларуси: Минская область (Region 1), Гомельская область (Region 2), Брестская область (Region 3), Витебская область (Region 4), Гродненская область (Region 5)

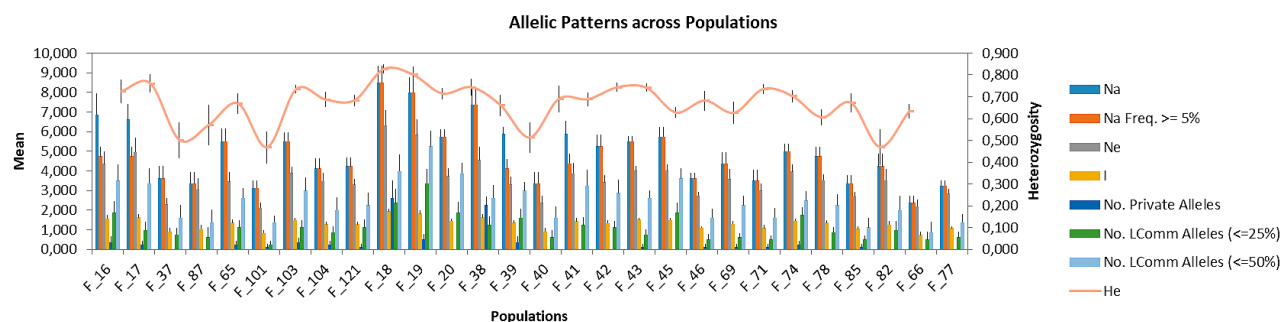


Рис. 6. Результаты анализа STR-локусов пчел, принадлежащих эволюционной линии М, собранных в Беларуси (Минская, Гомельская, Брестская, Витебская и Гродненская области) и России (Чувашия, Башкирия)

пуляций Балеарских островов изучали по восьми полиморфным STR-локусам. Была выявлена низкая вариабельность, что ожидаемо для островных популяций [23]. Полиморфные STR-локусы активно используются при изучении генетической структуры автохтонных популяций медоносной пчелы в различных регионах Российской Федерации: популяций гибридных пчел Томской области и популяций *A. m. mellifera* Пермского края [24], Республики Башкортостан [25], Архангельской и Владимирской областей, Красноярского края и Республики Татарстан; популяций *A. m. carpatica* Республики Адыгея; популяций *A. m. caucasica* Орловской области и Краснодарского края, гибридных пчел Новосибирской области. Результаты анализа генетической структуры популяции медоносной пчелы в южной части Башкортостана были проведены с использованием панели из пяти микросателлитных праймеров [25].

Анализ генетических дистанций показал, что пчелосемьи 39, 42, 43, 37, 40 и пчелосемьи 19, 45, 18, 46, 20 генетически близки, что совпадает с их регионами обитания и сбора — Гомельская и Витебская области соответственно. Наибольшим генетическим подобием обладали образцы 16, 17, у которых было общее место сбора — Минская область, Минский район, д. Раубичи (рис. 7А). В одном кластере оказались образцы из Гродненской области — 101, 121, 103, 104.

Анализ генетических дистанций образцов Беларуси, Чувашии и Башкирии показал, что наибольшей генетической близостью характеризовались образцы 18, 16, 17, 69, 45, 71, 19, 74, что соответствует региону обитания — северная часть Беларуси и Чувашия; образцы 37, 40, 39, 43 также группировались вместе на филогенетическом дереве, что соответствует региону обитания — Гомельская область Беларуси (рис. 7Б).

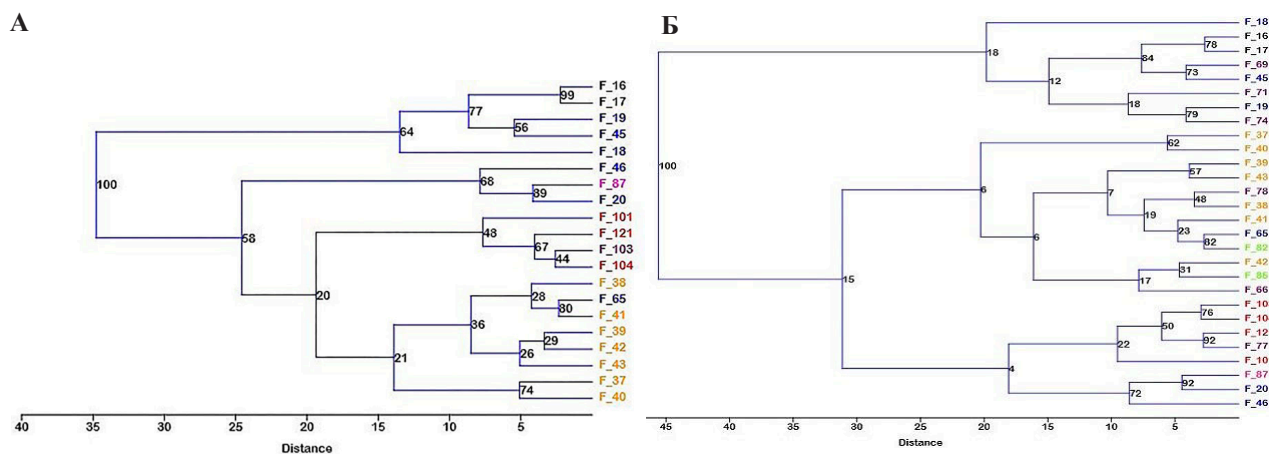


Рис. 7. Генетические расстояния между исследуемыми пчелосемьями, собранными на территории Беларуси (А), на территории Беларуси, Чувашии, Башкирии (Б)

Установление чистоты и гибридности исследуемых пчелосемей

Анализ структуры пчелосемей с использованием моделирования в STRUCTURE v.2.3.4 обнаружил достоверную разницу между образцами и показал предпочтительное разделение на два кластера (K) (группы) ($K = 2$) (рис. 8). Данная программа использует методы Монте-Карло по схеме марковской цепи (Markov Chain Monte Carlo, MCMC), что позволяет минимизировать неравновесие Харди – Вайнберга и неравенство по сцеплению локусов в кластерном анализе индивидов. Ко-

личество кластеров находилось в диапазоне от 1 до 10.

Точность консолидированности (принадлежности к тому или иному кластеру) большинства особей в группе 1 (93 особи из 155, включенных в анализ) варьировала в пределах от 70,8 до 98,6%. Аналогичный показатель был для особей в группе 2 (62/155) — от 69,7 до 98,8%. Зафиксирована 21 особь, вероятность отнесения которой к тому или иному кластеру колебалась от 50,1 до 68,6%, что характерно для пчелосемей с высокой степенью метизации.

Установлено, что пчелосемьи 37, 39, 40, 41, 42, собранные в Гомельской области, и пчелосемьи 18, 19, 20, собранные в Витебской области, являются генетически однородными, т. е. без примеси генетического материала другого подвида пчел. Следовательно, данные семьи можно считать чистопородными. Остальные семьи имели значительную степень гибридности, причина которой интрогрессия чужеродного генетического материала как результат неконтролируемого завоза и разведения медоносных пчел разных подвидов на территории Беларуси (рис. 9).

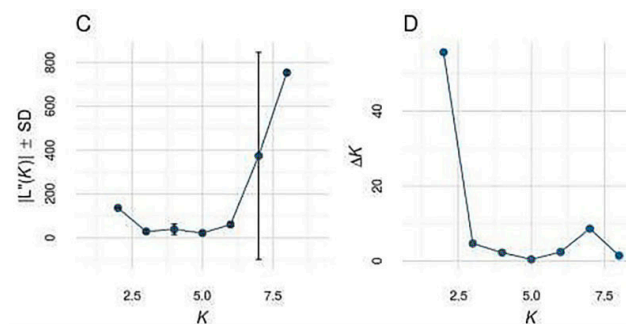


Рис. 8. Результаты моделирования в STRUCTURE v.2.3.4

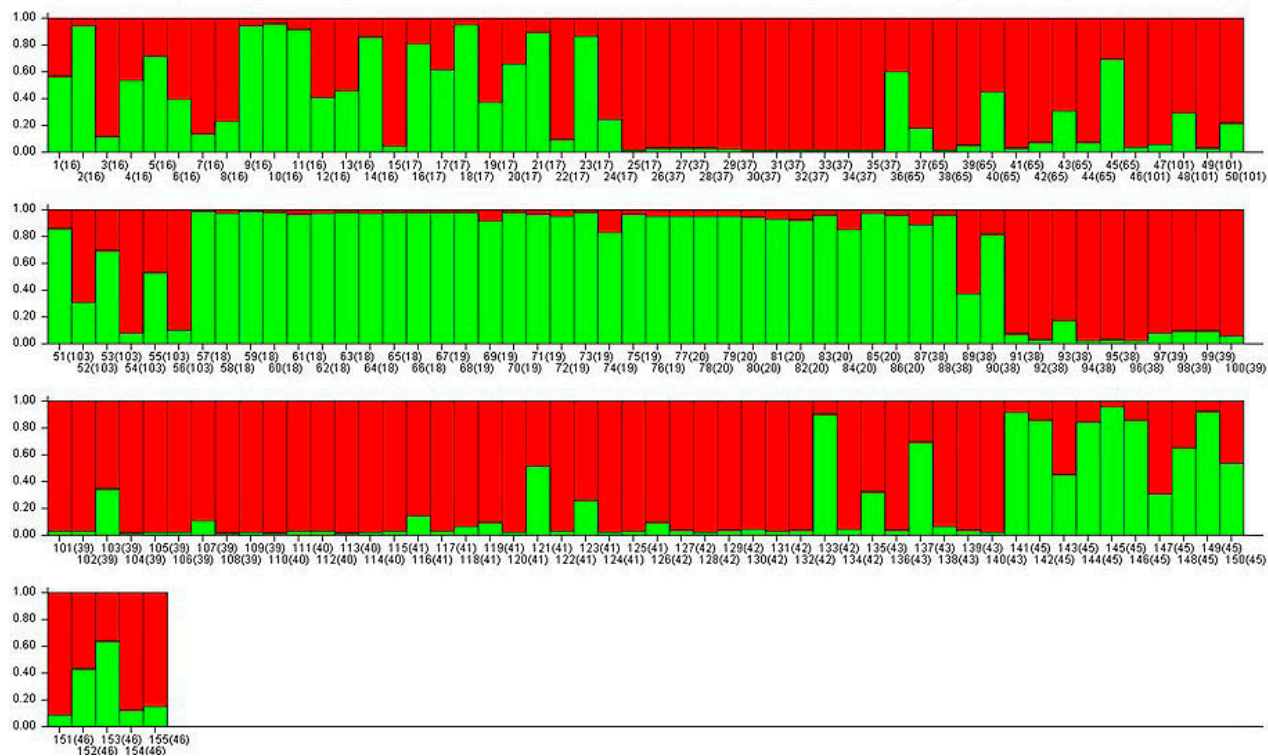


Рис. 9. Результаты моделирования в STRUCTURE v.2.3.4

Заключение

Цель исследований заключалась в достоверной идентификации пчел, относящихся к аборигенному подвиду темная лесная пчела, с помощью молекулярно-генетических маркеров; проведении популяционно-генетических исследований и оценке степени метизации пчелосемей указанного подвида. В задачи исследования входило проведение комплексного анализа как митохондриальной ДНК, так и STR-локусов ядерной ДНК, с тем чтобы определить генетическую структуру и филогенетические связи пчел подвида *A. m. mellifera*, собранных на территории Беларуси.

Впервые с помощью методов молекулярно-генетического анализа подтверждено существование темной лесной пчелы на территории Беларуси. Генетическая структура анализируемых популяций темной лесной пчелы отличалась в зависимости от территории обитания. Обнаружено три варианта комбинаций Р и Q элементов в локусе COI-COII: PQ, PQQ, PQQQ. Для образцов Гомельской области характерна комбинация PQQ (800 п. н.), для образцов, собранных на территории других областей, типичных комбинаций не обнаружено. Анализ STR-локусов показал высокое генетическое разнообразие. Выявлено 179 аллелей размером от 65 до 340 п. н. В среднем количество аллелей на локус составило 22,375, что является высоким значением; количество эффективных аллелей колебалось от 11,492 до 3,404 в зависимости от локуса и в среднем составило 6,829. Среднее значение индекса Шеннона — 2,230, что указывает на достаточно высокое генетическое разнообразие изучаемых образцов *A. m. mellifera*. Для образцов популяции Гомельской и Витебской областей характерно наибольшее количество и выявленных, и уникальных аллелей по сравнению с популяциями других областей Беларуси.

Филогенетический анализ выявил связь между пчелосемьями, обитающими на одной территории, а также генетическую близость образцов из северной части Беларуси (Витебская область) и Чувашии, что может свидетельствовать о подобию генотипов, обеспечивающих выживание в сходных условиях среды. Индивидуальный молекулярно-генетический анализ пчел позволил оценить степень метизации в пчелосемьях. Выявлены восемь

генетически однородных пчелосемей подвида *A. m. mellifera*. Данные семьи могут быть использованы для восстановления популяций темной лесной пчелы в контролируемых условиях национальных парков и заповедников.

Полученные результаты исследований являются новыми данными о генетической структуре популяций медоносных пчел *A. m. mellifera*, встречающихся на территории Республики Беларусь.

Список использованных источников

1. ООН: программа по окружающей среде [сайт]. – UNEP, 2025. – URL: <https://www.unep.org/ru> (дата обращения: 06.06.2025)
2. Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. – URL: www.ipbes.net (date of access: 06.06.2025)
3. Conserving diversity and vitality for honey bee breeding / M. D. Meixner, C. Costa, P. Kryger [et al.] // Journal of Apicultural Research. – 2010. – Vol. 49, № 1. – P. 85–92.
4. Biodiversity, conservation and current threats to European honeybees / P. De la Rúa, R. Jaffe, R. Dall'Olio [et al.] // Apidologie. – 2010. – Vol. 40. – P. 263–284. doi: 10.1051/apido/2009027
5. A review of methods for discrimination of honey bee populations as applied to European beekeeping / M. Bouga, C. Alaux, M. Bienkowska [et al.] // Journal of Apicultural Research. – 2011. – Vol. 50. – P. 51–81.
6. The influence of genetic origin and its interaction with environmental effects on the survival of *Apis mellifera* L. colonies in Europe / R. Büchler, C. Costa, F. Hatjina [et al.] // Journal of Apicultural Research. – 2014. – Vol. 53. – P. 205–214.
7. Старостенко, Е. В. Какие пчелы лучшие? / Е. В. Старостенко. – Минск : Урожай, 1969. – 100 с.
8. The dark European honey bee *Apis mellifera mellifera* in Lithuania: data on mitotype diversity of native bee population / S. Orlovskytė, E. Budrys, V. Skrodenyte-Arbaciauskiene, L. Blazyte-Cereskiene // Journal of Apicultural Research. – 2025. – Vol. 64, № 3. doi: 10.1080/00218839.2024.2327125
9. Nei, M. Analysis of gene diversity in subdivided populations / M. Nei // Proc. Natl. Acad. Sci. U.S.A. – 1973. – Vol. 70, № 12. – P. 3321–3323.

10. Peakall, R. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research — an update / R. Peakall, P. E. Smouse // *Bioinformatics*. – 2012. – Vol. 28, № 19. – P. 2537–2539.
11. Pritchard, J. K. Inference of population structure using multilocus genotype data / J. K. Pritchard, M. Stephens, P. Donnelly // *Genetics*. – 2000. – Vol. 155, № 2. – P. 945–959.
12. Evanno, G. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study / G. Evanno, S. Regnaut, J. Goudet // *Mol. Ecol.* – 2005. – Vol. 14, № 8. – P. 2611–2620.
13. Earl, D. A. Structure harvester: a website and program for visualizing structure output and implementing the Evanno method // *Conservation Genetics Resources*. – 2012. – Vol. 4, № 2. – P. 359–361.
14. Tamura, K. MEGA11: Molecular Evolutionary Genetics Analysis Version 11 / K. Tamura, G. Stecher, S. Kumar // *Molecular Biology and Evolution*. – 2021. – Vol. 38, iss. 7. – P. 3022–3027. doi: 10.1093/molbev/msab120
15. Николенко, А. Г. Полиморфизм локуса COI-COII митохондриальной ДНК *Apis mellifera* L. на Южном Урале / А. Г. Николенко, А. В. Поскряков. – *Генетика*. – 2002. – № 4. – С. 458–462.
16. Genetic biodiversity of the West European honeybee (*Apis mellifera mellifera* and *Apis mellifera iberica*). II. Mitochondrial DNA / L. Garnery, P. Franck, E. Baudry [et al.]. – *Genetics Selection and Evolution*. – 1998. – Vol. 30. – P. 31–47.
17. Bee genetics and breeding / Sylvester H.A. Biochemical genetics. In: Rinderer T, editor. – Orlando, Florida: Academic Press. – 1986. – P. 177–203.
18. The mitochondrial genome of the Spanish honey bee, *Apis mellifera iberiensis* (Insecta: Hymenoptera: Apidae), from Portugal / L. Boardman, A. Eimanifar, R. Kimball [et al.] // *Mitochondrial Dna Part B*. – 2020. – Vol. 5, № 1. – P. 17–18. doi: 10.1080/23802359.2019.1693920
19. Genomic Analyses Reveal Demographic History and Temperate Adaptation of the Newly Discovered Honey Bee Subspecies *Apis mellifera sinixinyuan* n. ssp. / Ch. Chen, L. Zhiguang, Qi Pan [et al.] // *Molecular Biology and Evolution*. – 2016. – Vol. 33, iss. 5. – P. 1337–1348. doi: 10.1093/molbev/msw017
20. Генетическая характеристика медоносных пчел *Apis mellifera* L., разводимых на пасеках Беларуси // Е. В. Гузенко, А. И. Царь, В. Н. Кипень, В. А. Лемеш // *Пчеловодство холодного и умеренного климата: Материалы 5-й Международной научно-практической конференции. Москва-Псков. 19–20 октября 2021 г. Псков, 2021. – С. 48–52*
21. A simple test using restricted PCR-amplified mitochondrial DNA to study the genetic-structure of *Apis mellifera* L. / L. Garnery, M. Solignac, G. Celebrano, J.-M. Cornuet – *Experientia*. – 1993. – Vol. 49, Iss. 11. – P. 1016–1021.
22. Rajabi-Maham, H. Genetic diversity evaluation of Persian honeybees (*Apis mellifera meda*) in North West of Iran, using microsatellite markers / H. Rajabi-Maham, T. Ghasemi, Shahrokh Pashaei-rad // *Journal of Wildlife and Biodiversity*. – 2018. – № 2(2). – P. 37–46. doi: 10.22120/jwb.2018.86877.1026
23. Genetic structure of Balearic honeybee populations based on microsatellite polymorphism / P. De la Rúa, J. Galian, J. Serrano, R. F. A. Moritz // *Genetics Selection Evolution*. – 2003. – № 35. – P. 339–350. doi: 10.1186/1297-9686-35-3-339
24. Гатауллин, А. Р. Генетическая структура уральской популяционной системы темной лесной пчелы *Apis mellifera mellifera* L. // А. Р. Гатауллин, А. А. Каримова, А. Г. Николенко // *Биомика*. – 2016. – Т. 8, № 2. – С. 82–83.
25. Анализ генетической структуры популяций медоносной пчелы (*Apis mellifera* L.) / М. Д. Каскинова, Р. А. Ильясов, А. В. Поскряков, А. Г. Николенко // *Генетика*. – 2015. – Т. 51. № 10. – С. 1199–1202.

Е. В. Guzenko, A. I. Tsar

GENETIC STRUCTURE AND VARIABILITY OF *APIS MELLIFERA* *MELLIFERA* POPULATIONS IN THE REPUBLIC OF BELARUS

State Scientific Institution

“Institute of Genetics and Cytology of the National Academy of Sciences of Belarus”

27 Akademicheskaya St., 220072 Minsk, the Republic of Belarus

e-mail: e.guzenko@igc.by

The dark forest bee *A. m. mellifera* is the only known honeybee subspecies that has perfectly adapted to life in the cold climate of Northern Europe. For Belarus, the dark forest bee is a native subspecies, which was presumed to have become extinct as a result of hybridization. Until recently, there was no reliable data evidencing the existence of the dark forest bee in Belarus. The object of this study were apiary and wild-hive honeybees collected in the territory of the Republic. Using a set of molecular genetic markers, the native *A. m. mellifera* subspecies was identified. The STR loci analysis revealed the high genetic variability of the samples in general and especially those collected in Gomel and Vitebsk regions. Analysis of genetic distances established the existence of structural groupings. Individual molecular genetic analysis revealed a high degree of crossbreeding in bee colonies. Genetically homogeneous bee colonies of the subspecies *A. m. mellifera*, which may form the basis of a micropopulation for the restoration and conservation of native honey bees in Belarus, were identified.

Keywords: honey bee, *Apis mellifera mellifera*, COI-COII mtDNA locus, STR markers, population genetic structure, crossbreeding, genetic homogeneity.

Дата поступления в редакцию: 05 марта 2025 г.

М. Е. Михайлова¹, Е. В. Снытков¹, Н. И. Тиханович¹, Н. А. Камыш¹, Р. И. Шейко¹,
А. А. Сермягин², Н. А. Зиновьева²

СРАВНИТЕЛЬНЫЙ АНАЛИЗ ГЕНЕТИЧЕСКОЙ СТРУКТУРЫ КРАСНОГО БЕЛОРУССКОГО И ГОЛШТИНСКОГО СКОТА ОТЕЧЕСТВЕННОЙ СЕЛЕКЦИИ ПО МИКРОСАТЕЛЛИТНЫМ ЛОКУСАМ

¹Государственное научное учреждение

«Институт генетики и цитологии Национальной академии наук Беларуси»

Республика Беларусь, 220072, г. Минск, ул. Академическая, 27

e-mail: M.Mikhailova@igc.by

²Федеральное государственное бюджетное научное учреждение

«Федеральный исследовательский центр животноводства — ВИЖ им. академика Л. К. Эрнста»

Российская Федерация, 142132, Московская область, городской округ Подольск, поселок Дубровицы, д. 60

При помощи 12 STR-маркеров исследован генетический полиморфизм выборки аборигенного красного белорусского скота (101 гол.) и голштинского скота белорусской селекции (126 гол.). Выявлены генетические различия между выборками популяций красного белорусского и голштинского скота отечественной селекции. Установлено, что в выборке красного белорусского скота в среднем на локус (Na) пришлось $7,33 \pm 0,71$ аллеля, число эффективных аллелей (Ne) составило $3,76 \pm 0,41$. Наблюдаемая гетерозиготность (Ho), которая является показателем полиморфности популяции и характеризует гетерозиготность исследуемой группы, составила $0,69 \pm 0,04$, ожидаемая же гетерозиготность (He) составила $0,7 \pm 0,03$. Выявлено, что индекс фиксации (F) составил $0,001 \pm 0,039$, что указывает на низкую степень дифференциации и генетическую однородность популяции красного белорусского скота по исследованным нами STR-локусам, а также незначительный сдвиг в сторону близкородственного скрещивания (дефицит гетерозигот), т. к. индекс фиксации для локусов ETH10, TGLA126 и TGLA53 составил 0,128, 0,101 и 0,285 соответственно. Для выборки голштинского скота в среднем на локус пришлось $7,33 \pm 0,51$ аллеля, число эффективных аллелей составило $3,42 \pm 0,321$. Наблюдаемая гетерозиготность составила $0,68 \pm 0,04$, ожидаемая гетерозиготность — $0,67 \pm 0,04$. Установлено, что в отличие от красного белорусского скота, голштинский скот имеет отрицательный индекс фиксации, что указывает на избыток гетерозигот и низкий уровень инбридинга, выявлены приватные аллели, которые могут характеризовать генетическую структуру красного белорусского и голштинского скота.

Исследованные генетические маркеры имеют важное значение для селекции животных и сохранения генетических ресурсов, а также в перспективе могут стать основой для разработки целевых программ разведения с целью улучшения желаемых признаков путем предотвращения потери уникальных аллелей в молочных породах крупного рогатого скота, разводимых в Беларуси.

Ключевые слова: микросателлиты, красный белорусский скот, голштинская порода отечественной селекции, крупный рогатый скот, популяция, генетическое разнообразие.

Для цитирования: Сравнительный анализ генетической структуры красного белорусского и голштинского скота отечественной селекции по микросателлитным локусам / М. Е. Михайлова, Е. В. Снытков, Н. И. Тиханович [и др.] // Молекулярная и прикладная генетика: сб. науч. тр. / Ин-т генетики и цитологии НАН Беларуси; редкол.: А. В. Кильчевский (гл. ред.) [и др.]. — Минск, 2025. — Т. 39. — С. 54–64.

Введение

Основное поголовье молочного скота в Беларуси представлено голштинской породой крупного рогатого скота (КРС) белорусской

селекции, а также породной группой красного белорусского скота (КБС). Голштинская порода, выведенная в Северной Америке методом чистопородного разведения с использованием

голландского скота, считается одной из лучших в мире среди молочных пород. Эта порода обладает высокой молочной продуктивностью, однако для реализации генетического потенциала необходимо качественное полноценное питание — в противном случае продуктивность значительно снижается. В 20-х гг. ХХI столетия, в результате активной селекционной работы по улучшению существующей белорусской черно-пестрой породы за счет скрещивания с зарубежным высокопродуктивным голландским скотом, в Беларуси создана и апробирована голштинская порода молочного скота отечественной селекции. Животные этой породы характеризуются высоким уровнем продуктивности, выраженным молочным типом, хорошими воспроизводительными качествами [1].

Голштинская порода является наиболее распространенной среди молочного скота, потенциал продуктивности которой находится на уровне 12–14 тыс. кг молока за лактацию. Эта порода занимает лидирующее положение в Беларуси и обладает следующей продуктивностью: удой — 5000–7000 кг, жир 3,8–4,12%, белок — 3,2–3,4%.

Породная группа КБС отличается высокой приспособленностью животных к местным условиям, неприхотливостью, крепким здоровьем, хорошими воспроизводительными качествами, молочной продуктивностью, повышенным содержанием в молоке белка и жира, специфическим полиморфизмом белков крови и молока. Имеет следующую продуктивность (за 305 дней наивысшей лактации): удой — 7353 кг, жир — 4,67%, белок — 3,8% [2].

КБС является ценной отечественной популяцией, которая издавна разводилась на территории Республики Беларусь. Первое обследование красного скота Гродненской губернии было проведено агрономами Дедовым Н. А., Михно А. И., Ширяевым В. И. и Атаназевичем Я. И. в 1910 г. В 1911 г. в Гродненской губернии была издана первая племенная книга КБС, запись в которую проводилась на ежегодных выводках лучших животных [3]. В 1960 г. в Беларуси в результате многолетней селекционной работы в колхозах, совхозах и частных хозяйствах создана и официально утверждена качественно новая породная груп-

па КБС молочного направления, поголовье которой превысило 30 тыс. животных, а также приняты меры по дальнейшему совершенствованию породных качеств и увеличению поголовья КБС в республике. Известно, что к середине 70-х гг. ХХ в. в мире произошло резкое уменьшение численности локальных пород сельскохозяйственных животных, обладающих целым рядом ценных хозяйственно-полезных качеств, но не отличающихся высокой продуктивностью. Исчезло около 30 пород КРС. В этот же период начался процесс резкого сокращения и породной группы КБС, которая наряду с другими 12 аборигенными породами стран бывшего СССР попала на грань исчезновения [4, 5]. Вместе с этим важно подчеркнуть, что сохранение генофонда исчезающих пород является необходимым условием обогащения генетических ресурсов и увеличивает возможности создания новых пород сельскохозяйственных животных, отвечающих запросам общества и требованиям времени [6]. В этой связи в начале 2000-х гг. на государственном уровне было принято решение о сохранении генофонда и создании стада КБС. Генофондное стадо принадлежит УСП «Новый Двор-Агро» Свислочского района, Гродненской области. В 2021 г. поголовье КБС составляло 376 гол., а в 2025 г. — 3928 гол., из них коров насчитывалось 2059 гол.

Одной из основных задач при сохранении и воспроизводстве редких видов животных является оценка их генетического разнообразия. Согласно исследованиям Felius с соавторами (2011) [7], КБС был отнесен в группу пород КРС балтийского красного скота [5]. Ernst с соавторами (2020 г.) показали, что на протяжении всей своей истории КБС улучшали ангельнским и немецким красным скотом в конце XIX — начале XX вв., польским красным и датским красным — в 20–30-х гг., эстонским красным и латышским красным — в 50-е г. и снова датским красным скотом — в 90-х гг. XX в. [8]. За последние 30 лет в Республике Беларусь черно-пестрая и голштинская породы пришли на смену аборигенной красной белорусской популяции, и сегодня она близка к исчезновению. В предыдущих исследованиях нами впервые изучен КБС на полногеномном уровне [11]. Показано отличие КБС от других европейских таурино-

вых (пород красного корня). Эти результаты согласуются с результатами ранних исследований, основанных на других типах ДНК-маркеров, и отражают происхождение КБС и его недавнюю историю [5, 9, 10]. Результаты анализа главных компонент, анализа «присоединения соседей» (Neighbor-Joining) на основе F_{st} и анализа популяционной структуры в STRUSTRUCTURE четко отличают КБС от других европейских пород КРС.

Красный белорусский и отечественный голштинский скот подвергаются действию направленной селекции для увеличения удоев, улучшения качества молока и повышения экономической эффективности молочного животноводства. В связи с этим, целью нашего исследования является изучение генетической структуры КБС и голштинского скота белорусской селекции по микросателлитным локусам. Понимание генетической основы популяции может способствовать повышению эффективности программ сохранения и селекции, а также обогащению генетического потенциала молочных пород скота, разводимого в Беларуси.

Материалы и методы

В ходе исследования генетической структуры аборигенного КБС (101 гол., УСП «Новый Двор-Агро») и голштинского скота белорусской селекции (126 гол., СПК «Агро-

комбинат Снов») была использована панель из 12 высокополиморфных STR-маркеров (микросателлитов), рекомендованных Международным обществом генетики животных (ISAG, International Society for Animal Genetics): BM1818, BM1824, BM2113, ETH3, ETH10, ETH225, INRA023, SPS115, TGLA122, TGLA126, TGLA227 и TGLA53. Применение международной стандартной панели STR-локусов позволяет сравнивать рассчитываемые параметры генетических оценок между популяциями.

В качестве источника биологического материала был использован ушной выщип. Выделение ДНК проводили набором «Нуклеосорб» в комплектации «С» («Праймтех», Беларусь), согласно прилагаемой к набору инструкции. Образцы ДНК были растворены в стерильной бидистиллированной воде, концентрация ДНК в каждом образце составляла 20 нг/мкл. Реакционная смесь для одной ПЦР-реакции объемом 20 мкл содержала компоненты, описанные в таблице 1.

Реакционная смесь помещалась в стерильные пробирки объемом 0,2 мл для проведения ПЦР. Амплификация осуществлялась на автоматическом программируемом термоциклере «Bio-Rad C-1000» (США). Условия амплификации представлены в таблице 2.

Температура отжига для каждого STR-праймера рассчитывается отдельно. По окончании

Таблица 1

Состав реакционной смеси и концентрации используемых реактивов

Компоненты реакционной смеси для ПЦР-реакции	Конечная концентрация компонентов / объем
1x Tag-буфер с $(\text{NH}_4)_2\text{SO}_4$	1 x / 2 мкл
Смесь dNTP	2,0 мМ/мкл / 2 мкл
MgCl_2	1,5 мМ/мкл / 2 мкл
Праймер F	10 пМ/мкл / 1 мкл
Праймер R	10 пМ/мкл / 1 мкл
Taq-полимераза	0,07 Ед/мкл / 0,3 мкл
ДНК	(20 нг/мкл) / 1 мкл
H_2O	до 10,7 мкл
Общий объем	20 мкл

Таблица 2

Условия проведения полимеразной цепной реакции для фрагментного анализа ДНК

Метод	Режим проведения полимеразной цепной реакции
Полимеразная цепная реакция (ПЦР)	«Горячий старт»: 98 °С продолжительность — 2 мин, 30 циклов ПЦР: • Денатурация, 98 °С — 10 сек; • Отжиг праймеров, 61,5 °С — 1 мин; • Элонгация, 72 °С — 30 сек; Финальная элонгация 5 мин при 72 °С; 5 мин при 12 °С

реакции пробы хранят при температуре +4 °С (–20 °С для длительного хранения).

Для определения длин микросателлитных фрагментов использовали генетический анализатор (секвенатор) Applied Biosystems 3500 Genetic Analyzer (США). Расшифровка и документирование полученных результатов осуществлялось с помощью программного обеспечения GeneMapper 4.1.

В качестве стандарта при определении величины амплифицированного фрагмента ДНК используется размерный стандарт S450 (Синтол, Россия).

Степень генетической дифференциации популяций оценивали в программе Structure.2.3.4. Оптимальное число кластеров (K) определяли с помощью программы Structure Harvester, Past v.4.03, а также веб-приложения POPHELPER (<http://pophelper.com/>) и расширения для Microsoft Excel GenAIEx v.6.503.

Результаты и обсуждения

Для выборки КБС в среднем на locus (Na) пришлось $7,33 \pm 0,71$ аллеля (от 5 у локусов BM1824 и SPS115 до 13 у локуса TGLA53), число эффективных аллелей (Ne) составило $3,76 \pm 0,41$ (от 1,91 у локуса ETH3 до 6,43 у локуса TGLA53). Значение индекса разнообразия Шеннона (I), характеризующее меру генетического разнообразия изучаемой выборки, составило $1,45 \pm 0,11$, диапазон значений составил от 0,95 у локуса ETH3 до 2,1 у локуса TGLA53. Наблюдаемая гетерозиготность (Ho), которая является показателем полиморфности популяции и характеризует гетерозиготность исследуемой выборки, составила $0,69 \pm 0,04$ (от 0,44 у локуса ETH3 до 0,87 у локуса BM2113), ожидаемая же гетерозиготность (He), которая рассчитывается исходя из равновесия Харди – Вайнберга, составила $0,7 \pm 0,03$ (от 0,48

у локуса ETH3 до 0,85 у локуса TGLA53). Индекс фиксации (F), являющийся мерой дифференциации популяции, составил $0,001 \pm 0,039$ (от –0,18 у локуса SPS115 до 0,29 у локуса TGLA53), что указывает на низкую степень дифференциации внутри группы и генетическую однородность выборки по исследуемым нами локусам (табл. 3). Следует отметить, что F для локусов ETH10, TGLA126 и TGLA53 составил 0,128, 0,101 и 0,285 соответственно, что указывает на незначительный сдвиг в сторону близкородственного скрещивания (дефицит гетерозигот). Можно предположить, что сдвиг этого показателя в положительную сторону произошел из-за целенаправленного отбора животных в ходе селекционного процесса по ценным сельскохозяйственным признакам, что может косвенно указывать на ассоциацию этих STR-маркеров с локусами количественных признаков QTL.

Для выборки голштинского скота в среднем на locus пришлось $7,33 \pm 0,51$ аллеля (от 6 у локусов BM1818, BM1824, BM2113, ETH3 и TGLA126 до 12 у локуса TGLA53), число эффективных аллелей составило $3,42 \pm 0,321$ (от 1,58 у локуса SPS115 до 5,09 у локуса TGLA53). Аналогичные данные получены Н. И. Стрекозовым с соавторами и Н. С. Фуревой с соавторами на КРС сычевской, бурой швицкой и ярославской пород, число аллелей на locus составило 7,4 [12–15].

Значение индекса разнообразия Шеннона (I) составило $1,37 \pm 0,09$, диапазон значений — от 0,81 у локуса SPS115 до 1,82 у локуса TGLA53. Для микросателлитных маркеров значения I считаются высокими, если они не менее 1,5. Это подтверждает достаточно высокое генетическое разнообразие в изученной выборке. Наблюдаемая гетерозиготность составила $0,68 \pm 0,04$ (от 0,38 у локуса SPS115

Таблица 3

Характеристика генетического разнообразия выборки красного белорусского скота

Locus	Na	Ne	I	Ho	He	F
BM1818	6,000	2,961	1,244	0,743	0,662	-0,121
BM1824	5,000	3,121	1,268	0,733	0,680	-0,078
BM2113	6,000	4,644	1,656	0,871	0,785	-0,110
ETH3	6,000	1,912	0,948	0,436	0,477	0,086
ETH10	7,000	3,254	1,338	0,604	0,693	0,128
ETH225	8,000	3,572	1,513	0,653	0,720	0,093
INRA023	6,000	4,679	1,592	0,832	0,786	-0,058
SPS115	5,000	2,132	1,045	0,624	0,531	-0,175
TGLA122	10,000	4,495	1,779	0,861	0,778	-0,108
TGLA126	6,000	2,281	1,019	0,505	0,562	0,101
TGLA227	10,000	5,577	1,920	0,842	0,821	-0,025
TGLA53	13,000	6,434	2,095	0,604	0,845	0,285
Среднее	7,333	3,755	1,451	0,692	0,695	0,001
Стандартное отклонение	0,711	0,410	0,107	0,042	0,034	0,039

до 0,84 у локуса TGLA122), ожидаемая гетерозиготность составила $0,67 \pm 0,04$ (от 0,37 у локуса SPS115 до 0,8 у локуса TGLA53). Индекс фиксации составил $-0,011 \pm 0,019$ (от $-0,096$ у локуса INRA023 до 0,105 у локуса

ETH3) (табл. 4). Стоит отметить, что в отличие от КБС, голштинский скот имеет отрицательный индекс фиксации, что указывает на избыток гетерозигот и низкий уровень инбридинга.

Таблица 4

Характеристика генетического разнообразия выборки голштинского скота

Locus	Na	Ne	I	Ho	He	F
BM1818	6,000	2,488	1,101	0,595	0,598	0,005
BM1824	6,000	3,284	1,363	0,746	0,695	-0,073
BM2113	6,000	3,657	1,413	0,754	0,727	-0,038
ETH3	6,000	2,519	1,102	0,540	0,603	0,105
ETH10	8,000	3,899	1,632	0,730	0,744	0,018
ETH225	7,000	2,917	1,233	0,651	0,657	0,010
INRA023	7,000	4,055	1,493	0,825	0,753	-0,096
SPS115	7,000	1,583	0,812	0,381	0,368	-0,034
TGLA122	8,000	4,506	1,711	0,841	0,778	-0,081
TGLA126	6,000	2,182	0,968	0,579	0,542	-0,070
TGLA227	9,000	4,854	1,811	0,778	0,794	0,020
TGLA53	12,000	5,090	1,819	0,722	0,804	0,101

Окончание таблицы 4

Locus	Na	Ne	I	Ho	He	F
Среднее	7,333	3,420	1,371	0,679	0,672	-0,011
Стандартное отклонение	0,512	0,321	0,097	0,039	0,037	0,019

На основании генетических дистанций для анализированных нами STR-локусов был построен график анализа главных компонент (PCA), который отражает различие между выборками красного белорусского и голштинского скота (рис. 1). На данном графике можно увидеть, что обе группы имеют генетические различия и при визуализации не смешиваются между собой, формируя два кластера.

С помощью программы STRUCTURE v.2.3.4 был осуществлен расчет критерия Q, который характеризует принадлежность каждого отдельного животного к конкретной популяции. Для расчета критерия Q использовали

следующие параметры: число предполагаемых кластеров составило от 1 до 10, длина burn-in периода — 50 000, модель марковских цепей Монте-Карло — 5000, число итераций — 10. Результат работы был проанализирован с использованием веб-приложения POPHELPER. Метод deltaK основан на анализе второй производной функции правдоподобия, полученной из результатов запуска программы STRUCTURE с разными значениями числа кластеров (K). Результат Евапо-анализа показал наличие в исследуемой выборке двух кластеров при deltaK равном 233,314 [16] (табл. 5).

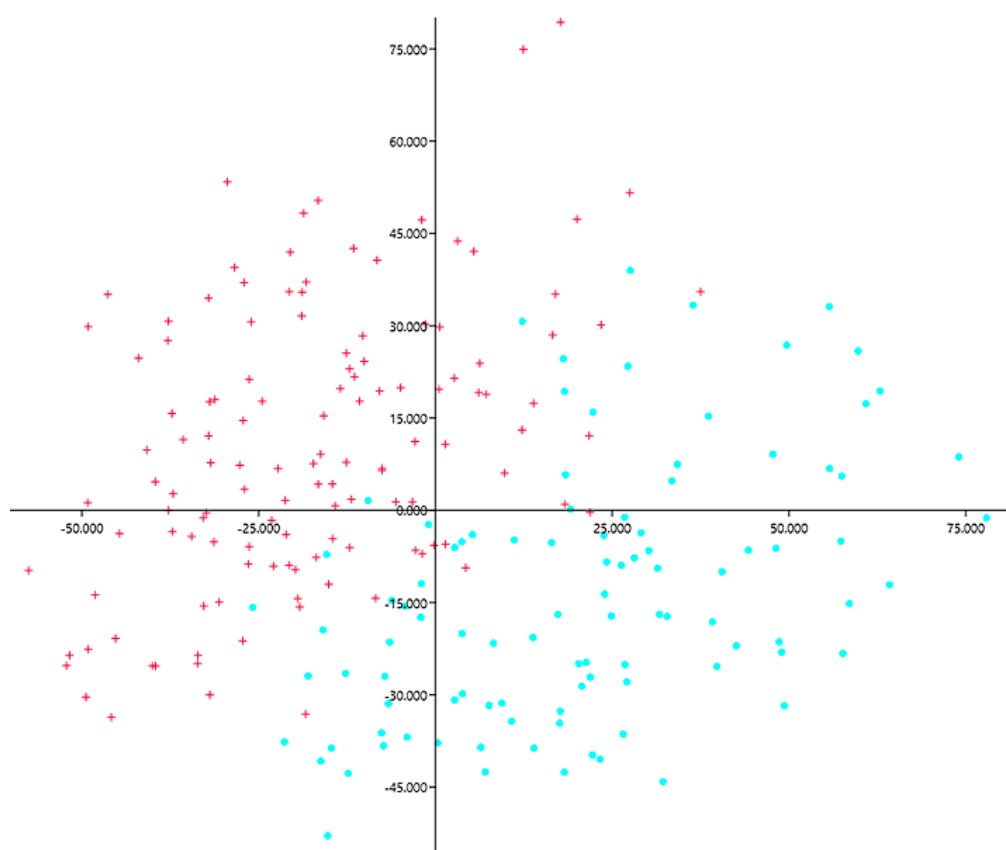


Рис. 1. График анализа главных компонент генетических дистанций между выборками красного белорусского и голштинского скота. Условные обозначения: синие точки — белорусский красный скот; красные крестики — голштинский скот

Таблица 5
 Результаты Евано-анализа для полученных в STRUCTURE v.2.3.4 данных микросателлитных локусов

	loci	ind	k	runs	elpdmean	elpdsd	elpdmin	elpdmax	lnk1	lnk1max	lnk1min	lnk2	lnk2max	lnk2min	deltaK
1	12	227	1	10	-8345,47	1,81	-8348,2	-8341,5	—	—	—	—	—	—	—
2	12	227	2	10	-7788,00	1,85	-7791,0	-7784,1	557,47	557,51	557,43	431,63	436,47	426,79	233,314
3	12	227	3	10	-7662,16	6,73	-7673,1	-7651,6	125,84	130,72	120,96	48,50	49,71	47,29	7,207
4	12	227	4	10	-7584,82	10,40	-7606,3	-7565,8	77,34	81,01	73,67	65,24	93,72	36,76	6,273
5	12	227	5	10	-7572,72	42,55	-7657,4	-7511,0	12,10	44,25	-20,05	43,18	65,50	20,86	1,015
6	12	227	6	10	-7517,44	32,72	-7588,2	-7482,5	55,28	65,11	45,45	69,49	78,72	60,26	2,124
7	12	227	7	10	-7531,65	32,12	-7610,0	-7499,2	-14,21	-13,61	-14,81	1,64	1,74	1,54	0,051
8	12	227	8	10	-7544,22	32,62	-7609,2	-7490,2	-12,57	-12,07	-13,07	107,18	176,39	37,97	3,286
9	12	227	9	10	-7663,97	102,33	-7905,7	-7561,4	-119,75	-50,04	-189,46	98,13	141,07	55,19	0,959
10	12	227	10	10	-7685,59	75,56	-7841,6	-7590,0	-21,62	5,15	-48,39	—	—	—	—

Таким образом, метод *deltaK* позволяет идентифицировать наиболее вероятное число генетических кластеров в анализе структурной популяционной генетики, что делает его очень полезным инструментом для интерпретации

генетических данных и выявления популяционной структуры [17].

На рисунке 2 приведен результат кластеризации каждого образца в веб-приложении POPHELPER для двух кластеров.

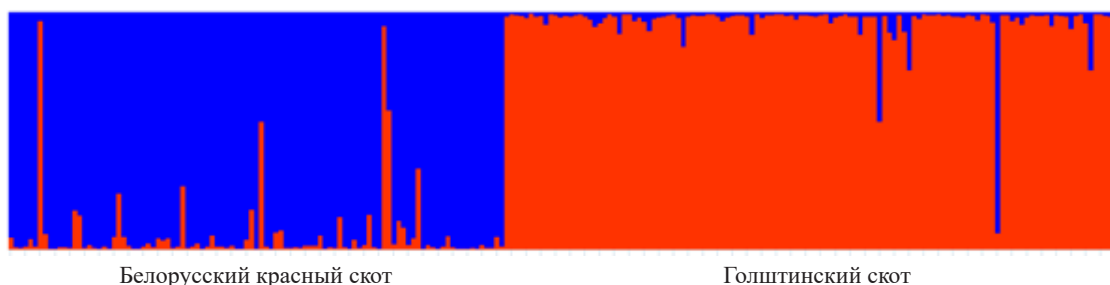


Рис. 2. Результаты кластеризации образцов выборок из популяций красного белорусского и голштинского скота, полученные с помощью программы Structure v.2.3.4. ($K = 2$). Синий цвет — белорусский красный скот, красный цвет — голштинский скот

График количественной кластеризации в программе STRUCTURE используется для визуализации результатов анализа генетической структуры популяций. На этом графике отражается вероятность принадлежности каждого индивида к различным генетическим кластерам, которые были обнаружены в ходе анализа. Таким образом, график количественной кластеризации позволяет оценить уровень генетической схожести или различий между популяциями и индивидами.

Исходя из представленного графика, очевидно, что проанализированные образцы чет-

ко разделяются на разные кластеры с высокой вероятностью принадлежности. Исключение составляют три образца из популяции КБС, которые с высокой вероятностью были отнесены к голштинскому скоту, а также два голштинских образца, которые, напротив, с высокой вероятностью были классифицированы как КБС.

Кроме того, для каждой из выборок наблюдаются особенности генетической структуры, которые выражены как в разнице частот аллелей, так и в наличии «приватных» аллелей, которые наблюдаются только в одной из исследуемых групп (табл. 6).

Таблица 6

Характеристика генетической структуры выборки популяций красного белорусского и голштинского скота по выявленным приватным аллелям STR-локусов

Выборка	Маркер	Аллель	Частота (%)
Белорусский красный скот	BM1818	258	5,4
	ETH3	121	0,5
	ETH225	154	1
	TGLA122	153	5
	TGLA122	165	1
	TGLA126	151	0,5
	TGLA227	85	13,4
	TGLA227	95	0,5
	TGLA53	164	5,4
	TGLA53	170	26,2
	TGLA53	180	1

Окончание таблицы 6

Выборка	Маркер	Аллель	Частота (%)
Голштинский скот	BM1818	270	2
	BM1824	172	0,4
	BM2113	131	0,4
	ETH3	139	0,4
	ETH10	209	11,1
	INRA023	198	0,8
	SPS115	250	0,4
	SPS115	258	0,4
	TGLA227	99	5,6
	TGLA53	154	1,6
	TGLA53	176	2

Анализ данных показывает, что в выборке КБС выявлено 11 «приватных» аллелей, распределенных по семи STR-локусам. Наиболее высокие частоты наблюдаются для аллелей TGLA53–170 (26,2%) и TGLA227–85 (13,4%), что указывает на их возможную значимость в формировании генетического профиля данной породы. Другие аллели, такие как TGLA53–164 и TGLA122–153, встречаются с частотой 5,4 и 5% соответственно, тогда как ряд аллелей (например, ETH3–121, TGLA126–151) имеют низкую частоту (0,5%) и не могут характеризовать генетическую структуру этой породы.

В выборке голштинского скота обнаружено 11 «приватных» аллелей, ассоциированных с девятью различными STR-локусами. Наиболее заметным является аллель ETH10–209, частота которого составляет 11,1%, что делает его наиболее распространенным среди «приватных» вариантов в этой группе. Также относительно высокую частоту имеет аллель TGLA227–99 (5,6%). Остальные «приватные» аллели, включая BM1818–270 (2%) и TGLA53–176 (2%), встречаются реже, преимущественно с частотой менее 2%.

Таким образом, полученные данные подтверждают наличие значимых различий в генетической структуре между красным белорусским и голштинским скотом. Наличие специфических приватных аллелей в каждой из пород, а также различия в их частотном распределении, отражают процессы генетической дивергенции, обусловленные, вероят-

но, селекционным отбором, происхождением популяций и историческим развитием пород. Эти особенности могут быть использованы для идентификации происхождения животных, мониторинга генетического разнообразия и разработки программ сохранения породных признаков.

Заключение

Анализ микросателлитных маркеров выявил генетические различия между выборками популяций красного белорусского и голштинского скота отечественной селекции. Установлено, что в выборке красного белорусского скота в среднем на locus (Na) пришлось $7,33 \pm 0,71$ аллеля (от 5 у локуса BM1824 до 13 у локуса TGLA53), число эффективных аллелей (N_e) составило $3,76 \pm 0,41$ (от 1,91 у локуса ETH3 до 6,43 у локуса TGLA53). Наблюдаемая гетерозиготность (H_o), которая является показателем полиморфности популяции и характеризует гетерозиготность исследуемой группы, составила $0,69 \pm 0,04$, ожидаемая же гетерозиготность (H_e), которая рассчитывается исходя из равновесия Харди – Вайнберга, составила $0,7 \pm 0,03$. Индекс фиксации (F), являющийся мерой дифференциации популяции, составил $0,001 \pm 0,039$, что указывает на низкую степень дифференциации внутри группы и генетическую однородность выборки по исследуемым нами локусам. Следует отметить, что F для локусов ETH10, TGLA126 и TGLA53 составил 0,128, 0,101 и 0,285 соответственно, что указывает на незначительный сдвиг в сторону

близкородственного скрещивания (дефицит гетерозигот).

Также нами были выявлены «приватные» аллели, которые могут характеризовать генетическую структуру красного белорусского и голштинского скота. Так, высокочастотными «приватными» аллелями для красного скота являются TGLA53–170 (26,2%) и TGLA227–85 (13,4%), в то же время для голштинского скота характерен аллель ETH10–209, уникальный для этой породы.

Таким образом, изучение генетической структуры пород молочного скота является важным элементом в современной племенной работе, которая направлена на оптимизацию генетического потенциала поголовья с учетом как селекционных целей, так и необходимости сохранения редких и специфических аллелей, характерных для красного белорусского и голштинского скота отечественной селекции.

Исследование выполнено при частичной поддержке грантов Белорусского республиканского фонда фундаментальных исследований Б23РНФ-060 и Б23КУБ-02.

Список использованных источников

1. Голштинская порода молочного скота отечественной селекции : материалы Междунар. науч.-практ. конф., посвященной 120-летию со дня рождения профессора О. А. Ивановой, Витебск, 3–5 нояб. 2021 г. / Витебская государственная академия ветеринарной медицины, Самаркандский институт ветеринарной медицины; Н. И. Гавриченко (гл. ред.) [и др.]. – Витебск : ВГАВМ, 2021. – 186 с.
2. Концепция селекционно-племенной работы по созданию белорусской красной молочной породы скота / А. И. Портной, И. П. Шейко, В. Н. Тимошенко [и др.] // Зоотехническая Наука Беларуси. – 2025. – Т. 60, № 1. – С. 3–11.
3. Гайко, А. А. Красный белорусский скот / А. А. Гайко, С. И. Тузов, М. П. Гринь. – Минск : Ураджай, 1968. – 142 с.
4. Государственная племенная книга красного белорусского скота. – Минск, 1967. – 223 с.
5. Генетические ресурсы сельскохозяйственных животных в России и сопредельных странах / сост. Эрнст Л. К. и др. – СПб. : Всерос. НИИ генетики и разведения с.-х. животных, 1994. – 469 с.
6. Коронец, И. Н. Сохранение генофонда красного белорусского скота как исчезающей локальной породы / И. Н. Коронец, Н. В. Климец, Ж. И. Шеметовец // Разведения і генетика тварин. – 2015. – № 49. – С. 232–235.
7. On the Breeds of Cattle — Historic and Current Classifications / M. Felius, P. A. Koolmees, B. Theunissen [et al.] // Diversity. – 2011. – Vol. 3, № 4. – P. 660–692.
8. Ernst, L. K. Farm Animal Genetic Resources in Russia and Neighboring Countries / L. K. Ernst, N. G. Dmitriev, I. A. Paronyan // St. Petersburg: All-Russian research institute of animal genetics and breeding. – 1994. – 469 p.
9. The genetic structure of cattle populations (*Bos taurus*) in northern Eurasia and the neighbouring Near Eastern regions: implications for breeding strategies and conservation / M. Li, I. Tapio, J. Vilkkilä [et al.] // Mol. Ecol. – 2007. – Vol. 16, № 18. – P. 3839–3853.
10. Maternal and paternal genealogy of Eurasian taurine cattle (*Bos taurus*) / J. Kantanen, C. J. Edwards, D. G. Bradley [et al.] // Heredity. – 2009. – Vol. 103, № 5. – P. 404–415.
11. Genome-wide SNP analysis clearly distinguished the Belarusian Red cattle from other European cattle breeds / N. A. Zinovieva, I. P. Sheiko, A. V. Dotsev [et al.] // Anim. Genet. – 2021. – Vol. 52, № 5. – P. 720–724.
12. Генетическая характеристика созданных типов скота бурой швицкой и сычевской пород с использованием полиморфизма микросателлитных локусов / Н. И. Стрекозов, Н. А. Зиновьева, П. В. Горелов [и др.] // С.-х. биология. – 2009. – Т. 44, № 2. – С. 10–15.
13. Сравнительная характеристика аллелофонда крупного рогатого скота трех родственных пород черно-пестрого корня, разводимых в Республике Казахстан / Т. Н. Карымсаков, Е. А. Гладырь, С. Д. Нурбаев [и др.] // Молочное и мясное скотоводство. – 2017. – № 3. – С. 11–14.
14. Характеристика аллелофонда сычевской породы крупного рогатого скота по ДНК микросателлитам / Д. Н. Кольцов, В. В. Волкова, Е. А. Гладырь [и др.] // Достижения науки и техники АПК. – 2012. – № 8. – С. 56–57.
15. Полиморфизм микросателлитных локусов крупного рогатого скота герефордской породы различных эколого-генетических генераций / Т. А. Седых, Е. А. Гладырь, И. Ю. Дол-

матова [и др.] // Аграрный вестник Северного Кавказа. – 2014. – № 3(15). – С. 121–128.

16. An overview of STRUCTURE: applications, parameter settings, and supporting software / L. Porras-Hurtado, Y. Ruiz, C. Santos [et al.] // Front. Genet. – 2013. – Vol. 4, № 96.

17. Evanno, G. Detecting the number of clusters of individuals using the software STRUCTURE : a simulation study / G. Evanno, S. Regnaut, J. Goudet // Mol. Ecol. – 2005. – Vol. 14, № 8. – P. 2611–2620.

M. E. Mikhailova¹, E. V. Snytkov¹, N. I. Tikhanovich¹, N. A. Kamysh¹, R. I. Sheiko¹, A. A. Sermyagin²,
N. A. Zinovyeva²

COMPARATIVE ANALYSIS OF THE GENETIC STRUCTURE OF THE RED BELARUSIAN AND HOLSTEIN CATTLE OF DOMESTIC BREEDING BASED ON MICROSATELLITE LOCI

¹State Scientific Institution

“Institute of Genetics and Cytology of the National Academy of Sciences of Belarus”

27 Akademicheskaya St., 220072 Minsk, the Republic of Belarus

e-mail: M.Mikhailova@igc.by

²Federal State Budgetary Scientific Institution

“Federal Research Center for Animal Husbandry named after Academy Member L. K. Ernst”

60 Dubrovitsy settlement, Podolsk city district, 142132 Moscow region, Russian Federation

Using 12 STR markers, the genetic polymorphism of the sampling of Belarusian aboriginal red cattle (101 heads) and Holstein cattle of Belarusian breeding (126 heads) was studied. Genetic differences between the samples of the Belarusian Red and Holstein cattle populations of domestic breeding were studied. It was found that in the Belarusian red cattle sampling, on average, there were 7.33 ± 0.71 alleles per locus (N_a); the number of effective alleles (N_e) accounted for 3.76 ± 0.41 . The observed heterozygosity (H_o), which is an indicator of population polymorphism and characterizes the heterozygosity of the studied group, was 0.69 ± 0.04 , while the expected heterozygosity (H_e) was 0.7 ± 0.03 . It was revealed that the fixation index (F) was 0.001 ± 0.039 , which indicates a low degree of differentiation and genetic homogeneity of the Belarusian red cattle population by the STR loci we have studied, as well as a slight shift towards inbreeding (heterozygote deficiency) since the fixation index F for the loci ETH10, TGLA126 and TGLA53 was 0.128, 0.101 and 0.285 respectively. For the Holstein cattle sampling, on average, there were 7.33 ± 0.51 alleles per locus; the number of effective alleles was 3.42 ± 0.321 . The observed heterozygosity was 0.68 ± 0.04 ; the expected heterozygosity was 0.67 ± 0.04 . It was established that, unlike red Belarusian cattle, Holstein cattle have a negative fixation index, which indicates an excess of heterozygotes and a low level of inbreeding. Private alleles that can characterize the genetic structure of Belarusian red and Holstein cattle were identified. The studied genetic markers are of great importance for animal breeding and conservation of genetic resources, and in the future, they may become the basis for the development of targeted breeding programs to improve desired traits by preventing the loss of unique alleles in dairy cattle breeds being bred in Belarus.

Keywords: microsatellites, Belarusian red cattle, Holstein breed of domestic breeding, cattle, population, genetic diversity.

Дата поступления в редакцию: 20 февраля 2025 г.

И. И. Саванович¹, И. А. Малеваная¹, Л. Э. Макарина-Кибак², М. Д. Амелянович²,
П. М. Морозик²

ГЕНЕТИЧЕСКИЕ МАРКЕРЫ В ПЕРСОНАЛИЗАЦИИ ПРИЕМА КОФЕИНА ДЛЯ ПОВЫШЕНИЯ СКОРОСТИ РЕАКЦИИ СПОРТСМЕНОВ

¹Государственное учреждение

«Республиканский научно-практический центр спорта»

Республика Беларусь, 220062, г. Минск, ул. Нарочанская, 8

²Государственное научное учреждение

«Институт генетики и цитологии Национальной академии наук Беларуси»

Республика Беларусь, 220072, г. Минск, ул. Академическая, 27

e-mail: P.Marozik@igc.by

Персонализация нутритивной поддержки является ключевым фактором высоких спортивных результатов. Кофеин остается одним из наиболее распространенных эргогенных средств, однако его эффект существенно варьирует. Целью исследования явилась оценка вклада вариантов генов *CYP1A1*, *CYP1A2*, *AHR* и *ADORA2A* в индивидуальную реакцию на кофеин у спортсменов. Показано, что варианты rs2472297, rs762551 и rs4410790 ассоциированы с ускорением реакции при физической нагрузке после приема кофеина, а вариант rs5751876, ассоциированный с повышенной тревожностью, продемонстрировал улучшение времени реакции при физической нагрузке после приема кофеина, что, вероятно, отражает особенности нервной регуляции при адаптации к постоянным физическим нагрузкам. Генетическая модель скорости метаболизма кофеина показала высокую прогностическую точность (AUC = 0,92). Полученные результаты подтверждают целесообразность применения генетического тестирования при разработке персонализированных стратегий использования кофеина для оптимизации спортивной работоспособности.

Ключевые слова: кофеин, спортивная генетика, полиморфизм генов, *CYP1A2*, *ADORA2A*, *AHR*, время реакции.

Для цитирования: Генетические маркеры в персонализации приема кофеина для повышения скорости реакции спортсменов / И. И. Саванович, И. А. Малеваная, Л. Э. Макарина-Кибак [и др.] // Молекулярная и прикладная генетика: сб. науч. тр. / Ин-т генетики и цитологии НАН Беларуси; редкол.: А.В. Кильчевский (гл. ред.) [и др.]. – Минск, 2025. – Т. 39. – С. 65–73.

Введение

Кофеин является одним из наиболее изученных и широко используемых эргогенных средств в спорте: около 75% профессиональных спортсменов используют кофеин перед или во время соревнований, стремясь повысить результативность выступлений [1]. Его положительное влияние на физическую активность и выносливость, концентрацию и скорость реакции, уровень бодрости и устойчивость к утомлению неоднократно было подтверждено в многочисленных исследованиях. Оптимальный эффект кофеина обычно наблюдается при дозировках в диа-

пазоне 2–6 мг/кг массы тела при приеме за 30–60 мин до тренировки или соревнований, что позволяет повысить выносливость в среднем на 2–4% [2]. Эти цифры могут показаться скромными, однако в условиях современного спорта, среди всего прочего, они могут определять исходы соревнований.

Физиологическое действие кофеина обусловлено комплексом центральных и периферических механизмов. На уровне центральной нервной системы он выступает конкурентным антагонистом аденозиновых рецепторов A1 и A2a, что приводит к повышению нейрональной возбудимости, улучшению концентрации

внимания и снижению субъективного восприятия усталости. Что касается периферического действия, то кофеин способствует мобилизации свободных жирных кислот, влияет на кальциевый обмен в мышечных клетках и повышает сократительную способность мышечных волокон [3]. После поступления в организм кофеин быстро всасывается и метаболизируется преимущественно в печени с участием фермента CYP1A2, на долю которого приходится до 95% его биотрансформации. В результате образуются три активных метаболита — параксантин, теобромин и теofilлин, каждый из которых вносит вклад в суммарный физиологический эффект.

Индивидуальный ответ на прием кофеина существенно различается у разных индивидов: у одних он вызывает выраженный прирост производительности, у других — отсутствие эффекта или даже снижение результативности. Эти различия обусловлены как фармакокинетикой (скоростью метаболизма), так и фармакодинамикой (чувствительностью рецепторов) [4]. Именно поэтому в последние годы активно развивается направление спортивной нутригеномики, изучающей влияние полиморфизма генов на реакцию атлетов на различные нутриенты. В этом аспекте наибольший интерес представляют полиморфные варианты генов *CYP1A1*, *CYP1A2*, гена-регулятора *AHR* и гена *ADORA2A*, ассоциированного с привыканием к кофеину и выраженностью физиологического ответа на его употребление [5].

Ген *CYP1A2* кодирует основной фермент, катализирующий первичное деметилирование кофеина. Наиболее изученный полиморфный вариант rs762551 связывают с различиями в активности фермента. Аллель А ассоциирован с более высокой транскрипционной активностью гена. Это означает, что производится больше функционального фермента CYP1A2 в печени. Следовательно, у носителей этого аллеля кофеин быстрее инактивируется (они являются «быстрыми» метаболиторами). Аллель С, в свою очередь, связан с пониженной транскрипционной активностью или сниженной стабильностью мРНК, что приводит к меньшему количеству функционального фермента или его пониженной активности. В результате скорость метаболизма кофеина замедляется (носители аллеля С являются

«медленными» метаболиторами) [6].

Ген *AHR* (aryl hydrocarbon receptor) кодирует арилгидрокарбоновый рецептор — белок, который функционирует как лиганд-активируемый транскрипционный фактор, контролирующий экспрессию ряда ферментов биотрансформации, включая CYP1A2. Активированный белок AHR связывается с регуляторными участками гена *CYP1A2*, стимулируя его транскрипцию и, как следствие, увеличивая производство фермента CYP1A2 в печени. Вариант rs4410790 находится в регуляторной области *AHR*, может модифицировать структуру самого гена и менять уровень экспрессии CYP1A2 и, следовательно, скорость метаболизма кофеина [4].

Полиморфный вариант rs2472297 локализован в межгенном пространстве, а именно между генами *CYP1A1* и *CYP1A2* (15q24.1). Считается, что rs2472297 функционирует как регуляторный элемент (энхансер или промотор) гена *CYP1A2* — основного фермента, метаболизирующего кофеин в печени. Аллель Т связан с повышенной экспрессией активного фермента CYP1A2, что способствует ускоренному выведению кофеина из организма.

Особое внимание уделяется гену *ADORA2A*, кодирующему аденозиновый рецептор A2a, который играет ключевую роль в реализации центральных эффектов кофеина. Полиморфный вариант rs5751876 (1976C > T) связан с индивидуальными различиями в чувствительности к эффектам кофеина [4]. Для спортсменов это особенно важно, поскольку чрезмерная тревожность или нарушение концентрации могут нивелировать потенциальный положительный эффект от приема кофеина. Хотя точный функциональный механизм не всегда однозначен, ключевые гипотезы связывают этот вариант с изменением доступности (стабильность, плотность экспрессии на поверхности нейронов) и аффинности/активности (незначительными структурными изменениями) рецепторов A2a [7].

Несмотря на очевидную биологическую значимость этих маркеров, экспериментальные исследования демонстрируют выраженную неоднозначность результатов. Например, в отношении rs762551 одни исследования [8, 9] показывают, что эргогенный эффект кофеина в дозах 4–6 мг/кг более выражен

у «быстрых» метаболизаторов (носители генотипа A/A), в то время как другие работы [10] не выявляют связи между генотипом и показателями мощности или выносливости. Подобная противоречивость характерна и для исследований ассоциации гена *ADORA2A*. Систематические обзоры указывают, что эта гетерогенность обусловлена множеством неучтенных факторов, включая различия в дозировках, времени приема, уровне подготовки спортсменов, а также недостаточную статистическую мощность выборок [11]. Кроме того, зачастую игнорируется полигенный характер реакции (т. е. взаимодействие *CYP1A2*, *AHR* и *ADORA2A*) и такие важные составляющие, как привычное потребление кофеина и курение.

В связи с этим сохраняющаяся актуальность использования кофеина спортсменами, в сочетании с противоречивостью имеющихся литературных данных, требует более глубокого и комплексного подхода при включении его в схему фармакологического обеспечения спортивной подготовки. Целью настоящей работы является оценка ассоциации полиморфных вариантов генов *CYP1A1* (rs2472297), *CYP1A2* (rs762551), *AHR* (rs4410790) и *ADORA2A* (rs5751876) с объективными показателями психофизиологической реакции и производительности у спортсменов циклических видов спорта.

В рамках поставленной цели ключевой проблемой исследования являлось определение того, как исследуемые генетические локусы влияют на спортивные показатели. Было необходимо установить распределение генотипов и оценить, как каждый из них, а также их аллельные комбинации, влияют на показатели времени реакции после физической нагрузки с приемом кофеина. Особое внимание уделялось выявлению таких полигенных комбинаций, которые могут либо усиливать, либо, наоборот, ослаблять ожидаемый эргогенный эффект кофеина, что позволит перейти от оценки одиночных маркеров к пониманию комплексной генетической детерминации спортивного ответа.

Материалы и методы

Когорта исследования

В кросс-секционное исследование с функци-

ональным тестированием и генотипированием были включены спортсмены циклических видов, мужского и женского пола в возрасте 18 лет и старше. Общая когорта составила 104 человека, соответствующих критериям включения и подписавших информированное согласие участника и согласие на обработку персональных данных.

Исследование было рассмотрено и утверждено на заседании комитета по биоэтике Государственного научного учреждения «Институт генетики и цитологии Национальной академии наук Беларуси» (протокол заседания от 08.06.2021 № 2).

Молекулярно-генетическое исследование

В качестве биологического материала использовали ДНК, выделенную из лейкоцитов крови. Выделение ДНК осуществлялось на многофункциональной автоматической станции Eppendorf epMotion 5075 с использованием наборов реагентов для выделения нуклеиновых кислот на магнитных частицах «Арт НК Магнит» (ООО «АртБиоТех», Беларусь) в соответствии с инструкцией производителя. Концентрация и чистота выделенной ДНК оценивалась на флуориметре Qubit, при этом минимально допустимой для анализа считалась концентрация 5 нг/мкл. Для сохранения образцов использовались рекомендованные режимы: при температуре +2...+8 °C до 35 суток, при температуре –16 °C и ниже — для более длительного хранения.

Генотипирование выполняли методом полимеразной цепной реакции в режиме реального времени с использованием специфических праймеров и зондов, меченных флуорофорами. В исследование были включены полиморфные варианты генов, играющих ключевую роль в метаболизме кофеина и чувствительности к нему: *CYP1A2* (rs762551), *CYP1A1* (rs2472297), *AHR* (rs4410790) и *ADORA2A* (rs5751876). Реакции ставились в объеме 10 мкл на системе Bio-Rad CFX96. Для каждого из исследуемых локусов были подобраны оптимальные температурно-временные условия амплификации, а также определены пороговые значения флуоресценции каналов FAM и ROX, по которым проводилась интерпретация результатов. Для обеспечения достоверности анализа использовались отрицательные и положительные контрольные образцы.

Функциональное тестирование

Функциональное тестирование включало оценку среднего времени реакции спортсменов в трех состояниях: в покое, после физической нагрузки без приема кофеина и после нагрузки с приемом кофеина в стандартной дозе. Измерения проводились с учетом фармакокинетики вещества (пик биодоступности через 30–45 мин, период полувыведения 3–5 ч). Время реакции регистрировалось с использованием специализированного зрительно-моторного анализатора и компьютерной программы, фиксирующей интервал от подачи зрительного сигнала до моторного ответа (нажатия кнопки). Измерялись три типа реакций: простая зрительно-моторная, реакция различения и реакция на движущийся объект (РДО). Показатели выражались в миллисекундах. Для РДО диапазон от –20 до +50 мс рассматривался как отражающий сбалансированность процессов возбуждения и торможения. Изменение времени реакции (Δt_p) рассчитывалось для оценки эффекта физической нагрузки и кофеина.

Статистическая обработка

Вся статистическая обработка результатов и построение графиков осуществлялись с использованием языка программирования R (<http://r-project.org>). Для оценки различий между количественными и качественными показателями индивидов использовали линейную и логистическую регрессии соответственно.

Кроме того, проводился анализ гаплотипов с оценкой их частоты и связи с изменением времени реакции.

ROC-анализ

Для ROC-анализа с расчетом AUC, чувствительности и специфичности в качестве бинарного исхода рассматривалась вероятность положительного ответа на кофеин. Различия между группами считались статистически значимыми при $p < 0,05$.

Результаты и их обсуждение

В исследовании были проанализированы полиморфные варианты генов, участвующих в метаболизме кофеина: *CYP1A1* (rs2472297), *CYP1A2* (rs762551), *AHR* (rs4410790) и *ADORA2A* (rs5751876). Анализ среднего значения времени реакции на движущийся объект (t_p), измеренного в состоянии покоя, а также после физической нагрузки без употребления кофеина, не выявил статистически достоверной ассоциации ни для одного из исследованных локусов (табл. 1, представлены значения для указанного генотипа относительно референсного генотипа). Это указывает на то, что данные генетические маркеры не оказывают значимого влияния на базовую скорость психомоторной реакции или ее ухудшение, вызванное только физической усталостью.

Однако после приема кофеина анализ абсолютных показателей среднего значения t_p выявил статистически значимую ассоциацию для

Таблица 1

Показатели среднего времени реакции на движущийся объект (t_p)

Ген, вариант	Генотип	после нагрузки без употребления кофеина		после нагрузки с употреблением кофеина	
		t_p , мс (95% ДИ)	p	t_p , мс (95% ДИ)	p
<i>CYP1A1</i> rs2472297	A/A	–4,5 (–21,2...12,2)	0,86	1,6 (–15,3...18,5)	0,78
<i>CYP1A2</i> rs762551	T/T	11,7 (–15,1...38,5)	0,7	–5,6 (–34,6...23,5)	0,37
<i>AHR</i> rs4410790	G/G	5,9 (–14,0...25,8)	0,38	23,6 (2,4...44,8) 21,8 (1,9...41,6)*	0,11 0,038
<i>ADORA2A</i> rs5751876	T/T	–13,3 (–30,9...4,3)	0,26	–25,3 (–47,4...–3,2)	0,022

Примечание. * — рецессивная модель наследования рассчитана для генотипов T/T + G/T vs. G/G

двух полиморфных вариантов (табл. 1). Для варианта *AHR* rs4410790 установлена ассоциация для рецессивной модели наследования: носители генотипа G/G имели достоверное увеличение t_p по сравнению с носителями генотипов T/T + G/T (разница 21,8 мс; 95% ДИ 1,9–41,7; $p = 0,038$). Поскольку продукт гена *AHR* регулирует количество фермента CYP1A2, данное наблюдение может указывать на то, что генотип G/G ассоциирован с менее эффективным метаболизмом кофеина, приводя к более длительному присутствию активных метаболитов в крови и, возможно, их влиянию на психомоторные функции.

Для полиморфного варианта rs5751876 гена *ADORA2A* выявлена статистически значимая ассоциация для общей модели наследования. Среди носителей генотипа T/T было выявлено значимое сокращение t_p (–25,3 мс; 95% ДИ –47,4...–3,2; $p = 0,022$). Этот генотип ассоциирован с высокой чувствительностью к кофеину, что в данном случае привело к наиболее выраженному стимулирующему эффекту, проявляющемуся в ускорении времени реакции. Такое направление эффекта можно рассматривать как физиологически правдоподобное, но требующее дальнейшего подтверждения. Предположительно, данное расхождение можно объяснить тем, что регулярная физическая тренировка сопровождается снижением тонуса аденозинергической системы и более эффективной регуляцией нейротрансмиттеров, в связи с чем кофеин может не вызывать типичных проявлений тревожно-

сти, а напротив обеспечивать оптимальный уровень нейрональной активации, улучшая когнитивную скорость реакции на стимулы. Кроме того, известно, что проявления «гиперчувствительности» у носителей T/T-генотипов чаще возникают при высоких дозах кофеина (>5–6 мг/кг массы тела), тогда как в данном исследовании использовалась умеренная доза (около 3 мг/кг).

Для полиморфных вариантов генов цитохрома P450 не было выявлено статистически значимой ассоциации, однако носители генотипа T/T варианта *CYP1A2* rs762551 имеют тенденцию к улучшению t_p после приема кофеина, что соответствует гипотезе о более быстром метаболизме и снижении риска неблагоприятных эффектов у этой группы исследуемых.

Для точной оценки вклада генетических факторов в эффективность метаболизма кофеина был проведен анализ разницы средних значений времени реакции (Δt_p) между состоянием после нагрузки с кофеином и исходным состоянием (без нагрузки, табл. 2). Данный показатель отражает чистый эффект воздействия кофеина, исключая влияние базовой психомоторной скорости.

Для полиморфного варианта *CYP1A2* rs762551 обнаружена статистически значимая ассоциация генотипа T/T с максимальным улучшением Δt_p ($\beta = -28,7$; 95% ДИ –55,8...–1,7, $p = 0,03$). Этот результат полностью соответствует классической модели: быстрый метаболизм кофеина позволяет быстро достичь пика его действия и мини-

Таблица 2

Показатели разницы средних значений времени реакции (Δt_p) после нагрузки с кофеином и исходным состоянием

Ген, вариант	Δt_p , нагрузка без кофеина, мс	p	Δt_p , нагрузка с кофеином, мс	p
<i>CYP1A1</i> rs2472297	A/A, $\beta = 4,9$ (–11,9...21,8)	0,69	A/A, $\beta = -12,7$ (–24,3...–1,1)	0,03
<i>CYP1A2</i> rs762551	T/T, $\beta = -3,8$ (–17,5...9,8)	0,21	T/T, $\beta = -28,7$ (–55,8...–1,7)	0,03
<i>AHR</i> rs4410790	G/G, $\beta = -3,5$ (–23,1...16,0)	0,89	G/G, $\beta = 17,7$ (–0,2...35,6)	0,089
<i>ADORA2A</i> rs5751876	T/T, $\beta = 0,5$ (–17,9...19,0)	0,85	T/T, $\beta = -4,5$ (–24,8...15,8)	0,81

мизировать риск накопления, который может проявиться в форме тревожности или тремора. Аналогичные эффекты были получены в других исследованиях — например, у быстрых метаболизаторов (генотип Т/Т) кофеин улучшал результаты на 10–40 км велоэргометрии, тогда как у медленных метаболизаторов (С/С) он мог даже ухудшать показатели [9].

Менее выраженная зависимость установлена для полиморфного варианта *CYP1A1* rs2472297 — среди носителей генотипа А/А,

Δt_p была также статистически значимо ниже по сравнению с носителями референсного генотипа G/G ($\beta = -12,7$; 95% ДИ $-24,3 \dots -1,1$, $p = 0,03$).

Анализ чистого эргогенного эффекта кофеина (разница показателей после нагрузки с кофеином относительно исходного состояния) позволил установить четкую ген-дозовую зависимость для локусов *CYP1A1* rs2472297 (рис. 1А) и *CYP1A2* rs762551 (рис. 1Б).

Этот эффект проявляется в том, что вели-

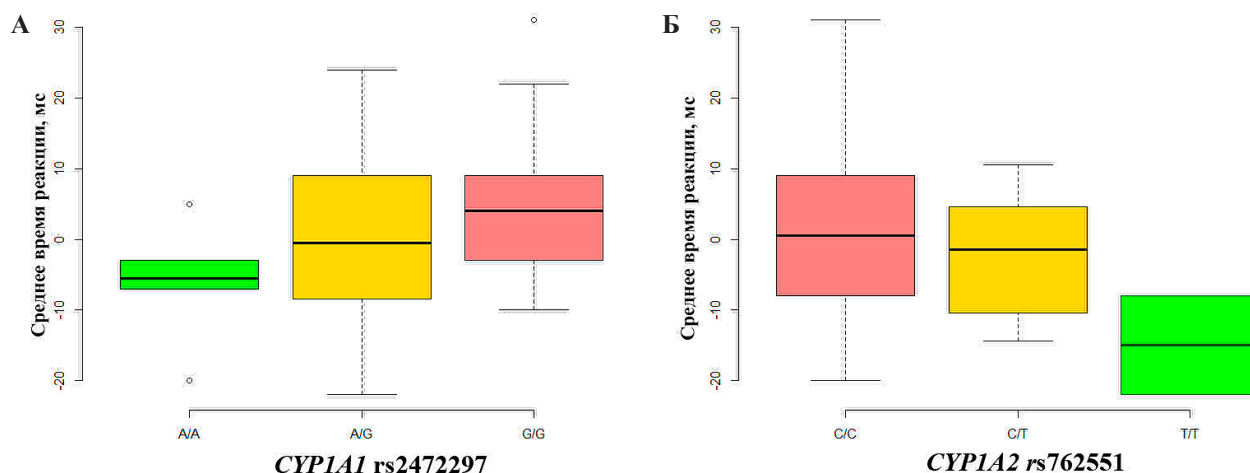


Рис. 1. Распределение Δt_p после нагрузки и употребления кофеина в зависимости от генотипа по вариантам *CYP1A1* rs2472297 (А) и *CYP1A2* rs762551 (Б)

чина Δt_p находится в прямой зависимости от количества «быстрых» аллелей метаболизма кофеина, присутствующих в генотипе атлета. Носители гомозиготных генотипов, ассоциированных с быстрым метаболизмом (rs2472297 А/А и rs762551 Т/Т), демонстрировали максимальное значение ускорения времени реакции, носители альтернативных гомозиготных генотипов G/G и C/C показывали минимальный эффект, а гетерозиготные генотипы для обоих локусов обеспечивали промежуточный показатель ответа. Наличие классической модели наследуемости служит дополнительным сильным функциональным доказательством того, что полиморфные варианты оказывают существенное влияние на активность фермента *CYP1A2* и, следовательно, на скорость выведения кофеина, что в итоге определяет интенсивность его психо-стимулирующего действия.

Отдельный интерес представляют результа-

ты анализа аллельных комбинаций (гаплотипов), что позволяет проводить комплексную интерпретацию сразу по четырем генетическим маркерам (рис. 2).

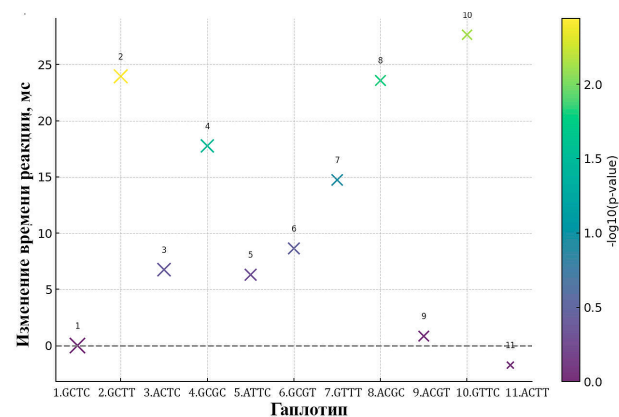


Рис. 2. Влияние аллельных комбинаций генов *CYP1A1*-*CYP1A2*-*AHR*-*ADORA2A* на изменение времени реакции после приема кофеина.

Всего выявлено 11 сочетаний аллелей с частотой встречаемости выше 3%, часть из которых достоверно влияла на показатели времени реакции. Так, гаплотип G-C-T-T был связан с увеличением времени реакции на 23,9 мс ($p = 0,0036$), а гаплотип G-C-G-C — на 17,8 мс ($p = 0,031$). Выраженные статистически значимые ассоциации наблюдаются также для комбинаций G-T-T-C ($p = 0,0076$) и A-C-G-C ($p = 0,017$), демонстрирующих увеличение времени реакции в среднем на 23,6–27,7 мс. При этом для полиморфного варианта *AHR* rs4410790, индивидуально демонстрировавшего лишь пограничную значимость, при учете гаплотипов его влияние усиливалось. Эти результаты показывают, что эффект кофеина определяется не одним локусом, а комбинацией нескольких, регулирующих как фармакокинетику, так и фармакодинамику, что в целом согласуется с представлениями о полигенной природе фармакогенетического ответа [12] и позволяет сформировать индивидуальный профиль ответа спортсмена.

Анализ вклада отдельных вариантов генов, связанных с метаболизмом кофеина, в фенотипический ответ показал, что различия между генотипами обладают очень высокой величиной эффекта для трех изученных маркеров, что свидетельствует об их потенциальной клинической или физиологической значимости. Так, для гена *CYP1A1* величина эффекта Cohen's d составила $\approx 1,14$ (высокая), для гена *CYP1A2* $\approx 1,63$ (очень высокая), для гена *ADORA2A* $\approx 1,11$ (высокая).

Для оценки диагностической и прогностической значимости совокупности выявленных генетических маркеров *AHR* rs4410790, *CYP1A1* rs2472297, *CYP1A2* rs762551, ассоциированных со скоростью метаболизма кофеина, был проведен ROC-анализ (рис. 3). Полиморфный вариант *ADORA2A* rs5751876 был исключен, поскольку он ассоциирован преимущественно не со скоростью метаболизма, а с чувствительностью к кофеину.

Модель оценки скорости метаболизма кофеина обладает высокой прогностической ценностью ($AUC = 0,917$; 95% ДИ 0,87–0,96), а также характеризуется высоким уровнем чувствительности (98,4%) при средней специфичности (75,7%) и высокой общей точности классификации (90,1%). Средняя

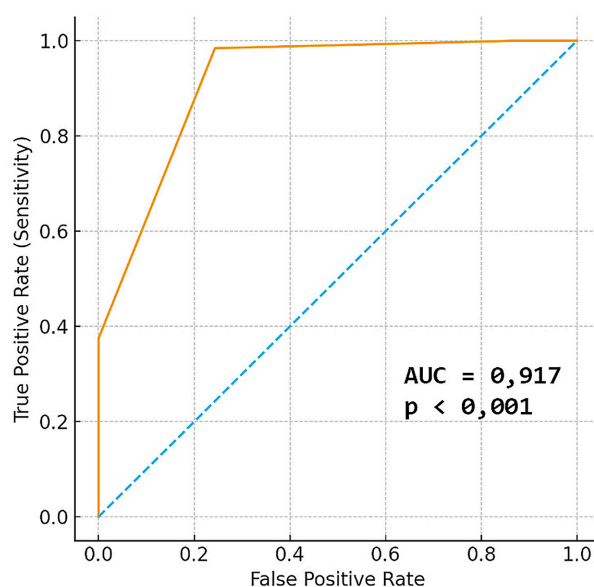


Рис. 3. ROC-кривая, характеризующая предсказательную способность скорости метаболизма кофеина по маркерам *AHR* rs4410790, *CYP1A1* rs2472297 и *CYP1A2* rs762551

специфичность свидетельствует о том, что при выбранном оптимальном пороге модель практически безошибочно определяет лиц с быстрым метаболизмом (чувствительность $\approx 98\%$), однако часть индивидов со средним и медленным типами метаболизма ($\sim 24\%$) может быть ошибочно классифицирована как «быстрые». Включение дополнительных генетических маркеров, вероятно, позволит еще больше повысить ее прогностическую ценность и специфичность.

Заключение

Полученные результаты однозначно подтверждают, что генетический профиль спортсмена оказывает значимое влияние на реакцию организма при приеме кофеина, которая носит полигенный характер. Полиморфные варианты генов, связанные с метаболизмом кофеина (*CYP1A1*, *CYP1A2*, *AHR*) и чувствительностью к нему (*ADORA2A*), не влияют на базовую психомоторную реакцию, но определяют индивидуальный психофизиологический ответ на прием кофеина. Исследование вносит вклад в понимание комплексного взаимодействия генов метаболизма и чувствительности к кофеину, что критически важно для разработки научно обоснованных и персонализи-

рованных рекомендаций по использованию кофеина в спорте, обеспечивая максимальный эргогенный эффект при одновременной минимизации рисков побочных эффектов. Включение генетического тестирования при составлении плана фармакологического обеспечения спортивной подготовки позволит минимизировать риски побочных эффектов и повысить эффективность подготовки.

Список использованных источников

1. Del Coso, J. Prevalence of caffeine use in elite athletes following its removal from the world anti-doping agency list of banned substances / J. Del Coso, G. Muñoz, J. Muñoz-Guerra // *Appl. Physiol. Nutr. Metab.* – 2011. – Vol. 36. – P. 555–561.
2. International society of sports nutrition position stand: caffeine and exercise performance / N. S. Guest, T. A. VanDusseldorp, M. T. Nelson [et al.] // *J. Int. Soc. Sports Nutr.* – 2021. – Vol. 18, iss. 1. – P. 1.
3. The long way to establish the ergogenic effect of caffeine on strength performance: an overview review / V. Giráldez-Costas, J. Del Coso, A. Mañas, J. José Salinero // *Nutrients* – 2023. – Vol. 15, iss. 5. – P. 1178.
4. Genome-wide meta-analysis identifies regions on 7p21 (*AHR*) and 15q24 (*CYP1A2*) as determinants of habitual caffeine consumption / M.C. Cornelis, K. L. Monda, K. Yu [et al.] // *PLoS Genet.* – 2011. – Vol. 7, iss. 4. – P. e1002033.
5. PharmGKB summary: caffeine pathway / C. F. Thorn, E. Aklillu, E. M. McDonagh [et al.] // *Pharmacogenet. Genomics* – 2012. – Vol. 22, iss. 5. – P. 389–395.
6. Effects of *CYP1A2* and *ADORA2A* genotypes on the ergogenic response to caffeine in professional handball players / A. Muñoz, Á. López-Samanes, M. Aguilar-Navarro [et al.] // *Genes* – 2020. – Vol. 11, iss. 8. – P. 933.
7. Effect of *ADORA2A* gene polymorphism and acute caffeine supplementation on hormonal response to resistance exercise: a double-blind, crossover, placebo-controlled study / M. R. Rahimi, E. A. Semenova, G. John [et al.] // *Nutrients* – 2024. – Vol. 16, iss. 12. – P. 1803.
8. The influence of a *CYP1A2* polymorphism on the ergogenic effects of caffeine / C. J. Womack, M. J. Saunders, M. K. Bechtel [et al.] // *J. Int. Soc. Sports Nutr.* – 2012. – Vol. 9, iss. 1. – P. 7.
9. Caffeine, *CYP1A2* genotype, and endurance performance in athletes / N. Guest, P. Corey, J. Vescovi, A. El-Sohemy // *Med. Sci. Sports Exerc.* – 2018. – Vol. 50, iss. 8. – P. 1570–1578.
10. *CYP1A2* genotype variations do not modify the benefits and drawbacks of caffeine during exercise: a pilot study / J. J. Salinero, B. Lara, D. Ruiz-Vicente [et al.] // *Nutrients*. – 2017. – Vol. 9, iss. 3. – P. 269.
11. Does ergogenic effect of caffeine supplementation depend on *CYP1A2* genotypes? A systematic review with meta-analysis / J. Wang, L. Dewi, Y. Peng [et al.] // *J. Sport Health Sci.* – 2024. – Vol. 13, iss. 4. – P. 499–508.
12. Pharmacogenomics polygenic risk score for drug response prediction using PRS-PGx methods / S. Zhai, H. Zhang, D. V. Mehrotra, J. Shen // *Nat. Commun.* – 2022. – Vol. 13. – P. 5278.

I. I. Savanovich¹, I. A. Malyovanaya¹, L. E. Makaryna-Kibak², M. D. Ameliyanovich², P. M. Marozik²

GENETIC MARKERS IN PERSONALIZING CAFFEINE INTAKE TO BOOST REACTION TIME IN ATHLETES

¹State Institution

“Republican Scientific and Practical Sport Center”

8 Narochanskaya St., 220062 Minsk, the Republic of Belarus

²State Scientific Institution

“Institute of Genetics and Cytology of the National Academy of Sciences of Belarus”

27 Akademicheskaya St., 220072 Minsk, the Republic of Belarus

e-mail: P.Marozik@igc.by

Personalization of nutritional support is a key factor in achieving high athletic performance. Caffeine remains one of the most widely used ergogenic aids; however, its effect varies substantially among individuals. The aim of this study was to assess the contribution of *CYP1A1*, *CYP1A2*, *AHR*, and *ADORA2A* gene variants to individual responses to caffeine in athletes. It was shown that variants rs2472297, rs762551, and rs4410790 are associated with faster reaction times following physical exertion with caffeine, and the rs5751876 variant, associated with increased anxiety, demonstrated improved reaction time during physical exertion after caffeine intake, likely reflecting specific neural regulation during adaptation to regular physical activity. The genetic model of caffeine metabolism rate showed high predictive accuracy (AUC = 0.92). The results obtained support the feasibility of using genetic testing in developing personalized caffeine strategies to optimize athletic performance.

Keywords: caffeine, sports genetics, gene polymorphism, *CYP1A2*, *ADORA2A*, *AHR*, reaction time.

Дата поступления в редакцию: 10 сентября 2025 г.

Н. В. Никитченко¹, И. Ю. Бакутенко¹, А. М. Чичко², Е. В. Сечко², Н. И. Рябоконт¹

МИКРОСАТЕЛЛИТНЫЕ ЛОКУСЫ ГЕНОВ *PSMA6*, *FAM177A1* И *PRORP* КАК ГЕНЕТИЧЕСКИЕ МАРКЕРЫ ПРЕДРАСПОЛОЖЕННОСТИ К РАЗВИТИЮ ВОСПАЛИТЕЛЬНОГО ПРОЦЕССА РАЗЛИЧНОЙ ЭТИОЛОГИИ В СУСТАВАХ У ДЕТЕЙ

¹Государственное научное учреждение
«Институт генетики и цитологии Национальной академии наук Беларуси»
Республика Беларусь, 220072, г. Минск, ул. Академическая, 27

e-mail: n.ryabokon@igc.by

²Учреждение образования
«Белорусский государственный медицинский университет»
Республика Беларусь, 220116, г. Минск, пр. Дзержинского, 83

У педиатрических пациентов с воспалительным процессом в суставах различной этиологии, а именно с диагнозом ювенильный идиопатический артрит (ЮИА) и с суставным синдромом без аутоиммунной этиологии, а также в группе клинического контроля (без аутоиммунных и хронических воспалительных заболеваний) проведено генотипирование полиморфизма хромосомной области 14q13.2–14q13 по пяти микросателлитным локусам генов *PSMA6*, *FAM177A1* и *PRORP*: HSMS006 (TG)_n, HSMS602 (CAA)_n(A)_m, HSMS701 (AC)₅AT(AC)_n, HSMS702 (TG)_n и HSMS801 (CA)_n. При анализе частот полиморфных микросателлитных повторов у пациентов с воспалением суставов различной этиологии по сравнению с контролем было выделено пять аллелей генов *FAM177A1* и *PRORP*: 170 п. н. HSMS602, 138 п. н. HSMS701, 124 п. н. HSMS702, 137 и 151 п. н. HSMS801, — которые ассоциированы с риском воспаления суставов в детском возрасте. Сравнительный анализ групп с ЮИА и суставным синдромом не аутоиммунной этиологии позволил обнаружить еще один аллель гена *PSMA6* (189 п. н. HSMS006), который связан с риском ЮИА и, таким образом, возможно, с возникновением аутоиммунной компоненты при воспалении суставов у детей.

Ключевые слова: микросателлитные локусы, гены *PSMA6*, *FAM177A1* и *PRORP*, суставной синдром, ювенильный идиопатический артрит, воспалительный процесс, фрагментный анализ.

Для цитирования: Микросателлитные локусы генов *PSMA6*, *FAM177A1* и *PRORP* как генетические маркеры предрасположенности к развитию воспалительного процесса различной этиологии в суставах у детей / Н. В. Никитченко, И. Ю. Бакутенко, А. М. Чичко [и др.] // Молекулярная и прикладная генетика: сб. науч. тр. / Ин-т генетики и цитологии НАН Беларуси; редкол.: А.В. Кильчевский (гл. ред.) [и др.]. — Минск, 2025. — Т. 39. — С. 74–83.

Введение

Микросателлитные повторы составляют около 3% человеческого генома и представляют собой участки повторяющихся олигонуклеотидных мотивов или tandemных повторов. Микросателлиты состоят из 5–30 повторов, каждый из которых содержит 1–13 нуклеотидов, из-за чего их еще называют короткими tandemными повторами (Short Tandem Repeat — STR) [1]. Они распространены по всему геному и расположены чаще всего в нетранскрибируемых и нетранслируемых

участках, в то же время являются важными компонентами последовательностей, регулирующих экспрессию генов: энхансеров, сайленсеров и инсуляторов. Микросателлиты также вовлечены в ряд других процессов, протекающих в клетке: в альтернативный сплайсинг, рекомбинацию и репарацию ДНК. Таким образом, очевидно, что микросателлиты причастны к различным аспектам функционирования генома. Изучение этих процессов важно для понимания механизмов регуляции генов и их роли в развитии различных патологий.

Известно, что микросателлиты вовлечены в патогенез около 50 заболеваний человека, большинство из которых являются тяжелыми дегенеративными [2].

Основу полиморфизма микросателлитов составляют различия в числе повторяющихся единиц. Показано, что большинство микросателлитов, ассоциированных с патологиями у человека (саркоидоз, болезнь Альцгеймера, диабет 1-го и 2-го типов и т. д.), располагаются в генах, кодирующих ферменты; при этом во многих случаях прослеживается влияние длины микросателлита на функции гена [3]. В некоторых случаях патологическое увеличение количества нуклеотидных повторов в нестабильных участках генома может приводить к развитию болезней экспансии нуклеотидных повторов. Эти заболевания обусловлены нарушением функции протеинов, и наиболее известными из них являются спинальная и бульбарная мышечная атрофия (болезнь Кеннеди), а также синдром ломкой X-хромосомы (синдром Мартина – Белл). В целом, в настоящее время обнаружено более 30 патологий, характеризующихся экспансией нуклеотидных повторов [2, 4].

Т. Sjakste с соавторами установлена ассоциация микросателлитных повторов в отдельных локусах слабоизученной хромосомной области 14q13.2–14q13 с гипертиреозом (болезнь Базедова) [5] и ювенильным идиопатическим артритом (ЮИА) [6] в латвийской популяции, а также диабетом 2-го типа в латвийской и финской популяциях [7], что делает исследования в данном направлении многообещающими. В указанной хромосомной области авторами изучались пять микросателлитных локусов трех генов: протеасомного гена *PSMA6* (Proteasome Subunit Alpha Type 6) и прилегающих к нему генов *FAM177A1* (Family With Sequence Similarity 177 Member A1, также известный как *C14orf24*) и *PRORP* (Protein Only RNase P Catalytic Subunit, известный как *KIAA0391*).

Ген *FAM177A1* кодирует белок, участвующий в ингибировании сигнального пути интерлейкина-1 β (IL-1 β), мощного провоспалительного цитокина [8].

Ген *PRORP* кодирует каталитическую субъединицу митохондриальной РНКазы Р, необходимую для процессинга митохондриальных

тРНК и, следовательно, для нормальной функции митохондрий [9, 10]. Митохондриальный биогенез зависит как от ядерного, так и от митохондриального генома, и дисбаланс в их экспрессии может привести к врожденным ошибкам метаболизма, воспалению и старению [11]. Hong с соавторами провели метаанализ ассоциаций по всему геному и выявили более 241 локуса восприимчивости к воспалительным заболеваниям кишечника, а также идентифицировали новые ассоциации, в частности, rs57275892 в новом локусе *PRORP/PSMA6* [12].

Ген *PSMA6* отвечает за синтез субъединицы альфа 6 протеасомы 20S. Протеасома — это многобелковый комплекс, предназначенный для разрушения дефектных или короткоживущих белков в клетке и, таким образом, для сохранения клеточного гомеостаза. Это крупномолекулярная мультикаталитическая протеаза, которая способна гидролизовать белки до более мелких пептидов как по убиквитинзависимому, так и по убиквитин-независимому механизму [13]. В работе Wang J. с соавторами [14] описывается повышенная концентрация протеасом при отдельных системных аутоиммунных заболеваниях, таких как системная красная волчанка и ревматоидный артрит, связывая это с усилением симптомов. Авторы объясняют это наблюдение вовлеченностью протеасом в регуляцию активации NF- κ B — антиапоптотического и провоспалительного регулятора экспрессии цитокинов, которые играют важные роли в развитии воспалительных и аутоиммунных заболеваний.

Имеющиеся данные [8–14] дали нам основание предполагать, что нестабильность микросателлитных повторов хромосомной области 14q13.2–14q13 может быть связана с различными воспалительными процессами соединительной ткани, включая суставы. Воспаление суставов (суставной синдром) в детском возрасте является симптомом ряда заболеваний, включая неспецифические артриты, артралгию, реактивную артропатию, синовит, ЮИА и другие. В целом генетическая природа воспаления суставов слабо изучена, что объясняется вовлечением большого числа генов в возникновение воспалительной реакции и в патогенез перечисленных заболеваний [15].

Целью представленной работы был поиск связи между микросателлитной изменчивостью хромосомной области 14q13.2–14q13, включающей гены *PSMA6*, *FAM177A1* и *PRORP*, и воспалительными заболеваниями суставов у детей.

Материалы и методы

Участники исследования

Исследование проводилось в соответствии с этическими принципами Хельсинкской декларации. Оно одобрено комитетом по этике учреждения здравоохранения «2-я городская детская клиническая больница» (г. Минск), являющегося клинической базой 1-й кафедры детских болезней учреждения образования «Белорусский государственный медицинский университет». Пациенты были включены в исследование при прохождении ими терапии в 2012–2021 гг. в учреждении здравоохранения «2-я городская детская клиническая больница» после подписания добровольного информированного согласия на участие в научных исследованиях. За пациентов в возрасте до 16 лет информированное согласие подписали их родители.

Общее количество участников исследования составило 663 человека в возрасте от 1 года до 17 лет, которые были включены в различные группы исследования в соответствии с диагнозом. Из них 368 человек без аутоиммунных и хронических воспалительных заболеваний составили группу клинического контроля, 182 пациента вошли в группу с системным аутоиммунным ревматическим забо-

леванием, а именно с диагнозом ювенильный идиопатический артрит (ЮИА), и 113 человек — с суставным синдромом без аутоиммунной этиологии. Среди пациентов с ЮИА преобладающим диагнозом (более 50%) был олигоартрит (M08.4 в соответствии с Международной классификацией болезней МКБ-10), а в группе с суставным синдромом — артрит неуточненный (M13.9). Воспалительный процесс у большинства пациентов в основных группах исследования затрагивал один или несколько (до трех) суставов.

Базовые демографические характеристики групп исследования представлены в таблице 1. Распределение по возрасту и полу в группах исследования соответствовало демографическим характеристикам этих заболеваний в белорусской популяции (неопубликованные данные).

Выделение и исследование ДНК

Геномную ДНК выделяли из образцов периферической цельной крови стандартным фенол-хлороформным методом.

Генотипирование образцов ДНК проведено в хромосомной области 14q13.2–14q13 по пяти микросателлитным локусам: HSMS006 (TG)_n гена *PSMA6*; HSMS602 (CAA)_n (A)_m гена *FAM177A1*; HSMS701 (AC)₅AT(AC)_n, HSMS702 (TG)_n и HSMS801 (CA)_n гена *PRORP*. Генотипирование основано на амплификации участков ДНК с последующей проверкой наличия продуктов ПЦР путем электрофореза в 1%-ом агарозном геле и на фрагментном анализе ПЦР-продуктов. Использованы праймеры с флуоресцентными метками и следующие условия амплификации, подробно описанные

Таблица 1

Демографические характеристики групп исследования

Группы исследования	Число пациентов, n (%)	Пол, м / ж, n (%)	Возраст на момент обследования, годы Среднее значение ± стандартное отклонение
Контроль клинический	368 (55,51)	205 / 163 (55,7 / 44,3)	14 ± 2,48
Ювенильный идиопатический артрит	182 (27,45)	66 / 116 (36,3 / 63,7)	9 ± 4,95
Суставной синдром без аутоиммунной этиологии	113 (17,04)	63 / 50 (55,8 / 44,2)	9 ± 5,16
Все участники исследования	663 (100,0)	334 / 329 (50,4 / 49,6)	12 ± 4,66

в работе [5]: денатурация при 94 °С в течение 5 мин; 35 циклов амплификации — 30 сек денатурации при 94 °С; 30 сек отжига при 55 °С (60 °С для HSMS006); 30 сек элонгация при 72 °С, 10 мин при 72 °С; охлаждение до 4 °С.

Фрагментный анализ продуктов амплификации проводился методом автоматического капиллярного электрофореза с полихромным лазерным сканированием на анализаторе «ABI PRISM 3500» (Applied Biosystems, США); использовались капилляры длиной 50 см и 7%-й раствор полимера POP-7 (Applied Biosystems).

Подготовка реакционной смеси выполнялась в соответствии с инструкцией фирмы-производителя генетического анализатора.

Анализ длины фрагментов проводился с помощью пакета программ GeneMapper v.4.1 (Microsatellite Analysis Getting Started Guide). Для каждого микросателлитного локуса выбирались один или два пика с максимальной концентрацией фрагментов, свидетельствующих соответственно о гомозиготном или гетерозиготном состоянии микросателлита (рис. 1).

Уточнение длины фрагментов выполнено

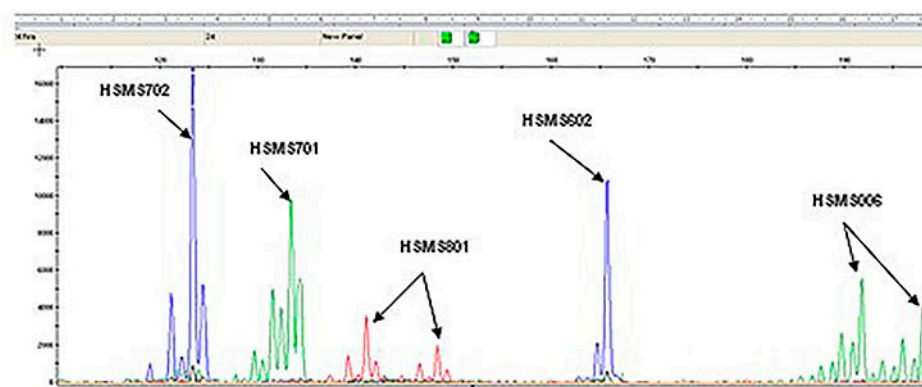


Рис. 1. Пример электрофореграмм исследуемых микросателлитных локусов HSMS702, HSMS701, HSMS801, HSMS602 и HSMS006 с указанием стрелками фрагментов, имеющих максимальную концентрацию (пик): одиночная стрелка — один пик при гомозиготном состоянии, двойная стрелка — два пика при гетерозиготном состоянии микросателлита

на том же приборе путем секвенирования по Сэнгеру выбранных образцов ДНК, полученных от гомозиготных особей. Пример секвенограммы представлен на рис. 2.

Статистический анализ

Статистическую обработку данных проводили с использованием пакетов программ

STATISTICA 7.0 (Statsoft), MS Excel 2010. Об ассоциации аллелей с заболеванием судили по величине показателя отношения шансов (OR — odd ratio) и 95% доверительного интервала (95% CI — 95% confidence interval). Различия считались статистически значимыми при $p < 0,05$.

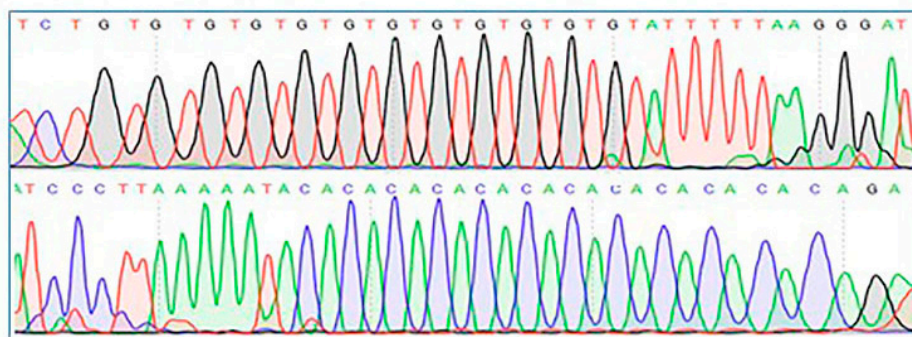


Рис. 2. Фрагмент секвенограммы микросателлитного локуса HSMS702, содержащего 12 TG повторов, при использовании прямого (верхний рисунок) и обратного праймеров (нижний рисунок)

Результаты и обсуждения

Проведенные исследования показали, что диапазон длин микросателлитных локусов на исследуемом хромосомном участке 14q13.2–14q13 у детей белорусской популяции был следующим: 171–195 п. н. HSMS006 (диапазон повторов $(TG)_{15}-(TG)_{26}$), 166–177 п. н. HSMS602 $((CAA)_7(A)_6)-(CAA)_{10}(A)_8$), 126–142 п. н. HSMS701 $((AC)_5AT(AC)_{14}-(AC)_5AT(AC)_{22})$, 124–150 п. н. HSMS702 $((TG)_{12}-(TG)_{24})$ и 129–155 п. н. HSMS801 $((CA)_{11}-(CA)_{24})$. При этом обнаружена высокая вариабельность полиморфных аллелей каждого микросателлитного локуса: в среднем около 10 аллелей на локус в каждой группе исследования, — что характерно для микросателлитов в целом и соответствует известным литературным данным по этому хромосомному участку в латвийской и финской популяциях [3, 5–7].

Полный перечень аллелей в нашем исследовании был следующий:

171, 173, 175, 177, 179, 181, 183, 185, 187, 189, 191, 193 и 195 п. н. HSMS006;

166, 167, 168, 169, 170, 171, 174 и 177 п. н. HSMS602;

126, 128, 130, 132, 134, 136, 138, 140 и 142 п. н. HSMS701;

124, 126, 128, 130, 132, 134, 136, 138, 142, 144, 148 и 150 п. н. HSMS702;

129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 151, 153 и 155 п. н. HSMS801.

Сравнительный анализ частот встречаемых аллелей между исследуемыми группами пациентов выявил несколько особенностей. Так, из всего разнообразия обнаруженных аллелей было отмечено восемь (табл. 2), которые имели значимо более высокие или низкие частоты при суставном синдроме и/или при ЮИА по сравнению с группой клинического контроля. При этом существенные различия между группами с ЮИА и суставным синдромом отсутствовали. Это позволило объединить данные этих групп по 7 (из 8) аллелям, которые имели тенденции в одном направлении (протективные или рискованные), и тем самым повысить статистическую значимость для большинства полученных результатов в объединенной группе, которая имеет одну общую составляющую — воспаление суставов. Таким образом, были выявлены семь микросателлитных аллелей (табл. 2), которые с высокими уров-

нями значимости (от 0,006 и менее) были ассоциированы с воспалительным процессом суставов различной этиологии. При этом пять аллелей: 170 п. н. HSMS602, 138 п. н. HSMS701, 124 п. н. HSMS702, 137 и 151 п. н. HSMS801 — были связаны с повышенным риском, и только два аллеля: 167 п. н. HSMS602 и 136 п. н. HSMS701 — имели протективное значение по отношению к воспалению суставов у детей.

В отличие от перечисленных аллелей, представленных в табл. 2, один аллель 130 п. н. HSMS702 показал свое протективное значение в отношении ЮИА ($OR = 0,26 [0,08-0,86]$, $p = 0,019$), но не в группе суставного синдрома без аутоиммунной этиологии, что не позволило объединить эти данные в одну группу и, скорее всего, указывает на связь этого аллеля только с ЮИА. Однако низкая частота встречаемости этого аллеля у пациентов с ЮИА требует подтверждения полученных результатов на большей выборке пациентов.

Кроме того, было обнаружено два аллеля (табл. 3): 189 п. н. HSMS006 и 141 п. н. HSMS801, — частота которых значимо ($p < 0,05$) отличалась у пациентов с аутоиммунным ревматическим заболеванием ЮИА по сравнению с суставным синдромом без аутоиммунной компоненты. Частота аллеля 189 п. н. HSMS006 была существенно выше в группе с ЮИА. Эти данные, по-видимому, указывают на связь аллеля 189 п. н. HSMS006 с возникновением аутоиммунной реакции при воспалительном процессе в суставах и на то, что он может быть использован в качестве маркера повышенного риска возникновения ЮИА у пациентов с выявленным суставным синдромом. И наоборот, в отношении аллеля 141 п. н. HSMS801, показавшего протективное действие как по сравнению с суставным синдромом (табл. 3), так и с клиническим контролем ($OR = 0,32 [0,17-0,62]$, $p = 0,0022$), можно предположить его использование в качестве маркера пониженного риска ЮИА (табл. 3).

Микросателлитные повторы $(TG)_n$ HSMS006 расположены в интроне 6 протеасомного гена *PSMA6*, и их связь с аутоиммунными заболеваниями (гипертиреозом, диабетом 2-го типа и ЮИА) была установлена в латвийской и частично финской популяциях [3, 5, 6]. В наших предыдущих работах было выявлено, что гено-

Таблица 2
Анализ частот аллелей микросателлитных повторов на хромосомном участке 14q13.2–14q13 у пациентов с диагнозом ЮИА и суставным синдромом без аутоиммунной компоненты по сравнению с группой клинического контроля (представлены данные, статистически значимо отличающиеся от группы клинического контроля)

Название микросателлита, нуклеотидный мотив	Аллели, п. н.	Клинический контроль (общее N = 736), n (%)	Суставной синдром без аутоиммунной этиологии (общее N = 226)			ЮИА (общее N = 364)			Объединенные данные по суставному синдрому без аутоиммунной этиологии и ЮИА (общее N = 590)		
			n, %	OR [95% CI]	p	n, %	OR [95% CI]	p	n, %	OR [95% CI]	p
HSMS602, (CAA) _n (A) _m	167	215 (29,21)	47 (20,80)	0,64 [0,44–0,91]	0,013	75 (20,60)	0,63 [0,47–0,85]	0,002	122 (20,68)	0,63 [0,49–0,82]	0,0004
	170	58 (7,88)	33 (14,60)	2,0 [1,27–3,15]	0,004	50 (13,74)	1,86 [1,25–2,78]	0,003	88 (14,17)	2,05 [1,44–2,91]	0,00006
HSMS701, (AC) ₅ AT(AC) _n	136	223 (30,38)	37 (16,37)	0,45 [0,3–0,66]	0,00002	63 (17,31)	0,48 [0,35–0,66]	0,000002	100 (16,95)	0,47 [0,36–0,61]	0,000000
	138	98 (13,32)	51 (22,57)	1,9 [1,3–2,77]	0,001	63 (17,31)	1,36 [0,97–1,92]	0,085	114 (19,32)	1,55 [1,16–2,09]	0,004
HSMS702, (TG) _n	124	383 (52,04)	140 (61,95)	1,5 [1,11–2,04]	0,009	215 (59,07)	1,33 [1,03–1,71]	0,029	355 (60,17)	1,39 [1,12–1,73]	0,003
	130	23 (3,13)	6 (2,65)	0,85 [0,34–2,10]	0,826	3 (0,82)	0,26 [0,08–0,86]	0,019	–	–	–
HSMS801, (CA) _n	137	82 (11,14)	40 (17,70)	1,72 [1,14–2,59]	0,011	57 (15,66)	1,48 [1,03–2,13]	0,042	97 (16,55)	1,57 [1,14–2,15]	0,006
	151	35 (4,76)	25 (11,06)	2,49 [1,46–4,26]	0,001	27 (7,42)	1,6 [0,96–2,70]	0,094	52 (8,81)	1,94 [1,24–3,01]	0,004

Примечание. n — количество аллелей в группах; п/ж начертанием выделены статистически значимые результаты

Таблица 3

Сравнительный анализ частот микросателлитных аллелей на хромосомном участке 14q13.2–14q13 у пациентов с диагнозом ЮИА и суставным синдромом без аутоиммунной компоненты (представлены статистически значимые результаты)

Название микросателлита, нуклеотидный мотив	Аллели, п. н.	Суставной синдром без аутоиммунной этиологии (общее N = 226), n (%)	ЮИА (общее N = 364)		
			n (%)	OR [95% CI]	P
HSMS006, (TG) _n	189	20 (8,85)	55 (15,11)	1,83 [1,07-3,15]	0,031
HSMS801, (CA) _n	141	16 (7,08)	11 (3,02)	0,41 [0,19-0,90]	0,026

Примечание. n — количество аллелей в группах; п/ж начертанием выделены статистически значимые результаты

тип G/C rs1048990 гена *PSMA6* ассоциирован с риском возникновения ЮИА в целом, особенно у девочек, а также с олигоартикулярной формой ЮИА (или олигоартритом) в частности [16, 17]. Эти данные подтверждают наличие тесной связи между изменчивостью гена *PSMA6* и риском возникновения аутоиммунной реакции в различных органах.

Существует предположение о влиянии полиморфной изменчивости (длины) микросателлитных повторов протеасомного гена *PSMA6* на функциональную активность самого гена и, соответственно, на предрасположенность к патологии [3]. Одним из возможных механизмов его участия в воспалительных и аутоиммунных реакциях допускается регуляция активности фактора транскрипции NF-κB, играющего ключевые роли в этих реакциях. Активная форма NF-κB представляет собой гетеродимер, состоящий из субъединицы p65 и субъединицы p50. Последняя представлена в цитозоле клетки в неактивной форме, а именно p105, полипептидного предшественника p50. Кроме того, преобразованные p50 и p65 поддерживаются в цитозоле в виде неактивного комплекса с ингибирующим белком IκB. Сигналы воспаления активируют NF-κB в результате инициации сигнального пути на полную деградацию IκB и стимулируют также превращение p105 в p50. Таким образом, для сигнала, индуцирующего активацию NF-κB, требуется два протеолитических события, оба определяемые убиквитин-протеасомным путем. Повышенная активность протеасомного гена *PSMA6* может усиливать убиквитин-протеасомный путь, который активирует фактор

транскрипции NF-κB, и наоборот, низкая экспрессия гена может вызывать снижение активности NF-κB [3, 18, 19].

Фактор транскрипции NF-κB играет плеiotропную роль в иммунной системе и, в частности, участвует в миграции иммунных эффекторных клеток, обеспечивая секрецию провоспалительных цитокинов. Неудивительно, что генетические варианты, влияющие на функцию эффекторов или на потерю функции негативных регуляторов пути NF-κB, ответственны за аутоиммунные или воспалительные заболевания. Можно сказать, что NF-κB действует на перекрестке многих сигнальных путей. Неправильная или чрезмерная активация NF-κB может привести к воспалительным заболеваниям и раку [18, 20]. Хорошо известно также, что активация NF-κB тесно связана с патогенезом ревматоидного артрита — тяжелого аутоиммунного заболевания у взрослых, которое характеризуется хроническим воспалением синовиальной ткани и, как следствие, повреждением костей и хрящей [20, 21].

Четыре других микросателлитных повтора, исследуемых в нашей работе, также локализованы в интронах [3]. Первый повтор (CAA)_n HSMS602 — в гене *FAM177A1*, кодирующего белок, локализованный в комплексе Гольджи, участвующий в ингибировании сигнального пути интерлейкина-1β (IL-1β), мощного провоспалительного цитокина. Изменение экспрессии или функции *FAM177A1* влияет на его способность ингибировать TRAF6-Ubc13-зависимую передачу сигнала от рецептора IL-1β [8, 22]. Протективный аллель 167 п. н. микросателлита HSMS602 может способствовать

поддержанию оптимального уровня экспрессии *FAM177A1*, что обеспечивает адекватное сдерживание IL-1 β -опосредованного воспаления. Если *FAM177A1* эффективно подавляет этот провоспалительный путь, то аллели, поддерживающие или усиливающие эту функцию, логично должны снижать риск развития воспалительных заболеваний. И наоборот, микросателлитные аллели, снижающие активность гена *FAM177A1*, могут способствовать развитию провоспалительной цепочки событий. К последним, возможно, относится аллель 170 п. н. микросателлита HSMS602 (табл. 2).

Три оставшихся интронных повтора (AC)_mAT(AC)_n HSMS701, (TG)_n HSMS702 и (CA)_n HSMS801 локализованы в гене *PRORP*, кодирующем эндонуклеазную субъединицу Р митохондриального рибонуклеазного комплекса, ответственного за созревание транспортной РНК [23]. Интронные микросателлиты способны влиять на процессы сплайсинга мРНК, ее стабильность или участвовать в формировании регуляторных некодирующих РНК [24]. Это подтверждается существованием естественного химерного некодирующего транскрипта *PRORP-PSMA6*, обнаруженного в 2022 г. Сквозной транскрипт представляет собой мРНК и вряд ли будет производить белковый продукт, но может иметь регуляторное значение [25].

Тем не менее, накопленных данных недостаточно, чтобы объяснить роль генов *PSMA6*, *FAM177A1* и *PRORP* в возникновении воспалительных и/или аутоиммунных реакций. Их изучение имеет особое значение, поскольку этот регион неоднократно связывался с предрасположенностью к различным аутоиммунным и воспалительным заболеваниям, включая ЮИА [5, 6]. Близость этих генов повышает вероятность их совместного наследования в виде гаплотипов, а также возможность существования общих регуляторных элементов, координирующих их экспрессию. Эта область генома активно изучается, т. к. она может привести к пониманию новых фундаментальных механизмов межгенных взаимодействий и их влияния на многофакторные заболевания [23], а также к разработке новых методов диагностики. Результаты нашей работы указывают на ассоциацию исследуемых микросателлитов генов *FAM177A1* и *PRORP* с воспалительным

процессом в суставах при различных формах суставного синдрома, включая аутоиммунное заболевание ЮИА. Эти результаты находятся в соответствии и расширяют данные, полученные Т. Sjakste с соавторами [3], которые впервые описали микросателлитную изменчивость на хромосомном участке 14q13.2–14q13 в связи с ЮИА в латвийской популяции.

Заключение

Таким образом, проведенное исследование представило описание микросателлитных локусов на хромосомном участке 14q13.2–14q13 у детей, проживающих на территории Беларуси. Впервые была установлена ассоциация микросателлитных повторов в генах *FAM177A1* и *PRORP* с суставным синдромом различной этиологии, свидетельствуя об их вовлеченности в воспалительный процесс суставов в детском возрасте. Кроме того, обнаружено, что микросателлитный аллель (189 п. н. HSMS006), локализованный в гене *PSMA6*, ассоциирован с риском ЮИА, но не суставного синдрома без аутоиммунной этиологии, что, по-видимому, указывает на его участие в возникновении аутоиммунной реакции при воспалительном процессе в суставах. Полученные данные раскрывают новые функциональные роли генов *PSMA6*, *FAM177A1* и *PRORP* и могут служить основой для разработки дополнительных критериев оценки предрасположенности к воспалительным и аутоиммунным реакциям в суставах у детей.

Работа проведена в рамках задания 2.28 подпрограммы «Геномика» ГПНИ «Фундаментальные основы биотехнологий» в 2012–2014 гг. и мероприятия 6.4 НТП Союзного государства «ДНК-идентификация» в 2017–2021 гг. Фрагментный анализ выполнен в центре коллективного пользования «Геном» Государственного научного учреждения «Институт генетики и цитологии Национальной академии наук Беларуси» с использованием прибора «ABI PRISM 3500» (Applied Biosystems, США). Авторы выражают благодарность сотрудникам центра за обслуживание прибора, а также медицинским работникам Учреждения здравоохранения «2-я городская клиническая больница», участвующим в сборе биологического материала.

Список использованных источников

1. Initial sequencing and analysis of the human genome / E. S. Lande, L. M. Linton, B. Birren [et al.] // *Nature*. – 2001. – Vol. 409, № 6822. – P. 860–921.
2. Khristich, A. N. On the wrong DNA track: Molecular mechanisms of repeat-mediated genome instability / A. N. Khristich, S. N. Mirkin // *Journal of Biological Chemistry*. – 2020. – Vol. 295, № 13. – P. 4134–4170.
3. Sjakste, T. Functional significance of microsatellite markers / T. Sjakste, N. Paramonova, N. Sjakste // *Medicina (Kaunas)*. – 2013. – Vol. 49, № 12. – P. 505–509.
4. Нестабильность микросателлитных повторов при неопухолевых заболеваниях легких: взаимосвязь с воспалением / К. А. Сычевская, М. В. Ерохина, С. К. Кравченко, Л. Н. Лепеха // *Практическая пульмонология*. – 2021. – № 1. – С. 52–60.
5. Microsatellite genotyping of chromosome 14q13.2–14q13 in the area of proteasomal gene *PSMA6* and association with Graves' disease in the Latvian population / T. Sjakste, J. Eglite, A. Sochnevs [et al.] // *Immunogenetics*. – 2004. – Vol. 56, № 4. – P. 238–243.
6. Identification of a novel candidate locus for juvenile idiopathic arthritis at 14q13.2 in the Latvian population by association analysis with microsatellite markers / T. Sjakste, I. Trapina, I. Rumba-Rozenfelde [et al.] // *DNA and Cell Biology*. – 2010. – Vol. 29, № 9. – P. 543–551.
7. Association of microsatellite polymorphisms of the human 14q13.2 region with type 2 diabetes mellitus in Latvian and Finnish populations / T. Sjakste, M. Kalis, I. Poudziunas [et al.] // *Annals of Human Genetics*. – 2007. – Vol. 71, Part 6. – P. 772–776.
8. FAM177A1 Inhibits IL-1 β -Induced Signaling by Impairing TRAF6-Ubc13 Association / B. W. Liao, H. Y. Zhang, W. T. Du [et al.] // *The Journal of Immunology*. – 2021. – Vol. 207, № 12. – P. 3090–3097.
9. Reinhard, L. Structure of the nuclease subunit of human mitochondrial RNase P / L. Reinhard, S. Sridhara, B. M. Hällberg // *Nucleic Acids Research*. – 2015. – Vol. 43, № 11. – P. 5664–5672.
10. Bi-allelic variants in the mitochondrial RNase P subunit PRORP cause mitochondrial tRNA processing defects and pleiotropic multi-system presentations / I. Hochberg, L. A. M. De-main, J. Richer [et al.] // *The American Journal of Human Genetics*. – 2021. – Vol. 108, № 11. – P. 2195–2204.
11. Cytosolic *N6AMT1*-dependent translation supports mitochondrial RNA processing / M. M. Foged, E. Recazens, S. Chollet [et al.] // *Proceedings of the National Academy of Sciences*. – 2024. – Vol. 121, № 47. – e2414187121.
12. Immunochip Meta-Analysis of Inflammatory Bowel Disease Identifies Three Novel Loci and Four Novel Associations in Previously Reported Loci / M. Hong, B. D. Ye, S. K. Yang [et al.] // *Journal of Crohn's and Colitis*. – 2018. – Vol. 12, № 6. – P. 730–741.
13. Сорокин, А. В. Протеасомная система деградации и процессинга белков / А. В. Сорокин, Е. Р. Ким, Л. П. Овчинников // *Успехи биологической химии*. – 2009. – Т. 49. – С. 3–76.
14. Wang, J. The ubiquitin-proteasome system and its role in inflammatory and autoimmune diseases / J. Wang, M. A. Maldonado // *Cellular and Molecular Immunology*. – 2006. – Vol. 3, № 4. – P. 255–261.
15. Osteoarthritis: trauma vs disease / G. Jiménez, J. Cobo-Molinos, C. Antich, E. López-Ruiz // *Advances in Experimental Medicine and Biology*. – 2018. – Vol. 1059. – P. 63–83.
16. Полиморфные варианты генов *PSMA6* и *PSMC6* как факторы предрасположенности к ювенильному идиопатическому артриту в белорусской популяции / И. Ю. Бакутенко, И. Д. Гилевская, Н. В. Никитченко [и др.] // *Молекулярная и прикладная генетика: сб. науч. тр.* – 2019. – Т. 27. – С. 39–45.
17. Polymorphism of proteasomal genes can be a risk factor for systemic autoimmune diseases in children / I. Y. Bakutenko, I. D. Hileuskaya, N. V. Nikitchenko [et al.] // *Journal of Pediatric Genetics*. – 2021. – Vol. 10, № 2. – P. 98–104.
18. NF- κ B: At the borders of autoimmunity and inflammation / L. Barnabei, E. Laplantine, W. Mbongo [et al.] // *Frontiers in Immunology*. – 2021. – Vol. 12. – 716469.
19. Inflammatory responses and inflammation-associated diseases in organs / L. Chen, H. Deng, H. Cui [et al.] // *Oncotarget*. – 2017. – Vol. 9, № 6. – P. 7204–7218.
20. Valin, A. The role of the transcriptional regulation of stromal cells in chronic inflammation / A. Valin, J. L. Pablos // *Biomolecules*. – 2015. – Vol. 5, № 4. – P. 2723–2757.

21. Firestein, G. S. NF- κ B: holy grail for rheumatoid arthritis / G. S. Firestein // *Arthritis and Rheumatology*. – 2004. – Vol. 50. – P. 2381–2386.
22. VPS13B is localized at the interface between Golgi cisternae and is a functional partner of FAM177A1 / B. Ugur, F. Schueder, J Shin [et al.] // *Journal of Cell Biology*. – 2024. – Vol. 223, № 12. – e202311189.
23. Rossmann, W. Discovery, structure, mechanisms, and evolution of protein-only RNase P enzymes / W. Rossmann, Ph. Giegé, R. K. Hartmann // *Journal of Biological Chemistry*. – 2024. – Vol. 300, № 3. – 105731.
24. Sjakste, T. Structural and functional significance of microsatellites / T. Sjakste, N. Paramonova, N. Sjakste // *Biopolymers and Cell*. – 2016. – Vol. 32, № 5. – P. 334–346.
25. An integrative database of all non-coding RNA (ncRNA) human genes. – URL: <https://www.genecards.org/cgi-bin/carddisp.pl?gene=PRORP-PSMA6#publications> (date of access: 11.06.2025).

N. V. Nikitchenko¹, I. Y. Bakutenko¹, A. M. Tchichko², E. V. Sechko², N. I. Ryabokon¹

MICROSATELLITE LOCI OF THE *PSMA6*, *FAM177A1* AND *PRORP* GENES AS GENETIC PREDISPOSITION MARKERS TO THE DEVELOPMENT OF THE INFLAMMATORY PROCESS OF VARIOUS ETIOLOGY IN CHILDREN'S JOINTS

¹State Scientific Institution

“Institute of Genetics and Cytology of the National Academy of Sciences of Belarus”

27 Akademicheskaya St., 220072 Minsk, the Republic of Belarus

e-mail: n.ryabokon@igc.by

²Belarusian State Medical University

83 Dzerzhinsky Ave., Minsk 220116, the Republic of Belarus

In pediatric patients with an inflammatory process in joints of various etiology, namely with a diagnosis of juvenile idiopathic arthritis (JIA) and the articular syndrome of non-autoimmune etiology, as well as in the clinical control group (without autoimmune and chronic inflammatory diseases), genotyping of the polymorphism of the chromosomal region 14q13.2–14q13 was performed by 5 microsatellite loci of the *PSMA6*, *FAM177A1* and *PRORP* genes: HSMS006 (TG)_n, HSMS602 (CAA)_n(A)_m, HSMS701 (AC)₅AT(AC)_n, HSMS702 (TG)_n и HSMS801 (CA)_n. When analyzing the frequencies of polymorphic microsatellite repeats in patients with the inflammation of joints of various etiology compared to the control, 5 alleles of the *FAM177A1* and *PRORP* genes were identified: 170 bp HSMS602; 138 bp HSMS701; 124 bp HSMS702; 137 and 151 bp HSMS801 associated with joint inflammation risk in childhood. The comparative analysis of groups with the JIA and the articular syndrome of non-autoimmune etiology made it possible to detect one more allele of the *PSMA6* gene (189 bp HSMS006), which is associated with JIA risk and, thus, possibly with the emergence of an autoimmune component in the inflammation of joints in children.

Keywords: microsatellite loci, *PSMA6*, *FAM177A1* and *PRORP* genes, articular syndrome, juvenile idiopathic arthritis, joint inflammation, fragment analysis.

Дата поступления в редакцию: 16 июня 2025 г.

Е. В. Волочник¹, Д. Р. Капуза¹, И. В. Пахомова¹, О. С. Вшивкова¹, Т. В. Шман²

ИЗМЕНЕНИЯ КАРИОТИПА ЛЕЙКЕМИЧЕСКИХ КЛЕТОК ПРИ РЕЦИДИВЕ ОСТРОГО ЛИМФОБЛАСТНОГО ЛЕЙКОЗА ИЗ В-КЛЕТОК ПРЕДШЕСТВЕННИКОВ У ПАЦИЕНТОВ ДЕТСКОГО И МОЛОДОГО ВЗРОСЛОГО ВОЗРАСТА

¹Государственное учреждение

«Республиканский научно-практический центр детской онкологии, гематологии и иммунологии»
Республика Беларусь, 223053, Минская обл., Минский р-н, д. Боровляны, ул. Фрунзенская, 43
e-mail: valochnikalena@gmail.com

²Минский научно-практический центр хирургии, трансплантологии и гематологии
Республика Беларусь, 220087, г. Минск, ул. Семашко, 8

При рецидиве острого лимфобластного лейкоза из В-клеток предшественников (ВП-ОЛЛ) в сравнении с первично диагностированным лейкозом наблюдается изменение структуры кариотипа как за счет уменьшения частоты встречаемости благоприятных *ETV6::RUNX1* (+) ($p < 0,001$) и высоко-гипердиплоидных ($p < 0,01$) кариотипов, так и увеличения частоты неблагоприятных кариотипов, таких как *KMT2A* (+) ($p < 0,001$), *BCR::ABL1* (+) ($p < 0,01$), гипоплоидного (≤ 44) ($p < 0,01$). Кроме того, при рецидиве группы ВП-ОЛЛ без специфических хромосомных aberrаций возрастает количество низко-гипердиплоидных ($p = 0,07$), сложных ($p < 0,05$) и моносомных кариотипов ($p < 0,05$), при этом в ~90% случаев моносомный кариотип сочетается со сложным кариотипом. Показано, что при рецидиве заболевания не наблюдается увеличения случаев с хромосомным дисбалансом, однако происходит изменение структуры кариотипа за счет увеличения количества структурных aberrаций ($p < 0,05$). Также возрастает количество сложных кариотипов с ≥ 3 структурными aberrациями в группе «B-other» ($p < 0,01$) и группе с высокой гипердиплоидией ($p < 0,05$), а в группе с гипоплоидией (≤ 44) не происходит усложнения кариотипа. Были выявлены дополнительные хромосомные aberrации, преимущественно несбалансированные, ассоциированные с цитогенетическими группами при первичной диагностике и рецидиве заболевания. По наличию хромосомного дисбаланса, сложного кариотипа и неклональных изменений выявлен наиболее стабильный вариант кариотипа — *KMT2A*+, и наименее стабильный — *ETV6::RUNX1*+ ВП-ОЛЛ.

Ключевые слова: острый лимфобластный лейкоз, рецидив, дополнительные хромосомные aberrации, сложный кариотип, моносомный кариотип, несбалансированные aberrации.

Для цитирования: Изменения кариотипа лейкемических клеток при рецидиве острого лимфобластного лейкоза из В-клеток предшественников у пациентов детского и молодого взрослого возраста / Е. В. Волочник, Д. Р. Капуза, И. В. Пахомова [и др.] // Молекулярная и прикладная генетика: сб. науч. тр. / Ин-т генетики и цитологии НАН Беларуси; редкол.: А. В. Кильчевский (гл. ред.) [и др.]. — Минск, 2025. — Т. 39. — С. 84–99.

Введение

Острый лимфобластный лейкоз (ОЛЛ) — генетически гетерогенное заболевание, связанное с накоплением генетических изменений в гемопоэтических клетках-предшественниках. За последнее десятилетие классификация ВОЗ была дополнена новыми подтипами лейкозов, и в настоящее время, в зависимости от наличия специфических генетических изменений, выделяют 12 вариантов острого лимфобластного лейкоза из В-клеток предше-

ственников (ВП-ОЛЛ) [1]. Данные aberrации при ОЛЛ имеют не только диагностическое значение, но и являются независимым фактором прогноза заболевания и используются для стратификации пациентов по группам риска. Помимо специфических клонообразующих aberrаций, дополнительно выделяют «вторичные» генетические изменения, которые не входят в классификацию ВП-ОЛЛ, однако могут быть ассоциированы с клинико-биологическими особенностями лейкемических клеток

и их можно рассматривать в качестве возможных предикторов прогноза.

В последние десятилетия использование современных протоколов лечения и риск-стратификации на основании наличия молекулярно-генетических маркеров, иммунологических и клинико-гематологических особенностей позволило повысить показатели общей выживаемости пациентов с ОЛЛ до 80–90% [2, 3]. Однако, несмотря на это, в 20–35% случаев с ОЛЛ выявляют рецидивы заболевания [4, 5], которые являются самой частой причиной неудач в лечении ОЛЛ у детей и молодых взрослых. Так, показатели 5-летней общей выживаемости и бессобытийной выживаемости у детей и молодых взрослых с рецидивом ОЛЛ в Республике Беларусь составили $42 \pm 5\%$ и $38 \pm 5\%$ соответственно [6]. Худший исход терапии рецидивов связан, в том числе, с тем, что лейкоэмические клетки при рецидиве более устойчивы к химиотерапии, чем при первичной диагностике, за счет наличия дополнительных генетических нарушений, обуславливающих селективное преимущество клеток с данными изменениями при лекарственной терапии. Поэтому оптимальной стратегией будет профилактика рецидива, что требует выявления дополнительных факторов прогноза рецидива на этапе первичной диагностики ОЛЛ с интенсификацией терапии и/или возможным проведением трансплантации гемопоэтических стволовых клеток в первой линии [4, 6].

Причиной генетических изменений, в том числе хромосомных aberrаций (ХА), является генетическая нестабильность, благодаря которой с большей частотой, чем в норме, происходят генетические изменения, являющиеся источником для отбора. При стандартной цитогенетической диагностике невозможно оценить хромосомные изменения в динамике, однако возможно определить наличие и количество клональных и неклональных ХА, в том числе несбалансированных (в результате которых образуется дисбаланс большого массива генов). Таким образом, хромосомная нестабильность на цитогенетическом уровне при клинической диагностике проявляется как совокупность структурных (СХА) и числовых хромосомных aberrаций (ЧХА). Наличие большого количества ХА, в том чис-

ле сложной субклональной структуры, может определять повышенную адаптивную приспособленность лейкоэмических клеток к лекарственной терапии и, как следствие, — выход в рецидив. В настоящее время описано достаточно много дополнительных хромосомных aberrаций (ДХА), ассоциированных как с первичным лейкозом, так и с рецидивом. Например, aberrации с перестройками генов *PAX5* [7, 8], *CRLF2* [9–11], *CREBBP* [7], *IKZF1* [7, 12], *CDKN2A/2B* [7, 12, 13], *IGH* [8], *TCF3* [14], *ETV6* [8, 12], или несбалансированные структурные aberrации — *dup(X)*, *dup(1q)* [12], *del(7p)/7–* [12], *abn(17p)/del(17p)* [8, 12] и другие цитогенетические изменения, многие из которых ассоциированы с плохим исходом заболевания и повышенным риском развития рецидива.

Среди ВП-ОЛЛ без специфических клонообразующих транслокаций выделяют два варианта кариотипа с большим количеством анеусомий: высокую гипердиплоидию (ВГ) с модальным числом хромосом (МЧХ) равным 51–65 и гиподиплоидию с $\text{МЧХ} \leq 44$. В данных кариотипах, за счет большого количества ЧХА, одновременно изменяется доза генов у большого массива генов. ВГ кариотип ассоциируется с благоприятным исходом терапии [15–17], однако, несмотря на это, примерно у 20% пациентов с ВГ наблюдается рецидив [18], что подчеркивает важность выявления дополнительных прогностических факторов в этой группе. Так, описаны более благоприятные aberrации, такие как сочетание трисомий 4, 10, 17 и/или 18 хромосом [7, 19]. Были определены и менее благоприятные варианты ВГ, такие как изменения в хромосомах 5, 6, 7, 10 и 17, выступающие как предикторы рецидива ВГ ВП-ОЛЛ [7], или наличие трисомий 5 и 20, как ухудшающих прогноз [7, 19]. Изучение наличия СХА при ВГ кариотипе, как неблагоприятного прогностического маркера, в ранних исследованиях не подтвердило его клиническую значимость [19, 20]. Однако некоторые дальнейшие исследования показывают, что при рецидиве чаще встречаются структурные изменения (в основном криптические гемизиготные делеции) и что структурные изменения и мутации RTK-RAS являются вторичными по отношению к ВГ паттерну [18]. Возможно влияние

генного дисбаланса, внесенного в кариотип благодаря несбалансированным ХА, нивелируется за счет анеуплоидий при ВГ.

Сложный кариотип (СК) определяется как вариант кариотипа с ≥ 3 несвязанными ХА и является общепризнанным прогностическим фактором у пациентов с ОМЛ, однако влияние СК на прогноз для пациентов с ОЛЛ изучен хуже. Так, в исследованиях А. V. Moorman с соавторами (2007) был выделен вариант СК с ≥ 5 любыми ХА, как фактор плохого прогноза у взрослых пациентов с ОЛЛ, включенных в исследование MRC UKALL XII/ECOG [21–23]. А при стратификации, предложенной NILG-ALL (Northern Italy Leukemia Group), СК кариотип расценивался как ≥ 3 любых ХА, не ассоциированных с плоидиями [21, 24]. Недостаточно изученным вариантом кариотипа при ОЛЛ у детей является и моносомный кариотип (МК) (≥ 2 аутосомных моносомий или единичная аутосомная моносомия в сочетании как минимум с одной СХА). По некоторым данным, МК кариотип у взрослых пациентов не был ассоциирован с худшим прогнозом [24, 25]. А в исследовании S. S. Kenderian с соавторами для взрослых В-ОЛЛ показано, что показатели общей выживаемости у пациентов с МК несколько хуже, однако МК кариотип не является независимым прогностическим маркером и большинство пациентов (86%) входили в группу высокого риска, а в 63% случаев содержали признаки СК [26].

Таким образом, изучение цитогенетических aberrаций, их сочетаний и динамика изменений кариотипа при рецидиве в сравнении с первичной диагностикой, позволят выявить новые значимые aberrации для диагностики и прогноза заболевания.

Цель работы — проанализировать структуру цитогенетических нарушений и изменение кариотипа лейкоэмических клеток при рецидиве ВП-ОЛЛ в сравнении с первично диагностированным лейкозом у пациентов детского и молодого взрослого возраста.

Материалы и методы

В исследовании был проанализирован кариотип 1002 первично диагностированных пациентов с ВП-ОЛЛ, получавших лечение с 2002 по 2020 гг. на базе государственного учреждения «Республиканский научно-практиче-

ский центр детской онкологии, гематологии и иммунологии» (далее — Центр), и 171 случай костномозгового рецидива, выявленного до 01.01.2025 г. у данных пациентов. Соотношение по полу (мужской:женский) при первичной диагностике и рецидиве заболевания составило 1,15:1 и 1,44:1 соответственно. Медиана возраста пациентов при первичной диагностике составила 4,7 года (интервал от 1 месяца до 29 лет) и 9,4 года (интервал от 5 месяцев до 32 лет) при рецидиве. Медиана наблюдения от наступления ремиссии до развития рецидива составила 2,5 года (интервал от 2 месяцев до 15 лет). У всех пациентов и/или их законных представителей было получено информированное согласие на обследование и лечение по клиническим протоколам, принятым в Центре.

Наличие специфических хромосомных перестроек с образованием химерных генов *E2A::PBX1*, *ETV6::RUNX1*, *BCR::ABL1* и *KMT2A::v* подтверждалось методами флуоресцентной *in situ* гибридизации (FISH) и ПЦР. Кариотипирование проводилось на метафазных пластинках ночной культуры костного мозга с использованием метода GTG (англ. G-bands by trypsin using Giemsa) и, при необходимости уточнения структуры aberrаций, с использованием дополнительных локус-специфических, центромерных и полнохромосомных FISH зондов. Регистрация данных проводилась в соответствии с рекомендациями международной системы цитогеномной номенклатуры человека (ISCN) [27]. Для статистической обработки данных использовали программу Statistica 7.0, значение *p*-value определяли при помощи критерия χ^2 и критерия Фишера.

Результаты и обсуждение

В связи с тем, что до конца 2019 г. в Центре при первичной диагностике наличие ВГ и гипоплоидного (≤ 44) кариотипов определялись методом дифференциального окрашивания (GTG), в некоторых случаях методом FISH — только на наличие трисомий 4, 10 и 17, то при оценке структуры ВП-ОЛЛ по цитогенетическим группам в данной работе были выделены два условных блока: с наличием специфических ХА (*t*(12;21)/*ETV6::RUNX1*, *t*(9;22)/*BCR::ABL1*, *t*(1;19)/*E2A::PBX1* и *t*(11;v)/

KMT2A::v) и без них. Чтобы избежать искажения частоты встречаемости ВГ и гипоплоидных кариотипов, данные варианты оценивались только при эффективном стандартном цитогенетическом исследовании (СЦИ). Кариотип «В-other» определялся как ВП-ОЛЛ без специфических ХА, ВГ кариотипа и гипоплоидии (≤ 44).

Во всех случаях было проведено исследование на наличие транслокаций *t(12;21)/ETV6::RUNX1*, *t(1;19)/E2A::PBX1*, *t(9;22)/BCR::ABL1* и *t(11;v)/KMT2A::v* методами молекулярной и/или цитогенетической диагностики. Изменение структуры рекуррентных хромосомных транслокаций при рецидиве заболевания показано на рисунке 1.

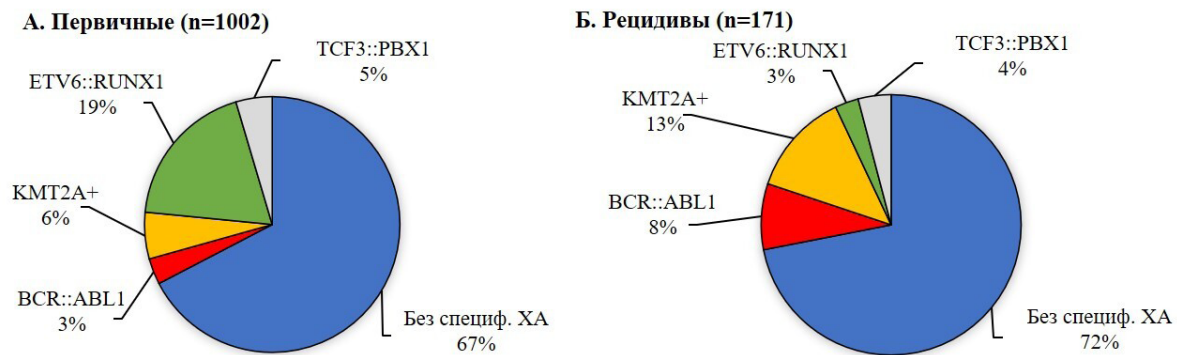


Рис. 1. Частота встречаемости специфических хромосомных транслокаций *t(12;21)/ETV6::RUNX1*, *t(11;v)/KMT2A::v*, *t(9;22)/BCR::ABL1* и *t(1;19)/E2A::PBX1* при первичной диагностике (А) и рецидивах (Б) ВП-ОЛЛ

Анализ кариотипа лейкомиических клеток при рецидиве ожидаемо показал изменение структуры цитогенетических изменений, с увеличением частоты встречаемости известных для ВП-ОЛЛ неблагоприятных кариотипов с рекуррентными транслокациями *t(11;v)/KMT2A::v* ($p < 0,001$) и *t(9;22)/BCR::ABL1* ($p < 0,01$), и уменьшением частоты благоприятного — с *t(12;21)/ETV6::RUNX1* ($p < 0,001$).

Анализ МЧХ, наличие дисбаланса и ДХА был проведен в 623 случаях с первично диагностированным ВП-ОЛЛ и 117 случаях рецидива заболевания с эффективным СЦИ (наличие достаточного количества информативных метафазных пластинок). К клональным изменениям относили наличие СХА или трисомий в ≥ 2 метафазных пластинках или моносомий

в ≥ 3 метафазах [27].

Анеуплоидии

При изучении кариотипов по МЧХ, в различных генетических подгруппах показано, что группы со специфическими ХА в основном представлены псевдодиплоидным кариотипом, а ВГ или гипоплоидные (≤ 44) кариотипы выявляются крайне редко, и их роль в прогнозе заболевания не определена. Большое количество случаев при первичной диагностике и рецидиве, а также статистически значимые различия для некоторых вариантов анеуплоидий наблюдались в группе ВП-ОЛЛ без специфических aberrаций (табл. 1).

При рецидиве ВП-ОЛЛ без специфических aberrаций было показано увеличение частоты встречаемости всех вариантов гипоплоидий

Таблица 1

Распределение по МЧХ в группе ВП-ОЛЛ без специфических aberrаций при манифестации и рецидиве заболевания

Группа МЧХ (количество хромосом)	Первичные n (%)	Рецидивы n (%)	χ^2 , p-value
Нормальный кариотип (46, N)	39 (9,7)	8 (9,3)	0,9
Псевдодиплоидия (46, abn)	83 (20,7)	18 (20,9)	0,96

Окончание таблицы 1

Группа МЧХ (количество хромосом)	Первичные n (%)	Рецидивы n (%)	χ^2 , p-value
Гипоплоидия (≤ 45)	26 (6,5)	15 (17,4)	0,0009
Гипоплоидия (=45)	17 (4,2)	8 (9,3)	0,054
Гипоплоидия (≤ 44)	9 (2,2)	7 (8,1)	0,005
Низкая гипердиплоидия (47–50)	35 (8,7)	13 (15,1)	0,071
Высокая гипердиплоидия (51–65)	211 (52,6)	31 (36,0)	0,005
Другие варианты	7 (1,7)	1 (1,2)	0,69
Всего*	401	86	

Примечание. * — количество ВП-ОЛЛ без специфических аберраций с успешным СЦИ; выделены группы МЧХ основываясь на классификацию, предложенную А. Моогман [8, 19, 21] и использующуюся в ВОЗ-классификации [1]

($p < 0,001$) (в частности, за счет гипоплоидного варианта с МЧХ ≤ 44 , $p = 0,005$) и уменьшение частоты ВГ кариотипов ($p = 0,005$). Также при рецидиве заболевания чаще выявлялся НГ кариотип, а соотношение случаев с нормальными, псевдодиплоидными кариотипами или гиперплоидией с МЧХ > 66 хромосом не изменялось.

Хромосомный дисбаланс и ДХА

Причиной цитогенетических нарушений является хромосомная нестабильность, при которой увеличивается частота структурных и числовых нарушений кариотипа опухолевых клеток по отношению к нормальным клеткам. При СЦИ хромосомная нестабильность может проявляться как в виде неклональных аберраций, так и клональных ДХА. При наличии хромосомного дисбаланса увеличивается вероятность хромосомных нарушений, что влечет за собой появление новых аберраций и клональную эволюцию под действием отбора. Однако очень высокий уровень нестабильности кариотипа также может повлечь за собой элиминацию клеток. В нашей работе была проведена оценка уровня дисбаланса (как фактора нестабильности кариотипа) для основных цитогенетических подгрупп ВП-ОЛЛ при первичной диагностике и рецидиве заболевания.

Анализ хромосомных аберраций показал наличие большого количества случаев с несбалансированным кариотипом (имеются ХА, в результате которых изменяется количественный состав генов) во всех группах ВП-ОЛЛ

как при первичной диагностике, так и при рецидиве заболевания (рис. 2А). Также была проведена оценка структурного состава дисбаланса при первичной диагностике и рецидиве заболевания (рис. 2Б).

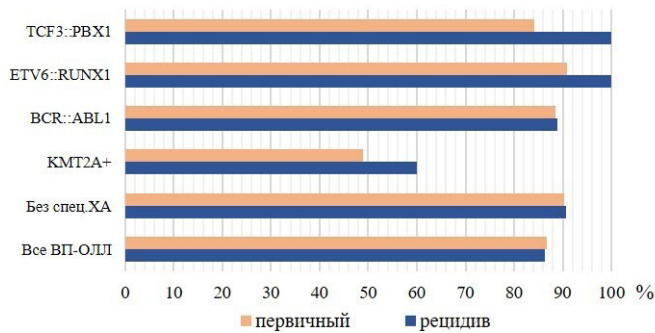
Состав дисбаланса определялся тем, какими ХА вызван дисбаланс:

- 1) нет дисбаланса (нормальный кариотип или наличие только сбалансированных ХА);
- 2) только СХА (наличие несбалансированных ХА, без числовых);
- 3) только ЧХА (наличие числовых изменений в отсутствие несбалансированных СХА);
- 4) смешанный тип (наличие ЧХА и несбалансированных СХА);
- 5) дисбаланс с неясным типом (выявлены изменения МЧХ, но из-за неполного разбора кариотипа неизвестно, имеются ли СХА).

Во всех группах ВП-ОЛЛ выявлена высокая частота встречаемости несбалансированного кариотипа как при первичной диагностике (84–91%), так и при рецидиве заболевания (86–100%), за исключением КМТ2А+ ВП-ОЛЛ (49 и 60%, $p < 0,001$). Однако при рецидиве заболевания ни в одной из групп не наблюдалось статистически значимого увеличения дисбаланса (рис. 2А) ($p > 0,05$).

При анализе структурного состава дисбаланса показано, что при рецидиве заболевания в сравнении с первичными лейкозами увеличивается частота дисбаланса за счет СХА ($p < 0,05$) (как в изолированном виде, так и в сочетании с ЧХА), кроме групп

А. Частота встречаемости несбалансированного кариотипа



Б. Структура дисбаланса

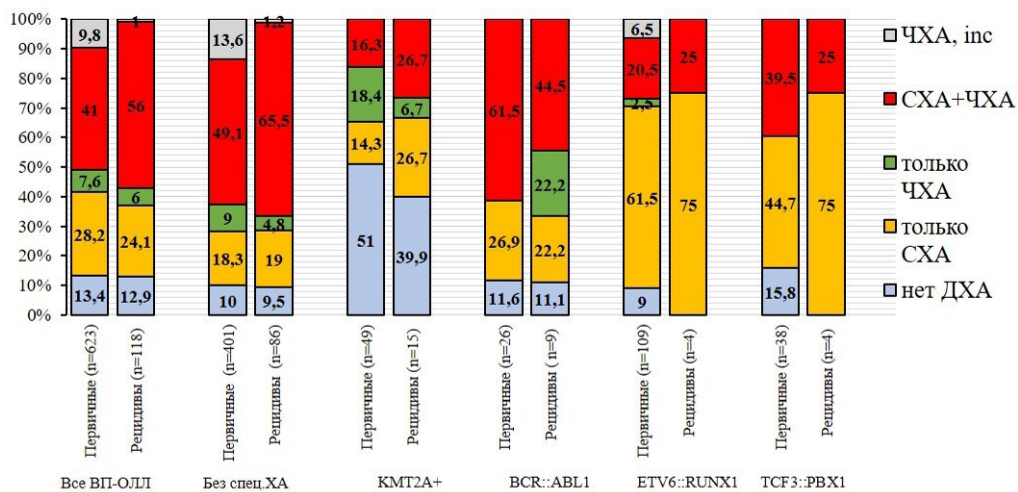


Рис. 2. Частота встречаемости несбалансированного кариотипа (А) и структура дисбаланса (Б) в различных цитогенетических группах ВП-ОЛЛ при первичной диагностике и рецидивах заболевания. inc (incomplete karyotype) — не полностью разобранный кариотип, на рисунке показаны случаи с установленными числовыми aberrациями, но не известно наличие СХА

BCR::ABL1+ и *KMT2A*+ ВП-ОЛЛ. При этом статистически значимо при рецидиве увеличивается именно количество случаев по смешанному типу (несбалансированные СХА + ЧХА) для всех ВП-ОЛЛ (41 против 56%; $p < 0,01$).

В группе ВП-ОЛЛ без специфических ХА имеется высокий уровень дисбаланса на момент первичной диагностики, который не изменяется при рецидиве заболевания. Однако при рецидиве происходит изменение качественного состава кариотипа — статистически значимо увеличивается количество случаев с СХА (как изолированными, так и в сочетании с числовыми) и имеются различия по наличию дисбаланса по смешанному типу (49,1 против 65,5%; $p < 0,01$) (рис. 2Б). По всей видимости, в данной группе числовые aberrации не играют значимой роли и являются ней-

тральными ХА, в отличие от СХА, в большей части несбалансированных, которые приводят к генному дисбалансу с потерей или увеличением количества значимых для патогенеза генов. Например, aberrации с вовлечением короткого плеча 9-й хромосомы приводят к потере генов *CDKN2A/B*, *JAK2* или реаранжировкам гена *PAX5*, что подтверждается данными FISH анализа.

Тип дисбаланса определяется и коррелирует с наличием несбалансированных ДХА, однако в группах со специфическими ХА дисбаланс определяют не только несбалансированные ДХА, но и несбалансированный вариант драйверной ХА (например, при *TCF3::PBX1*+ ВП-ОЛЛ, где дисбаланс в более чем 70% обеспечивается *der(19)t(1;19)*). Соответственно, в данных подгруппах уровень дисбаланса мо-

жет быть выше уровня ДХА. Возможен и обратный вариант, когда частота выявляемых ДХА выше, чем уровень дисбаланса, что происходит в случае, если ДХА представлены сбалансированными ХА (псевдодиплоидный кариотип со сбалансированными СХА). Данные об изменении структуры ДХА при первичной диагностике и рецидиве заболевания представлены на рисунке 3.

В структуре кариотипа у первично диагностированных ВП-ОЛЛ и при рецидиве заболевания наиболее часто выявляют СХА, как в сочетании с числовыми, так и без них. Наименьшее количество ДХА представлено в группе *KMT2A*+ ВП-ОЛЛ (хотя при рецидиве происходит прирост СХА) в сравнении с другими группами ВП-ОЛЛ ($p < 0,01$), за исключением *TCF3::PBX1*+ ВП-ОЛЛ, который имеет промежуточное значение по наличию

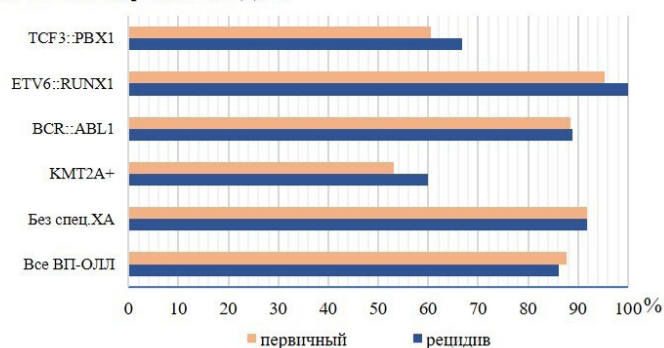
ДХА (рис. 3А). Изменение структуры ДХА при рецидиве во всех группах происходит за счет снижения количества случаев изолированных ЧХА и увеличения количества случаев с СХА (как изолированных, так и по смешанному типу) (рис. 3Б).

Таким образом, при рецидиве заболевания в большинстве случаев выявляются и, по-видимому, играют значимую роль в развитии и прогрессии заболевания именно структурные ДХА, подавляющее количество которых несбалансированные.

Паттерны ХА

Помимо общей характеристики структуры кариотипа по типам аберраций, были изучены конкретные ХА, входящие в состав кариотипа. Повторяющиеся ХА, ассоциированные с основными цитогенетическими подгруппами ВП-ОЛЛ, в том числе отдельно для гипо-

А. Частота встречаемости ДХА



Б. Типы ДХА в кариотипе лейкоэмического пула клеток

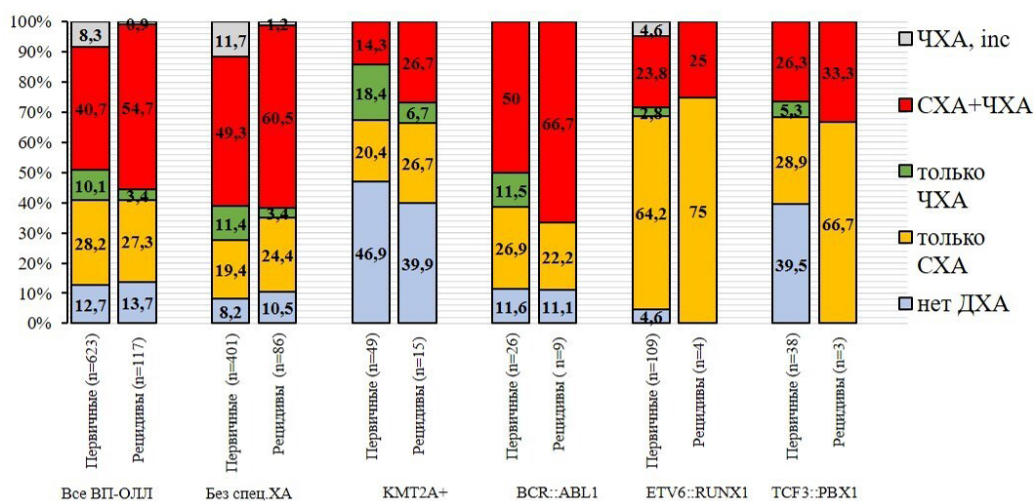


Рис. 3. Частота встречаемости (А) и структура ДХА (Б) при первичной диагностике и рецидивах заболевания в различных цитогенетических группах ВП-ОЛЛ. inc (incomplete karyotype) — не полностью разобранный кариотип, на рисунке показаны случаи с установленными числовыми аберрациями, но не известно наличие СХА

плоидных (≤ 44) и ВГ кариотипов, при первичной диагностике и при рецидиве представлены в таблице 2.

В группе *KMT2A*+ ВП-ОЛЛ при рецидиве заболевания увеличивается частота встречаемости аберраций 7-й хромосомы, с потерей ге-

Таблица 2

Частота встречаемости ДХА, выявляемых при первичной диагностике и рецидиве заболевания в цитогенетических группах ВП-ОЛЛ

Группы ВП-ОЛЛ	ДХА	Первичные n (%)	Рецидивы n (%)	χ^2 , p-value
ВП-ОЛЛ со специфическими ХА				
t(11;v)/ <i>KMT2A::v</i>	t1q	0/49 (0)	2/15 (13,3)	0,009
	del4q	2/49 (4,1)	2/15 (13,3)	0,19
	der6q	5/49 (10,2)	1/15 (6,7)	0,68
	del/t7p	2/49 (4,1)	3/15 (20)	0,044
	der11q (интактная)	4/49 (8,2)	1/15 (6,7)	0,85
	-9	0/49 (0)	2/15 (13,3)	0,009
	-10	3/49 (6,1)	0/15 (0)	0,32
	+X	5/49 (10,2)	1/15 (6,7)	0,68
	+8	3/49 (6,1)	0/15 (0)	0,32
t(9;22)/ <i>BCR::ABL1</i>	der1q	4/26 (15,4)	1/9 (11,1)	0,75
	imb7p/7q	3/26 (11,5)	2/9 (22,2)	0,42
	del9p	1/26 (3,8)	2/9 (22,2)	0,089
	+der(22)t(9;22)	8/26 (30,8)	1/9 (11,1)	0,24
	-2	1/26 (3,8)	2/9 (22,2)	0,089
	-7	4/26 (15,4)	1/9 (11,1)	0,75
	-9	5/26 (19,2)	0/9 (0)	0,15
	-17	4/26 (15,4)	0/9 (0)	0,21
	+5	4/26 (15,4)	2/9 (22,2)	0,69
	+21	2/26 (7,4)	3/9 (33,3)	0,058
t(12;21)/ <i>ETV6::RUNX1</i>	del6q	28/98 (28,6)	1/4 (25)	0,87
	del9p	6/98 (6,1)	0/4 (0)	0,61
	del11q	10/98 (10,2)	0/4 (0)	0,5
	der12p/loss12p (интактная)	55/106 (51,9)*	1/4 (25)	0,29
	-13	5/98 (5,1)	0/4 (0)	0,64
	+16	6/98 (6,1)	0/4 (0)	0,61
	+21	9/98 (9,2)	0/4 (0)	0,52
t(1;19)/ <i>E2A::PBX1</i>	del6q	7/38 (18,4)	0/3 (0)	0,34
	der9p	9/38 (23,7)	0/3 (0)	0,27
	der(19)t(1;19)**	28/38 (73,7)	4/4 (100)	0,23
	-13	4/38 (10,5)	0/3 (0)	0,49
	+X	4/38 (10,5)	0/3 (0)	0,49
	+4	4/38 (10,5)	0/3 (0)	0,49
	+6	4/38 (10,5)	0/3 (0)	0,49
	+21	7/38 (18,4)	0/3 (0)	0,34

Окончание таблицы 2

Группы ВП-ОЛЛ	ДХА	Первичные п (%)	Рецидивы п (%)	χ^2 , p-value
ВП-ОЛЛ без специфических ХА				
ВГ (51–65) ^а	t1q	15/211 (7,1)	5/32 (15,6)	0,1
	dup1q	25/211 (11,8)	1/32 (3,1)	0,13
	der4q/loss4q	8/211 (3,8)	2/32 (6,3)	0,51
	del6q	43/211 (20,4)	7/32 (21,9)	0,5
	abn9p/loss9p	13/211 (6,2)	7/32 (21,9)	0,057
	del9q	7/211 (3,3)	3/32 (9,4)	0,1
	abn12p/loss12p	11/211 (5,2)	1/32 (3,1)	0,61
	i(7)(q10)	3/211 (1,4)	2/32 (6,3)	0,073
	der7/loss7p22	10/211 (4,7)	5/32 (15,6)	0,017
	t14q	4/211 (1,9)	1/32 (3,1)	0,64
	abn17p/loss17p	10/211 (4,7)	5/32 (15,6)	0,017
	i(17)(q10)	7/211 (3,3)	5/32 (15,6)	0,0028
гипоплоидия (≤ 44) ^а	del6q	1/9 (11,1)	3/7 (42,9)	0,19
	t7q	3/9 (33,3)	1/7 (14,3)	0,39
«B-other» (без ВГ и гипо (≤ 44) кариотипов)	+X	10/180 (5,6)	3/47 (6,4)	0,82
	-7	4/180 (2,2)	3/47 (6,4)	0,14
	-9	6/180 (3,3)	4/47 (8,5)	0,12
	-13	5/180 (2,8)	2/47 (4,3)	0,6
	-14	3/180 (1,7)	3/47 (6,4)	0,073
	-15	5/180 (2,8)	3/47 (6,4)	0,23
	-20	6/180 (3,3)	4/47 (8,5)	0,12
	+21	13/180 (7,2)	4/47 (8,5)	0,76
	delXp	7/180 (3,9)	4/47 (8,5)	0,189
	t1q	5/180 (2,8)	5/47 (10,6)	0,019
	del4q	3/180 (1,7)	3/47 (6,4)	0,073
	del6q	33/180 (18,3)	6/47 (12,8)	0,36
	t7q	8/180 (4,4)	4/47 (8,5)	0,26
	der9p/loss9p	53/180 (29,4)	13/47 (27,7)	0,81
	der12p/loss12p	27/180 (15)	5/47 (10,6)	0,44
	t14q	17/180 (9,4)	8/47 (17)	0,13
	der17p/loss17p	8/180 (4,4)	3/47 (6,4)	0,34
	t19p	5/180 (2,8)	6/47 (12,8)	0,0045
	der21q/amp21q	3/180 (1,7)	2/47 (4,3)	0,28

Примечание. * — наличие делеции 12p с вовлечением интактной копии *ETV6* при первичных ВП-ОЛЛ с t(12;21)/*ETV6::RUNX1*+ было возможно определить в 106 случаях за счет данных полученных методом FISH;

** — der(19)t(1;19) не относится к ДХА, однако является характерным маркером, поэтому включен в паттерн группы der(19)t(1;19) ВП-ОЛЛ; а — для ВГ и гипоплоидий оценивались только СХА

нетического материала хромосомного региона 7p22 ($p < 0,05$). Также были выявлены повторяющиеся случаи моносомии 9-й хромосомы

и несбалансированных транслокаций с вовлечением хромосомного региона 1q (приведшие к удвоению материала 1q23→1qter), не выяв-

ляемые при первичной диагностике.

Для рецидивов *BCR::ABL1* + ВП-ОЛЛ наиболее частой аберрацией явилась трисомия 21-й хромосомы, частота которой возрастает при рецидиве ($p = 0,058$). При этом дополнительная дериватная хромосома $+der(22)t(9;22)$, по некоторым источникам описываемая как неблагоприятный маркер, хотя и выявлялась в 30% первичных Ph⁺ ВП-ОЛЛ, однако при рецидиве заболевания была обнаружена всего в 11% случаев.

В группе с гипоплоидией (≤ 44) при первичной диагностике и рецидиве заболевания выявлено небольшое количество повторяющихся ХА ($del6q$ и $t7q$), частота встречаемости которых не изменялась при рецидиве ($p > 0,05$). При том, что как минимум одна структурная аберрация детектирована у 66,7% (6/9) первичных лейкозов и 85,7% (6/7) рецидивов. Большинство СХА при гипоплоидии (≤ 44) случайны или представлены маркерной хромосомой неизвестного происхождения (44,4 и 57,1% соответственно).

При рецидиве ВГ кариотипа наблюдается увеличение частоты встречаемости несбалансированных аберраций с вовлечением хромосомных регионов $9p/loss9p$ ($p = 0,057$), $7p/loss7p22$ ($p < 0,05$), $17p/loss17p$ ($p = 0,017$) (в том числе $i(17)(q10)$, $p < 0,01$). Аберрации $9p$ представлены разнообразными вариантами аберраций: терминальными делециями короткого плеча — $dic(9;12)$, $dic(9;20)$, $i(9)(q10)$, и сложными перестройками с вовлечением короткого и длинного плеча 9-й хромосомы. При подтверждении делеции $9p$ методом FISH выявляли наличие гомо- и гетерозиготных делеций $9p21/CDKN2A/B$. Аберрации $7p$ были представлены во всех случаях в виде несбалансированных транслокаций с неизвестной хромосомой, в результате чего происходила потеря терминального региона короткого плеча.

Для группы «B-other» (без ВГ и гипоплоидии ≤ 44) характерно большое количество повторяющихся ДХА, частота встречаемости которых существенно не изменяется в сравнении с первично диагностированными лейкозами. Значимое увеличение частоты встречаемости показано только для транслокаций с вовлечением $1q$ ($p < 0,05$) и $19p$ ($p < 0,01$). Транслокации с вовлечением длинного пле-

ча 1-й хромосомы были разнообразны, однако в 60% случаев при рецидиве входили в состав несбалансированных перестроек, сопровождающихся увеличением генетического материала $1q$, и в одном случае — в виде $dup(1)(q25q42)$. Все выявленные при рецидиве аберрации 19-й хромосомы ($n = 6$) были в виде транслокаций с вовлечением хромосомного региона $19p13$, при этом в четырех случаях подтверждена реаранжировка гена *TCF3* (*TCF3::PBX1*-отрицательные) методом FISH, в одном — делеция гена *TCF3*, и еще в одном случае сохранялись обе интактные копии гена *TCF3*. Из четырех случаев с реаранжировкой гена *TCF3* только в одном была выявлена транслокация $t(17;19)$ с образованием химерного гена *TCF3::HLF* (в виде сбалансированной реципрокной транслокации), в остальных трех — в виде несбалансированных перестроек с неизвестным геном-партнером. Аберрации 14-й хромосомы в основном были представлены в виде моносомии 14-й хромосомы и транслокаций $14q$ (в 75% случаев — с участием хромосомного региона $14q32/IGH+$). Однако в структуре рецидивов не произошло изменение частоты встречаемости таких аберраций, как $derXp/CRLF2+$, $del6q$, $der9p/loss9p$, $der12p/loss12p$ и $der17p/loss17p$ ($p > 0,1$).

Таким образом, были изучены паттерны ХА, характерные для разных подгрупп ВП-ОЛЛ и для большинства выявленных ДХА, не происходит значимого изменения частот при рецидиве, т. е. данные аберрации являются общими и для первичного лейкоза, и для рецидива заболевания. При этом имеются аберрации, характерные почти для всех цитогенетических групп, такие как: $del6q$ (кроме *BCR::ABL1* +) или $abn9p/del9p$ (кроме *KMT2A* + и гипоплоидии (≤ 44)). Кроме того, выявлены аберрации, ассоциированные только с одной-двумя группами — $dup1q$ (ВГ), $t7q$ (гипоплоидии (≤ 44) и «B-other»), $i(17)(q10)$ (ВГ), $der21q/iAMP21$ («B-other»). Возможно, ДХА, ассоциированные с определенной цитогенетической группой, имеют значение в патогенезе именно для данного варианта лейкоза (содержат гены, изменение экспрессии которых значимо для патогенеза), что требует дальнейшего изучения. Для ряда выявленных ХА возрастает частота встречаемости при рецидиве заболевания. Это происходит как за счет развития рециди-

ва у пациентов, в кариотипе которых при первичной диагностике имелась данная ХА, так и за счет приобретения данного маркера при рецидиве заболевания, что может указывать на их адаптивную роль под действием лекарственной селекции. Так, например, в группе ВГ при первичной диагностике из семи случаев с *i(17)(q10)* два развили рецидив, и еще в трех случаях при рецидиве была выявлена *i(17)(q10)*, при первичной диагностике которых был выставлен ВГ кариотип без данного маркера (еще в двух случаях не было данных СЦИ при первичной диагностике).

Количество СХА и варианты СК

Анализ количества СХА для всех ВП-ОЛЛ показал, что наблюдается увеличение количества структурных aberrаций при рецидиве, при этом значимо увеличивается количество случаев с тремя и более СХА (первично — 164/612, 26,7% и при рецидиве — 49/117, 41,9%; $p = 0,001$). При оценке изменения количественного состава по цитогенетическим группам показано, что увеличение количества СХА наблюдается в группах с *KMT2A+*, *BCR::ABL1+* и ВП-ОЛЛ без специфических ХА, однако статистически значимые различия получены только в последней группе. При наличии нескольких субклонов с раз-

ным количеством СХА, учитывались результаты с наибольшим количеством изменений. Так как в понятие СК входит наличие любых трех несвязанных aberrаций [27], но в разных литературных данных прогностическое значение для разных вариантов СК различается [21–24], то в нашей работе были определены частоты встречаемости для двух вариантов СК. Первый вариант — кариотип с наличием любых ≥ 3 несвязанных хромосомных aberrаций и как минимум одной структурной перестройкой ($СК \geq 1СХА$), второй — только с ≥ 3 несвязанными СХА ($СК \geq 3СХА$) (табл. 3).

Для группы ВП-ОЛЛ без специфических ХА было показано значимое увеличение СК кариотипа не только для $СК \geq 1СХА$ варианта, но и для $СК \geq 3СХА$ ($p < 0,01$). Причем значимые различия в типах кариотипа получены для группы ВП-ОЛЛ без специфических ХА в подгруппах ВГ и «В-other». Усложнение структуры при рецидиве гипоплоидных (≤ 44) кариотипов не наблюдалось ($p > 0,1$).

Если анализировать только структурные aberrации, то при рецидиве ВП-ОЛЛ без специфических ХА происходит увеличение частоты встречаемости $СК \geq 3СХА$ (23,7 и 45,3%; $p < 0,001$), при этом наблюдается увеличение во всех подгруппах с 3, 4 и $\geq 5СХА$ ($p < 0,05$).

Таблица 3

Частота встречаемости СК в цитогенетических группах

Цитогенетические группы	СК $\geq 1СХА$		СК $\geq 3СХА$	
	первичные n (%)	рецидивы n (%)	первичные n (%)	рецидивы n (%)
<i>t(12;21)/ETV6::RUNX1</i>	54/98 (55,1)	0/4 (0)*	41/98 (41,8)	0/4 (0)
<i>t(1;19)/E2A::PBX1</i>	18/38 (47,4)	2/3 (66,7)	13/38 (34,2)	2/3 (66,7)
<i>t(9;22)/BCR::ABL1</i>	18/26 (69,2)	6/9 (66,7)	8/26 (30,8)	3/9 (33,3)
<i>t(11;v)/KMT2A::v</i>	11/49 (22,5)	6/15 (40)	7/49 (14,3)	5/15 (33,3)
ВП-ОЛЛ без специфических aberrаций	196/401 (48,9)	57/86 (66,3)*	95/401 (23,7)	39/86 (45,3)*
ВГ (51–65)	136/211 (64,5)	29/32 (90,6)*	55/211 (26,1)	14/32 (43,8)*
Гипоплоидия (≤ 44)	6/9 (66,7)	5/7 (71,4)	3/9 (33,3)	4/7 (57,1)
«В-other» (без ВГ и гипоплоидия (≤ 44) кариотипов)	54/181 (29,8)	23/47 (48,9)*	37/181 (20,4)	20/47 (42,6)*

Примечание. * — p -value $< 0,05$

Разделение группы ВП-ОЛЛ без специфических ХА на подгруппы с ВГ, гипоплоидным (≤ 44) и «B-other» показало, что усложнение структурного состава ($\geq 3\text{СХА}$) происходит в подгруппе «B-other» (20,4 и 42,6%, $p < 0,01$) и подгруппе с ВГ (26,1 и 43,8%, $p < 0,05$), в то время как в случаях с гипоплоидией (≤ 44) не выявлено усложнение структуры (33,3 и 57,1%, $p > 0,1$). По данным А. V. Moorman с соавторами как прогностический маркер для ВП-ОЛЛ выделяют вариант СК с любыми ≥ 5 аберрациями [19, 21]. Однако увеличение при рецидиве количества случаев с $\geq 3\text{СХА}$ и уменьшение изолированных ЧХА позволяет рассматривать именно кариотип СК $\geq 3\text{СХА}$ (без учета числовых) как потенциальный маркер прогноза для групп «B-other» и ВГ, что требует дальнейшего изучения.

В других цитогенетических группах не выявлено значимое увеличение случаев в обоих вариантах СК. При рецидиве $ETV6::RUNX1$ + ВП-ОЛЛ не выявлено ни одного случая с СК, несмотря на то, что при первичной диагностике вариант СК $\geq 3\text{СХА}$ выявлялся чаще в сравнении с другими группами ($p < 0,05$). При рецидиве $t(1;19)/E2A::PBX1$ + ВП-ОЛЛ также не наблюдалось чрезмерного увеличения количества ДХА — в двух случаях выявлено три несвязанных СХА, при этом случаи с $\geq 4\text{СХА}$ при рецидиве заболевания отсутствовали.

Моносомный кариотип

Одним из показателей сложности кариотипа является МК, который имеет неблагоприятное прогностическое значение при МДС и ОМЛ. Однако данный вариант кариотипа слабо изучен при ОЛЛ, и в основном данные получены для взрослых пациентов. В нашей работе была проведена оценка встречаемости МК при первичной диагностике и рецидиве как по отдельности, так и в сочетании с СК кариотипом. Выделены четыре группы МК/СК кариотипов определяющие сложность кариотипа: МК(-)/СК(-) — отсутствуют СК и МК; МК(-)/СК(+) — отсутствует МК, имеется СК; МК(+)/СК(-) — есть МК, отсутствует СК; МК(+)/СК(+) — в наличии СК и МК. В нашем исследовании МК определялся только в группе «B-other» как в наиболее многочисленной группе. Выявлено, что при рецидиве «B-other» увеличивалась частота встречаемости МК (в 15,5% (28/181) первичных и в 29,8%

(14/47) при рецидиве, $p = 0,024$). При этом в ~90% выявлялся МК(+)/СК(+) вариант при СК $\geq 1\text{СХА}$ ($p < 0,05$), что согласуется с данными литературы [26].

Таким образом, возможно рассматривать вариант кариотипа МК(+)/СК(+) для группы «B-other» как потенциальный маркер прогноза. Если не будет показана эффективность МК как независимого прогностического маркера, то имеет смысл определение конкретных ХА, входящих в состав изучаемой группы.

Неклональные аберрации

Единичные ХА или комплекс ХА, с не подтвержденной методом FISH клональностью, а также случайные разрывы хромосом с ацентрическим фрагментом, относили к неклональным изменениям. Анализ выявленной частоты неклональных аберраций в различных подгруппах показал отсутствие значимых различий между первичным заболеванием и рецидивом, как в целом среди ВП-ОЛЛ (9,1%, 50/547 и 12,7%, 14/115 соответственно), так и при разбивке на группы. Несмотря на отсутствие значимых различий между первичной диагностикой и рецидивом заболевания, в группах наблюдается гетерогенность. Так, при первичной диагностике наименьшее количество неклональных аберраций выявлено в группе $t(11;v)/KMT2A$ + ВП-ОЛЛ (1/49, 2%), наибольшее — с $t(12;22)/ETV6::RUNX1$ + ВП-ОЛЛ (18/98, 18,4%), а при рецидиве заболевания частота неклональных изменений в данных группах составила 13,3% (2/15) и 25% (1/4) соответственно.

Остается открытым вопрос о том, следует ли учитывать при оценке сложного кариотипа только данные одного субклона или проводить учет всех ХА, выявляемых в пуле лейкоэмических клеток, а также значение комплексных перестроек и можно ли их наличие приравнивать с сложному кариотипу. Представляет интерес оценка изменений кариотипа с точки зрения ХА не только как носителей конкретных единичных генетических изменений, но и как источника одновременного изменения экспрессии большого количества генов вследствие изменения дозы генов, поскольку ХА, выявляемые при помощи GTG метода, имеют большой размер. Дальнейшее накопление данных об изменении кариотипа при рецидиве под действием лекарственной селекции,

учет данных о субклональной структуре, наличии комплексных перестроек и использование молекулярно-цитогенетических методов для уточнения структуры маркерных хромосом позволит лучше изучить значение изолированных aberrаций (в том числе редких) и их сочетаний для исхода терапии и расчета риска развития рецидива.

Заключение

В группе рецидивов по сравнению с группой первично диагностированных ВП-ОЛЛ наблюдается изменение структуры кариотипа за счет уменьшения частоты встречаемости *ETV6::RUNX1* + ($p < 0,001$) и ВГ ($p < 0,01$) кариотипов, увеличения *KMT2A* + ($p < 0,001$), *BCR::ABL1* + ($p < 0,01$), гипоплоидии (≤ 44) ($p < 0,01$) и НГ ($p = 0,07$).

Были определены паттерны ХА для основных цитогенетических групп при первичной диагностике и рецидиве заболевания. Выявлены ХА, ассоциированные с определенными цитогенетическими группами (определяются чаще в конкретной группе при первичной диагностике и сохраняются при рецидиве заболевания): с ВГ — *dup1q, i(17)(q10)*, с гипоплоидией ≤ 44 — *t7q*, с группой «B-other» — *t7q, der12p/loss12p, t14q32, der21q/iAMP21*, с *ETV6::RUNX1* + — *del12p/ETV6* (интактная); с *BCR::ABL1* + ВП-ОЛЛ — дополнительный *der(22)t(9;22)*, а также ХА общие для большинства групп — *del6q* и *abn9/del9p*. Определены ХА, частота встречаемости которых значительно увеличивается при рецидиве заболевания: в группе «B-other» — *t1q, t19p, KMT2A* + — *t1q, abn7/del7p22, -9* и при ВГ кариотипе — *abn7/del7p22, abn9/del9p21, abn17/loss17p, i(17)(q10)*.

Частота встречаемости несбалансированных хромосомных aberrаций высока при первичной диагностике (80–90%) и не меняется при рецидиве заболевания во всех группах ВП-ОЛЛ, кроме *KMT2A* + ВП-ОЛЛ (49 и 60%) в сравнении с другими ВП-ОЛЛ ($p < 0,001$). При рецидиве заболевания выявлено изменение структуры дисбаланса за счет увеличения количества СХА ($p < 0,05$) (как в изолированном виде, так и в сочетании с ЧХА) и снижения изолированных ЧХА, во всех группах, кроме *BCR::ABL1* + и *KMT2A* + ВП-ОЛЛ. При этом наиболее сильное изменение структуры

наблюдается при рецидиве ВП-ОЛЛ без специфических транслокаций, где значительно увеличивается количество случаев с хромосомным дисбалансом по смешанному типу (несбалансированные СХА + ЧХА) (49,1 против 65,5%; $p < 0,01$). Анализ количества хромосомных aberrаций в пределах одного лейкоэмического субклона показал увеличение количества aberrаций при рецидиве заболевания во всех группах, кроме *ETV6::RUNX1* +, в том числе за счет увеличения количества случаев с ≥ 3 СХА. При этом наибольшее усложнение структуры выявлено в группе ВП-ОЛЛ без специфических транслокаций ($p < 0,05$ для обоих вариантов $СК \geq 1$ СХА и $СК \geq 3$ СХА), в частности, для подгрупп с ВГ и «B-other» ВП-ОЛЛ. В группе «B-other» также наблюдается увеличение частоты встречаемости МК ($p < 0,05$), при этом в большинстве случаев (~90%) выявлена ассоциация с СК. Таким образом, по наличию хромосомного дисбаланса, СК и неклональных изменений кариотипа были выявлены группы с наиболее (*KMT2A* +) и наименее (*ETV6::RUNX1* +) стабильными вариантами кариотипа. Так, при первичной диагностике в группе *KMT2A* + дисбаланс, $СК \geq 3$ СХА и неклональные изменения выявлены в 49, 14,3 и 2%, а при *ETV6::RUNX1* + — в 90,8, 41,8 и 18,4% соответственно.

Список использованных источников

1. The 5th edition of the World Health Organization Classification of Haematolymphoid Tumours: Lymphoid Neoplasms / R. Alaggio, C. Amador, I. Anagnostopoulos [et al.] // *Leukemia*. – 2022. – Vol. 36. – P. 1720–1748.
2. Improved Survival for Children and Adolescents With Acute Lymphoblastic Leukemia Between 1990 and 2005: A Report From the Children's Oncology Group / S. P. Hunger, X. Lu, M. Devidas [et al.] // *J. Clin. Oncol.* – 2012. – Vol. 30. – P. 1663–1669.
3. Efficacy and toxicity of dexamethasone vs methylprednisolone-long-term results in more than 1000 patients from the Russian randomized multicentric trial ALL-MB 2002 / A. Karachunskiy, J. Roumiantseva, S. Lagoiko [et al.] // *Leukemia*. – 2015. – Vol. 29, № 9. – P. 1955–1958.
4. How I treat relapsed childhood acute lymphoblastic leukemia / F. Locatelli, M. Schrappe, M. E. Bernardo, S. Rutella // *Blood*. – 2012. –

Vol. 120, № 14. – P. 2807–2816.

5. Достижения в терапии острого лимфобластного лейкоза у молодых взрослых Республики Беларусь / Н. В. Мигаль, О. И. Быданов, А. С. Лелей [и др.] // Гематология. Трансфузиология. Восточная Европа. – 2017. – Т. 3, № 3. – С. 352–364.

6. Результаты лечения рецидива острого лимфобластного лейкоза у детей и молодых взрослых в Республике Беларусь с 2005 по 2017 г. / О. В. Карась, Н. В. Мигаль, О. В. Алейникова // Гематология. Трансфузиология. Восточная Европа. – 2017. – Т. 3, № 3. – С. 329–340.

7. Genomic determinants of outcome in acute lymphoblastic leukemia / T. C. Chang, W. Chen, C. Qu [et al.] // J. Clin. Oncol. – 2024. – Vol. 42, № 29. – P. 3491–3503.

8. Prognostic effect of chromosomal abnormalities in childhood B-cell precursor acute lymphoblastic leukaemia: results from the UK Medical Research Council ALL97/99 randomised trial / A. V. Moorman, H. M. Ensor, S. M. Richards [et al.] // Lancet Oncol. – 2010. – Vol. 11, № 5. – P. 429–438.

9. Deregulated Expression of Cytokine Receptor Gene, CRLF2, Is Involved in Lymphoid Transformation in B-cell Precursor Acute Lymphoblastic Leukemia / L. J. Russell, M. Capasso, I. Vater [et al.] // Blood. – 2009. – Vol. 114, № 13. – P. 2688–2698.

10. Rearrangement of CRLF2 Is Associated With Mutation of JAK Kinases, Alteration of IKZF1, Hispanic/Latino Ethnicity, and a Poor Outcome in Pediatric B-progenitor Acute Lymphoblastic Leukemia / R. C. Harvey, C. G. Mullighan, I. M. Chen [et al.] // Blood. – 2010. – Vol. 115, № 26. – P. 5312–5321.

11. Molecular Characterization of Acute Lymphoblastic Leukemia With High CRLF2 Gene Expression in Childhood / J. Schmäh, B. Fedders, R. Panzer-Grümayer [et al.] // Ped. Blood Cancer. – 2017. – Vol. 64, № 10. – P. e26539. doi: 10.1002/pbc.26539

12. Integrated Genomic Analysis of Chromosomal Alterations and Mutations in B-Cell Acute Lymphoblastic Leukemia Reveals Distinct Genetic Profiles at Relapse. / M. Forero-Castro, A. Montaña, C. Robledo [et al.] // Diagnostics. – 2020. – Vol. 10, № 7. – P. 455. doi: 10.3390/diagnostics10070455.

13. Zhang, W. Prognostic significance of CD-

KN2A/B deletions in acute lymphoblastic leukaemia: a meta-analysis / W. Zhang, P. Kuang, T. Liu // Ann. Med. – 2019. – Vol. 51, № 1. – P. 28–40.

14. TCF3 gene rearrangements in pediatric B-cell acute lymphoblastic leukemia – a single center experience / E. Zerkalenkova, Y. Menchits, A. Borkovskaia [et al.] // Int. J. Lab. Hematol. – 2023. – Vol. 45, № 4. – P. 533–540.

15. Inaba, H. Pediatric acute lymphoblastic leukemia / H. Inaba, C. G. Mullighan // Haematologica. – 2020. – Vol. 105, № 11. – P. 2524–2539.

16. Paulsson, K. High hyperdiploid childhood acute lymphoblastic leukemia / K. Paulsson, B. Johansson // Genes Chromosomes Cancer. – 2009. – Vol. 48, № 8. – P. 637–660.

17. Long-term results of the AIEOP-ALL-95 Trial for Childhood Acute Lymphoblastic Leukemia: insight on the prognostic value of DNA index in the framework of Berlin–Frankfurt–Muenster based chemotherapy / M. Arico, M. G. Valsecchi, C. Rizzari [et al.] // J. Clin. Oncol. – 2008. – Vol. 26, № 2. – P. 283–289.

18. Relapsed childhood high hyperdiploid acute lymphoblastic leukemia: presence of preleukemic ancestral clones and the secondary nature of microdeletions and RTK-RAS mutations / J. Davidsson, K. Paulsson, D. Lindgren [et al.] // Leukemia. – 2010. – Vol. 24. – P. 924–931.

19. Outcome heterogeneity in childhood high-hyperdiploid acute lymphoblastic leukemia / A. V. Moorman, S. M. Richards, M. Martineau [et al.] // Blood 2003. – Vol. 102, № 8. – P. 2756–2762.

20. Prognostic impact of karyotypic findings in childhood acute lymphoblastic leukaemia: a Nordic series comparing two treatment periods / E. Forestier, B. Johansson, G. Gustafsson [et al.] // Br. J. Haematol. – 2000. – Vol. 110, № 1. – P. 147–153.

21. Adult Leukaemia Working Party, Medical Research Council/National Cancer Research Institute. Karyotype is an independent prognostic factor in adult acute lymphoblastic leukemia (ALL): analysis of cytogenetic data from patients treated on the Medical Research Council (MRC) UKALLXII/Eastern Cooperative Oncology Group (ECOG) 2993 trial / A. V. Moorman, C. J. Harrison, G. A. Buck [et al.] // Blood. – 2007. – Vol. 109, № 8. – P. 3189–3197.

22. In adults with standard-risk acute lymphoblastic leukemia, the greatest benefit is achieved

from a matched sibling allogeneic transplantation in first complete remission, and an autologous transplantation is less effective than conventional consolidation/maintenance chemotherapy in all patients: final results of the International ALL Trial (MRC UKALL XII/ECOG E2993) / A. H. Goldstone, S. M. Richards, H. M. Lazarus [et al.] // *Blood*. – 2008. – Vol. 111, № 4. – P. 1827–1833.

23. Impact of cytogenetics on the outcome of adult acute lymphoblastic leukemia: results of Southwest Oncology Group 9400 study / V. Pullarkat, M. L. Slovak, K. J. Kopecky [et al.] // *Blood*. – 2008. – Vol. 111, № 5. – P. 2563–2572.

24. PETHEMA Group, Spanish Society of Hematology. Prognostic significance of complex karyotype and monosomal karyotype in adult patients with acute lymphoblastic leukemia treated with risk-adapted protocols / C. Motlló, J. M. Ribera, M. Morgades [et al.] // *Cancer*. – 2014. – Vol. 120, № 24. – P. 3958–3964.

25. Acute Leukemia Committee of the CIBMTR. Impact of cytogenetic abnormalities on outcomes of adult Philadelphia-negative acute lymphoblastic leukemia after allogeneic hematopoietic stem cell transplantation: a study by the Acute Leukemia Working Committee of the Center for International Blood and Marrow Transplant Research / A. Lazaryan, M. Dolan, M. J. Zhang [et al.] // *Haematologica*. – 2020. – Vol. 105, № 5. – P. 1329–1338.

26. Monosomal karyotype in Philadelphia chromosome-negative acute lymphoblastic leukemia / S. S. Kenderian, A. Al-Kali, N. Gangat [et al.] // *Blood Cancer J*. – 2013. – Vol. 3, № 7. – P. e122. doi: 10.1038/bcj.2013.21.

27. McGowan-Jordan, J. ISCN 2020: An international system for human cytogenomic nomenclature (2020) (Cytogenetic and Genome Research) / eds. J. McGowan-Jordan, R. J. Hastings, S. Moore : S Karger Ag, 2020. – 170 p.

A. V. Valochnik¹, D. R. Kapuza¹, I. V. Pakhomava¹, V. S. Vshyukova¹, T. V. Shman²

KARYOTYPE CHANGES OF LEUKEMIC CELLS IN CHILDREN AND YOUNG ADULTS WITH RELAPSED B-CELL PRECURSOR ACUTE LYMPHOBLASTIC LEUKEMIA

¹Belarussian Research Center for Pediatric Oncology, Hematology and Immunology
43 Frunzenskaya St., v. Borovlyany 223053, Minsk region, the Republic of Belarus
e-mail: valochnikalena@gmail.com

²Minsk Scientific and Practical Center for Surgery, Transplantology and Hematology
8 Semashko St., Minsk 220087, the Republic of Belarus

In relapsed B-cell precursor acute lymphoblastic leukemia (BCP-ALL), compared to the primary diagnosed BCP-ALL, changes in the karyotype structure were observed due to both a decreased frequency of favourable *ETV6::RUNX1* + ($p < 0.001$) and high hyperdiploid ($p < 0.01$) karyotypes and an increased frequency of unfavourable karyotypes such as *KMT2A* + ($p < 0.001$), *BCR::ABL1* + ($p < 0.01$), and hypodiploidy (≤ 44) ($p < 0.01$). In addition to that, the relapsed BCP-ALL without recurrent chromosomal aberrations shows increased frequencies of low hyperdiploid ($p = 0.07$), complex ($p < 0.05$) and monosomal ($p < 0.05$) karyotypes, and notably, the monosomal karyotype goes with a complex karyotype in ~90% of cases at that. It was shown that with the relapsed disease, no increase in chromosomal imbalance cases is observed; however, a change in the karyotype structure occurs due to an increase in the number of structural aberrations ($p < 0.05$). The frequency of complex karyotypes with ≥ 3 structural aberrations also increases in the “B-other” ($p < 0.01$) and high hyperdiploidy ($p < 0.05$) groups, while no karyotype complexity progression occurs in the hypodiploid (≤ 44) group. Additional chromosomal aberrations (predominantly unbalanced) associated with cytogenetic groups at a primary diagnosis and a relapse were identified. With regard to the presence of chromosomal imbalance, a complex karyotype and non-clonal changes, the most stable karyotype variant *KMT2A* + and the least stable *ETV6::RUNX1* + BCP-ALL were identified.

Keywords: acute lymphoblastic leukemia, relapse, additional chromosomal aberrations, complex karyotype, monosomal karyotype, unbalanced aberrations.

Дата поступления в редакцию: 24 июня 2025 г.

Е. В. Снытков, М. Д. Амелянович

ПОИСК ПОЛИМОРФНЫХ ЛОКУСОВ В ДИФФЕРЕНЦИАЛЬНО ЭКСПРЕССИРУЕМЫХ ГЕНАХ, АССОЦИИРОВАННЫХ С РАЗВИТИЕМ ТУБЕРКУЛЕЗА

Государственное научное учреждение
«Институт генетики и цитологии Национальной академии наук Беларуси»
Республика Беларусь, 220072, г. Минск, ул. Академическая, 27
e-mail: evsnytkov@gmail.com

Статья представляет собой исследование, направленное на поиск генетических маркеров, ассоциированных с развитием туберкулеза. Авторами использованы данные РНК-секвенирования (RNA-seq) из публичного репозитория GEO (проект GSE222001), проведен биоинформатический анализ экспрессии генов и последующий анализ геномных вариантов (SNVs) в значимо дифференциально экспрессируемых генах. Цель исследования — выявить генетические маркеры, ассоциированные с развитием туберкулезной инфекции, для совершенствования диагностики, прогноза и лечения. В ходе выполнения работы выявлено, что позиции Chr14:65051895 и Chr14:65051897 гена *MAX*, Chr14:65002523 гена *FNTB*, а также Chr14:35763984, Chr14:35763994, Chr14:35928234, Chr14:35803547, Chr14:35788303, Chr14:35817409, Chr14:35761582, Chr14:35761594, Chr14:35802733, Chr14:35928244, Chr14:35819153, Chr14:35876309, Chr14:35893835, Chr14:35793522, Chr14:35803537 и Chr14:35833886 гена *PHF20* связаны с развитием туберкулеза.

Ключевые слова: туберкулез, *Mycobacterium tuberculosis*, генотипирование *in silico*, транскриптом, иРНК, bcftools, RNA-seq.

Для цитирования: Снытков, Е. В. Поиск полиморфных локусов в дифференциально экспрессируемых генах, ассоциированных с развитием туберкулеза / Е. В. Снытков, М. Д. Амелянович // Молекулярная и прикладная генетика: сб. науч. тр. / Ин-т генетики и цитологии НАН Беларуси; редкол.: А. В. Кильчевский (гл. ред.) [и др.]. — Минск, 2025. — Т. 39. — С. 100–111.

Введение

Наряду с положительными изменениями эпидемиологических показателей и снижением заболеваемости туберкулезом, в том числе и в Республике Беларусь, которое наблюдается в последние годы, сохраняется влияние на динамику эпидемиологического процесса и характер течения туберкулеза (ТБ) ряда неблагоприятных тенденций и факторов. К ним относятся: увеличение числа ВИЧ-инфицированных лиц; широкое применение иммуносупрессивных, в том числе генно-инженерных биологических препаратов; повышение уровня миграции населения; пандемия коронавирусной инфекции COVID-19 и др. Основные усилия в настоящее время сосредоточены на раннем выявлении активного ТБ современными молекулярно-генетическими, иммунологическими, лучевыми метода-

ми и контролируемой химиотерапии с учетом лекарственной устойчивости возбудителя. Кроме того, ключевым решением для достижения целевых показателей ВОЗ (всемирная организация здравоохранения) по ликвидации туберкулеза является совершенствование диагностики и лечения латентной туберкулезной инфекции (ЛТИ) — резервуара будущего туберкулеза.

ЛТИ определяется как состояние стойкого иммунного ответа макроорганизма на антигены микобактерий туберкулеза (МБТ) при отсутствии клинических проявлений активного туберкулеза [1]. При инфицировании или латентной инфекции МБТ находятся в макроорганизме в покоящемся, или «дормантном», состоянии, которое может сохраняться таковым всю жизнь либо перейти в клинически манифестное заболевание при размножающихся

МБТ и ослаблении иммунных реакций [2].

Материалы и методы

С целью выявления генетических маркеров, потенциально ассоциированных с развитием туберкулеза, были использованы данные, размещенные в свободном доступе в базе Gene Expression Omnibus (GEO) — общедоступный репозиторий данных функциональной геномики, поддерживающий отправку данных в соответствии с набором руководящих принципов «Минимальная информация об эксперименте с микрочипами» (Minimum Information About a Microarray Experiment, MIAME). В качестве основы для дальнейшего исследования была выбрана работа автора Cui Hua Liu «Transcriptional profiling of NK cells from tuberculosis patients» (GSE222001), в которой определялись профили экспрессии с помощью высокопроизводительного секвенирования [3]. В исследовании GSE222001 авторами был использован биологический материал 5 здоровых лиц (образцы NC-rep1, NC-rep2, NC-rep3, NC-rep4, NC-rep5), 5 лиц с ЛТИ (образцы LTBI-rep1, LTBI-rep2, LTBI-rep3, LTBI-rep4, LTBI-rep5) и 5 лиц с активной формой туберкулезной инфекции (образцы TB-rep1, TB-rep2, TB-rep3, TB-rep4, TB-rep5). Все участники не имели других значительных заболеваний в ме-

дицинском анамнезе, таких как инфекция вируса иммунодефицита человека (ВИЧ), рак и диабет. Пациенты с активной и латентной формой туберкулеза имели положительный тест T-SPOT.TB. Лица с активной и латентной формами туберкулеза были набраны в Пекинском торакальном госпитале — специализированной туберкулезной больнице в Китае. Все здоровые контрольные лица были набраны в Институте микробиологии Китайской академии наук [4]. Полученные в результате секвенирования транскриптомные данные были нормализованы методом TPM (Transcripts Per Million), так как он, в отличие от RPKM (Reads Per Kilobase per Million), осуществляет нормализацию данных сначала по длине транскрипта, а затем — по общему числу транскриптов, из-за чего сумма всех TPM в образце равна 1 млн, что облегчает сравнение между образцами и обуславливает корректность такого сравнения в случае, если суммарная экспрессия образцов сильно различается. Информация об исходных данных, входящих в исследование GSE222001, представлена в таблице 1.

Для каждого гена, данные об экспрессии которого были определены в исследовании, нами осуществлено сравнение средних значений экспрессии между группами «активный

Таблица 1

Информация об исходных данных, входящих в исследование GSE222001

Проект	Эксперимент	Образец	Пол	Возраст
PRJNA915785	SRX18861724	LTBI-rep5	М	54
	SRX18861723	LTBI-rep4	Ж	29
	SRX18861722	LTBI-rep3	М	31
	SRX18861721	LTBI-rep2	Ж	41
	SRX18861720	LTBI-rep1	М	20
	SRX18861719	TB-rep5	М	33
	SRX18861718	TB-rep4	М	18
	SRX18861717	TB-rep3	Ж	55
	SRX18861711	TB-rep2	Ж	48
	SRX18861710	TB-rep1	М	29
	SRX18861716	NC-rep5	М	53
	SRX18861715	NC-rep4	Ж	53
	SRX18861714	NC-rep3	М	24
	SRX18861713	NC-rep2	Ж	21
	SRX18861712	NC-rep1	М	23

туберкулез» и «контрольная группа» с помощью U-критерия Манна – Уитни. Поправка на множественность сравнений не вводилась, вместо этого при расчете р-уровня использовали точный критерий, реализованный в SPSS v.20. Для локусов, для которых точный р-уровень по результатам U-критерия Манна – Уитни оказался менее 0,01, дополнительно провели сравнение уровней экспрессии с использованием критерия ANOVA. Для дальнейшего анализа использовали только те локусы, для которых р-уровень (ANOVA) был $<0,01$, и для критерия Левена $p > 0,05$. Для каждого гена, показавшего статистически значимые различия при сравнении групп по уровню экспрессии, был создан файл с расширением *.fa, который использовался при запросах к sra-blast в качестве референса (также дополнительно взяты в исследуемую последовательность гена относительно небольшие участки на 3'- и 5'-концах последовательности ДНК). Для каждого результата секвенирования, приведенного в таблице 1 (колонка «Эксперимент»), был создан файл с расширением *.sam, содержащий прочтения, входящие в референсную последовательность. В дальнейшем с использованием команды mpileup программного обеспечения bcftools (<https://github.com/samtools/bcftools>) на основе полученных нами sam-файлов были созданы файлы с расширением *.vcf, которые включали в себя позиции, для которых была выявлена вариабельность по отношению к референсной последовательности, за исключением инсерций и делеций.

На следующем этапе работы было проведено исследование генотипического состава в позициях наиболее перспективных генов с использованием результатов полногеномного секвенирования, не вошедших в предыдущий этап, отобранных в базе SRA (Sequence Read Archive). Для этого анализа использованы 99 экспериментов:

Австралия: SRX28263276, SRX28263277;

Дания: ERX14408767, ERX14408768;

Швеция: ERX14344449, ERX14344450, ERX14344469, ERX14344470, SRX23416714, SRX23416715, SRX23416716, SRX23416717, SRX23416718, SRX23416719, SRX23416720, SRX23416721, SRX23416722, SRX23416723;

Европа: ERX14286624, ERX14286625, ERX14286626, ERX14286627, ERX14286628,

ERX14286629, ERX14286631, ERX14286632, ERX14286633, ERX14286634, ERX14286635, ERX14286636, ERX14286637, ERX14286638, ERX14286639;

Китай: SRX24404586, SRX24404587, SRX24404588, SRX24404589, SRX24404590, SRX24404591, SRX24404592, SRX24404593, SRX24404594, SRX24404595, SRX24404596, SRX24404597, SRX24404598, SRX24404599, SRX24503353, SRX24503354, SRX24503355, SRX24503356, SRX24503357, SRX24503358, SRX26366571, SRX27161711, SRX27161715, SRX27161717, SRX27161718, SRX27161719, SRX27161723, SRX27161727, SRX27161728, SRX27161730, SRX27185158, SRX27185160, SRX27185161, SRX27283030, SRX27283031, SRX27283032, SRX27283033;

Япония: DRX206987, DRX206988, DRX206989, DRX206990, DRX413736, DRX413737, DRX413738, DRX424865, DRX424866, DRX424867, DRX424868, DRX566955, DRX566956, DRX566957, DRX571587, DRX571588, DRX571589, DRX578365, DRX578366, DRX578367, DRX578368, DRX578369, DRX578370, DRX616162, DRX616163, DRX616164, DRX616165, DRX616166, DRX616167.

Эта же информация в дальнейшем использовалась при поиске позиций, имеющих одинаковый генотипический паттерн и которые потенциально входят в одну группу сцепления.

Результаты и обсуждения

В ходе применения U-критерия Манна – Уитни и ANOVA при сравнении средних значений экспрессии генов между группами «активный туберкулез» и «контрольная группа» нами были определены гены, показывающие наибольшие статистически значимые различия. Отобранные таким образом гены приведены в таблице 2.

В результате сравнения генотипического состава вариабельных позиций между исследуемыми группами было выявлено, что в паре сравнения «активная форма туберкулеза / контрольная группа» для генетических вариантов генов *HBPI*, *LRRC27*, *SLPI*, *HPS6*, *RBM15B* и псевдогена *ENSG00000266155* статистически значимые различия отсутствуют. Для гена *MAX* были выявлены 3 позиции (Chr14:65051897, Chr14:65051895

Таблица 2

Список генов, показавших статистически значимые различия при сравнении групп «активный туберкулез» и «контроль»

Ensembl ID	Ген	Локализация*	Среднее АТ**	Среднее НК***	p-уровень#
ENSG00000148814	<i>LRRC27</i>	Chr10:132330063-132381508	5,97	3,74	0,009
ENSG00000105856	<i>HBP1</i>	Chr7:107169003-107202522	26,85	36,74	0,009
ENSG00000124107	<i>SLPI</i>	Chr20:45252239-45254564	1,52	0,13	0,008
ENSG00000025293	<i>PHF20</i>	Chr20:35772015-35950370	42,93	60,13	0,009
ENSG00000125952	<i>MAX</i>	Chr14:65006101-65102695	75,89	100,12	0,009
ENSG00000266155	—	Chr17:64936485-64940229	0,88	0,0	0,005
ENSG00000166189	<i>HPS6</i>	Chr10:102065349-102068036	47,54	57,79	0,009
ENSG00000259956	<i>RBM15B</i>	Chr3:51391285-51397908	64,02	69,73	0,009
ENSG00000257365	<i>FNTB</i>	Chr14:64986895-65062650	2,73	5,04	0,028

Примечание. * — позиции представлены в соответствии с версией геномной сборки человека GRCh38.p14; ** — среднее значение экспрессии в группе лиц с активной формой туберкулеза; *** — среднее значение экспрессии в группе лиц, вошедших в контрольную группу; # — U-критерия Манна – Уитни (двухсторонняя асимптотическая значимость)

и Chr14:65002523), генотипический состав которых статистически значимо различается в группах. Стоит отметить, что позиции Chr14:65051897 и Chr14:65051895, по данным RefSeq, локализуются в области генов *MAX* (антисмысловая цепь) и *FNTB*, в то же время позиция Chr14:65002523 выходит за пределы гена *MAX* и локализуется в области гена *FNTB*, находясь от исследуемого гена *MAX* на расстоянии в 3578 нуклеотидов, и далее она будет рассматриваться как часть гена *FNTB* (это связано с тем, что позиция Chr14:65002523 попала в дополнительный участок, включенный в референс при его формировании). При сравнении уровней экспрессии по этому гену также были выявлены статистически значимые различия, а средний уровень экспрессии в группе с активным туберкулезом и контрольной группе составил 2,73 и 5,04 соответственно. Согласно аннотации dbSNP, позиции Chr14:65051897 и Chr14:65051895 имеют аннотированные SNV. Так, в позиции Chr14:65051897 находится rs2062622202, в котором отмечается возможная замена С > Т, однако частота минорного аллеля Т в ряде исследований составила от 0,00003 до 0,000013 [3], в то же время нами в исследовании было выявлено, что в группе здоровых пациентов генотип G/T наблюдается в 100%

случаев (4 образца), а в группе с активным туберкулезом генотип G/G — в 80% случаев и G/T — в 20% случаев (4:1). В группе с латентным туберкулезом соотношение генотипов G/G к G/T составило 40 к 60% (2:3).

В позиции Chr14:65051895 аннотирован rs143210184 с возможной заменой С > А / С > Т, при этом частота минорного аллеля А была указана как 0, частота минорного аллеля Т — от 0,0003 до 0,0202 [5]. При этом нами выявлены следующие генотипы: в контрольной группе G/A — 100% (4 образца), в группе лиц с активным туберкулезом G/G — 100% (5 образцов), в группе с латентной формой туберкулеза G/G — 60% (3 образца) и G/A — 40% (2 образца).

Для позиции Chr14:65002523 отсутствует информация о наличии какого-либо SNV (табл. 3).

По итогу проведенной нами работы были определены позиции, потенциально связанные с развитием туберкулеза. Так, при анализе генотипического состава по генам, имеющим статистически значимо отличающийся уровень экспрессии у людей, больных туберкулезом, а также здоровых лиц, было выявлено, что позиции Chr14:65051895 и Chr14:65051897 гена *MAX*, а также Chr14:65002523 гена *FNTB*, имеют статистически значимо различающийся-

Таблица 3

Ожидаемые в соответствии с аннотацией dbSNP популяционные, а также полученные нами в ходе изучения генов *MAX* и *FNTB* частоты аллелей для всех групп

Ген	Позиция/ ОНП	Аллель	Популяционные частоты [6]	Аллель	Полученные частоты [ДИ]*
<i>MAX, FNTB</i>	Chr14:65051895/ rs143210184	C	0,9798	G	0,786 [0,63, 0,94]
		T	0,0202	A	0,214 [0,06, 0,37]
<i>MAX, FNTB</i>	Chr14:65051897/ rs2062622202	C	0,99997	G	0,714 [0,55, 0,88]
		T	0,00003	T	0,286 [0,12, 0,45]
<i>FNTB</i>	Chr14:65002523/ —	—	—	G	0,192 [0,04, 0,34]
		—	—	C	0,808 [0,66, 0,96]

Примечание. * — доверительный интервал

ся генотипический состав. При этом можно заметить, что в образцах контрольной группы для этих позиций характерно наличие гетерозигот, для группы лиц с активным туберкулезом — гомозигот (за исключением одного из образцов позиции Chr14:65051895), для латентной формы туберкулеза характерен смешанный генотипический состав (табл. 4).

При расширении исследования нами было выявлено, что позиция Chr14:65051897 не имеет вариабельности, т. е. в рандомизированной выборке, состоящей из 99 результатов полногеномного секвенирования ДНК, генотип по этой позиции составил G/G. Для позиции Chr14:65051895 была выявлена вариабельность, более того, также были выяв-

Таблица 4

Распределение генотипов при исследовании транскриптома генов *MAX* и *FNTB*

Группа	Chr14:65051897*	Chr14:65051895**	Chr14:65002523***
Контрольная	—	—	—
	G/A	G/T	G/C
	G/A	G/T	G/C
	G/A	G/T	G/C
	G/A	G/T	—
Активный туберкулез	G/G	G/G	C/C
	G/G	G/G	C/C
	G/G	G/G	C/C
	G/G	G/G	C/C
	G/G	G/T	C/C
Латентный туберкулез	G/A	G/T	G/C
	G/G	G/G	C/C
	G/G	G/G	C/C
	G/A	G/T	G/C
	G/G	G/T	C/C

Примечание. * — $p = 0,008$; ** — $p = 0,018$; *** — $p = 0,048$

лены позиции с одинаковым генотипическим паттерном по отношению к этой позиции: Chr14:65051955 (*FNTB*, *MAX*, rs996248162), Chr14:64999363 (*FNTB*) и Chr14:65037848 (*FNTB*, *MAX*, rs2062248289 — делеция участка в 12 нуклеотидов CCAGCTCACTGC/C), которые расположены от исходной позиции на 60, 52 532 и 14 047 нуклеотидов соответственно.

Для позиции Chr14:65002523 также были выявлены позиции с одинаковым генотипическим паттерном: Chr14:65002526 (*FNTB*) и Chr14:64999355 (*FNTB*), отстоящие от целевой позиции на 3 и 3168 нуклеотидов соответственно. Предположительно, данные позиции могут находиться в одной группе сцепления. Все отмеченные позиции находятся в зонах, которым соответствуют зоны активного сплайсинга в молекулах иРНК (RNA-seq intron-spanning reads, aggregate (filtered), NCBI Homo sapiens Annotation Release 110) [7] и, таким образом, вероятно, влияют на экспрессию соответствующих генов.

Из приведенных выше данных видно, что гены *MAX* и *FNTB* демонстрируют согласованное наследование определенных аллелей в некоторых позициях, а так как эти гены наследуются вместе, то их совместное влияние на иммунный ответ может быть сильнее, чем по отдельности.

Ген *MAX* (ассоциированный с МҮС фактор Х) — это ген, кодирующий белок, который относится к семейству факторов транскрипции, тем самым регулируя экспрессию генов, участвующих в пролиферации, апоптозе клеток. На текущий момент имеются данные, что МҮС, с которым связывается белковый продукт гена *MAX*, является ключевым регулятором многих фундаментальных клеточных процессов в иммунных клетках, в первую очередь метаболических программ. Показано, что он связывается с промоторной областью широкого спектра активных генов для дальнейшего стимулирования их экспрессии, идентифицируя МҮС как глобальный транскрипционный усилитель [8, 9]. МҮС, являясь важнейшим фактором транскрипции, регулирует экспрессию многочисленных генов как в клетках врожденного, так и адаптивного иммунитета, управляя их активацией, пролиферацией, поляризацией и последующими функциональными событиями, вызываемыми

этими клетками [10].

Ген *FNTB* обеспечивает активность активатора ацетилтрансферазы; связывание ферментов, а также связывание ионов цинка. Способствует активности фарнезилтрансферазы белка. Участвует в фарнезилировании белка и регуляции движения микротрубочек. Входит в состав комплекса, ассоциированного с микротрубочками, и комплекса фарнезилтрансферазы белка. Показано, что данный ген демонстрирует уникальные паттерны сплайсинга в макрофагах, инфицированных штаммом *Mycobacterium tuberculosis* H37Rv, по сравнению с неинфицированным контролем [11]. Кроме того, отмечено повышение экспрессии этого гена при бактериальном гломерулонефрите [12], что указывает на его участие в иммунном ответе на бактериальные инфекции.

При исследовании гена *PHF20* были выявлены 16 позиций, генотипический состав которых статистически различается для контрольной группы и группы пациентов с активным туберкулезом: Chr20:35763984, Chr20:35763994, Chr20:35928234, Chr20:35803547, Chr20:35788303, Chr20:35817409, Chr20:35761582, Chr20:35761594, Chr20:35802733, Chr20:35928244, Chr20:35819153, Chr20:35876309, Chr20:35893835, Chr20:35793522, Chr20:35803537 и Chr20:35833886. В таблице 5 приведены данные по частотам аллелей для обозначенных выше хромосомных позиций.

В таблице 6 приведены генотипы 16 позиций для каждого из образцов, изученных в рамках исследования транскриптома.

При расширенном исследовании на 79 результатах полногеномного секвенирования не было выявлено позиций, которые имели бы аналогичный представленным позициям генотипический паттерн.

Ген *PHF20* обеспечивает связывание ДНК и ионов цинка. Предполагается, что он участвует в регуляции транскрипции РНК-полимеразой II. Показано, что уровень экспрессии этого гена связан с инфицированием клеток устойчивым к аутофагии штаммом *Mycobacterium tuberculosis* [13]. Кроме того, имеются данные о том, что этот ген задействован в трансактивации NFκB при остром воспалении [14].

Таблица 5

Наблюдаемые и ожидаемые в соответствии с аннотацией dbSNP частоты при изучении гена *MAX*

Ген	Позиция / ОНП	Аллель	Популяционные частоты [6]	Аллель	Полученные частоты [ДИ]*
	Chr20:35763984 / rs2515846267	A	0,99934	A	0,318 [0,12, 0,51]
		C	0,00066	C	0,682 [0,49, 0,88]
	Chr20:35763994 / rs2515846275	T	0,99979	T	0,286 [0,12, 0,45]
		A	0,00021	A	0,714 [0,55, 0,88]
<i>PHF20</i>	Chr20:35928234	—	—	T	0,714 [0,55, 0,88]
		—	—	A	0,286 [0,12, 0,45]
<i>PHF20</i>	Chr20:35803547 / rs1452695636	A	0,999985	A	0,733 [0,57, 0,89]
		G	0,000015	T	0,267 [0,11, 0,43]
<i>PHF20</i>	Chr20:35788303 / rs2515867267	A	0,99996	A	0,682 [0,49, 0,88]
		G	0,00004	G	0,318 [0,12, 0,51]
<i>PHF20</i>	Chr20:35817409	—	—	T	0,308 [0,13, 0,49]
		—	—	A	0,692 [0,51, 0,87]
	Chr20:35761582	—	—	T	0,731 [0,56, 0,9]
		—	—	C	0,269 [0,1, 0,44]
	Chr20:35761594 / rs1360384092	A	0,999993	A	0,269 [0,1, 0,441]
		T	0,000007	T	0,731 [0,56, 0,9]
<i>PHF20</i>	Chr20:35802733	—	—	T	0,731 [0,56, 0,9]
		—	—	A	0,269 [0,1, 0,44]
<i>PHF20</i>	Chr20:35928244	—	—	G	0,75 [0,58, 0,92]
		—	—	T	0,25 [0,08, 0,42]
<i>PHF20</i>	Chr20:35819153 / rs2042125567	A	—	A	0,75 [0,59, 0,91]
		C	—	T	0,25 [0,09, 0,41]
<i>PHF20</i>	Chr20:35876309 / rs1792401924	G	0,999993	G	0,75 [0,58, 0,92]
		A	0,000007	A	0,25 [0,08, 0,42]
<i>PHF20</i>	Chr20:35893835	—	—	T	0,75 [0,58, 0,92]
		—	—	C	0,25 [0,08, 0,42]
<i>PHF20</i>	Chr20:35793522 / rs1445047045	A	1	A	0,767 [0,62, 0,92]
		T	0	T	0,233 [0,08, 0,38]
<i>PHF20</i>	Chr20:35803537	—	—	C	0,786 [0,62, 0,96]
		—	—	A	0,214 [0,04, 0,38]
<i>PHF20</i>	Chr20:35833886 / rs1348172939	C	0,99994	C	0,727 [0,54, 0,91]
		T	0,00006	T	0,273 [0,09, 0,46]

Примечание. * — доверительный интервал

Также при расширенном изучении выборки случайным образом отобранных результатов полногеномного секвенирования нами были определены частоты аллелей, которые приведены в таблице 7.

По приведенным в таблице 7 данным вид-

но, что ряд полиморфных локусов по частотам аллелей, входящих в их состав, значительно отличается от популяционных данных, приведенных в таблицах 3 и 4. Так, частоты аллелей в позициях Chr20:35763984, Chr20:35763994, Chr20:35803547,

Таблица 6

Распределение генотипов при исследовании генотипического состава гена *PHF20*

Группа#	Chr14:35763984*	Chr14:35763994*	Chr14:35928234*	Chr14:35803547*	Chr14:35788303**	Chr14:35817409**	Chr14:35761582***	Chr14:35761594***
НК	–	–	T/T	A/A	–	–	–	–
	C/C	A/A	T/T	A/A	A/G	A/A	T/T	T/T
	C/C	A/A	–	A/A	G/G	A/A	T/T	T/T
	C/C	A/A	T/T	A/A	A/G	A/A	T/T	T/T
	C/C	A/A	T/T	A/A	A/G	A/A	T/T	T/T
АТ	A/C	A/T	A/T	A/T	–	–	–	–
	A/C	A/T	A/T	A/T	A/A	T/T	T/C	A/T
	A/C	A/T	A/T	A/T	A/A	A/T	T/C	A/T
	A/C	A/T	A/T	A/T	A/A	A/T	T/C	A/T
	A/C	A/T	A/T	A/T	A/A	A/T	T/C	A/T
ЛТИ	–	A/A	T/T	A/A	–	A/A	T/T	T/T
	A/C	A/T	A/T	A/T	A/G	A/T	T/C	A/T
	A/C	A/T	A/T	A/T	A/G	A/T	T/C	A/T
	–	A/A	T/T	A/A	–	A/A	T/T	T/T
	–	A/T	A/T	A/T	A/A	A/T	T/C	A/T

Группа#	Chr14:35802733***	Chr14:35928244***	Chr14:35819153****	Chr14:35876309****	Chr14:35893835****	Chr14:35793522****	Chr14:35803537****	Chr14:35833886****
НК	–	G/G	–	A/G	–	A/A	C/C	T/C
	T/T	G/G	A/A	A/G	T/C	A/A	C/C	T/C
	T/T	–	A/A	A/G	T/C	A/A	C/C	T/C
	T/T	G/G	A/A	A/G	T/C	A/A	C/C	T/C
	T/T	G/G	A/A	G/G	T/C	A/A	C/C	T/C
АТ	–	G/T	A/T	–	T/T	A/A	A/C	C/C
	A/T	G/T	A/T	G/G	T/C	A/T	A/C	C/C
	A/T	G/T	A/T	G/G	T/T	A/T	A/C	T/C
	A/T	G/T	A/A	G/G	T/T	A/T	A/C	C/C
	A/T	–	A/T	G/G	T/T	A/T	C/C	C/C

Окончание таблицы 6

Группа#	Chr14:35802733***	Chr14:35928244***	Chr14:35819153****	Chr14:35876309****	Chr14:35893835****	Chr14:35793522****	Chr14:35803537****	Chr14:35833886****
ЛТИ	T/T	G/G	A/A	–	T/T	A/A	C/C	–
	A/T	G/T	A/T	G/G	T/T	A/T	A/C	–
	A/T	G/T	A/T	A/G	T/C	A/T	A/C	C/C
	T/T	G/G	A/A	A/G	–	A/A	C/C	–
	A/T	–	A/T	–	–	A/T	–	–

Примечание. # — НК — контрольная группа, АТ — активный туберкулез, ЛТИ — латентный туберкулез.

* — $p = 0,008$; ** — $p = 0,018$; *** — $p = 0,029$; **** — $p = 0,048$

Таблица 7

Распределение генотипов при исследовании расширенной выборки

Позиция / ОНП	Аллель	Частоты в расширенной выборке [ДИ]	Позиция / ОНП	Аллель	Частоты в расширенной выборке [ДИ]
Chr14:65051895 / rs143210184	G	0,905 [0,86, 0,95]	Chr20:35761582	T	0,608 [0,52, 0,7]
	A	0,095 [0,05, 0,14]		C	0,392 [0,3, 0,48]
Chr14:65051897 / rs2062622202	G	1,00	Chr20:35761594 / rs1360384092	A	0,569 [0,47, 0,66]
	T	0,00		T	0,431 [0,34, 0,53]
Chr14:65002523 / –	G	0,122 [0,07, 0,17]	Chr20:35802733	T	0,707 [0,62, 0,79]
	C	0,878 [0,83, 0,93]		A	0,293 [0,21, 0,38]
Chr20:35763984 / rs2515846267	A	0,410 [0,33, 0,49]	Chr20:35928244	G	0,689 [0,61, 0,77]
	C	0,590 [0,51, 0,67]		T	0,311 [0,23, 0,39]
Chr20:35763994 / rs2515846275	T	0,429 [0,34, 0,51]	Chr20:35819153 / rs2042125567	A	0,788 [0,71, 0,86]
	A	0,571 [0,49, 0,66]		T	0,212 [0,14, 0,29]
Chr20:35928234	T	0,702 [0,61, 0,79]	Chr20:35876309 / rs1792401924	G	0,865 [0,79, 0,94]
	A	0,298 [0,21, 0,39]		A	0,135 [0,06, 0,21]
Chr20:35803547/ rs1452695636	A	0,804 [0,74, 0,87]	Chr20:35893835	T	0,409 [0,33, 0,49]
	T	0,196 [0,13, 0,26]		C	0,591 [0,51, 0,67]
Chr20:35788303 / rs2515867267	A	0,345 [0,26, 0,43]	Chr20:35793522 / rs1445047045	A	0,776 [0,71, 0,85]
	G	0,655 [0,57, 0,74]		T	0,224 [0,15, 0,29]
Chr20:35817409	T	0,33 [0,24, 0,42]	Chr20:35803537	C	0,757 [0,69, 0,83]
	A	0,67 [0,58, 0,76]		A	0,243 [0,17, 0,31]
			Chr20:35833886 / rs1348172939	C	0,875 [0,82, 0,93]
				T	0,125 [0,07, 0,18]

Chr20:35788303, Chr20:35761594, Chr20:35876309, Chr20:35793522, Chr20:35833886 значительно отличаются от ожидаемых частот в соответствии с популяционными данными, из-за чего для этих позиций мы наблюдаем повышенную частоту альтернативных аллелей. Однако для позиции Chr14:65051897 наблюдается полное совпадение по частоте референсного аллеля, которая на расширенной выборке составила 100% при ожидаемой частоте 97,98%. Для контроля качества секвенирования и исключения артефактов, связанных с ошибками чтения нуклеотидов, мы сформировали FASTQ-файлы по каждому образцу и проанализировали точность считывания в исследуемых локусах. Несоответствие полученных аллельных частот с популяционными может быть связано с низкой мощностью для редких вариантов в проектах, которые включены в агрегирование, шум в оценках частот аллелей, некоторые ложноположительные результаты, неслучайный сбор данных в выборках, платформах и популяциях, а также использование импутированных генотипов [15, 16].

Заключение

Проведенное исследование позволило выявить значимые генетические маркеры, ассоциированные с развитием туберкулеза. В нашем исследовании, основанном на анализе транскриптомных данных, выявлены гены, экспрессия которых статистически значимо различается между группами пациентов с активным туберкулезом и здоровыми лицами. Среди них особое внимание заслуживают *MAX*, *FNTB* и *PHF20* — гены, вовлеченные в регуляцию иммунного ответа, клеточной пролиферации и воспалительных процессов. Для ряда полиморфных локусов в этих генах (в частности, Chr14:65051895, Chr14:65051897, Chr14:65002523 и 16 позиций в гене *PHF20*) обнаружены статистически значимые различия в генотипическом составе между группами, что указывает на их потенциальную ассоциацию с развитием туберкулеза.

Полученные данные свидетельствуют о возможной роли этих генетических вариантов в модуляции иммунного ответа на *Mycobacterium tuberculosis* и открывают перспективы для разработки молекулярно-ди-

агностических панелей, направленных на определение риска прогрессирования туберкулезной инфекции. Учитывая согласованное наследование аллелей в генах *MAX* и *FNTB*, а также их функциональную взаимосвязь с ключевыми иммунными путями, представляется целесообразным дальнейшее изучение их совместного влияния на патогенез туберкулеза. Обнаруженные генетические особенности расширяют понимание генетической предрасположенности к туберкулезу и могут быть использованы для дальнейшей разработки новых биомаркеров для дифференциальной диагностики активной и/или латентной форм туберкулеза, а также идентификации мишеней для персонализированной терапии.

Данная работа вносит вклад в понимание молекулярных основ туберкулеза, выделяя такие гены, как *MAX*, *FNTB* и *PHF20* как потенциальные ключевые генетические факторы в иммунном ответе на инфицирование *M. tuberculosis*. Полученные результаты открывают новые направления для разработки стратегий ранней диагностики и таргетной терапии, учитывающие в том числе и генетические особенности пациентов.

Список использованных источников

1. World Health Organization. Global tuberculosis report 2015 [Электронный ресурс] / World Health Organization. — Geneva: WHO, 2015. — 192 p. — ISBN 978-92-4-156505-9. — Режим доступа: <https://apps.who.int/iris/handle/10665/191102>. — Дата доступа: 20.08.2025.
2. Скрининг детей и подростков на туберкулезную инфекцию в России — прошлое, настоящее, будущее / В. А. Аксенова, Л. А. Барышникова, Н. И. Клевно [и др.] // Туберкулез и болезни легких. — 2019. — № 97(9). — С. 59–67. doi: 10.21292/2075-1230-2019-97-9-59-67
3. National Center for Biotechnology Information. dbSNP: Short Genetic Variations [Электронный ресурс]. — Режим доступа: https://www.ncbi.nlm.nih.gov/snp/rs2062622202#frequency_tab. — Дата доступа: 20.08.2025.
4. LILRB1-HLA-G axis defines a checkpoint driving natural killer cell exhaustion in tuberculosis / J. Wang, Q. Chai, Z. Lei [et al.] // EMBO molecular medicine, 2024. — Vol. 16(8). — P. 1755–1790. doi: 10.1038/s44321-024-00106-1

5. National Center for Biotechnology Information. dbSNP: Short Genetic Variations [Электронный ресурс]. – Режим доступа: <https://www.ncbi.nlm.nih.gov/snp/rs143210184>. – Дата доступа: 20.08.2025.
6. National Center for Biotechnology Information. dbSNP: Short Genetic Variations [Электронный ресурс]. – Режим доступа: <https://www.ncbi.nlm.nih.gov/snp/>. – Дата доступа: 20.08.2025.
7. National Center for Biotechnology Information. Genome Data Viewer [Электронный ресурс]. – Режим доступа: <https://www.ncbi.nlm.nih.gov/gdv>. – Дата доступа: 20.08.2025.
8. c-Myc is a universal amplifier of expressed genes in lymphocytes and embryonic stem cells / Z. Nie, G. Hu, G. Wei [et al.] // *Cell*. – 2012. – Vol. 151(1). – P. 68–79. doi: 10.1016/j.cell.2012.08.033
9. Transcriptional Amplification in Tumor Cells with Elevated c-Myc / C. Y Lin, J. Lovén, P. B Rahl [et al.] // *Cell*. – 2012. – Vol. 151(1). – P. 56–67. doi: 10.1016/j.cell.2012.08.026.
10. Gnanaprakasam, JN R. MYC in Regulating Immunity: Metabolism and Beyond / J N R. Gnanaprakasam, R. Wang // *Genes*. – 2017. – Vol. 8(3):88. doi: 10.3390/genes8030088
11. Kalam, H. Alternate splicing of transcripts shape macrophage response to *Mycobacterium tuberculosis* infection / H. Kalam, M. F. Fontana, D. Kumar // *PLoS pathogens*. – 2017. – Vol. 13(3):e1006236. doi: 10.1371/journal.ppat.1006236
12. Identifying functional subtypes of IgA nephropathy based on three machine learning algorithms and WGCNA / H. Ren, W. Lv, Z. Shang [et al.] // *BMC medical genomics*. – 2024. – Vol. 17:61. doi: 10.1186/s12920-023-01702-9
13. Lysosome repositioning as an autophagy escape mechanism by *Mycobacterium tuberculosis* Beijing strain / T. Laopanupong, P. Prombutara, P. Kanjanasirirat [et al.] // *Scientific reports*. – 2021. – Vol. 11:4342. doi: 10.1038/s41598-021-83835-4
14. Proteomic dissection of LPS-inducible, PHF8-dependent secretome reveals novel roles of PHF8 in TLR4-induced acute inflammation and T cell proliferation / Ö. Erdoğan, L. Xie, Li Wang [et al.] // *Scientific reports*. – 2016. – Vol. 6:24833. doi: 10.1038/srep24833
15. A map of human genome variation from population-scale sequencing / The 1000 Genomes Project Consortium // *Nature*. – 2010. – Vol. 467. – P. 1061–1073. doi: 10.1038/nature09534.
16. The mutational constraint spectrum quantified from variation in 141,456 humans / K. J. Karczewski, L. C. Francioli, G. Tiao [et al.] // *Nature*. – 2020. – Vol. 581. – P. 434–443. doi: 10.1038/s41586-020-2308-7

E. V. Snytkov, M. D. Amelyanovich

SEARCH FOR POLYMORPHIC LOCI IN DIFFERENTIALLY EXPRESSED GENES ASSOCIATED WITH TUBERCULOSIS DEVELOPMENT

State Scientific Institution

“Institute of Genetics and Cytology of the National Academy of Sciences of Belarus”

27 Akademicheskaya St., 220072 Minsk, the Republic of Belarus

e-mail: evsnytkov@gmail.com

The article presents a study aimed at identifying genetic markers associated with tuberculosis development. The authors utilized RNA sequencing (RNA-seq) data from the public GEO repository (project GSE222001) and performed a bioinformatic analysis of gene expression followed by genomic variant (SNV) analysis in significantly differentially expressed genes. The study's objective was to identify genetic markers associated with tuberculosis infection development to improve diagnostics, prognosis, and treatment. The analysis revealed that the following genomic positions are associated with tuberculosis development: Chr14:65051895 and Chr14:65051897 in the *MAX* gene; Chr14:65002523 in the *FNTB* gene; and Chr20:35763984, Chr20:35763994, Chr20:35928234, Chr20:35803547, Chr20:35788303, Chr20:35817409, Chr20:35761582, Chr20:35761594, Chr20:35802733, Chr20:35928244, Chr20:35819153, Chr20:35876309, Chr20:35893835, Chr20:35793522, Chr20:35803537, and Chr20:35833886 in the *PHF20* gene.

Keywords: tuberculosis, *Mycobacterium tuberculosis*, *in silico* genotyping, transcriptome, mRNA, bcftools, RNA-seq.

Дата поступления в редакцию: 27 августа 2025 г.

Е. В. Гузова, М. А. Лиховец, А. А. Гавричков, И. С. Трусов, И. Н. Ильюшёнко,
Т. В. Романовская, В. В. Гринев

ИДЕНТИФИКАЦИЯ НЕАННОТИРОВАННЫХ БЕЛКОВ В КЛЕТКАХ ОСТРОГО МИЕЛОИДНОГО ЛЕЙКОЗА ЧЕЛОВЕКА С ПОМОЩЬЮ МУЛЬТИ-OMICS ПОДХОДА

Белорусский государственный университет
Республика Беларусь, 220030, г. Минск, пр. Независимости, 4
e-mail: grinev_vv@bsu.by

Поиск и идентификация белков, экспрессия которых присуща только раковым клеткам человека, является одним из современных вызовов для молекулярной биологии и биоинформатики. Вместе с тем в случае их обнаружения, структурно-функциональная организация таких белков представляет большой научный интерес. Кроме того, такие белки могут выступать в качестве диагностических и/или прогностических маркеров, а также мишеней для иммунотерапии и/или молекулярных терапевтиков нового поколения.

В настоящей работе изложены результаты биоинформатического исследования, основанного на последовательном анализе транскриптомных и протеомных данных и нацеленного на идентификацию полипептидов, экспрессирующихся в лейкозных клетках и имеющих менее 100% гомологии с известными белками человека. Детектировано 29 неканонических белков, кодируемых в общей сложности 23 генами, а также 33 новых белковых фрагмента. Выявлено четыре источника потенциально иммуногенных полипептидов.

Ключевые слова: лейкозные клетки человека, транскриптом, протеом, неоантигены, биоинформатический анализ.

Для цитирования: Идентификация неаннотированных белков в клетках острого миелоидного лейкоза человека с помощью мульти-omics подхода / Е. В. Гузова, М. А. Лиховец, А. А. Гавричков [и др.] // Молекулярная и прикладная генетика: сб. науч. тр. / Ин-т генетики и цитологии НАН Беларуси; редкол.: А. В. Кильчевский (гл. ред.) [и др.]. – Минск, 2025. – Т. 39. – С. 112–125.

Введение

Поиск и идентификация белков, экспрессия которых присуща только раковым клеткам человека, является одним из современных вызовов для молекулярной биологии и биоинформатики. Вместе с тем, в случае их обнаружения, структурно-функциональная организация таких белков представляет большой научный интерес. Кроме того, такие белки могут выступать в качестве диагностических и/или прогностических маркеров, а также мишеней для иммунотерапии и/или молекулярных терапевтиков нового поколения.

В частности, особый интерес представляют те опухоль-специфические белки, которые могут выступать в качестве антигенов для адаптивных TCR(CAR) T-клеток и противоопухолевых вакцин [1]. Одним из направлений адаптивной T-клеточной терапии явля-

ется генетическая модификация выделяемых из периферической крови цитотоксических Т-лимфоцитов, ведущая к приобретению ими противоопухолевой активности. Под таковой подразумевается способность к направленному распознаванию и последующему уничтожению тех клеток, на поверхности которых обнаруживаются специфические антигены.

TCR T-клетки экспрессируют доставленные с использованием, как правило, вирусных векторов T-клеточные рецепторы (TCRs, от англ. T cell receptors), чьи вариабельные α - и β -цепи обеспечивают специфическое взаимодействие с мишенью [2]. В качестве мишени может выступать антиген, локализованный на поверхности клетки в комплексе с молекулами главного комплекса гистосовместимости класса I (MHC I, от англ. major histocompatibility complex class I). Известно, что снижение экс-

прессии последних является одним из возможных механизмов ухода раковых клеток от иммунного надзора. В таком случае альтернативой TCR T-клеткам могут являться CAR T-клетки, которые экспрессируют химерные антигенные рецепторы (CARs, от англ. chimeric antigen receptors), способные распознавать поверхностные опухолевые антигены без участия молекул МНС [3].

Для создания вакцин активно используются дендритные клетки. Нагруженные антигенами и введенные в организм пациента, они мигрируют во вторичные лимфоидные ткани, где и происходит влекущая за собой запуск противоопухолевого иммунного ответа презентация целевых антигенов Т-лимфоцитам [4].

Все антигены, рассматриваемые в качестве потенциальных молекулярных мишеней, могут быть разделены на два класса: опухоль-ассоциированные и опухоль-специфические. Применение антигенов первого класса, встречающихся на поверхности как раковых, так и нормальных клеток, затруднено часто выявляемой токсичностью по отношению к последним [5]. По этой причине наибольший интерес представляют антигены второго класса, также называемые неоантигенами. Известно, что порождаться они могут генными мутациями, возникающими при опухолевой трансформации клеток, альтернативным сплайсингом молекул РНК, а также изменениями на уровне посттрансляционной модификации белков [6].

В существующих исследованиях, сфокусированных на поиске неоантигенов, в роли их источников рассматриваются главным образом генные мутации. Поскольку лейкозные клетки характеризуются низкой мутационной нагрузкой, количество детектированных опухоль-специфических антигенов остается относительно малым, и потому работы по их идентификации продолжают быть актуальными. Целью настоящего исследования явился поиск новых потенциально иммуногенных продуктов экспрессии генов, активных в лейкозных клетках человека.

Материалы и методы

В настоящем исследовании в качестве исходных данных выступили описанные в работе [7] результаты шести независимых повторов секвенирования транскриптома клеток мо-

дельной линии острого миелоидного лейкоза Kasumi-1, проведенного по методу парноконцевого RNA-Seq. Три библиотеки прочтений были получены для клеток, предварительно обработанных короткими интерферирующими РНК (киРНК), обеспечивающими нокдаун гибридного онкогена *RUNX1::RUNX1T1*, и еще три библиотеки — для клеток, трансфекция которых проводилась с использованием неактивных киРНК. Данные RNA-Seq, представленные в файлах формата FASTQ, проходили стандартную процедуру предобработки, после чего использовались для сборки полноразмерных молекул РНК.

Реконструкция транскриптома целевых клеток с опорой на эталонный геном осуществлялась с помощью программного пакета StringTie [8]. Исходно прочтения картировались относительно сборки GRCh38.p7 генома человека с использованием алгоритма subunc [9], после чего полученные BAM файлы наряду с генными аннотациями, содержащимися в базе данных Ensembl, передавались на вход StringTie. Запуск сборки проводился с использованием значений параметров, установленных по умолчанию. Исключение составил показатель «minimum isoform fraction», которому присваивалось значение, равное 0,05. Для фильтрации реконструированных полноразмерных молекул РНК была разработана оригинальная R-функция *filtrateTrans()*, исходный код которой, равно как и другие R-коды, созданные при выполнении описываемой работы, находится в открытом доступе в репозитории GitHub [10].

Сборка транскриптов без опоры на эталонный геном (*de novo*) проводилась при помощи ассемблера Trinity с использованием стандартных параметров [11]. В качестве входных данных выступили FASTQ файлы, содержащие процессированные библиотеки прочтений RNA-Seq. Для оценки качества сборки были использованы такие биоинформатические инструменты, как bowtie2 и BUSCO [12, 13]. С помощью bowtie2 прочтения, которые ранее были задействованы при работе с Trinity, выравнивались относительно реконструированных сборщиком последовательностей РНК. Возвращаемые результаты считались удовлетворительными в том случае, если число пар прочтений, выровненных конкордантно,

составляло не менее 70% от их общего количества. При помощи программы BUSCO оценивалась полнота реконструкции контрольных последовательностей ортологических генов в каждой отдельно взятой сборке. О высоком качестве последних свидетельствовала реконструкция свыше 80% последовательностей всех генов-ортологов BUSCO, экспрессирующихся в исследуемых клетках.

Идентификация в полноразмерных транскриптах главных открытых рамок считывания (mORFs, от англ. main open reading frames) осуществлялась с помощью функционала R/Bioconductor-пакета ORFhunter [14]. Поиск открытых рамок считывания, предшествующих mORFs (uORFs, от англ. upstream open reading frames), проводился с использованием пользовательской R-функции findUpstrORFs(). Для последующей трансляции *in silico* всех обнаруженных ORFs была применена функция translateORFs() из пакета ORFhunter.

Поиск новых полипептидов, основанный на реконструкции фрагментов зрелых молекул РНК, осуществлялся в соответствии с методикой, предложенной в работе [15]. Необходимые для анализа данные секвенирования транскриптомов медуллярных эпителиальных клеток тимуса (mTECs, от англ. medullary thymic epithelial cells) и миелоидных клеток-предшественников (MPCs, от англ. myelocytic precursor cells) (6 и 11 образцов соответственно) были взяты из базы данных Sequence Read Archive.

Отбор среди предсказанных *in silico* белковых последовательностей тех, которые действительно экспрессируются в клетках Kasumi-1, выполнялся с использованием нейронной сети DIA-NN [16]. Необходимые для анализа протеомные данные были получены на основе тотального белка, выделенного из указанных клеток, который подвергался трипсинизации и обработке йодоуксусной кислотой. Анализ проводился на базе Института биологических наук при Ньюкаслском университете (г. Ньюкасл, Великобритания) с помощью масс-спектрометра Orbitrap Fusion Lumos Tribrid. Создание эталонного протеома происходило путем объединения последовательностей всех белков человека, депонированных в базах данных UniProtKB/Swiss-Prot, Consensus CDS и NCBI RefSeq, с последующим удалением дубликатов и сохранением результа-

тов в формате FASTA. Итоговый FASTA-файл содержал в общей сложности 105 681 запись.

Биоинформатический анализ полученных результатов, подробно описанный в разделе «Результаты и обсуждение», осуществлялся в среде программирования R с помощью как оригинальных кодов собственной разработки, так и функционала пакетов, депонированных в хранилищах CRAN [17] или Bioconductor [18].

Результаты и обсуждение

Обобщенная схема разработанного биоинформатического конвейера, базирующегося на последовательной обработке транскриптомных и протеомных данных, изображена на рисунке 1. В рамках настоящего исследования с целью создания наиболее полного каталога новых белков и/или белковых фрагментов, экспрессирующихся в целевых клетках, было проведено три независимых цикла анализа. В их основе лежали соответственно три принципиально различающихся подхода к работе с библиотеками прочтений RNA-Seq. Ключевые этапы каждого анализа подробно рассмотрены ниже.

Поиск новых белков лейкозных клеток человека, основанный на реконструкции транскриптома с опорой на эталонный геном

Исходно данные RNA-Seq использовались для реконструкции полноразмерных молекул РНК с опорой на эталонный геном. В качестве сборщика в данном случае был выбран StringTie, хорошо зарекомендовавший себя среди отечественных и зарубежных пользователей. Общее количество реконструированных им транскриптов составило 215 688. Для снижения шума в полученных данных мы удалили из собранного транскриптома те молекулы РНК, чья экспрессия оказалась меньше 0,1 FPKM, общая длина — менее 45 нуклеотидов, длина хотя бы одного из экзонов и/или интронов — менее 25 и 50 нуклеотидов соответственно, либо транскрипты, которые оказались принадлежащими неканоническим хромосомам, а также для которых не удалось установить, какая из цепей ДНК их кодирует. Успешно прошли такую фильтрацию 52 564 (24,37%) молекулы РНК из собранных StringTie.

Для того чтобы классифицировать прошедшие фильтрацию РНК по потенциалу кодировать белки, мы использовали функционал

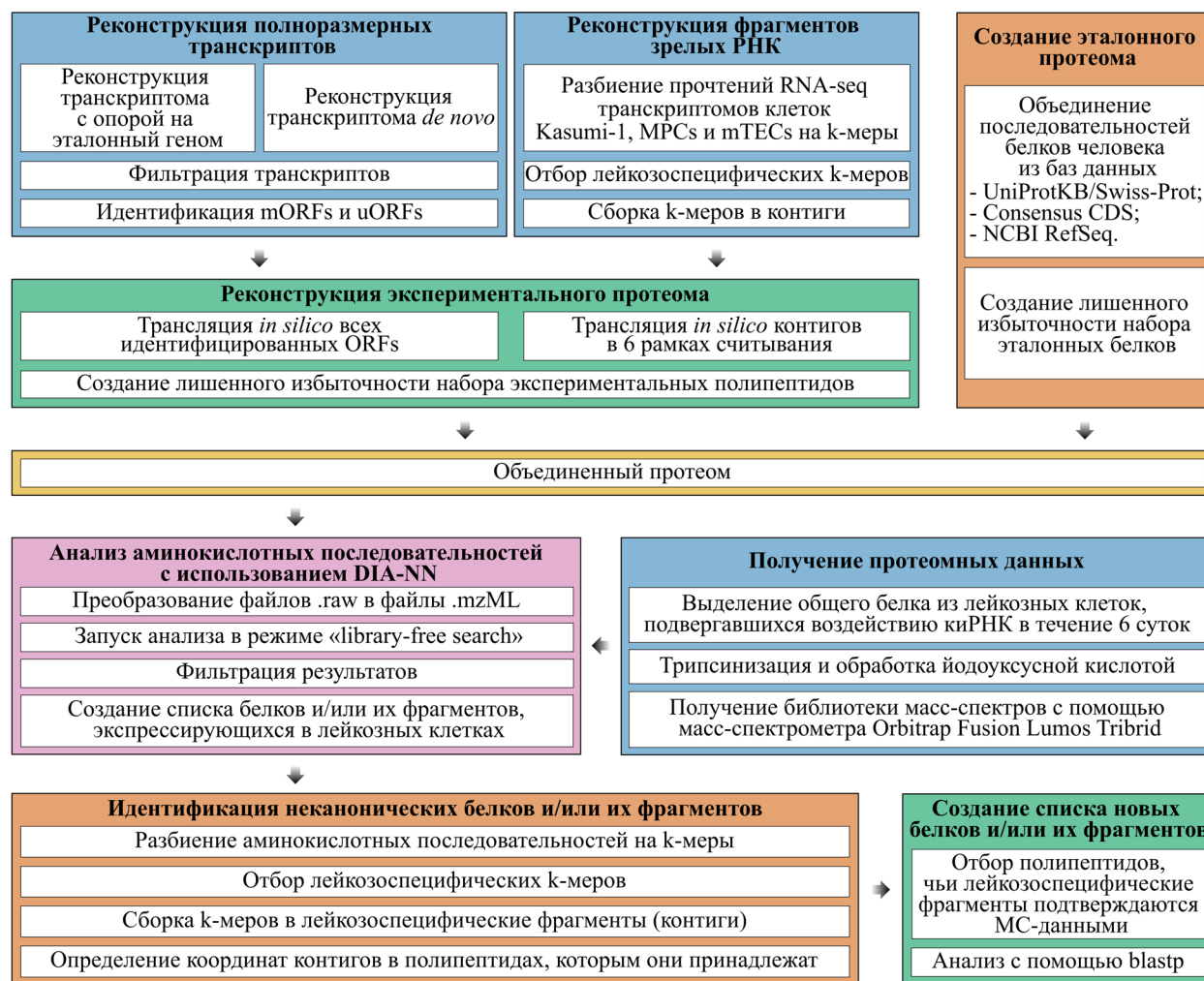


Рис. 1. Основные этапы идентификации новых полипептидов лейкозных клеток человека

R-пакета ORFhunter. Способность кодировать белок была выявлена у 31 041 последовательности (рис. 2А). Однако на сегодняшний день достоверно известно, что один транскрипт может нести несколько ORFs: помимо mORFs в его пределах также обнаруживаются малые ORFs (smORFs, от англ. small open reading frames), транслирующиеся с образованием микробелков [19]. Поэтому нами была разработана R-функция findUpstrORFs(), обеспечивающая поиск превалирующих smORFs, называемых uORFs. Они представляют собой smORFs, стартовый кодон которых локализуется в 5'-нетранслируемой области РНК (UTR, от англ. untranslated region), тогда как стоп-кодон может находиться как в 5'-UTR, так и в пределах mORF. С помощью указанной R-функции в молекулах РНК, собранных StringTie и прошедших фильтрацию, было

идентифицировано в общей сложности 67 527 uORFs, характеризующихся длиной 30 нуклеотидов и более.

Поскольку мы преследовали цель детектировать новые, неизвестные на сегодняшний день белки, то неотъемлемой частью разрабатываемого подхода являлся переход на протеомный уровень. Для этого последовательно все выявленные ORFs транслировались *in silico* при помощи функции translateORFs() из R-пакета ORFhunter. Таким образом, был сформирован первый экспериментальный протеом, включивший 73 192 уникальные последовательности белков (рис. 2Б). Поиск среди них, действительно синтезирующихся в клетках Kasumi-1, осуществлялся с использованием нейронной сети DIA-NN. На вход ей передавались масс-спектрометрические данные (МС-данные) в формате *.mzML, а также

FASTA-файл, содержащий последовательности белков экспериментального и эталонного протеомов.

DIA-NN запускалась в режиме «library-free search», не требующем наличия у пользователя библиотеки эталонных масс-спектров. Работая в таком режиме, программа осуществляет гидролиз *in silico* поступивших на вход аминокислотных последовательностей по сайтам расщепления трипсином. Для полученных пептидов длиной 7–30 аминокислотных остатков (а. о.) она генерирует библиотеку масс-спектров, которая в дальнейшем используется для поиска в МС-данных пептидов, подтверждающих экспрессию в клетках Kasumi-1 белков из переданного программе FASTA-файла.

Возвращаемые нейронной сетью результаты требуют дополнительной обработки. Для ее выполнения была разработана и применена R-функция `parseDIANNoutput()`, обеспечивающая фильтрацию данных по нескольким параметрам:

1) условная вероятность случайного события при идентификации пептида не должна превышать 0,01;

2) q-значение для пептидов, детектируемых в каждом специфическом запуске, должно быть не более 0,01;

3) количественное содержание пептида должно быть выше нуля не менее чем в 3 из 5 образцов одной экспериментальной группы (с нокдауном гибридного онкогена *RUNX1::RUNX1T1* и без него);

4) q-значение для белков, детектируемых

в каждом специфическом запуске, не должно быть выше 0,01;

5) число уникальных пептидов, подтверждающих экспрессию белка, должно быть не менее 2. Количество предсказанных белков, экспрессию которых удалось подтвердить МС-данными с помощью описанного подхода, составило 12 585. Из них 11 964 представляли собой продукты *in silico* трансляции mORFs (рис. 2В).

Непосредственно поиск новых, никем ранее не описанных изоформ белков осуществлялся в несколько этапов. На первом этапе проводилось сравнение детектированных белков с эталонными белками человека по методу k-меров: аминокислотные последовательности обоих наборов разбивались на k-меры длиной 8 а. о., после чего из набора k-меров экспериментальных белков исключались те, которые также обнаруживались среди k-меров эталонных белков. Таким образом был получен список лейкозоспецифических k-меров. Следующим шагом стала сборка последних в более протяженные фрагменты, которые далее, наряду с МС-пептидами, подтверждающими их экспрессию в изучаемых клетках, анализировались с помощью *blastp* [20]. По результатам этого анализа создавался каталог новых белков, в который включались только те из них, чьи лейкозоспецифические фрагменты, равно как и выравнивающиеся относительно них МС-пептиды, имели идентичность менее 100% по отношению к любым известным белкам человека. В данном случае их количество оказалось равным 14.

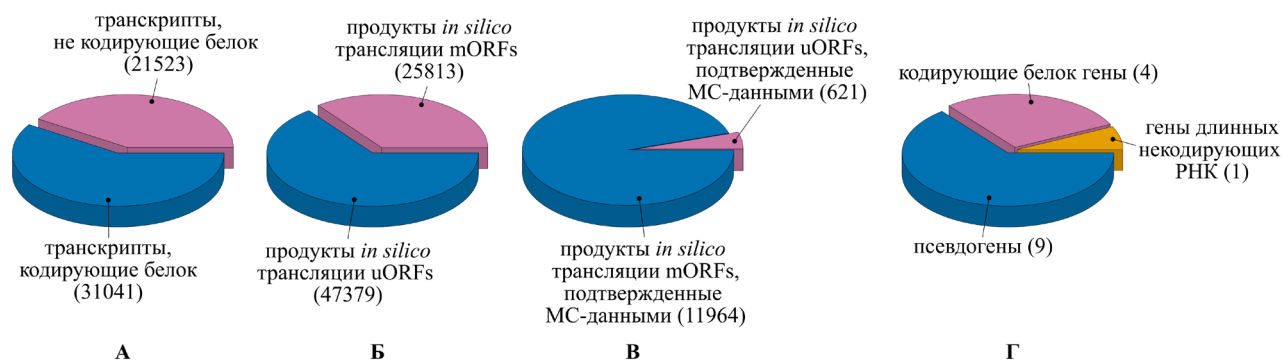


Рис. 2. Кодировующий потенциал транскриптов, реконструированных StringTie: А — распределение кодирующих и не кодирующих транскриптов; Б — структура экспериментального протеома; В — подтверждаемость МС-данными экспрессии экспериментальных белков в клетках Kasumi-1; Г — принадлежность генов, кодирующих новые белки, различным классам генетических элементов

Интересными оказались результаты анализа принадлежности генов, кодирующих новые белки, разным классам генетических элементов (рис. 2Г). Лишь 4 из 14 белков представляли собой продукты экспрессии белок-кодирующих генов, тогда как большая часть (9) кодировалась псевдогенами.

В качестве репрезентативного случая мы приводим на рисунке 3А схему реконструированного StringTie транскрипта псевдогена *WBP11P1*. Его функциональный ген-гомолог *WBP11* кодирует белок, участвующий в контроле сплайсинга молекул РНК [21]. По результатам биоинформатического анализа был детектирован белковый продукт экспрессии *WBP11P1* длиной 135 а. о., обладающий идентичностью 77% по отношению к белку WBP11 и содержащий участок неупорядоченной структуры неизвестной функциональности (рис. 4А).

Среди белок-кодирующих генов, для которых были обнаружены новые продукты экспрессии, оказались гены *YWHAZ*, *ARHGEF6*, *SPR* и *APEH*. Первые два кодируют белки, вовлеченные в регуляцию сигнальных путей [22, 23]. Ген *SPR* кодирует сепиаптерин-редуктазу, необходимую для образования тетрагидриоптерина [24]. Продуктом экспрессии гена *APEH* является фермент ацилпептидгидролаза, необходимый для нормальной деградации и рециркуляции N-ацетилированных белков [25].

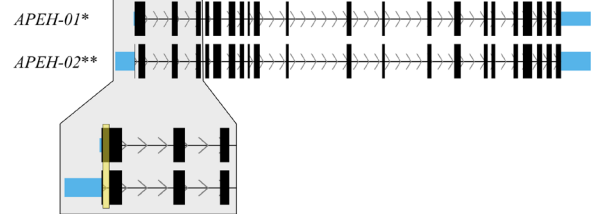
В случае каждого гена новый белок кодировался mORF, а его образование являлось следствием альтернативного сплайсинга соответствующих молекул пре-мРНК. Так, например, транскриптомным сборщиком была реконструирована неизвестная изоформа мРНК ге-

Транскрипт гена *WBP11P1*, реконструированный StringTie



А

Транскрипты гена *APEH*, реконструированные StringTie



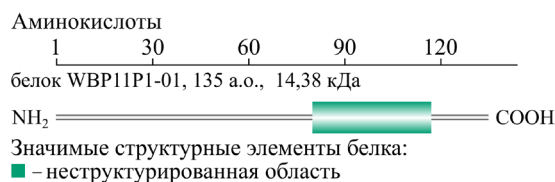
Б

Рис. 3. Структура реконструированных StringTie транскриптов генов *WBP11P1* (А) и *APEH* (Б). Знаками «*» и «**» отмечены транскрипты гена *APEH*, кодирующие описанную новую изоформу белка и канонический белок соответственно. Область сохраняемого интрона ограничена рамкой

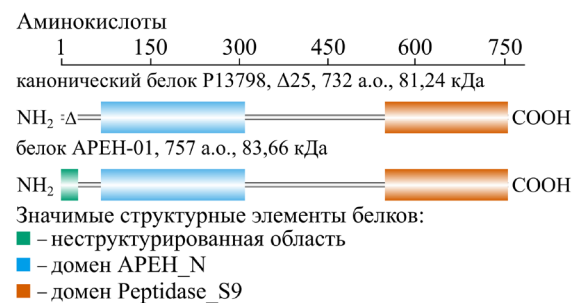
на *APEH*, несущая в своей структуре интрон, фланкированный первыми двумя экзонами (рис. 3Б). Интронная последовательность оказалась локализованной в области mORF, однако ее включение не сопровождается сдвигом рамки считывания. По этой причине кодируемый таким транскриптом белок отличается от уже известного тем, что вблизи N-конца содержит новый фрагмент длиной 25 а. о., который определяет появление участка неупорядоченной структуры (рис. 4Б).

Поиск новых белков лейкозных клеток человека, основанный на реконструкции транскриптома de novo

Сборка полноразмерных РНК с опорой на эталонный геном имеет некоторые недостатки. Таковыми являются: 1) невозможность учета вариативности нуклеотидных последователь-



А



Б

Рис. 4. Структура ранее неизвестных белков, кодируемых генами *WBP11P1* (А) и *APEH* (Б)

ностей генов и, соответственно, их РНК-продуктов; 2) относительно низкое разнообразие новых (структурно отличающихся от известных транскриптов) молекул РНК среди реконструируемых сборщиком последовательностей. Указанные особенности оказывают влияние на полноту восстановления лейкозного протеома и, как следствие, ограничивают репертуар детектируемых новых полипептидов. В то же время известно, что реконструкция транскриптома *de novo* лишена перечисленных недостатков. По этой причине, с целью расширить созданный на предыдущих этапах исследования каталог новых белков клеток Kasumi-1, дополнительно проводился анализ, базирующийся на сборке транскриптома целевых клеток с помощью сборщика Trinity.

Реконструированные Trinity молекулы РНК проходили фильтрацию по уровню экспрессии. Пороговое значение параметра устанавливалось равным 0,2 TPM. Полученный транскриптом включал 111 276 последовательностей, из которых 15 079, согласно результатам биоинформатического анализа, обладали кодирующим потенциалом (рис. 5А). С помощью функции findUpstrORFs() в их последовательностях удалось выявить 71 714 uORFs.

Путем трансляции *in silico* всех обнаруженных ORFs и последующего удаления повторяющихся аминокислотных последовательностей был получен набор белков, включивший в себя 54 500 последовательностей (экспериментальный протеом) (рис. 5Б), которые далее подвергались анализу, подробно описанному выше. В конечном счете удалось подтвердить, опираясь на МС-данные, экспрессию 3578

экспериментальных белков. При этом количество белков, кодируемых uORFs, как и при использовании первого подхода, описанного выше, было незначительным (рис. 5В). В целом, лимитировать детекцию могло несколько факторов. Один из них — их длина, которая оказывается значительно меньше, чем таковая у белков, кодируемых mORFs (рис. 6А), и требует более высокой чувствительности МС-анализа.

Среди белков с подтвержденной экспрессией 15 были классифицированы как новые. Один являлся продуктом *in silico* трансляции uORFs. Все белки принадлежали в общей сложности 10 белок-кодирующим генам и отсутствовали в списке новых аминокислотных последовательностей, который создавался по итогам анализа, основанного на сборке транскриптома с помощью StringTie.

Такие результаты, вероятно, были обусловлены сравнительно малой протяженностью молекул РНК, реконструированных Trinity (рис. 6Б). Это, в свою очередь, ожидаемо негативно отразилось на количестве транскриптов, у которых обнаруживался кодирующий потенциал, и, как следствие, на многообразии включенных в анализ аминокислотных последовательностей. Поскольку качество проведенной сборки *de novo* было высоким (о чем свидетельствуют результаты соответствующих тестов, см. «Материалы и методы»), по-видимому, реконструкция коротких последовательностей — это особенность алгоритма, используемого ассемблером Trinity.

Данные биоинформатического моделирования указывают на то, что образование трех из

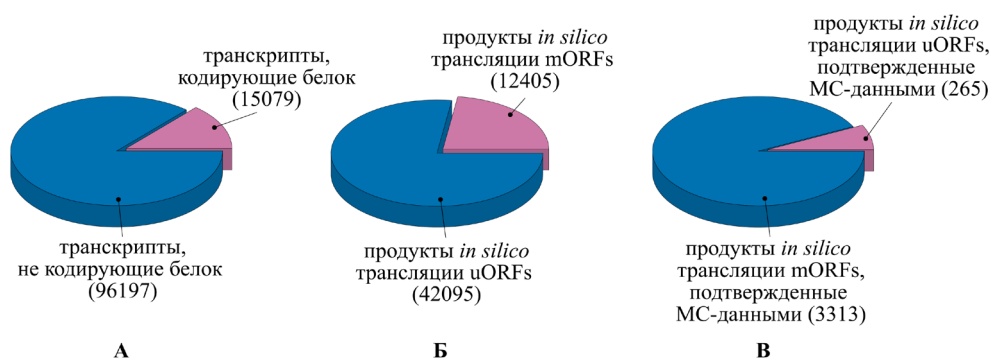


Рис. 5. Кодирующий потенциал транскриптов, реконструированных Trinity: А — распределение кодирующих и не кодирующих белок транскриптов; Б — структура экспериментального протеома; В — подтверждаемость МС-данными экспрессии экспериментальных белков в клетках Kasumi-1

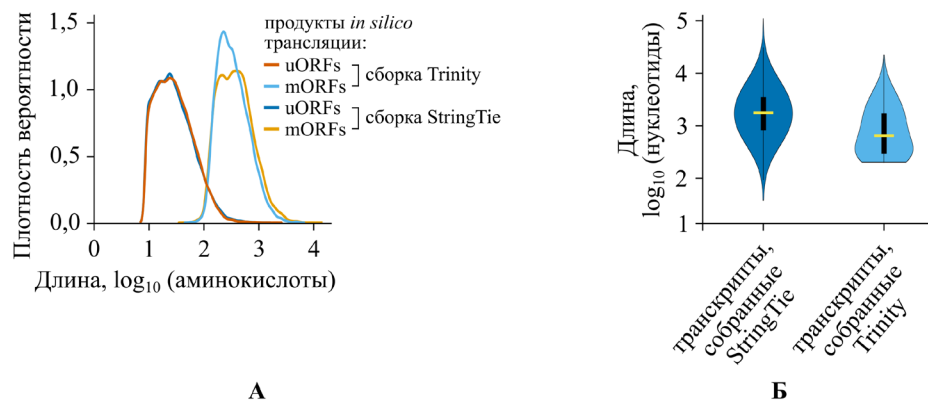


Рис. 6. Распределение длин продуктов *in silico* трансляции mORFs и uORFs (А), а также распределение длин транскриптов, реконструированных различными сборщиками (Б)

15 новых белков определялось альтернативным сплайсингом транскриптов соответствующих генов, 6 — генетической вариабельностью последних и еще 6 — совокупностью перечисленных факторов.

В качестве одного из примеров могут быть рассмотрены РНК/белковые продукты гена *PJA2*. Кодированный им канонический белок представляет собой убиквитин-лигазу, содержащую цинк-связывающий домен RING E3 и осуществляющую ковалентное субстрат-специфичное присоединение молекул убиквитина к белкам-мишеням [26]. Полученные данные свидетельствуют о том, что в ходе альтернативного сплайсинга первичных транскриптов рассматриваемого гена возможно образование ранее неизвестной изоформы мРНК (рис. 7), кодирующей короткий белок длиной 129 а. о. Причем интересным является то, что этот белок имеет принципиально новую первичную структуру, не обнаруживающую совпадений ни с одной другой, размещенной в публичных базах данных белков человека.

Помимо белков, порождаемых альтерна-

тивным сплайсингом РНК, были также обнаружены мутантные варианты, образование которых в клетках Kasumi-1 могло являться результатом возникновения несинонимичных однонуклеотидных замен в кодирующих областях соответствующих генов. Так, в гене *SDF2L1*, экспрессия которого активируется в ответ на стресс эндоплазматического ретикулума (ЭПР), выявлена транзикация NC_000022.11:g.21644096C > T [27]. Как следствие, продукт его трансляции, в норме локализующийся в просвете ЭПР и ответственный за контроль качества фолдинга белков, несет неконсервативную замену NP_071327.2:p.T196M (расстояние Грэнтэма (англ. Grantham distance) = 81) в пределах одного из доменов MIR.

Еще одна транзикация — NC_000019.10:g.45164900C > T — детектирована в последовательности гена *TRAPPC6A*, который кодирует субъединицу комплекса TRAPP, необходимого для транспорта белков от ЭПР к аппарату Гольджи [28]. В клетках Kasumi-1 с помощью МС-данных была подтверждена

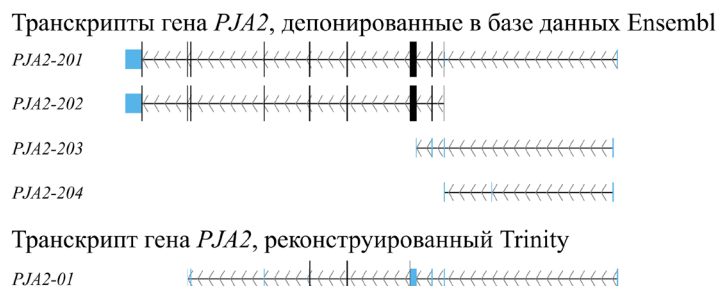


Рис. 7. Структура известных транскриптов гена *PJA2*, а также транскрипта, реконструированного Trinity

экспрессия двух изоформ одноименного белка протяженностью 173 и 159 а. о., содержащих консервативную замену NP_077013.1:p.V89M и NP_001257820.1:p.V75M (расстояние Грэнтэма = 21) соответственно.

Среди белков с измененной первичной структурой также оказались два продукта экспрессии гена *FAF1*. Первый — полноразмерный белок FAF1, участвующий в регуляции процессов клеточной пролиферации, апоптоза, воспаления, а также протеостаза [29]. Второй — его укороченная по N-концу на 87 а. о. изоформа. В последовательностях обоих белков обнаружена затрагивающая UBL2 домен консервативная замена NP_008982.1:p.L227V и XP_047298699.1:p.L140V (расстояние Грэнтэма = 32) соответственно, возникновение которой обусловлено трансверсией NC_000001.11:g.506555507G > C в рассматриваемом гене.

Поиск новых полипептидов лейкозных клеток человека, основанный на реконструкции фрагментов зрелых РНК

Очевидно, что реконструкция описанными способами полноразмерных молекул РНК не позволяет оценить все их многообразие. Кроме того, несовершенны и алгоритмы, используемые для выявления у реконструированных транскриптов кодирующего потенциала. Так или иначе, биоинформатическое предсказание экспрессии по крайней мере части действительно синтезирующихся в исследуемых клетках белков оказывается невозможным. В то же время известно, что при поиске молекулярных мишеней, пригодных для иммунотерапии лейкозов, интерес представляют пептиды длиной 8–16 а. о., способные образовывать стабильные комплексы с молекулами МНС I. В связи с этим может быть использован альтернативный подход, нацеленный на идентификацию непосредственно лейкозоспецифических белковых фрагментов. В его основе лежит частичная реконструкция зрелых молекул РНК, экспрессирующихся в изучаемых клетках. В данном исследовании она осуществлялась в соответствии с методикой, предложенной Grégory Ehx и соавторами [15].

Исходно предобработанные библиотеки прочтений, полученные в ходе секвенирования транскриптомов клеток Kasumi-1, разбивались на k-меры длиной 33 нуклеотида. Для

минимизации количества k-меров, содержащих ошибочно считанные основания, проводилось исключение последовательностей, встречающихся менее трех раз в каждом образце одной экспериментальной группы. Прошедшие фильтрацию k-меры далее были объединены в один лишенный избыточности набор лейкозных k-меров.

На следующем этапе в работу включались публичные данные RNA-Seq, полученные для выбранных в качестве контроля MPCs и mTECs. Библиотеки прочтений аналогичным образом последовательно разбивались на k-меры, после чего с их помощью выполнялось вычитание списка лейкозных k-меров. В результате такой отрицательной фильтрации список лейкозных k-меров был сокращен до 916 764 уникальных лейкозоспецифических последовательностей, что составило 1,41% от их исходного количества. При этом, как и следовало ожидать, наибольший вклад в вычитание был внесен k-мерами, принадлежащими MPCs (рис. 8А).

Следующим шагом стала сборка лейкозоспецифических k-меров в более протяженные контиги, которые затем транслировались *in silico* в шести рамках считывания (+0, +1 и +2 в обеих ориентациях). Созданный таким образом экспериментальный протеом содержал 618 910 полипептидов протяженностью 8 а. о. и более. Эти полипептиды, вероятнее всего, являются лишь фрагментами полноразмерных белков, что находит отражение в ха-

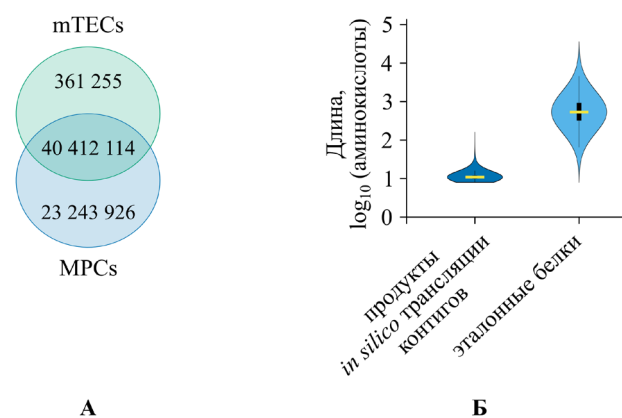


Рис. 8. k-меры mTECs и MPCs, для которых были обнаружены совпадения среди лейкозных k-меров (А), а также распределение длин экспериментальных полипептидов (Б)

рактуре распределения их длин относительно длин известных белков человека (рис. 8Б). По этой причине при фильтрации результатов поиска таких последовательностей с помощью DIA-NN число МС-пептидов, достаточное для подтверждения их экспрессии, было снижено до 1.

Количество полипептидов с подтвержденной экспрессией оказалось равным 318. При этом необходимо понимать, что отсутствие контекстной информации о структуре белка, которому принадлежал тот или иной экспериментальный полипептид, значительно ограничивало количество МС-пептидов, способных подтвердить его экспрессию. Также важно учесть, что 181 119 (29,26%) из подвергнутых анализу аминокислотных последовательностей характеризовались наличием сайтов, по которым возможно их расщепление трипсином с образованием набора слишком коротких или же, наоборот, слишком длинных пептидов, обнаружение которых в МС-данных невозможно. Более того, у 190 634 (30,8%) полипептидов этих сайтов вовсе не было обнаружено. Все перечисленное позволяет предположить, что среди реконструированных фрагментов РНК число последовательностей, обладающих кодирующим потенциалом, на самом деле может быть выше.

Из 318 полипептидов с подтвержденной экс-

прессией 33 были отнесены к новым (рис. 9). Точное определение генной принадлежности с помощью инструмента BLAT оказалось невозможным для 17 фрагментов зрелых РНК, кодирующих новые экспериментальные полипептиды, что было обусловлено, в частности, множественным либо частичным выравниванием отдельных последовательностей относительно различных участков генома. Среди оставшихся 16 фрагментов два картировались на участки генома, в которых перекрывалось несколько генов, и поэтому далее они будут рассматриваться как отдельные фрагменты. Таким образом, общее количество целевых последовательностей РНК с установленной генной принадлежностью оказалось равным 18. Примечательно, что образование пяти из них могло быть обусловлено активностью генов, которые, согласно аннотациям Ensembl, не обладают кодирующим потенциалом. Остальные 13 являлись фрагментами транскриптов белок-кодирующих генов. Среди них пять представляли собой последовательности ранее не описанных сохраняемых интронов, причем три из них содержали стоп-кодоны. Еще три реконструированных фрагмента РНК принадлежали нетранслируемым областям известных транскриптов генов *SCUBE1*, *ULBP1*, а также *KCNJ11* и, вероятно, транслировались с образованием микробелков в составе smORFs.



Рис. 9. Результаты анализа генной принадлежности фрагментов зрелых РНК, кодирующих новые экспериментальные полипептиды

Наконец, оставшиеся пять фрагментов принадлежали кодирующим областям полноразмерных молекул РНК и содержали однонуклеотидные замены (табл. 1). Среди них наиболее интересная затрагивает протоонкоген *PBXIP1*, высокая экспрессия которого ассоциирована с прогрессией ряда онкологических забо-

леваний. Обнаруженная трансверсия приводит к тому, что в пределах соответствующего белка, известного как регулятор эпителиально-мезенхимального перехода, а также пролиферации и миграции опухолевых клеток, происходит замена гидрофобного лейцина на гидрофильный глутамин [30].

Таблица 1

Однонуклеотидные замены, выявленные в кодирующих областях генов

Ген	Хромосома	Позиция в хромосоме	Вариант нуклеотидной замены	Вариант аминокислотной замены	Расстояние Грэнтэма
<i>PBXIP1</i>	1	154946463	A > T	L > Q	113
<i>FKBP5</i>	6	35637137	C > T	G > S	56
<i>STIM1</i>	11	4091617	C > T	T > I	89
<i>IRAK4</i>	12	43772969	G > A	R > Q	43
<i>WIZ</i>	19	15425333	C > T	R > H	29

Заключение

Представленный конвейер, предполагающий последовательный биоинформатический анализ транскриптомных и протеомных данных, позволил идентифицировать неизвестные на сегодняшний день белковые продукты экспрессии генов, активных в модельных лейкозных клетках человека. На основе полученных результатов может быть выделено несколько их источников: 1) вариабельность нуклеотидных последовательностей генов; 2) альтернативный сплайсинг РНК; 3) smORFs; 4) трансляционная активность РНК-продуктов генов, аннотированных как некодирующие (в частности, псевдогенов).

В то же время открытым остался вопрос относительно того, являются обнаруженные белковые продукты одним из вариантов нормы либо же их образование представляет собой следствие протекания патологического процесса. Для того чтобы ответить на него, требуется проведение аналогичного анализа транскриптомных и протеомных данных, полученных для нормальных клеток крови человека. В случае обнаружения белков/белковых фрагментов, синтез которых присущ только лейкозным клеткам, одним из направлений дальнейших исследований может стать дополнение описанного биоинформатического конвейера алгоритмом, который позволит прове-

сти среди них поиск неоантигенов, пригодных для иммунотерапии лейкозов.

Авторы признательны доктору Хосе Марин-Рубио из Института биологических наук при Ньюкаслском университете (г. Ньюкасл, Великобритания), а также профессору Олафу Хайденрайху из Центра детской онкологии им. Принцессы Максими (г. Утрехт, Нидерланды) за помощь в получении оригинальных протеомных данных для клеток Kasumi-1.

Эта работа получила финансовую поддержку Белорусского республиканского фонда фундаментальных исследований в рамках проекта «Разработка комплексного подхода по идентификации новых, зависимых от дифференциального сплайсинга, иммуногенных пептидов лейкозных клеток» (грант № M23-056, № гос. регистрации 20231102, сроки выполнения 2023–2025 гг.).

Список использованных источников

1. Immunotherapy approaches for hematological cancers / O. L. Lanier, E. Perez-Hertero, A. P. D'Andrea [et al.] // iScience. – 2022. – Vol. 25, № 11. – Art. 105326.

2. TCR-T cell therapy: current development approaches, preclinical evaluation, and perspectives on regulatory challenges / E. A. Golikova, A. A. Alshevskaya, S. Alrhoun [et al.] // Jour-

nal of Translational Medicine. – 2024. – Vol. 22, № 1. – Art. 897.

3. Finding potential targets in cell-based immunotherapy for handling the challenges of acute myeloid leukemia / A. H. Kheirkhah, S. Habibi, M. H. Yousefi [et al.] // *Frontiers in Immunology*. – 2024. – Vol. 15. – Art. 1460437.

4. Lee, K.-W. Dendritic cell vaccines: a shift from conventional approach to new generations / K.-W. Le, J. W. P. Yam, X. Mao // *Cells*. – 2023. – Vol. 12, № 17. – Art. 2147.

5. Uncovering the tumor antigen landscape: What to know about the discovery process / S. Feola, J. Chiaro, B. Martins, V. Cerullo // *Cancers (Basel)*. – 2020. – Vol. 12, № 6. – Art. 1660.

6. Neoantigens: promising targets for cancer therapy / N. Xie, G. Shen, W. Gao [et al.] // *Signal Transduction and Targeted Therapy*. – 2023. – Vol. 8, № 1. – Art. 9.

7. RUNX1/RUNX1T1 mediates alternative splicing and reorganises the transcriptional landscape in leukemia / V. V. Grinev, F. Barneh, I. M. Ilyushonak [et al.] // *Nature Communications*. – 2021. – Vol. 12, № 1. – Art. 520.

8. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads / M. Pertea, G. M. Pertea, C. M. Antonescu [et al.] // *Nature Biotechnology*. – 2015. – Vol. 33, № 3. – P. 290–295.

9. Liao, Y. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads / Y. Liao, G. K. Smyth, W. Shi // *Nucleic Acids Research*. – 2019. – Vol. 47, № 8. – Art. 47.

10. GitHub. – URL: <https://github.com/VGrinev/TranscriptomicFeatures> (date of access: 18.02.2025).

11. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data / M. G. Grabherr, B. J. Haas, M. Yassour [et al.] // *Nature Biotechnology*. – 2011. – Vol. 29, № 7. – P. 644–652.

12. Langmead, B. Fast gapped-read alignment with Bowtie 2 / B. Langmead, S. L. Salzberg // *Nature Methods*. – 2013. – Vol. 9, № 4. – P. 357–359.

13. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes / M. Manni, M. R. Berkeley, M. Seppey [et al.] // *Molecular Biology and Evolution*. – 2021. – Vol. 38, № 10. – P. 4647–4654.

14. ORFhunter: An accurate approach to the automatic identification and annotation of open reading frames in human mRNA molecules / V. V. Grinev, M. M. Yatskou, V. V. Skakun [et al.] // *Software Impacts*. – 2022. – Vol. 12, № 2. – Art. 100268.

15. Atypical acute myeloid leukemia-specific transcripts generate shared and immunogenic MHC class-I-associated epitopes / G. Ehx, J.-D. Larouche, Ch. Durette [et al.] // *Immunity*. – 2021. – Vol. 54, № 4. – P. 737–752.

16. DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput / V. Demichev, Ch. B. Messner, S. I. Vernardis [et al.] // *Nature Methods*. – 2020. – Vol. 17, № 1. – P. 41–44.

17. The comprehensive R archive network. – URL: <https://cran.r-project.org/> (date of access: 18.02.2025).

18. Bioconductor: Open source software for Bioinformatics. – URL: <https://www.bioconductor.org/> (date of access: 18.02.2025).

19. A high-resolution map of human RNA translation / S. P. Chothani, E. Adami, A. A. Widjaja [et al.] // *Molecular Cell*. – 2022. – Vol. 82, № 15. – P. 2885–2899.

20. BLAST. – URL: <https://blast.ncbi.nlm.nih.gov/Blast.cgi>. (date of access: 18.02.2025).

21. Heterozygous loss of WBP11 function causes multiple congenital defects in humans and mice / E. M. M. A. Martin, A. Enriquez, D. B. Sparrow [et al.] // *Human Molecular Genetics*. – 2020. – Vol. 29, № 22. – P. 3662–3678.

22. YWHAZ variation causes intellectual disability and global developmental delay with brain malformation / R.-P. Wan, Zh.-G. Liu, X.-F. Huang [et al.] // *Human Molecular Genetics*. – 2023. – Vol. 32, № 3. – P. 462–472.

23. Clinical implication and potential function of ARHGEF6 in acute myeloid leukemia: An in vitro study / K. Li, H. Wang, Ch. Yang [et al.] // *PLoS One*. – 2023. – Vol. 18, № 4. – Art. 283934.

24. Sepiapterin reductase: characteristics and role in diseases / Y. Wu, P. Chen, L. Sun [et al.] // *Journal of Cellular and Molecular Medicine*. – 2020. – Vol. 24, № 17. – P. 9495–9506.

25. Acyl peptide enzyme hydrolase (APEH) activity is inhibited by lipid metabolites and peroxidation products / K. Tyler, Sh. Geilman, D. M. Bell [et al.] // *Chemico-Biological Interactions*. – 2021. – Vol. 348, № 1. – Art. 109639.

26. Ubiquitylation of MFHAS1 by the ubiquitin ligase praja2 promotes M1 macrophage polarization by activating JNK and p38 pathways / J. Zhong, H. Wang, W. Chen [et al.] // *Cell Death & Disease*. – 2017. – Vol. 8, № 5. – Art. 2763.
27. Hanafusa, K. SDF2-like protein 1 (SDF2L1) regulates the endoplasmic reticulum localization and chaperone activity of ERdj3 protein / K. Hanafusa, I. Wada, N. Hosokawa // *Journal of Biological Chemistry*. – 2019. – Vol. 294, № 50. – P. 19 335–19 348.
28. A missense mutation in TRAPPC6A leads to build-up of the protein, in patients with a neurodevelopmental syndrome and dysmorphic features / H. Sh. Mohamoud, S. Ahmed, M. Jelani [et al.] // *Scientific Reports*. – 2018. – Vol. 8, № 1. – Art. 2053.
29. Park, G. A novel function of FAF1, which induces dopaminergic neuronal death through cell-to-cell transmission / G. Park, B.-S. Kim, E. Kim // *Cell Communication and Signaling*. – 2020. – Vol. 18, № 1. – Art. 133.
30. Khumukcham, S. S. Two decades of a protooncogene HPIP/PBXIP1: uncovering the tale from germ cell to cancer / S. S. Khumukcham, B. Manavathi // *Biochimica et Biophysica Acta (BBA) – Reviews on Cancer*. – 2021. – Vol. 1876, № 1. – Art. 188576.

K. V. Huzava, M. A. Likhavets, A. A. Gavrichkov, I. S. Trusau, I. M. Ilyushonak,
T. V. Ramanouskaya, V. V. Grinev

IDENTIFICATION OF UNANNOTATED PROTEINS IN HUMAN ACUTE MYELOID LEUKEMIA CELLS USING A MULTI-OMICS APPROACH

Belarusian State University
4 Nezavisimosti Ave., 220030 Minsk, the Republic of Belarus
e-mail: grinev_vv@bsu.by

The search for and identification of proteins whose expression is unique to human cancer cells is one of the modern challenges of molecular biology and bioinformatics. At the same time, when such proteins are discovered, their structural and functional organization is of great scientific interest. Furthermore, such proteins may serve as diagnostic and/or prognostic markers as well as targets for immunotherapy and/or next-generation molecular therapeutic agents.

This paper presents the results of a bioinformatics study based on the sequential analysis of transcriptomic and proteomic data and aimed at identifying polypeptides expressed in leukemia cells and that are of less than 100% homology with known human proteins. Twenty-nine non-canonical proteins encoded by a total of 23 genes, as well as 33 new protein fragments, were detected. The number of identified sources of potentially immunogenic polypeptides was 4.

Keywords: human leukemia cells, transcriptome, proteome, neoantigens, bioinformatics analysis.

Дата поступления в редакцию: 09 апреля 2025 г.

И. С. Казловский^{1,2}, Н. В. Поклонская¹, И. В. Бельская¹, Т. В. Амвросьева¹, А. И. Зинченко²

МЕТОД СОЗДАНИЯ ГЕНЕТИЧЕСКИХ КОНСТРУКЦИЙ, СОДЕРЖАЩИХ ГЕНЫ ПОЛНОРАЗМЕРНЫХ КАПСИДНЫХ БЕЛКОВ НОРОВИРУСОВ ЧЕЛОВЕКА

¹Республиканский центр гигиены, эпидемиологии и общественного здоровья

Республика Беларусь, 220114, г. Минск, ул. Филимонова, 23

e-mail: kazlouski.illia@gmail.com

²Государственное научное учреждение

«Институт микробиологии Национальной академии наук Беларуси»

Республика Беларусь, 220084, г. Минск, ул. Купревича, 2

Нуклеотидные последовательности, кодирующие полноразмерные капсидные белки норовирусов человека геновариантов GI.3, GII.4 и GII.17, клонированы в состав вектора pColdI. Методами молекулярной биологии подтверждена корректная ориентация вставок в конструкциях pCold-GI.3-VP1, pCold-GII.4-VP1 и pCold-GII.17-VP1, а также отсутствие спонтанных мутаций в их последовательностях. Полученные плазмидные клоны в дальнейшем будут использованы для интеграции в бактериальные клетки.

Ключевые слова: норовирус, капсидный белок, VP1, pColdI, рекомбинантная ДНК.

Для цитирования: Метод создания генетических конструкций, содержащих гены полноразмерных капсидных белков норовирусов человека / И. С. Казловский, Н. В. Поклонская, И. В. Бельская [и др.] // Молекулярная и прикладная генетика: сб. науч. тр. / Ин-т генетики и цитологии НАН Беларуси; редкол.: А.В. Кильчевский (гл. ред.) [и др.]. – Минск, 2025. – Т. 39. – С. 126–137.

Введение

Норовирусы человека (НоВ) являются ведущими этиологическими агентами острого гастроэнтерита (ОГЭ) не бактериальной природы, вызывая как спорадические, так и групповые случаи заболевания во всех возрастных группах [1]. Ежегодно норовирусная инфекция (НоВИ) приводит к более 200 тыс. смертей и миллионам госпитализаций, что подчеркивает необходимость разработки эффективных методов диагностики и профилактики [2].

Золотым стандартом диагностики НоВИ является ОТ-ПЦР, обладающая высокой чувствительностью и специфичностью [3]. Однако метод применяется преимущественно в стационарах, требует специального оборудования, квалифицированного персонала и дорогих реагентов. Между тем часть случаев НоВИ протекает вне стационара, но остается значимой с эпидемиологической точки зрения из-за длительного вирусовыделения и риска заражения окружающих. Для таких ситуаций предпочтительны экспресс-тесты [4], осно-

ванные на выявлении норовирусного антигена с помощью специфических антител.

Если технология получения поликлональных антител к норовирусам достаточно проста и изучена, то получение антигенного материала остается сложной задачей. Из-за некультивируемости НоВ невозможно применять методы классической вирусологии [5], поэтому для получения белков используют генно-инженерные подходы. Существуют системы экспрессии поверхностного белка НоВ с антигенными свойствами, наиболее распространенная из которых — с применением бакуловирусов и клеток насекомых Sf9 [6, 7]. Однако технология трудоемкая, требует длительного времени и сложной очистки. Альтернативой является экспрессия в клетках млекопитающих (HEK293) [8], но она отличается высокой себестоимостью, что ограничивает ее применение.

Наиболее доступным способом получения рекомбинантных белков остается экспрессия в *Escherichia coli*. Однако ранее она давала

крайне низкий выход белка либо вообще не позволяла синтезировать интересующие белки HoB [9]. В связи с этим разрабатываются новые подходы, включая экспрессию при «холодовом шоке» — метод, основанный на индукции белкового синтеза при 15 °С с одновременным снижением общей трансляции и активности протеаз, что обеспечивает выход целевого белка до 60% [10].

Капсидный белок VP1 HoB способен к самосборке в вирусоподобные частицы (ВПЧ), обладающие высокой стабильностью и иммуногенностью. Несмотря на отсутствие зарегистрированных вакцин против HoBI, большинство кандидатных препаратов включает белок VP1, экспрессированный в клетках млекопитающих или дрожжей [2]. Чаще всего используется белок VP1 генотипа GII.4 — наиболее распространенного среди патогенных для человека норовирусов. Он остается доминирующим уже несколько десятилетий, на его долю приходится более 80% вспышек. Высокая изменчивость GII.4 обуславливает регулярное появление новых эпидемических вариантов [11, 12].

В Беларуси HoB занимают второе место среди возбудителей острых кишечных инфекций (после ротавирусов) и лидируют по числу вспышек групповой заболеваемости. В этих условиях разработка доступных и эффективных технологий получения компонентов для диагностики и вакцинопрофилактики HoBI является актуальной задачей.

Данная работа посвящена получению генетических конструкций, несущих в себе нуклеотидные последовательности, кодирующие полноразмерные капсидные белки и основные полипептиды HoB геновариантов, эндемичных для Республики Беларусь (GII.4, GII.17 и GI.3) для последующей разработки простой и доступной биотехнологии получения норовирусных антигенов.

Материалы и методы

Образцы биологического материала (фекалии, $n = 7$), заведомо содержащие РНК HoB известных генотипов (GII.4, GII.17, GI.3) были взяты для исследований из банка лаборатории инфекций с природным резервуаром РЦГЭиОЗ.

Для клонирования и последующей наработ-

ки плазмидных векторов-носителей использовали коммерческий штамм *E. coli* XL1blue («Novagen», США) и плазмиду pJET1.2 («Novagen»).

Для создания векторов экспрессии в бактериальных клетках VP1 HoB разных генотипов использовали плазмиды pET42a(+) («Novagen») и pColdI («Takara Bio», Япония).

Для оценки перспективы использования различных систем экспрессии для получения HoB-полипептидов проводили:

- анализ нуклеотидных последовательностей, кодирующих VP1 HoB разных генотипов, с целью определения частоты встречаемости редких кодонов для организма-реципиента целевой конструкции, содержащей фрагмент генома HoB с использованием сервиса «GenScript Rare Codon Analysis Report»;
- анализ аминокислотных последовательностей HoB VP1 с использованием сервисов «Predicted antigenic peptides» (поиск антигенных детерминант) и «ScanProsite tool» (локализация дисульфидных связей).

Дизайн олигонуклеотидных праймеров для клонирования фрагментов, кодирующих VP1 методом ПЦР, проводили на основе анализа последовательностей OPC2 HoB с использованием ПО Primer3 (v. 0. 4. 0), принимая во внимание такие параметры, как вырожденность, T_m , GC%, ΔG , GC клэмп, образование димеров и шпилек.

РНК выделяли с помощью набора «НК-экстра» («РЦГЭиОЗ», Беларусь). Первичную апробацию разработанных олигонуклеотидов для накопления фрагментов 1-го раунда (клонирования в вектор pJET1.2) проводили в реакционной смеси объемом 25 мкл с добавлением 1х «5x-ArtMix» реагента для проведения ОТ-ПЦР в одной пробирке и 20 пмоль фланкирующих праймеров. В качестве матрицы использовали РНК со значениями порогового цикла реакции (C_t) в пределах 17–26 («ОКВИ-ПЦР», РНПЦ эпидемиологии и микробиологии, Беларусь). Режим амплификации был адаптирован под рекомендуемые параметры, предложенные производителем ферментного препарата. Температура отжига соответствовала наиболее низкой из теоретически рассчитанных (54 °С — для клонирования фрагмента, кодирующего VP1 HoB GII.17).

Электрофорез выполняли в агарозном геле

с ТАЕ-буфером и бромистым этидием (0,5 мкг/мл) при 3–4 В/см. Загрузочный буфер содержал бромфеноловый синий, ксиленианол (по 0,5%), глицерин (50%) и ЭДТА (0,1 М, pH 8,0). Гели сканировали в УФ-свете ($\lambda = 280$ нм).

Секвенирование проводили с GenomeLab DTCS Kit («Beckman Coulter») на CEQ 8000, анализ — в MEGA7.0 с использованием стандартных праймеров к вектору pJET1.2: pET1.2 reverse sequencing prime — 5'-GACGGTACCTTTTAGCTACAAGAA-3', pET1.2 forward sequencing primer — 5'-CGACTCACTATAGGGAGAGCGGC-3'.

Ампликоны и векторные конструкции очищали с GeneJET Kit («Thermo Fisher»), элюирование — в ТЕ-буфере. Концентрацию ДНК определяли на Qubit («Invitrogen») с красителем ZUBR Green I («Праймтех», Беларусь).

Лигирование фрагментов в вектор pJET1.2 осуществляли с использованием коммерческой лигазы бактериофага T4 («New England Biolabs», США) в соответствии с прилагаемой инструкцией.

Для верификации VP1-кодирующей последовательности в составе полученных плазмид проводили накопление ДНК-продуктов в реакционной смеси, содержащей по 20 пмоль прямого и обратного праймеров («Праймтех», Беларусь), с добавлением 1 ед. Pfu-полимеразы «Flash» («АртБиоТех», Беларусь) и 0,01 мкг ДНК.

Приготовление химически компетентных клеток *E. coli* XL1Blue и их трансформацию осуществляли с помощью набора Roti Transform согласно инструкции производителя («Carl Roth», Германия). Все клеточные суспензии после трансформации высевались на твердую агаризованную питательную среду LB с добавлением селективного антибиотика (ампициллин — для pColdI-дочерних векторов и канамицин для векторов, основанных на системе pET42a) до конечной концентрации 100 мкг/мл. При проведении контрольной трансформации селективным антибиотиком выступал ампициллин с конечной концентрацией в среде 100 мкг/мл.

ДНК выросших клеток-трансформантов подвергали ПЦР-анализу для подтверждения наличия в ней целевых фрагментов генов. В состав реакционной смеси входили 2,5х ПЦР-буфер (специфический для используе-

мого фермента), по 10 пмоль фланкирующих праймеров (прямой комплементарен последовательности вектора, обратный — вставке), 1 ед. высокоточной Flash-полимеразы («АртБиоТех», Беларусь).

Амплификацию целевых фрагментов с полученных векторов-носителей на основе pJET1.2 (10 нг), содержащих норовирусные гены капсидных белков, проводили в реакционной смеси, адаптированной для Flash-полимеразы (представлена ранее), содержащей по 20 пмоль фланкирующих праймеров для проведения ПП-ПЦР. Температура отжига данных праймеров была подобрана для амплификации целевых продуктов при одном температурном режиме (55–56 °C). Время элонгации соответствовало расчетному для получения полного гена VP1 (1 мин).

Выделение векторов экспрессии, содержащих VP1-кодирующие фрагменты генома NoV разных генотипов, проводили методом щелочного лизиса [13].

Для статистической обработки данных использовали методы, принятые в биологии [14]. Большинство приведенных в работе экспериментальных данных представляют усредненные величины 3–7 опытов. Доверительный интервал рассчитывали по критерию Хи-квадрат с поправкой Йейтса.

Результаты и обсуждение

Для выбора стратегии клонирования целевых генов мы проанализировали нуклеотидные последовательности, кодирующие VP1 норовирусов генотипов GI.4, GI.17 и GI.3, на наличие редких кодонов относительно предполагаемых систем экспрессии, чтобы оценить потенциальные сложности их экспрессии в выбранных системах синтеза белка (табл. 1).

Значения параметра частоты встречаемости редких кодонов для потенциальных продуцентов находились в пределах оптимума (<30%), тем не менее показатели прогнозирования уровня экспрессии гена на основе его кодового состава (ИАК) были ниже приемлемых (<0,8). Характеристика целевых нуклеотидных мишеней по данным показателям теоретически свидетельствовала об ожидаемом подавлении экспрессии в выбранных белок-синтезирующих системах и необходимости подбора специальных условий для получения

Таблица 1

Анализ нуклеотидных последовательностей, кодирующих VP1 НоВ разных генотипов

Генотипы НоВ	Длина НК-последовательности, п. о.	Доля редких кодонов, %	Индекс адаптации кодонов (ИАК)
GI.4	1648	13	0,60
GI.17	1644	14	0,60
GI.3	1656	12	0,61
Референс	—	<30	0,8–1

целевых полипептидов: использование штаммов *E. coli* BL21-Rosetta (с дополнительными тРНК для редких кодонов), вымывание из телец включений 8М мочевиной, добавление предельных концентраций индуктора синтеза белка (до 8мМ в среде), использование кодон-оптимизированных нуклеотидных последовательностей (НП), разработка адаптированной бесклеточной системы синтеза белка.

Анализ аминокислотных последовательностей VP1 НоВ разных генотипов при использовании алгоритма предсказания антигенных детерминант (АГД) Коласкара и Тонгаонкара, базирующемся на первичной структуре белка, позволил обнаружить 25 АГД локализованных в основном капсидном белке НоВ GI.4, 23 АГД — в VP1 НоВ GI.17 и 19 АГД — в VP1 НоВ GI.3. Для экспрессии были выбраны последовательности 541 аминокислотных (АМК) остатков (GI.17 и GI.4) и 545 АМК остатков (GI.3) с рейтингом антигенности формируемых полипептидов — 1,030–1,032. Поскольку НоВ является безоболочечным вирусом, анализ локализации сайтов гликозилирования

не проводили. Дисульфидных связей в кандидатных полипептидах обнаружено не было, в связи с чем при отсутствии основных посттрансляционных модификаций для получения потенциальных антигенов могут быть использованы как бактериальные, так и дрожжевые системы экспрессии белка.

В связи с предполагаемыми сложностями в получении полноразмерных капсидных белков НоВ разных генотипов оценивали перспективы наработки более коротких полипептидов. Основным капсидный белок НоВ включает N-концевой оболочечный (S) и C-концевой выступающий (P) домены. S-домен формирует оболочку вокруг вирусной РНК, в то время как домены Р димеризуются, образуя выступы на поверхности капсида. Домен Р содержит субдомены Р1 и Р2. Находящийся на поверхности капсида Р2 субдомен гипервариабелен и содержит участок связывания с рецептором клетки-хозяина. Исходя из доменной структуры полноразмерного белка, анализировали антигенные свойства каждой функциональной единицы (табл. 2).

Таблица 2

Поиск потенциальных антигенов НоВ в пределах структурных единиц VP1

Параметр	GI.4		GI.17		GI.3	
	АМК остатки	ИА	АМК остатки	ИА	АМК остатки	ИА
Полноразмерный VP1	1-541	1,032	1-541	1,031	1-545	1,030
S-домен (N-концевой)	3-274 3-231*	1,044 1,046*	14-274 103-260*	1,042 1,058*	2-295 72-289*	1,033 1,049*
P-домен (C-концевой)	311-531 443-531*	1,027 1,055*	304-531 345-509*	1,030 1,037*	306-531 396-531*	1,028 1,040*

Примечание. АМК-последовательность с лучшим рейтингом антигенности в пределах анализируемой структурной единицы

По анализу спектра АГД, локализованных в пределах структурных единиц VP1 НоВ, были выбраны альтернативные полипептиды, предположительно обладающие антигенными свойствами. Отличаясь большей консервативностью полиаминокислотной цепи, фрагменты S-домена VP1 характеризовались и более высоким индексом антигенности (ИА), что обуславливает их предпочтительное использование в качестве антигенного компонента. Тем не менее, в соответствии с функцией Р-домена, являющегося мишенью для связывания нейтрализующими антителами (АТ), для тест-систем, направленных на изучение противовирусного иммунитета, стоит использовать белковые фрагменты, относящиеся к данной структурной единице.

На основании результатов анализа множественного выравнивания последовательностей фрагмента генома НоВ, кодирующего VP1, построенном в MEGA 7.0, установлено, что НП анализируемого региона разных генотипов вируса имеют низкую степень сходства, в связи с чем для НоВ GII.4, GII.17, GI.3 были разработаны три уникальных для каждого генотипа комплекта праймеров. Дизайн праймеров осуществляли на основании локальных нуклеотидных последовательностей норовирусов, полученных авторами статьи в рамках эпидемиологического надзора в Республике Беларусь и частично депонированных в GenBank (ON533460.1, GII.4; MZ269081.1, GII.17); также учитывались внутренние, недепонированные нуклеотидные последовательности. Разработку праймеров проводили с помощью ПО Primer3 (v.0.4.0) в соответствии со следующей стратегией клонирования:

- накопить фрагмент генома НоВ, фланкирующий кодирующий регион-мишень из проб биологического материала (фекалии);
- использовать продукты амплификации, полученные на первом этапе, в качестве матрицы для гнездовой ПЦР;
- интегрировать ПЦР-продукты второго раунда в промежуточный плазмидный вектор pJET1.2;
- использовать векторы pJET1.2_VP1 НоВ разных генотипов для клонирования белок-кодирующих фрагментов с целью их переноса в экспрессирующие векторы.

В первую очередь такая стратегия клонирования была обусловлена выбором сложной и неоднородной по составу матрицы пробы биологического материала; АТ-богатым регионом, фланкирующим мишень для клонирования, что теоретически обуславливает возможность неспецифичного связывания при условии присутствия множества ДНК/РНК матриц в пробе; смещением равновесия ПЦР-реакции в сторону образования продуктов с низкой молекулярной массой. Создание векторов-носителей фрагмента генома НоВ разных генотипов упрощает процедуру переноса мишени в разные экспрессирующие плазмиды за счет исключения использования биологической матрицы, насыщенной побочной ДНК/РНК.

В результате оценки таких параметров, как вырожденность, T_m , GC%, ΔG , GC клэмп, образование димеров и шпилек, разработано три комплекта праймеров и зондов для создания векторов-носителей фрагментов генома НоВ разных генотипов pJET1.2_VP1 НоВ (табл. 3).

Разработанные праймеры были проверены

Таблица 3

Схема олигонуклеотидных праймеров для клонирования и интеграции фрагментов VP1 НоВ в вектор pJET1.2

Генотип вируса	Название праймера	Последовательность, 5'→3'	Аmplифицируемый фрагмент, п. о.	Температура отжига, °C
1 раунд				
GII.4	GII.4_5081-F	GTGAATGAAGATGGCGTCGAGTGA	1776	59
	GII.4_6856-R	GCTGCTAAATTGGAAGGATGCTTGT		59
GII.17	GII.17_5007-F	GTCAATGTTCAGATGGATGAGGTT	1835	56
	GII.17_6841-R	AGGAAGCAGACTGGAGTTGTT		54

Окончание таблицы 3

Генотип вируса	Название праймера	Последовательность, 5'→3'	Аmplифицируемый фрагмент, п. о.	Температура отжига, °C
GI.3	GI.3_5338-F	TGATTTGAGTTTGTGGACAGGAGAT	1796	57
	GI.3_7133-R	CTTGCTGGAACCGCTGATGTT		57
2 раунд				
GII.4	GII.4_5085-F	ATGAAGATGGCGTCGAGTG	1671	53
	GII.4_6755-R	CAGAGCCAAGGACATCAGAT		53
GII.17	GII.17_5079-F	TGTGAATGAAGATGGCGTCGAATGA	1763	60
	GII.17_6841-R	AGGAAGCAGACTGGAGTTGTTTGTT		59
GI.3	GI.3_5373-F	TGCCCCGATTATGTAAATGATGATGG	1761	56
	GI.3_7133-R	CTTGCTGGAACCGCTGATGT		57

в BLAST на отсутствие значительной степени сходства и возможности неспецифического отжига на нуклеотидных последовательностях других организмов. Основные характеристики выбранных олигонуклеотидов: содержание GC в пределах 40–55%, рейтинг образования целевого ПЦР-продукта — не менее 77,3%.

В рамках создания экспрессирующих векторов праймеры были адаптированы для интеграции фрагментов, кодирующих VP1 НоВ, методом продолжительной перекрывающейся ПЦР (ПП-ПЦР). Использование данного мето-

да позволяет проводить безлигзную сборку конструкции вектор-мишень (рис. 1). Праймеры для ПЦР-амплификации фрагментов ОРС2 НоВ подбирали таким образом, чтобы после их амплификации 5'- и 3'-концы ПЦР-продукта содержали последовательности, комплементарные вектору (в табл. 4 комплементарные участки в праймерах подчеркнуты). В свою очередь, праймеры для линейаризации вектора, при встраивании в плазмиду, позволяли добавлять на 3'-конце фрагмента последовательность нуклеотидов, обеспечивающую наличие

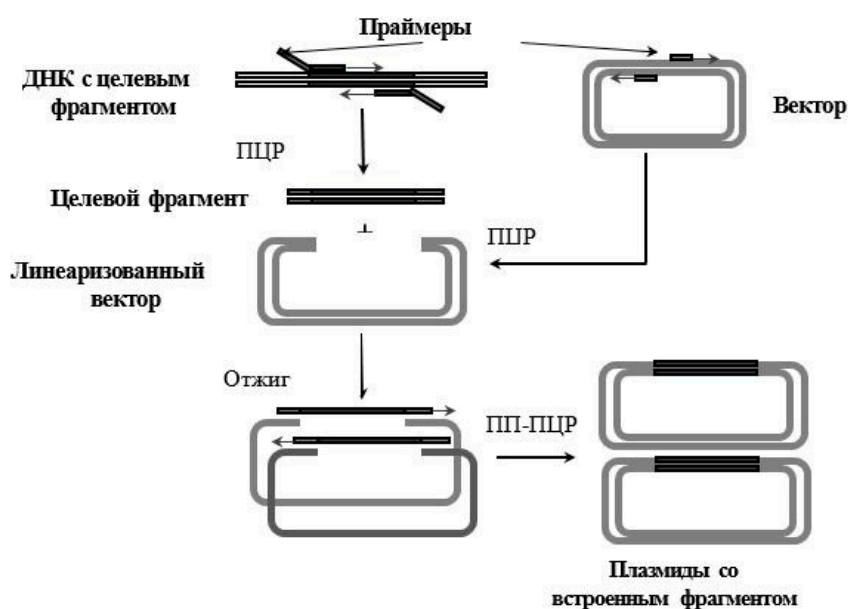


Рис. 1. Схема получения экспрессирующих векторов с фрагментами VP1 НоВ методом ПП-ПЦР

Таблица 4

Схема олигонуклеотидных праймеров клонирования фрагментов генома НоВ разных генотипов, кодирующих VP1 в составе экспрессирующих векторов

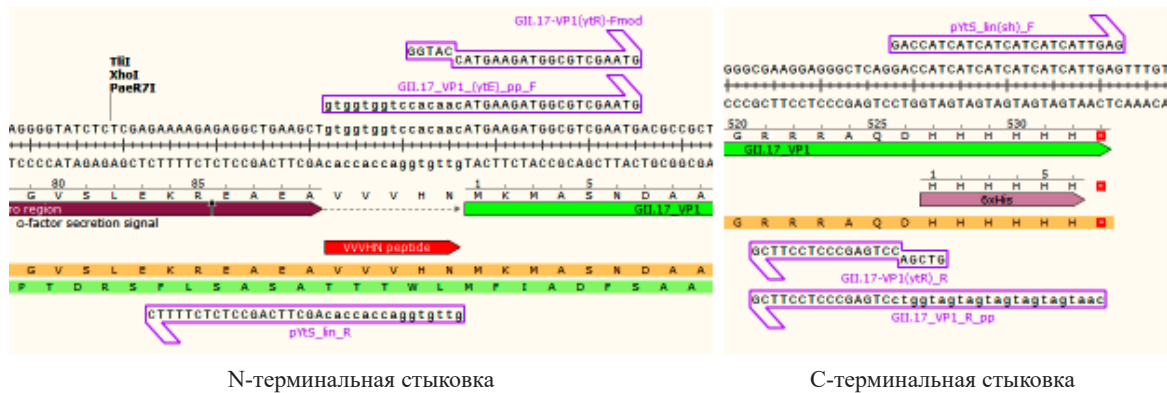
Вектор	Генотип НоВ/ назначение	Ориентация праймера	Последовательность, 5'→3'
pET42a	GII.4_VP1	прямой	<u>gtggtggtccacaacatgaagatggcgctcgag</u>
		обратный	<u>ggtgatggtgatgctctactacacgtctacgcc</u>
	GII.4_VP1Nd	обратный	<u>gtggtgatggtgatgctcgggtggcctaggagc</u>
	GII.4_VP1Cd	прямой	<u>gtggtggtccacaacatgtctcaaaattggaacaactatgacc</u>
	GII.17_VP1	прямой	<u>gtggtggtccacaacatgaagatggcgctcgaatg</u>
		обратный	<u>ggtgatggtgatgctcctgagccctccttcgc</u>
	GII.17_VP1Nd	обратный	<u>gtggtgatggtgatgctcgcttgaactcaccatcaag</u>
	GII.17_VP1Cd	прямой	<u>gtggtggtccacaacatgctgcaaacctcaatggtacaac</u>
	GI.3_VP1	прямой	<u>gtggtggtccacaacatgatgatggcgctaaaggac</u>
		обратный	<u>ggtgatggtgatgctctgatctcctgattcctagcct</u>
	GI.3_VP1Nd	обратный	<u>gtggtgatggtgatgctcaaccgtccccgc</u>
	GI.3_VP1Cd	прямой	<u>gtggtggtccacaacatgttaactgagattgatggcacc</u>
pColdI	GII.4	обратный	<u>gagcatcaccatcaccaccaccactaattg</u>
		прямой	<u>gttgtggaccaccacatattgtatatctctctt</u>
	GII.17	прямой	<u>gtggtggtccacaacatgaagatggcgctcgaatg</u>
		обратный	<u>gtgcttttaagcagagattacatttatactacacgtctacgcc</u>
	GI.3	прямой	<u>gtggtggtccacaacatgatgatggcgctaaaggac</u>
		обратный	<u>gtgcttttaagcagagattacattttatgatctcctgattccta</u>
	линеаризация	прямой	<u>ataggtaatctctgcttaaaagcacag</u>
		обратный	<u>gttgtggaccaccaccctaccttcgatatgatgatgatg</u>

гексагистидинового олигопептида на С-конце целевого полипептида. При проведении ПП-ПЦР вставка с вектором образуют смесь линейных конкатемеров, содержащих различное количество повторов вставка-вектор. При трансформации данной смесью линейных молекул в клетках-реципиентах происходит рекомбинация и формирование кольцевой молекулы-плазмиды, несущей в себе вставку.

Адаптацию олигонуклеотидных праймеров для безлигазного клонирования проводили в отношении основных выбранных векторов: pET42a (+), pColdI (табл. 4). В связи с ограничениями метода ПП-ПЦР, связанного с процессивностью ДНК-полимеразы, для вектора pET42a (5,5 kb) были также разработаны комплекты праймеров, фланкирующие фрагменты генома НоВ, кодирующие N- и С-концевые домены VP1.

Для обеспечения рациональной стратегии создания экспрессирующих векторов в конструкции ввели нуклеотидную последовательность N-терминального пептидного линкера, который служил регионом стыковки вектора и вставки, что, с одной стороны, позволило разработать универсальные праймеры для линеаризации каждого вектора, а с другой стороны, допустило использование прямого олигонуклеотида, комплементарного НП фрагмента генома НоВ, для интеграции во все векторы, используемые в данном исследовании (рис. 2).

Таким образом, в ходе проведенных исследований была определена оптимальная стратегия клонирования и интеграции фрагментов генома НоВ разных генотипов, на основании которой разработаны комплекты праймеров для накопления фрагментов генома НоВ раз-



N-терминальная стыковка

С-терминальная стыковка

Рис. 2. Схема интеграции фрагмента, кодирующего VP1 NoB GII.17, в вектор дрожжевой экспрессии pPICZαA

ных генотипов, их интеграции в вектор переноса pJET1.2, а также клонирования в составе экспрессионных векторов для бактериальной (pET42a(+), pColdI) системы синтеза белка.

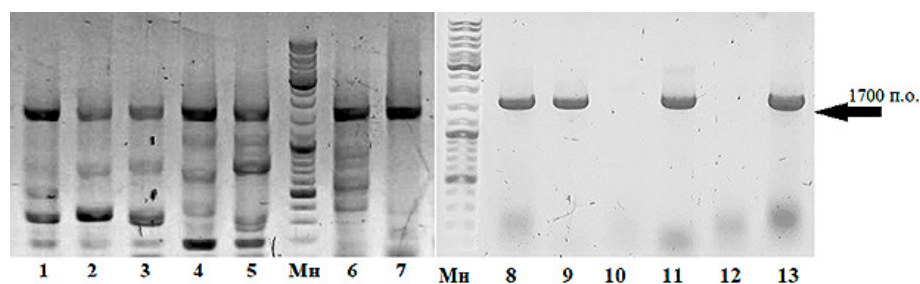
Присутствие неспецифических продуктов реакции (рис. 3, 1–7) свидетельствовало о необходимости проведения оптимизации амплификации. В первую очередь была подобрана подходящая температура отжига праймеров, которая составила 58 °C. Постановка реакции при более высокой температуре на данном этапе приводила к значительному снижению эффективности наработки целевого продукта. Для удаления избытка праймеров, способных связаться с побочной матрицей, их содержание в смеси было снижено до 15 пмоль. Концентрацию $MgCl_2$ в буфере оставили исходной, заявленной производителем (~4мМ), во избежание смещения равновесия реакции в сторону образования продуктов с низкой молекулярной массой. Изменение данных параметров проведения амплификации позволило значительно повысить специфичность накопления фрагментов генома NoB разных

генотипов (рис. 3).

Для одновременной проверки внутренних праймеров для клонирования в вектор-переноса и экспрессирующие векторы, а также соответствия полученных ПЦР-продуктов целевым фрагментам генома NoB разных генотипов, провели их секвенирование с использованием пар праймеров второго раунда. Результаты анализа нуклеотидных последовательностей, депонированных в GenBank, показали высокую степень их сходства (94,6–99,3%) с последовательностями генома NoB, изолированных из проб биологического материала.

Таким образом, в ходе проведенных работ получены фрагменты ДНК, содержащие последовательности NoB генотипов GII.4, GII.17, GI.3 для последующего клонирования.

На следующем этапе работы были сконструированы векторы переноса. Для этого коммерческий вектор pJET 1.2, содержащий ген резистентности к ампициллину в качестве селективного маркера, был линейаризован по сайту узнавания рестриктазой EcoRV (Eco32I). Вышеупомянутая эндонуклеаза по-

**Рис. 3.** Электрофореграмма образования продуктов реакции при клонировании фрагментов генома NoB разных генотипов. 1–7 — продукты амплификации до проведения оптимизации: 1–3 — GII.4, 4–5 — GI.3, 6–7 — GII.17; 8–13 — продукты амплификации после проведения оптимизации: 8–9 — GII.4, 10–11 — GI.3, 12–13 — GII.17

звolyет линеаризовать фрагменты по «тупым концам», что упрощает дальнейшее получение конечных плазмид, избегая обработки рестриктазами целевых фрагментов. Выбор вектора был обусловлен наличием вставки (около 1000 п. о.), при удалении которой исключается возможность самолигирования и, соответственно, рост «пустых» колоний. Анализ продуктов линеаризации вектора pJET 1.2 представлен на рисунке 4.

Полученные лигазные смеси (см. «Материалы и методы») трансформировали в компетентные клетки *E. coli* XL1-Blue; далее отбирали клоны для скрининга и верификации.

Результаты скрининга, представленные на рисунке 5, позволили отобрать колонии для выделения целевых плазмид. Верификацию правильности сборки векторных конструкций проводили секвенированием со специфичными праймерами к вектору pJET1.2 (см. «Материалы и методы»).

Из выбранных колоний клеток-трансформантов, после ночного культивирования (12 ч.) при 37 °С в жидкой питательной среде LB при добавлении ампициллина, были выделены плазмиды путем щелочного лизиса (рис. 6).

В результате были получены векторы переноса, содержащие фрагменты генома HoB разных генотипов, кодирующие VP1.

В рамках конструирования экспрессирующих векторов для получения белок-продуцирующих систем на основе прокариотических клеток были линеаризованы векторы pET42a и pColdI, а также получены фрагменты генома HoB разных генотипов, кодирующие VP1 или его N- или С-концевой домены для проведения безлигазной сборки конструкции. Результаты накопления фрагментов гена HoB, кодирующих разные полипептиды VP1, представлены на рисунке 7.

Продукты амплификации были интегрированы в экспрессирующие векторы, используемые в данном исследовании, методом ПП-ПЦР. Использование данного метода предполагало амплификацию смеси вектора и вставки (по 50 фмоль каждого) в реакционной смеси, адаптированной для Flash-полимеразы. Наличие комплементарных друг другу олигонуклеотидных последовательностей в составе вектора и вставки позволило провести сшивку без использования экзогенных праймеров.

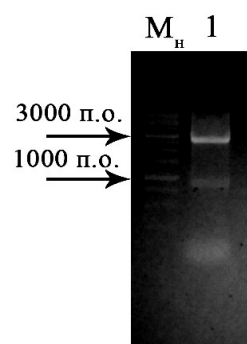


Рис. 4. Электрофореграмма продуктов линеаризации вектора pJet 1.2 (1) с использованием рестриктазы *EcoRV*

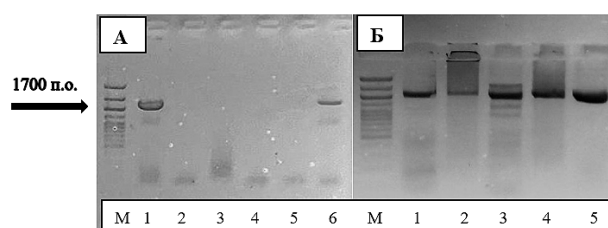


Рис. 5. Выявление специфических HoB-вставок при скрининге колоний: А — GI.3 и Б — GI.4 (1–3), GI.17 (4–5)

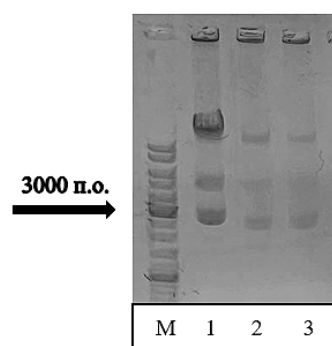


Рис. 6. Электрофореграмма векторов переноса, содержащих вставки генома HoB разных генотипов (1 — GI.17, 2 — GI.4, 3 — GI.3)

В результате проведения данного этапа были получены смеси линейных конкатемеров, содержащие слитый вектор и фрагмент гена VP1 HoB разных генотипов.

На следующем этапе проводили выделение и анализ векторов экспрессии, содержащих VP1-кодирующие фрагменты генома HoB. Из селективно выбранных и прошедших этап скрининга колоний клеток-трансформантов были выделены плазмиды методом щелочного

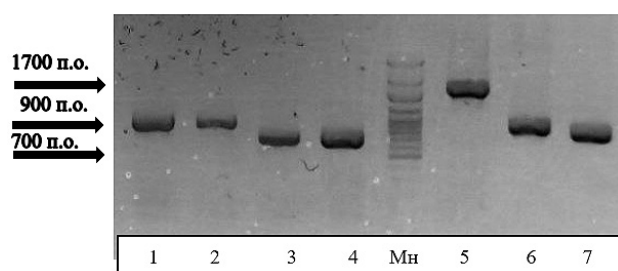


Рис. 7. Электрофореграмма продуктов амплификации фрагментов гена *VP1* НоВ на матрице векторов-носителей (pJET1.2). 1–2 — фрагменты генома НоВ GI.3, кодирующие С-концевой домен VP1; 3–4 — фрагменты генома НоВ GI.3, кодирующие N-концевой домен VP1; 5 — фрагмент генома НоВ GII.17, кодирующий полный VP1; 6, 7 — фрагменты гена *VP1* НоВ GII.17, кодирующие С- и N-концевые домены соответственно

лизиса. Характеристика изолированных плазмид, содержащих VP1-кодирующие фрагменты генома, представлена в таблице 5.

Для каждой из вышеописанных плазмид было получено несколько клонов, нуклеотидные последовательности которых были проанализированы секвенированием, начиная

от промотора до терминатора, с целью верификации отсутствия мутаций, сдвигов рамки считывания.

Представленные данные демонстрируют возможность получения плазмидных клонов экспрессирующих векторов на основе pET42a(+) и pColdI, содержащих нуклеотидные последовательности, кодирующие полноразмерные капсидные белки норовируса человека геновариантов GI.3, GII.4 и GII.17, с подтверждением корректной интеграции секвенированием. Дизайн праймеров выполнялся по актуальным последовательностям циркулирующих изолятов в Республике Беларусь. Безлигазная ПП-ПЦР-сборка с линейризацией векторов обеспечила воспроизводимую интеграцию вставок, что отражено в согласовании ожидаемых и полученных размеров ампликонов и плазмид.

Сконструированные векторы планируется использовать для получения рекомбинантных антигенов норовирусов и последующей оптимизации условий экспрессии и очистки, а также для прикладных задач серологической диагностики.

Таблица 5

Параметры плазмид, несущих целевые фрагменты НоВ разных генотипов

Название плазмиды	Размер кольцевой ДНК, п. о.	Соответствие изоляту NCBI	Расположение гистидинового тега	Теоретическая масса синтезируемого белка, кДа
pET42-GI.3-VP1	6656	MZ462931.1 99,10%	8His С-конец	59,8
pET42-GII.4-VP1	6644	OP901694.1 98,95%	8His С-конец	60,4
pET42-GII.17-VP1	6644	MK077711.1 98,82%	8His С-конец	60,6
pET42-GI.3-VP1Cd	5732	MZ462931.1 98,83%	8His С-конец	26,1
pET42-GI.3-VP1Nd	5852	MZ462931.1 98,96%	8His С-конец	32,9
pCold.I-GI.3-VP1	5998	MZ462931.1 99,10%	6His N-конец	61,0
pCold.I-GII.4-VP1	5986	OP901694.1 98,95%	6His N-конец	61,6
pCold.I-GII.17-VP1	5985	MK077711.1 98,82%	6His N-конец	61,8

Заключение

На основе коммерческих векторов pET42a(+) и pColdI созданы генетические конструкции, содержащие гены, кодирующие полноразмерные капсидные белки норовирусов человека геновариантов GI.3, GII.4 и GII.17. Полученные плазмидные ДНК в дальнейшем можно использовать для создания рекомбинантных штаммов-продуцентов целевых норовирусных белков.

Список использованных источников

1. Nordgren, J. Genetic Susceptibility to Human Norovirus Infection: An Update / J. Nordgren, L. Svensson // *Viruses*. – 2019. – Vol. 11, № 3 (226). – P. 1–19.
2. Tan, M. Norovirus Vaccines: Current Clinical Development and Challenges / M. Tan // *Pathogens*. – 2021. – Vol. 10, № 12. – P. 1641.
3. Robilotti, E. Norovirus / E. Robilotti, S. Deresinski, B. A. Pinsky // *Clin. Microbiol. Rev.* – 2015. – Vol. 28, № 1. – P. 134–164.
4. Yoon, S. H. Diagnostic Accuracy of Immunochromatographic Tests for the Detection of Norovirus in Stool Specimens: a Systematic Review and Meta-Analysis / S. H. Yoon, H. R. Kim, J. G. Ahn // *Microbiol. Spectrum*. – 2021. – Vol. 9, № 1. – P. e0046721.
5. Human Norovirus Cultivation in Nontransformed Stem Cell-Derived Human Intestinal Enteroid Cultures: Success and Challenges / M. K. Estes, K. Ettayebi, V. R. Tenge [et al.] // *Viruses*. – 2019. – Vol. 11, № 7 (638). – P. 1–12.
6. Protein Expression of the Human Norovirus Capsid Gene using the Baculovirus Expression System / J.-Y. Jin, C. J. Park, S.-W. Park, S.-Y. Paik // *Journal of Bacteriol. and Virol.* – 2011. – Vol. 41, № 3. – P. 183–187.
7. Heterologous expression of human norovirus GII.4 VP1 leads to assembly of T = 4 virus-like particles / J. M. Devant, G. Hofhaus, D. Bhella, G. S. Hansman // *Antiviral Res.* – 2019. – Vol. 168. – P. 175–182.
8. Structural heterogeneity of a human norovirus vaccine candidate / J. M. Devant, G. S. Hansman // *Virology*. – 2021. – Vol. 553. – P. 23–34.
9. Expression of Norovirus Virus-Like Particles in Different Systems / X.-J. Cheng, Y.-L. Wu, Y. Tanaka, W. Zhang // *ICCSET*. – 2014. – P. 537–540.
10. Expression and purification of norovirus virus like particles in *Escherichia coli* and their immunogenicity in mice / Y. Huo, X. Wan, T. Ling [et al.] // *Molec. Immun.* – 2018. – Vol. 93. – P. 278–284.
11. Ford-Siltz, L. A. Understanding the relationship between norovirus diversity and immunity / L. A. Ford-Siltz, K. Tohma, G. I. Parra // *Gut Microbes*. – 2021. – Vol. 13, № 1. – P. 1900994.
12. Генотипический состав циркулирующих в Республике Беларусь норовирусов / С. К. Лозюк [и др.] // *Современные проблемы инфекционной патологии человека* [Электронный ресурс]: сб. науч. тр. / М-во здравоохран. Респ. Беларусь. РНПЦ эпидемиологии и микробиологии; под ред. Л. П. Титова. – Минск: ГУ РНМБ, 2016. – Вып. 9. – С. 21–25.
13. Sambrook, J. Molecular Cloning: A Laboratory Manual / J. Sambrook, D. W. Russell // *CSH Protoc.* – 2001. – Vol. 1. – 2344 p.
14. Quirk, T. J. Correlation and simple linear regression / T. J. Quirk, M. H. Quirk, H. F. Horton // *Excel 2016 for biological and life sciences statistics: a guide to solving practical problems* / T. J. Quirk, M. H. Quirk, H. F. Horton. – Switzerland, 2016. – P. 109–152.

I. Kazlouski^{1,2}, N. Paklonskaya¹, I. Belskaya¹, T. Amvrosieva¹, A. Zinchenko²

METHOD FOR DESIGNING GENETIC CONSTRUCTS CONTAINING THE GENES OF FULL-LENGTH CAPSID PROTEINS OF HUMAN NOROVIRUSES

¹The Republican Center for Hygiene, Epidemiology and Public Health
23 Filimonov 23 St., Minsk 220114, the Republic of Belarus
e-mail: kazlouski.illia@gmail.com

²State Scientific Institution
“Institute of Microbiology of the National Academy of Sciences of Belarus”
2 Kuprevich St., Minsk 220084, the Republic of Belarus

Nucleotide sequences encoding the full-length capsid proteins of human noroviruses (gene variants GI.3, GII.4, and GII.17) were cloned into the pColdI vector. Molecular biology methods confirmed the correct orientation of the inserts in the constructs pCold-GI.3-VP1, pCold-GII.4-VP1, and pCold-GII.17-VP1, as well as the absence of spontaneous mutations in their sequences. The plasmid clones obtained will be further used for integration into bacterial cells.

Keywords: norovirus, capsid protein, VP1, pColdI, recombinant DNA.

Дата поступления в редакцию: 09 июня 2025 г.

О. Д. Левданский, А. М. Молчун

СРАВНИТЕЛЬНЫЙ АНАЛИЗ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ДЛЯ СБОРКИ МАЛЫХ ГЕНОМОВ *DE NOVO*

Государственное научное учреждение
«Институт генетики и цитологии Национальной академии наук Беларуси»
Республика Беларусь, 220072, г. Минск, ул. Академическая, 27
e-mail: o.liaudanski@igc.by

В рамках исследования были разработаны наборы скриптов для *de novo* сборки цитоплазматических геномов растений с помощью программ SPAdes, ABySS и NOVOPlasty. Проведена *de novo* сборка митохондриальных и хлоропластных геномов растений по данным высокопроизводительного секвенирования с высоким и низким количеством прочтений. Наилучшее качество полученных сборок продемонстрировала программа NOVOPlasty, однако ее работа сопряжена со значительными временными затратами. Из относительно более быстрых программ более высокие параметры эффективности сборки продемонстрировала SPAdes. Полученные результаты будут в дальнейшем использованы для составления методических рекомендаций для осуществления *de novo* сборки малых геномов по данным высокопроизводительного секвенирования.

Ключевые слова: сборка цитоплазматических геномов растений, NGS, ABySS, SPAdes, NOVOplasty.

Для цитирования: Левданский, О. Д. Сравнительный анализ программного обеспечения для сборки малых геномов *de novo* / О. Д. Левданский, А. М. Молчун // Молекулярная и прикладная генетика: сб. науч. тр. / Ин-т генетики и цитологии НАН Беларуси; редкол.: А. В. Кильчевский (гл. ред.) [и др.]. – Минск, 2025. – Т. 39. – С. 138–144.

Введение

Развитие технологий высокопроизводительного секвенирования (NGS) сделало их гораздо более доступными как для исследователей, так и в прикладном здравоохранении, что привело к значительному росту потребностей в обработке данных NGS. Такой масштабный спрос, а также разнообразие данных NGS в зависимости от исследуемого организма, способа пробоподготовки и самого секвенирования и преследуемых целей, привели к возникновению широкого спектра биоинформатических программных пакетов, позволяющих интерпретировать и анализировать получаемые данные. Любая биоинформатическая обработка данных NGS включает в себя ряд этапов, и в настоящее время для каждого из них, вне зависимости от типа анализа, существует несколько альтернативных программ, положительно зарекомендовавших себя и активно применяемых исследователями [1]. Поскольку отдельные программы основаны на различных алгоритмах, то и результат обработки данных

с их помощью может отличаться, обуславливая важность поиска оптимальных для решения различных задач наборов биоинформатических инструментов. Кроме того, учитывая неоднородность анализируемых данных по качеству и вариативность исследуемых объектов, актуальность разработки наиболее подходящих программных конвейеров (так называемых пайплайнов) для конкретных задач не вызывает сомнения [2].

В настоящее время одним из наиболее востребованных типов биоинформатической обработки данных является *de novo* сборка полных последовательностей различных малых геномов (вирусы, бактерии, цитоплазматические геномы растений и животных). В рамках исследования планируется провести обработку данных NGS полных последовательностей цитоплазматических геномов растений (митохондриальные и хлоропластные геномы ячменя (~525,6 и ~136,5 тыс. п. н. соответственно) и сои (~513,8 и ~152,2 тыс. п. н. соответственно)) с различным количеством полученных

прочтений, осуществить их *de novo* сборку и оценить качество сборки по основным параметрам (количество, суммарная и максимальная длина контигов и N50). Для проведения *de novo* сборки будут использованы наиболее широко применяемые для этих целей ассемблеры SPAdes и ABySS, а также программа NOVOPlasty, специализирующаяся на малых геномах. Последующий анализ основных параметров сборки позволит разработать методические рекомендации для *de novo* сборки малых геномов по результатам высокопроизводительного секвенирования.

Материалы и методы

В исследование были включены данные высокопроизводительного секвенирования 47 образцов хлоропластной и митохондриальной ДНК, полученные на приборе MiSeq (Illumina) с использованием набора реагентов Nextera XT DNA Library Preparation Kit (Illumina) согласно протоколу производителя.

Анализ качества прочтений с последующим триммингом некачественных оснований проводили в программах FastQC (v0.11.9) и Trimmomatic (v0.39) соответственно [3, 4]. *De novo* сборку осуществляли с помощью программ ABySS (v2.3.9) [5], SPAdes (v3.13.1) [6] и NOVOPlasty (v4.3.5) [7]. Количество потоков при запуске для программ ABySS и SPAdes — 72, NOVOPlasty распараллеливание выполнения не поддерживает. Параметры сборки вычислялись с помощью QUAST (v5.2.0) [8]. Обработка всех образцов

осуществлялась с применением одного и того же пайплайна. В качестве «seed» файла для программы NOVOPlasty использовались фрагменты гена *rbcL* для хлоропластного генома и гена *cytb* для митохондриального. В случаях, когда с вышеуказанными «seed» файлами не удавалось осуществить сборку, в их качестве использовались отдельные контиги, полученные с помощью SPAdes.

Все использованные для обработки и анализа данных скрипты доступны по ссылке [9].

Результаты и обсуждение

На рисунке 1 представлены количества прочтений, полученные для каждого из образцов. В качестве порогового для разделения на две группы было выбрано медианное значение, составившее около 300 тыс. прочтений. Образцы с количеством прочтений более 300 тыс. были отнесены к категории с высоким количеством прочтений (Выборка 1 — 24 образца), а образцы с меньшим количеством — к категории с низким количеством прочтений (Выборка 2 — 23 образца).

Качество прочтений во всех исследуемых образцах было проанализировано с помощью программы FastQC. После чего были проведены тримминг и фильтрация прочтений по качеству и длине с последующей повторной проверкой качества. Несмотря на то, что для задач текущего исследования фильтрация и тримминг прочтений не являются необходимыми, их проведение нужно для моделирования условий реальной обработки данных с целью

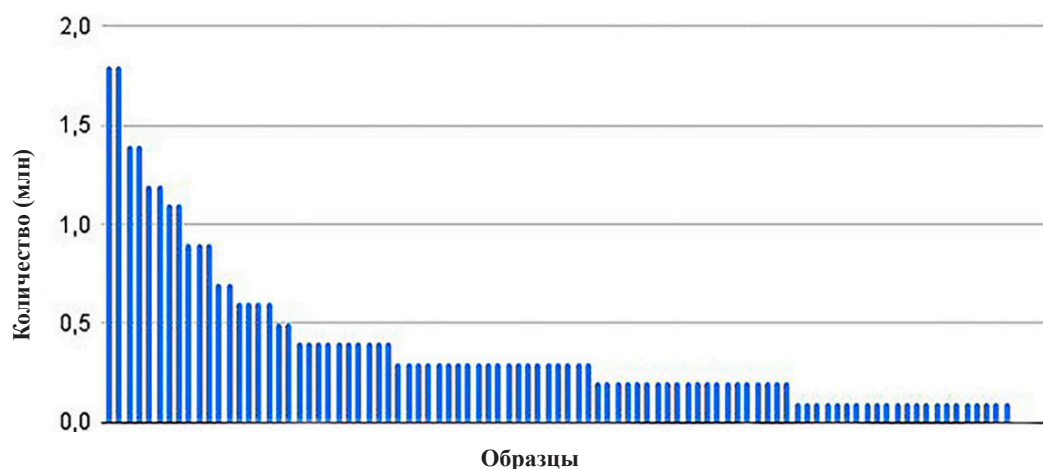


Рис. 1. Количество прочтений, полученное в результате высокопроизводительного секвенирования исследуемых образцов ДНК

анализа полученных контигов и скэффолдов. В результате работы программы FastQC было значительно улучшено среднее качество оснований и отфильтрованы последовательности, представленные большим числом копий, зачастую не имеющие отношения к целевой последовательности прочтения.

На рисунке 2 представлено распределение суммарной длины контигов, полученных с помощью программы ABySS. Основные параметры сборки приведены в таблице 1. Было выявлено значительное превышение среднего количества, средней длины наибольшего кон-

тига и средней суммарной длины в первой выборке по сравнению со второй. Такая разница во многом обусловлена двумя образцами, количество контигов в которых превысило 1200 и 2200 с суммарной длиной около 2 млн п. н. и 7,5 млн п. н. соответственно. Столь существенное превышение суммарной длины контигов над общей протяженностью целевой последовательности (около 660 тыс. п. н.) скорее всего обусловлено ошибками в работе программы. За вычетом этих двух образцов, лишь для двух образцов из первой выборки из оставшихся удалось достичь суммарной дли-

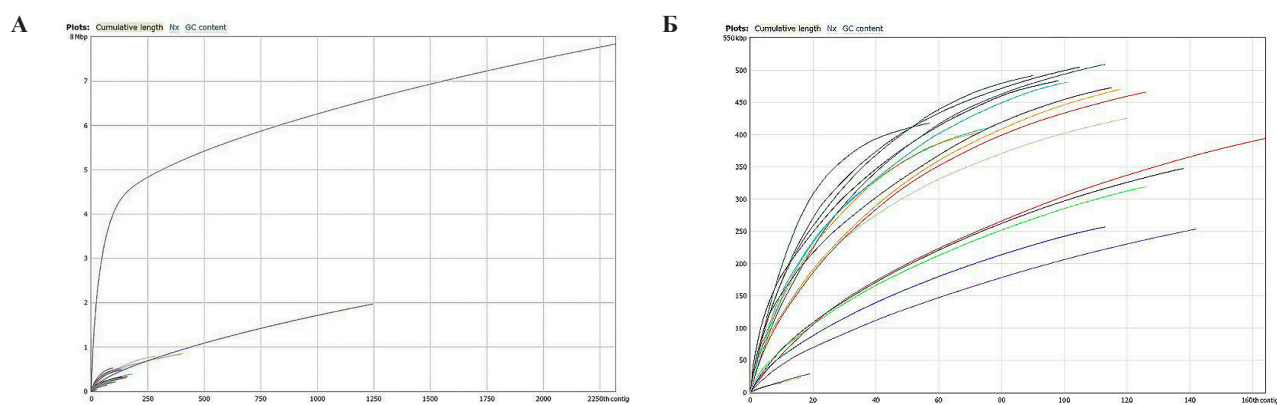


Рис. 2. Распределение суммарной длины контигов полученных с помощью программы ABySS для образцов с высоким (А) и низким (Б) количеством прочтений

Таблица 1

Основные показатели эффективности и времени *de novo* сборки геномов с помощью исследуемых программ (среднее значение $\pm$ стандартное отклонение)

Название программы	Качество прочтений	Количество контигов	Наибольший контиг (н. п.)	Общая длина (н. п.)	N50 (н. п.)	Время (мин)
SPAdes	Высокое	571,0 $\pm$ 986,5	117,8 $\pm$ 280,2	1736,5 $\pm$ 2979,4	14,8 $\pm$ 10,28	6,4 $\pm$ 10,5
	Низкое	68 $\pm$ 71,15	62,4 $\pm$ 30,9	470,2 $\pm$ 178,4	22,3 $\pm$ 13,8	0,9 $\pm$ 0,4
ABYSS	Высокое	258,6 $\pm$ 490,5	22,6 $\pm$ 42,4	716,9 $\pm$ 1532,7	3,8 $\pm$ 2,8	1,7 $\pm$ 1,5
	Низкое	91,9 $\pm$ 46,0	17,1 $\pm$ 11,6	342,4 $\pm$ 176,1	5,0 $\pm$ 3,2	0,8 $\pm$ 0,1
NOVOPlasty (митохондр.)	Высокое	32 $\pm$ 16,3	90,2 $\pm$ 46,8	908,7 $\pm$ 407,8	41,7 $\pm$ 20,3	7568,3 $\pm$ 3610,4
	Низкое	32,5 $\pm$ 23,9	124,6 $\pm$ 124,6	771,9 $\pm$ 397,9	94,9 $\pm$ 135,6	7101,9 $\pm$ 4671,3
NOVOPlasty (хлоропл.)	Высокое	7,7 $\pm$ 6,7	74,4 $\pm$ 36,9	292,2 $\pm$ 206,3	65,0 $\pm$ 37,5	1706,4 $\pm$ 1041,8
	Низкое	9,6 $\pm$ 6,7	82,0 $\pm$ 41,5	337,4 $\pm$ 179,9	67,2 $\pm$ 43,4	1570,4 $\pm$ 953,8

ны контигов, сопоставимой с длиной целевой последовательности. Для остальных, как и для всех образцов из выборки 2, суммарная длина не превышала 500 тыс. п. н.

На рисунке 3 представлено распределение суммарной длины контигов, полученных с помощью программы SPAdes. Как и в случае с ABySS, было показано значительное превышение аналогичных параметров для первой выборки относительно второй, обусловленное несколькими образцами (в данном случае тремя), количество контигов в которых превысило 2200, 2500 и 4000, с суммарной

длиной около 5, 4,5 и 14 млн п. н. соответственно. Однако, в отличие от ABySS, даже за исключением очевидных выбросов в некоторых образцах обеих выборок удалось достичь суммарной длины, сопоставимой с длиной целевой последовательности. В целом все анализируемые показатели для программы SPAdes значительно превышали таковые для ABySS. И если в отношении количества полученных контигов это скорее минус, то средняя длина и N50 однозначно свидетельствуют в пользу того, что сборка с помощью SPAdes является более эффективной.

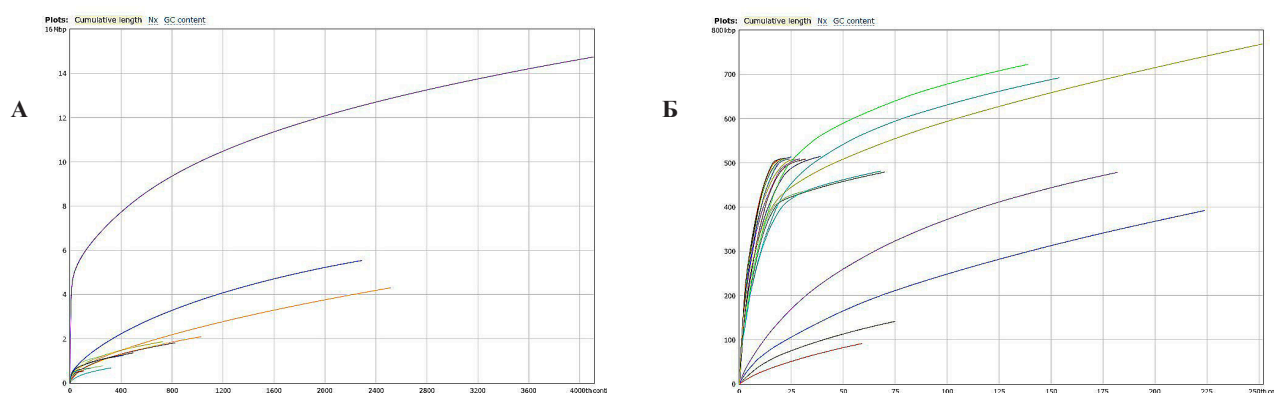


Рис. 3. Распределение суммарной длины контигов полученных с помощью программы SPAdes для образцов с высоким (А) и низким (Б) качеством прочтений

В силу особенностей работы программы NOVOPlasty, а именно необходимости для сборки так называемого «seed» файла, соответствующего целевой последовательности, сборка хлоропластного и митохондриального геномов осуществлялась отдельно. На рисунке 4 представлено распределение суммарной длины контигов мтДНК, полученных с помощью программы NOVOPlasty. В данном случае результаты выглядят гораздо более однородно по сравнению с программами ABySS и SPAdes. Суммарная длина в большинстве образцов превышает размер целевой последовательности, что позволяет потенциально рассчитывать на полное ее покрытие контигами. Значительно большие размеры контигов и значение N50 упростят дальнейшую сборку контигов в скэффолды. Кроме того, для трех образцов из второй выборки удалось собрать закольцованные последовательности, которые потенциально представляют собой полную це-

левую последовательность.

На рисунке 5 представлено распределение суммарной длины контигов хпДНК, полученных с помощью программы NOVOPlasty. Как и в случае с мтДНК, данные выглядят довольно однородно и предполагают потенциально полное покрытие целевой последовательности. Среднее количество контигов, не превышающее 10, и еще более близкие к размеру целевой последовательности значения N50 в значительной степени упростят дальнейшую сборку генома. Так же для четырех образцов из первой выборки и одного из второй удалось собрать закольцованные последовательности, которые потенциально представляют собой полную целевую последовательность.

Как видно из таблицы 1, NOVOPlasty продемонстрировал значительно более высокие показатели эффективности сборки по сравнению с SPAdes и ABySS, однако на обработку данных с его помощью уходит на порядки больше

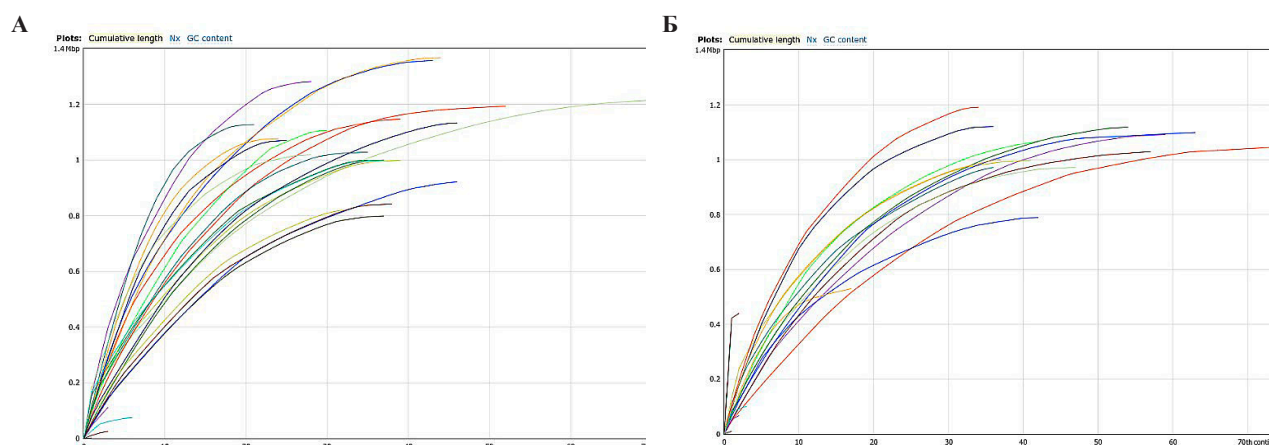


Рис. 4. Распределение суммарной длины контигов мтДНК полученных с помощью программы NOVOPlasty для образцов с высоким (А) и низким (Б) количеством прочтений

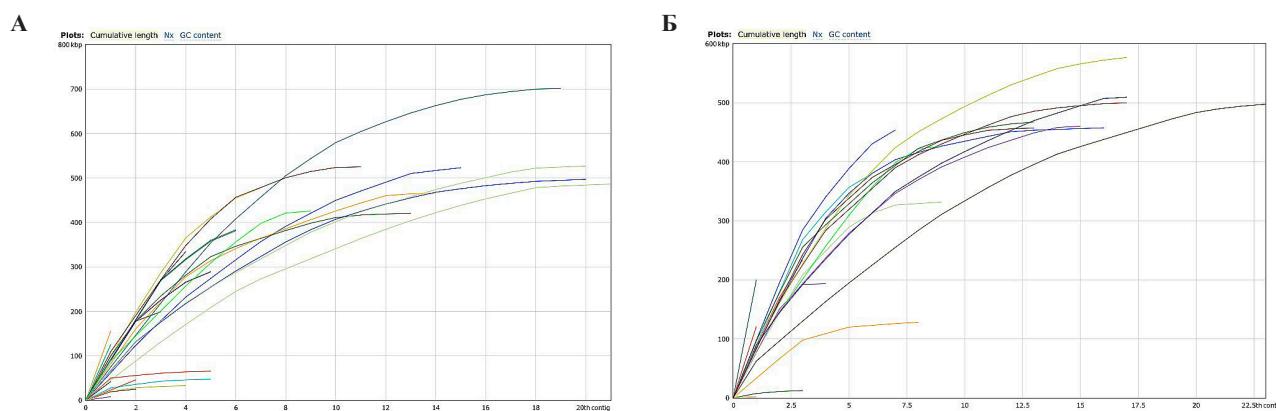


Рис. 5. Распределение суммарной длины контигов хпДНК полученных с помощью программы NOVOPlasty для образцов с высоким (А) и низким (Б) количеством прочтений

времени. Во многом это связано с невозможностью ее запуска в несколько потоков, что обусловило бы снижение разницы при запуске SPAdes и ABySS с использованием меньшего количества потоков, однако она все еще оставалась бы весьма значительной. Также следует учесть, что в таблице представлены параметры времени успешных запусков NOVOPlasty. При этом в процессе обработки с его помощью далеко не всегда с первого раза удастся подобрать подходящий «seed» файл, что приводит к необходимости повторного запуска программы. Таким образом, итоговое время обработки может в несколько раз превысить ожидаемое, исходя из таблицы. Если же говорить об относительных преимуществах SPAdes и ABySS, то первый продемонстрировал сравнительно более высокие показатели эффективности при

не слишком больших временных затратах.

Заключение

Таким образом, в рамках выполнения задания были разработаны наборы скриптов для *de novo* сборки цитоплазматических геномов растений с помощью программ SPAdes, ABySS и NOVOPlasty. Проведена *de novo* сборка митохондриальных и хлоропластных геномов растений по данным высокопроизводительного секвенирования с высоким и низким количеством прочтений. Наилучшее качество полученных сборок продемонстрировала программа NOVOPlasty, однако ее работа сопряжена со значительными временными затратами. Из относительно более быстрых программ более высокие параметры эффективности сборки продемонстрировал SPAdes.

Полученные результаты будут в дальнейшем использованы для составления методических рекомендаций для осуществления *de novo* сборки малых геномов по данным высокопроизводительного секвенирования.

Коллектив авторов выражает благодарность сотрудникам лаборатории нехромосомной наследственности Института генетики и цитологии НАН Беларуси Синявской М. Г. и Александрович В. В. за предоставленные данные высокопроизводительного секвенирования.

Работа выполнена в рамках ГПНИ «Биотехнологии-2» 2021–2025 гг., подпрограмма 2 «Геномика, эпигеномика, биоинформатика».

Список использованных источников

1. Koboldt, D. C. Best practices for variant calling in clinical sequencing / D. C. Koboldt // *Genome Medicine*. – 2020. – Vol. 12:91. – P. 1–13.
2. Pereira, R. Bioinformatics and Computational Tools for Next-Generation Sequencing Analysis in Clinical Genetics / R. Pereira, J. Oliveira, M. Sousa // *Journal of Clinical Medicine*. – 2019. – Vol. 9:132. – P. 1–30.
3. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline / G. A. Van der Auwera, M. O. Carneiro, C. Hartl [et al.] // *Curr Protoc Bioinformatics*. – 2014. – Vol. 11:1110. – P. 1–43.
4. Bolger, A. M. Trimmomatic: a flexible trimmer for Illumina sequence data / A. M. Bolger, M. Lohse, B. Usadel // *Bioinformatics*. – 2014. – Vol. 30. – № 15. – P. 2114–2120.
5. ABySS: A parallel assembler for short read sequence data / J. T. Simpson, K. Wong, S. D. Jackman [et al.] // *Genome Research*. – 2009. – Vol. 19. – P. 1117–1123.
6. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing / A. Bankevich, S. Nurk, D. Antipov [et al.] // *Journal Of Computational Biology*. – 2012. – Vol. 19. – P. 455–477.
7. Dierckxsens, N. NOVOPlasty: *de novo* assembly of organelle genomes from whole genome data / N. Dierckxsens, P. Mardulyn, G. Smits // *Nucleic Acids Research*. – 2017. – Vol. 45. – № 4:e18. – P. 1–9.
8. QUAST: quality assessment tool for genome assemblies / A. Gurevich, V. Saveliev, N. Vyahhi, G. Tesler // *Bioinformatics*. – 2013. – Vol. 29. – № 8. – P. 1072–1075.
9. Github. – URL: https://github.com/Sector-Bioinf/de_novo_assembly (дата обращения 15.09.2025).

A. D. Liaudanski, A. M. Molchun

COMPARATIVE ANALYSIS OF *DE NOVO* SMALL GENOME ASSEMBLY SOFTWARE

State Scientific Institution

“Institute of Genetics and Cytology of the National Academy of Sciences of Belarus”

27 Akademicheskaya St., 220072 Minsk, the Republic of Belarus

e-mail: o.liaudanski@igc.by

During the study, script sets for the *de novo* assembly of plant cytoplasmic genomes using the SPAdes, ABySS and NOVOPlasty programs were developed. The *de novo* assembly of plant mitochondrial and chloroplast genomes was performed using high-throughput sequencing with high- and low-output sequencing data. The best quality of the obtained assemblies was demonstrated by the NOVOPlasty program; however, its operation is associated with significant time costs. Among the relatively faster programs, SPAdes demonstrated higher assembly efficiency parameters. The obtained results will be further used to compile methodological recommendations for the *de novo* assembly of small genomes based on high-throughput sequencing data.

Keywords: plant cytoplasmic genome assembly, NGS, ABySS, SPAdes, NOVOPlasty.

Дата поступления в редакцию: 18 сентября 2025 г.

Я. И. Шейко, А. М. Слуквин, С. С. Ниязова, Р. И. Шейко

ИЗУЧЕНИЕ КЛЕТОЧНЫХ ПАРАМЕТРОВ С ПОМОЩЬЮ СИСТЕМЫ НЕЙРОСЕТЕВОГО АНАЛИЗА ИЗОБРАЖЕНИЙ КЛЕТОЧНЫХ КУЛЬТУР

Государственное научное учреждение
«Институт генетики и цитологии Национальной академии наук Беларуси»
Республика Беларусь, 220072, г. Минск, ул. Академическая, 27
e-mail: y.sheiko@igc.by

На основе нейросетевого алгоритма создана система, позволяющая сегментировать (распознавать) клетки на изображениях клеточных культур и измерять клеточные параметры, такие как площадь, периметр, форма, длина, ширина и яркость. После сегментации и измерения клеточных параметров система статистически обрабатывает данные с последующей оценкой достоверности различий в изучаемых выборках. В качестве объектов для апробации работы системы использовались изображения живых неокрашенных культур фибробластов человека и мыши, а также стандартных клеточных линий человека. В результате нейросетевого анализа изображений получены количественные данные, позволяющие сравнить различные клеточные культуры, а также выявить динамику изменений клеточных параметров в процессе роста или после воздействия биологически активных веществ в изучаемых культурах. Использование нейросетевого анализа позволило дополнить и уточнить полученные нами ранее данные исследования клеточных культур.

Ключевые слова: нейронные сети, машинное зрение, сегментация клеток, анализ изображений, клеточные культуры, клеточные параметры, видеомикроскопия.

Для цитирования: Изучение клеточных параметров с помощью системы нейросетевого анализа изображений клеточных культур / Я. И. Шейко, А. М. Слуквин, С. С. Ниязова, Р. И. Шейко // Молекулярная и прикладная генетика: сб. науч. тр. / Ин-т генетики и цитологии НАН Беларуси; редкол.: А. В. Кильчевский (гл. ред.) [и др.]. – Минск, 2025. – Т. 39. – С. 145–155.

Введение

В последние годы технологии на основе нейронных сетей (НС) активно внедряются в биомедицинских исследованиях. Обученные нейросети обрабатывают большие массивы данных быстрее человека и нередко превосходят его по качеству анализа. Эти особенности обуславливают широкое применение НС в различных исследовательских и прикладных медико-биологических областях, таких как медицинская диагностика и биоинформатика [1–5]. Развитие архитектур глубоких сетей, таких как U-Net для сегментации биомедицинских изображений, и совершенствование алгоритмов обучения способствовали переходу от экспериментальных разработок к практическому внедрению НС в лабораторную практику [6]. Одной из наиболее востребованных функций НС является распознавание изобра-

жений, что делает их мощным инструментом для анализа разнообразных медицинских снимков: рентгенограмм, УЗИ, МРТ, данных эндоскопических исследований, препаратов тканей и др. [7]. Особое значение эти методы приобретают в цитологических исследованиях, где точная сегментация и количественный анализ клеточных параметров критичны для оценки физиологического состояния культур, изучения патологий и скрининга фармакологических субстанций [8, 9]. Кроме того, автоматизация при оценке медико-биологических изображений сокращает в десятки раз время проведения анализа и существенно повышает воспроизводимость результатов [10, 11]. Точность результатов анализа на основе НС постоянно повышается благодаря увеличению объемов обучающих данных и совершенствованию математического аппарата нейросетей,

что делает методы компьютерного зрения перспективными в медико-биологических исследованиях, в том числе в исследованиях клеточных культур.

Целью настоящей работы является апробация разработанной нейросетевой системы на изображениях клеточных культур человека и мыши.

Материалы и методы

Объектами исследования служили изображения живых неокрашенных клеточных культур фибробластов человека и мыши, а также стандартных клеточных линий человека Hela и A549, культивируемых по стандартной методике [12]. Изображения клеточных культур получали с помощью метода прижизненной видеомикроскопии без использования фазового контраста и флуоресценции, что обусловило низкую контрастность клеток по отношению к фону и, как следствие, усложнило задачу для нейросетевого распознавания клеток.

Для сегментации таких сложных, слабоконтрастирующих с фоном объектов, как неокрашенные клетки, было апробировано более десятка нейронных сетей (например, DeepCell, TissueNet, Mesmer, ResNet50, Unet-VGG19 и др.). В результате была выбрана предобученная нейросеть Cellpose [13], представляющая собой сверточную сеть (CNN) с архитектурой, оптимизированной для универсальной сегментации биомедицинских изображений. Cellpose обучена на обширном наборе данных, включающем разнообразные типы клеток, что обеспечивает ее высокую адаптивность к различным образцам. Для адаптации к специфике исследуемых изображений проводилось дообучение нейросети на собственных изображениях клеток (34 изображения, количество клеток на каждом от 35 до 350, всего на изображениях около 4500 клеток). Для каждой клетки из обучающего набора была создана маска путем ее ручной разметки в программе Labelme. Для дообучения изображения были разбиты на группы: ~3600 клеток в обучающей выборке, ~450 в валидационной и ~450 в тестовой). Дообучение осуществлялось в течение 100 эпох с использованием оптимизатора AdamW (начальная скорость обучения 0,0001, весовое затухание 0,0001, размер батча 3).

Программный код нейросети Cellpose был

доработан для измерения клеточных параметров в сегментированных клетках: Area (площадь в пикселях²), Perimeter (периметр в пикселях), Roundness (округлость, рассчитываемая как $4\pi \times \text{Area} / \text{Perimeter}^2$, где 1 соответствует идеальному кругу), Feret max и Feret min (максимальный и минимальный диаметр Фере, в пикселях — длина и ширина), Feret_ratio (отношение Feret max к Feret min — степень продолговатости клетки), Luminance (фотометрическая величина яркости клетки от темного к светлому в градации серого цвета). После измерения клеточных параметров программно выполняется статистический анализ с выводом следующих результатов: среднее, минимум, максимум, стандартное отклонение, медиана, I и III квартили, коэффициент вариации. Доработанная нейросеть способна выявлять статистически значимые различия средних значений параметров при попарном сравнении двух выборок (как зависимых, так и независимых) по t-критерию Стьюдента при уровне значимости $<0,05$, что может быть использовано для оценки динамики изменений в одной культуре или при сравнении двух клеточных культур.

Результаты и обсуждение

Предобученная нейросеть Cellpose была дообучена на собственных клеточных изображениях эмбриональных фибробластов человека и мыши (рис. 1).

Нейросеть при анализе изображений создает маски инстанс-сегментации клеток (рис. 2). Инстанс-сегментация позволяет нумеровать каждый сегментированный объект, что дает возможность различать объекты, даже если они частично перекрываются. Таким образом,

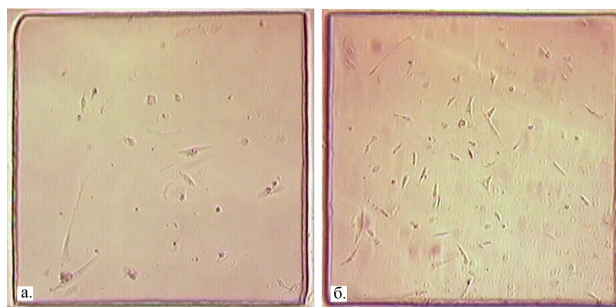


Рис. 1. Эмбриональные фибробласты человека (а) и мыши (б)



Рис. 2. Маски клеток, созданные нейросетью в результате инстанс-сегментации

сегментацию клеток можно проводить даже на изображениях плотных культур.

Маски клеток, созданные в процессе сегментации, имеют четко обозначенные контуры,

что позволяет измерять различные клеточные параметры, которые далее подвергаются статистической обработке (табл. 1).

Для апробации работы системы использовали собственные изображения культур фибробластов человека и мыши, а также стандартных клеточных линий, полученных с помощью компьютерной видеомикроскопии.

При сравнении клеточных параметров, полученных с помощью нейросетевого анализа изображений культур клеток человека разных лет исследований (для сравнения выбирались культуры на начальном этапе их роста — 1 пассаж), не выявлено статистически значимых отличий (табл. 2).

Так как в нескольких различных культурах фибробластов человека одинакового срока культивирования не выявлено различий в клеточных параметрах (значимые различия значе-

Таблица 1

Пример результата статистического анализа клеточных параметров сегментированных клеток на изображениях культуры фибробластов человека

	Area	Perimeter	Roundness	Feret_max	Feret_min	Feret_ratio	Luminance
mean	280,0	59,4	0,8	22,0	15,4	1,5	192,8
min	17,0	8,2	0,3	5,8	2,4	1,0	123,0
max	1805,0	204,8	3,2	67,8	50,9	8,5	209,7
std	284,3	31,7	0,2	11,1	7,7	0,7	10,9
median	199,0	56,3	0,9	20,5	14,3	1,3	195,2
Q. 25%	73,0	34,1	0,7	13,2	9,6	1,2	189,1
Q. 75%	353,0	72,5	1,0	26,6	19,4	1,6	200,4
CV%	101,5	53,4	27,7	50,6	50,3	47,3	5,7

Примечание. Результаты статистического анализа клеточных параметров: mean — среднее, min — минимум, max — максимум, std — стандартное отклонение, median — медиана, Q. 25% и Q. 75% — I и III квантили, CV% — коэффициент вариации

Таблица 2

Средние значения клеточных параметров сегментированных клеток в различных культурах фибробластов человека

№ выборки	Количество клеток	Area	Perimeter	Roundness	Feret_max	Feret_min	Feret_ratio	Luminance
1	65	235,5	58,1	0,8	21,1	13,1	1,6	193,7
2	233	280,0	59,4	0,8	22,0	15,4	1,5	192,8
3	92	294,4	70,7	0,7	29,0	13,6	2,1	202,3

ний параметров отмечаются системой после статистического анализа знаком «*»), то можно предположить, что клетки этих культур характеризуются статистически сходным паттерном параметров, что дает возможность сравнивать

эти клетки с другими типами клеток.

Подобное сравнение клеточных параметров в культурах фибробластов мыши также не выявило значимых отличий, кроме параметра Luminance — яркость клеток (табл. 3).

Таблица 3

Средние значения клеточных параметров сегментированных клеток в различных культурах фибробластов мыши

№ выборки	Количество клеток	Area	Perimeter	Roundness	Feret_max	Feret_min	Feret_ratio	Luminance
1	317	246,1	63,3	0,8	24,1	13,2	1,8	194,1
2	200	247,8	60,6	0,8	22,7	13,0	1,7	207,4*
3	167	256,2	66,3	0,7	25,7	13,1	1,9	194,0

Примечание. * — статистически значимое отличие значения параметра ($p < 0,05$) от предыдущего значения по t-критерию Стьюдента для независимых выборок

В таблице 3 одно значение (207,4 — Luminance в культуре фибробластов мыши) имеет статистически значимое отличие от предыдущей культуры фибробластов мыши. Однако следует обратить внимание, что этот параметр возможно учитывать только при обеспечении ряда одинаковых условий микрофотографирования: освещенность ростовой поверхности, цвет культуральных сред, фокусное расстояние объектива микроскопа и некоторых других. Указанные условия в различных экспериментах могут достаточно сильно различаться, поэтому не всегда удастся по этому параметру сделать определенные выводы.

С целью нахождения различий в культивируемых клетках человека и мыши было проведено сравнение параметров этих клеток, по-

лученных с помощью нейросетевого анализа (табл. 4).

Результаты статистического анализа, представленные в таблице 4, показали, что при сравнении культур фибробластов человека и мыши, помимо показателя яркости, статистически значимое отличие зарегистрировано в соотношении длины и ширины клетки — Feret_ratio (1,8 у фибробластов человека и 1,7 у фибробластов мыши). Чем больше значение этого параметра, тем более вытянута клетка, а у круглой клетки этот параметр равен 1. Таким образом, в данном сравнении выявлено, что клетки человека имеют более вытянутую форму. Эти данные согласуются с данными наших более ранних исследований [14], в которых сравнивались клеточные

Таблица 4

Сравнение средних значений клеточных параметров сегментированных клеток в культурах фибробластов человека и мыши

Количество клеток	Area	Perimeter	Roundness	Feret_max	Feret_min	Feret_ratio	Luminance
Параметры клеток в культурах фибробластов человека							
450	278,4	63,0	0,8	24,9	14,2	1,8	196,8
Параметры клеток в культурах фибробластов мыши							
905	261,7	62,5	0,8	23,2	14,0	1,7*	181,0*

Примечание. * — статистически значимое отличие значения параметра ($p < 0,05$) от предыдущего значения по t-критерию Стьюдента для независимых выборок

параметры (измеренные вручную) в культурах человека и мыши. Целью этих исследований была проверка данных литературы о наследовании формы клеток в клеточных поколениях. В результате было установлено, что форма фибробластов не сохраняется в клеточных поколениях. Вместе с тем было выявлено, что фибробласты человека значительно чаще принимают вытянутую (веретеновидную) форму (до 100%) по сравнению с фибробластами мыши (до 50%) (рис. 3).

Различие в доле времени пребывания клеток человека и мыши в веретеновидной форме, а также отсутствие митотической активности у крупных постаревших фибробластов человека в отличие от крупных фибробластов мыши, могут являться причинами резкого различия в частоте возникновения аномальных клеток и частоте спонтанной иммортализации в культурах фибробластов человека и мыши (в культуре фибробластов мыши данные явления возникают значительно чаще) [14].

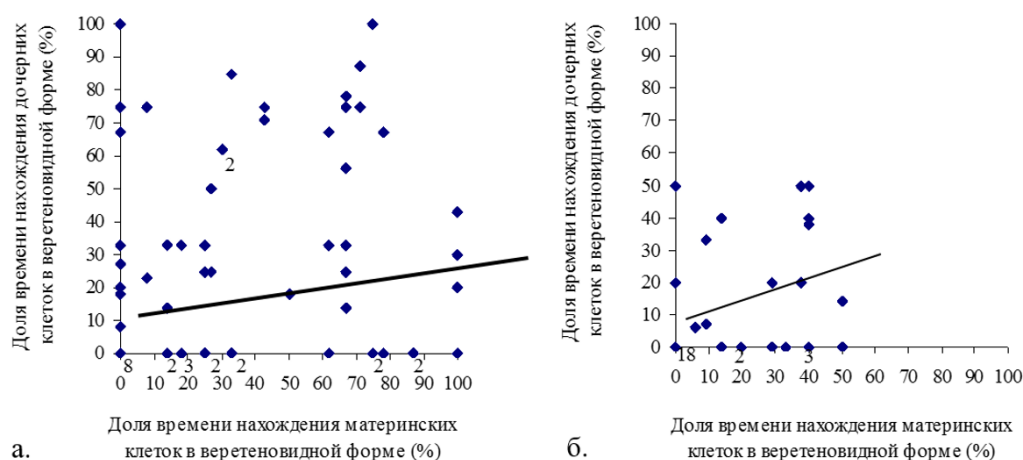


Рис. 3. Отсутствие зависимости доли времени пребывания в веретеновидной форме дочерних клеток от доли времени пребывания в веретеновидной форме материнских клеток: а — фибробласты человека, б — фибробласты мыши; цифры под точками показывают количество совпадающих парных сравнений «материнская – дочерняя клетка» во всей изученной клеточной популяции

Клеточные параметры культуры фибробластов человека также сравнивались с параметрами клеток стандартных раковых культур

человека Hela и A549. Сравнение выявило некоторые статистически значимые отличия, имеющие отношение к форме клеток (табл. 5).

Таблица 5

Сравнение средних значений клеточных параметров сегментированных клеток в культурах фибробластов человека и в линиях Hela и A549

Количество клеток	Area	Perimeter	Roundness	Feret_max	Feret_min	Feret_ratio	Luminance
Параметры клеток в культурах фибробластов человека							
450	278,4	63,0	0,8	24,9	14,2	1,8	196,8
Параметры клеток в культуре Hela							
88	259,5	58,3	0,9*	20,3*	14,5	1,4*	198,4
Параметры клеток в культуре A549							
733	285,5	67,2	0,7*	26,1	15,4	1,7	131,5*

Примечание. * — статистически значимое отличие значения параметра ($p < 0,05$) от соответствующего параметра фибробластов человека по t-критерию Стьюдента для независимых выборок

По данным таблицы 5, в культуре клеток Hela параметр «Округлость» (Roundness) (0,9 против 0,8 у нормальных фибробластов) выше, а параметры «Длина» (Ferret_max) (20,3 против 24,9 у нормальных фибробластов) и соотношение длины и ширины клетки (1,4 против 1,8 у нормальных фибробластов) ниже по сравнению с аналогичными показателями в культуре фибробластов человека. Таким образом, можно охарактеризовать клетки культуры Hela как более округлые и менее вытянутые по сравнению с клетками культуры нормальных фибробластов человека.

Клетки культуры A549 визуально отличаются от нормальных фибробластов более треугольной формой, что согласуется с более низким значением параметра «Округлость» (Roundness) (0,7 против 0,8 у нормальных фибробластов).

Наряду со сравнением различных культур по их клеточным характеристикам, также вызывает интерес изучение изменений клеток, происходящих в процессе роста одной культуры. В таблице 6 приведена динамика изменений средних значений параметров клеток в культуре фибробластов мыши на 1, 4 и 6-е сутки после посева. Клетки высевались на ростовую поверхность при крайне низкой плотности, что вызвало медленный рост числа клеток на начальных этапах культивирования. В первые сутки после посева в наблюдаемом участке ростовой поверхности были только две клетки.

По данным таблицы 6 в клеточной культуре фибробластов мыши зарегистрировано постепенное увеличение размеров клеток — на 6-е сутки увеличивается площадь (Area) — 356,9; периметр (Perimeter) — 71,9;

длина (Ferret_max) — 26,9 и ширина (Ferret_min) — 15,7 клеток (статистически значимые изменения). Изменение параметров формы (Округлость (Roundness) и соотношение длины и ширины клетки (Ferret_ratio)) статистически не достоверно. В данном анализе следует обратить внимание на статистически значимое изменение параметра Luminance. Так как анализируется один и тот же участок ростовой поверхности на протяжении относительно небольшого промежутка времени, то оптические условия имеют сходные значения в каждый момент наблюдений. Яркость клеток постепенно повышается, достигая на 6-е сутки статистически значимого отличия (199,7). Это может быть следствием уменьшения оптической плотности клеток, которая, в свою очередь, зависит от состава цитоплазмы и внутриклеточных компонентов. Однако в молодой культуре на протяжении относительно небольшого срока культивирования маловероятно накопление внутриклеточных изменений, способных приводить к изменению оптических характеристик клеток. Оптическая плотность также зависит от степени распластанности клеток — компактные округлые клетки имеют более высокую оптическую плотность (клетка менее прозрачна и выглядит темнее) и, наоборот, распластанная клетка становится визуально светлее. Таким образом, можно предположить, что повышение площади и периметра клеток при одновременном увеличении уровня яркости клеток (понижение оптической плотности) на протяжении 6-и суток роста клеточной культуры достигается не только за счет увеличения размеров клеток, но и за счет увеличения

Таблица 6

Изменение средних значений клеточных параметров сегментированных клеток в культуре фибробластов мыши на протяжении 6 суток

Сутки	Количество клеток	Area	Perimeter	Roundness	Feret_max	Feret_min	Feret_ratio	Luminance
1	2	149,0	48,4	0,8	18,6	9,7	1,8	174,0
4	19	160,8	52,2	0,7	17,8	11,8	1,6	187,1
6	57	356,9*	71,9*	0,8	26,9*	15,7*	1,7	199,7*

Примечание. * — статистически значимое отличие значения параметра ($p < 0,05$) от предыдущего значения по t-критерию Стьюдента для зависимых выборок

степени распластанности клеток.

Чтобы выяснить потенциальную возможность влияния на повышение уровня яркости клеток, изменение цвета культуральной среды (культуральная среда в процессе роста культуры может постепенно «выедается» и менять цвет) было произведено сравнение уровней яркости бесклеточных участков ростовой поверхности (табл. 7).

Как видно из таблицы 7 яркость бесклеточных участков на протяжении 6 суток наблюдений имеет очень близкие и статистически не различающиеся значения. Таким образом, оптические характеристики культуральной среды практически не влияют на изменение уровня яркости клеток на протяжении 6 суток наблюдения, что подтверждает предположение об увеличении степени распластанности клеток на начальных этапах культивирования. Явление распластанности может отражать функциональные особенности клеток — у распластанных клеток, по сравнению с компактными, более высокая площадь контакта как с подложкой, так и с культуральной средой.

Клеточные технологии имеют важное зна-

чение для оценки действия биологически активных веществ на клетки. При изучении действия рекомбинантного лактоферрина (ЛФ) человека на клеточные культуры в наших предыдущих исследованиях был выявлен дозозависимый ингибирующий эффект этого белка на спонтанно иммортализованную культуру из эмбриональных фибробластов человека [15]. Процесс иммортализации клеток является одним из этапов раковой трансформации, поэтому подобные клетки должны являться мишенью для иммунной системы. Лактоферрин — полифункциональный белок широкого спектра действия [16], в том числе и антиканцерогенного [17], играющий важную роль в иммунной системе. В проведенных нами исследованиях трехкратное введение лактоферрина в культуры иммортализованной линии эмбриональных фибробластов человека (введение ЛФ осуществлялось каждый раз после смены культуральной среды) приводило к уменьшению количества клеток, причем ингибирующий эффект возрастал при повышении концентрации препарата вплоть до полной гибели клеток [15].

Таблица 7

Сравнение средних значений уровня яркости бесклеточных участков ростовой поверхности в культуре фибробластов мыши

Сутки	Количество участков	Luminance
1	35	203,8
4	35	202,3
6	35	202,2

Для изучения характеристик клеток при различных концентрациях лактоферрина с помощью нейросетевой системы мы проанализировали соответствующие наборы изображений этого эксперимента (табл. 8).

По данным таблицы 8 обращает на себя внимание корреляция концентрации лактоферрина и размеров клеток (площадь и периметр), а также обратная корреляция количества выживших клеток от концентрации лактоферрина и размера клеток. Все изменения параметров площади и периметра статистически значимые. Подобные изменения размеров клеток имеют некоторое сходство со старением культуры. По мере старения культуры клетки

увеличиваются в размере и приобретают неправильную форму (распластанная со множеством выростов). В наших более ранних исследованиях [18] показано, что во фракции крупных клеток в спонтанно иммортализованной культуре фибробластов мыши в 3 раза выше доля окрашенных по SA- β -gal (ассоциированная со старением бета-галактозидаза, биомаркер клеточного старения) клеток по сравнению с фракцией мелких клеток. Таким образом, возможно, что действие лактоферрина на иммортализованную культуру имеет достаточно мягкий, растянутый во времени эффект — не немедленную гибель клеток, а постепенную утрату клетками иммортализи-

Таблица 8

Сравнение клеточных параметров в экспериментах с действием различных концентраций лактоферрина на иммортализированную линию эмбриональных фибробластов человека

	Количество клеток	Area	Perimeter	Roundness	Feret_max	Feret_min	Feret_ratio	Luminance
контроль	7803	439,9	79,1	0,9	27,9	20,3	1,4	136,5
10 мкг/мл	5935	564,3*	91,8*	0,8*	32,8*	23,8*	1,4	139,0*
100 мкг/мл	4094	778,8*	110,8*	0,8	40,4*	26,5*	1,5*	148,5*
1000 мкг/мл	27	1439,4*	140,9*	0,8	49,8*	27,9*	1,8	131,9*

Примечание. * — статистически значимое отличие значения параметра ($p < 0,05$) от предыдущего значения по t-критерию Стьюдента для независимых выборок

рованных свойств с последующим их старением и гибелью. В таблице 8 также обращает на себя внимание увеличение вытянутости клеток (Feret_ratio) при высоких концентрациях лактоферрина, однако статистически значимое отличие этого параметра (1,5) от предыдущего значения (1,4) отмечается только при концентрации 100 мкг/мл.

Еще одним потенциальным средством борьбы с трансформированными клетками являет-

ся препарат Деринат [19]. Воздействие на иммортализированную линию эмбриональных фибробластов человека препаратом Деринат в наших предыдущих экспериментах также приводило к дозозависимому ингибирующему эффекту.

Нейросетевой анализ наборов изображений одного из таких экспериментов представлен в таблице 9.

Как и в случае с лактоферрином, выявлена

Таблица 9

Сравнение клеточных параметров в экспериментах с действием различных концентраций препарата Деринат на иммортализированную линию эмбриональных фибробластов человека

	Количество клеток	Area	Perimeter	Roundness	Feret_max	Feret_min	Feret_ratio	Luminance
контроль	2447	305,9	66,0	0,9	23,7	16,5	1,4	172,5
0,1 мг/мл	971	513,7*	84,9*	0,9	30,2*	21,4*	1,4	175,2*
1 мг/мл	263	807,3*	100,9*	0,9	34,8*	26,8*	1,3*	179,3*

Примечание. * — статистически значимое отличие значения параметра ($p < 0,05$) от предыдущего значения по t-критерию Стьюдента для независимых выборок

статистически значимая корреляция размера клеток (площадь и периметр) иммортализированной культуры от концентрации препарата Деринат, а также обратная корреляция количества клеток от концентрации этого препарата и размеров клеток. Не исключено, что протестированные вещества вызывают быстрые изменения метаболизма клеток, приводящие к накоплению клеточных повреждений, подобно накоплению изменений, наблюдаемых при старении культуры. В отличие от действия на иммортализированную культуру лактофер-

рина, где показатель Feret_ratio растёт с концентрацией действующего вещества, при воздействии Дерината этот показатель немного снижается при максимальной концентрации препарата (1,3 при 1 мг/мл против 1,4 при 0,1 мг/мл и в контроле). Вероятно, это незначительное снижение не является отражением какого-либо биологического эффекта, тем не менее, возможно, отличие в тренде данного индекса отражает различные механизмы действия изучаемых препаратов.

Заключение

Разработанная система анализа изображений клеток позволяет за короткие временные промежутки, составляющие всего лишь десятки минут, получить данные о параметрах тысяч клеток. Подобные исследования без использования методов машинного зрения занимают недели, если не месяцы. Полученные цифровые данные легко поддаются статистической обработке, что позволяет выявлять динамику изменений в изучаемой культуре, а также сравнивать различные культуры между собой. В процессе апробации системы получена дополнительная уточняющая информация результатов предыдущих исследований клеточных культур. Так, выявленное различие параметра Feret_ratio в культурах фибробластов человека и мыши подтвердило полученные ранее данные о различии доли времени пребывания клеток этих культур в веретеновидной форме. Были установлены статистически значимые различия клеточных параметров при сравнении культур нормальных фибробластов человека и линий Hela и A549. По увеличению параметра Luminance в культуре фибробластов мыши на протяжении 6 суток был сделан вывод о возрастающей степени распластанности крупных клеток. Были обнаружены статистически значимые изменения параметров клеток в иммортализованной культуре при действии различных концентраций белка лактоферрина и препарата Деринат. Таким образом, разработанную нейросетевую систему можно использовать как в фундаментальных клеточных исследованиях, так и для решения практических задач, таких как оценка воздействия фармакологических или токсических веществ на клетки.

Список используемых источников

1. LeCun, Y. Deep learning / Y. LeCun, Y. Bengio, G. Hinton // *Nature*. – 2015. – Vol. 521, № 7553. – P. 436–444.
2. Krizhevsky, A. ImageNet classification with deep convolutional neural networks / A. Krizhevsky, I. Sutskever, G. E. Hinton // *Advances in Neural Information Processing Systems (NIPS)*. – 2012. – Vol. 25. – P. 1097–1105.
3. Dermatologist-level classification of skin cancer with deep neural networks / A. Esteva, B. Kuprel, R. Novoa [et al.] // *Nature*. – 2017. – Vol. 542, № 7639. – P. 115–118.
4. CheXNet: Radiologist-level pneumonia detection on chest radiographs using deep learning / P. Rajpurkar, J. Irvin, K. Zhu [et al.]. – arXiv. – 2017. – URL: <https://arxiv.org/pdf/1711.05225> (date of access: 01.07.2025).
5. Opportunities and obstacles for deep learning in biology and medicine / T. Ching, D. S. Himmelstein, B. K. Beaulieu-Jones [et al.] // *Journal of the Royal Society Interface*. – 2018. – Vol. 15, № 141. – Art. 20170387. – URL: <https://royalsocietypublishing.org/doi/10.1098/rsif.2017.0387> (date of access: 01.07.2025).
6. Ronneberger, O. Convolutional networks for biomedical image segmentation / O. Ronneberger, P. Fischer, T. Brox. – *Medical Image Computing and Computer-Assisted Intervention: Proceedings of the 18th International Conference, Munich, Germany, October 5–9, 2015*; eds.: N. Navab, J. Hornegger, W. M. Wells, A. Frangi. – Switzerland, 2015. – P. 234–241.
7. A survey on deep learning in medical image analysis / G. Litjens, T. Kooi, B. T. Bejnordi [et al.] // *Medical Image Analysis*. – 2017. – Vol. 42. – P. 60–88.
8. Shrestha, P. Efficient end-to-end learning for cell segmentation with machine generated weak annotations / P. Shrestha, N. Kuang, J. Yu // *Communications Biology*. – 2023. – Vol. 6. – URL: <https://www.nature.com/articles/s42003-023-04608-5> (date of access: 01.07.2025).
9. Wang, C. W. A survey of deep learning methods on cell instance segmentation / C. W. Wang, W. T. Lee, T. S. Su // *Neural Computing and Applications*. – 2025. – Vol. 37, № 17. – P. 11 195–11 264.
10. A multi-scale convolutional neural network for phenotyping high-content cellular images / W. J. Godinez, I. Hossain, S. E. Lazic [et al.] // *Bioinformatics*. – 2017. – Vol. 33, № 13. – P. 2010–2019.
11. Automatic Pancreas Segmentation Using Coarse-Scaled 2D Model of Deep Learning: Usefulness of Data Augmentation and Deep U-Net / M. Nishio, S. Noguchi, K. Fujimoto [et al.] // *Applied Sciences*. – 2020. – Vol. 10, № 3360. – URL: <https://doi.org/10.3390/app10103360> (date of access: 01.07.2025).
12. Human serum is an advantageous supplement for human dermal fibroblast expansion: clinical implications for tissue engineering of skin / A. L. Mazlyzam, B. S. Aminuddin, L. Saim, B. H. Ruszymah // *Archives of Medical Research*. – 2008. – Vol. 39,

№ 8. – P. 743–752.

13. Cellpose: a generalist algorithm for cellular segmentation / C. Stringer, T. Wang, M. Michaelos [et al.] // *Nature Methods*. – 2021. – Vol. 18. – P. 100–106.

14. Шейко, Я. И. Изучение клеточного фенотипа при старении и иммортализации клеток / Я. И. Шейко, О. В. Квитко, И. И. Конева. – Междунар. науч. конф. «От классических методов генетики и селекции к ДНК-технологиям», 2–5 октября 2007 года, Гомель, Республика Беларусь: 9-й съезд белорусского общества генетиков и селекционеров / НАН Беларуси [и др.]. – Гомель, 2007. – С. 144–145.

15. Исследование эффектов рекомбинантного лактоферрина человека на пролиферацию и апоптоз раковых и иммортализованных клеток / Я. И. Шейко, О. В. Квитко, И. И. Конева [и др.] // *Молекулярная и прикладная генетика: сб. науч. тр. / Институт генетики и цитологии НАН Беларуси; редкол.: А. В. Кильчевский (гл. ред.) [и др.]. – Минск, 2014. – Т. 18. – С. 77–83.*

16. Ward, P. P. Lactoferrin and host defense / P. P. Ward, S. Uribe-Luna, O. M. Conneely // *Biochemistry and Cell Biology*. – 2002. – Vol. 80, № 1. – P. 95–102.

17. Cancer prevention by bovine lactoferrin and underlying mechanisms – a review of experimental and clinical studies / H. Tsuda, K. Sekine, K. Fujita, M. Ligo // *Biochemistry and Cell Biology*. – 2002. – Vol. 80, № 1. – P. 131–136.

18. Активность бета-галактозидазы в стареющей и иммортализованной клеточных популяциях / О. В. Квитко, Я. И. Шейко, И. И. Конева, М. В. Анисович // *Весці НАН Беларусі. Серыя біял. навук.* – 2006. – № 4. – С. 53–57.

19. Биологические эффекты экзогенных полинуклеотидов и перспективы их использования в медицине / О. В. Квитко, А. С. Сапун, И. И. Конева [и др.] // *Молекулярная и прикладная генетика: сб. науч. тр. / Институт генетики и цитологии НАН Беларуси; редкол.: А. В. Кильчевский (гл. ред.) [и др.]. – Минск, 2012. – Т. 13. – С. 126–135.*

Ya. I. Sheiko, A. M. Slukvin, S. S. Niyazova, R. I. Sheiko

STUDY OF CELLULAR PARAMETERS USING NEURAL NETWORK ANALYSIS OF CELL CULTURE IMAGES

State Scientific Institution

“Institute of Genetics and Cytology of the National Academy of Sciences of Belarus”

27 Akademicheskaya St., 220072 Minsk, the Republic of Belarus

e-mail: y.sheiko@igc.by

Based on a neural network algorithm, a system that allows the segmentation (recognition) of cells in images of cell cultures and measures cellular parameters such as area, perimeter, shape, length, width, and brightness has been developed. After segmentation and measurement of cellular parameters, the system statistically processes data with a subsequent assessment of the reliability of differences in the studied samples. Images of live, unstained cultures of human and mouse fibroblasts, as well as standard human cell lines, were used as objects to test the system. As a result of the neural network image analysis, quantitative data were obtained, allowing the comparison of different cell cultures and revealing the dynamics of changes in cellular parameters during growth or after exposure to biologically active substances in the studied cultures. The use of neural network analysis made it possible to supplement and clarify the data we had previously obtained from the study of cell cultures.

Keywords: neural networks, machine vision, cell segmentation, image analysis, cell cultures, cellular parameters, video microscopy.

Дата поступления статьи: 26 августа 2025 г.

К ЮБИЛЕЮ АКАДЕМИКА КИЛЬЧЕВСКОГО АЛЕКСАНДРА ВЛАДИМИРОВИЧА, ГЛАВНОГО РЕДАКТОРА СБОРНИКА НАУЧНЫХ ТРУДОВ «МОЛЕКУЛЯРНАЯ И ПРИКЛАДНАЯ ГЕНЕТИКА»

С момента создания (2005 г.) и до настоящего времени главным редактором сборника научных трудов «Молекулярная и прикладная генетика» является известный ученый в области генетики, геномики, биотехнологии и селекции растений, доктор биологических наук, профессор, академик НАН Беларуси Александр Владимирович Кильчевский. Благодаря скрупулезной работе Александра Владимировича, курирующего выход каждого сборника, членов сформированной им редакционной коллегии и авторов публикаций, журнал стал популярным и востребованным среди ученых-генетиков Беларуси и зарубежья, стремящихся познакомить коллег с результатами своих исследований и закрепить за собой приоритетные права на результат.



17 августа 2025 г. Александр Владимирович Кильчевский отметил свой 70-летний юбилей. Он родился в городе Горки Могилёвской области в семье преподавателей Белорусской государственной сельскохозяйственной академии. Особая аура сообщества интеллектуальных, энергичных, любящих землю людей предопределила выбор профессии и научного будущего. В 1977 г. Александр Владимирович окончил Белорусскую государственную сельскохозяйственную академию и последующие 26 лет проработал в ней же. В 1982 г. он защитил кандидатскую диссертацию по специальности «генетика», а в 1988 г. возглавил первую в стране кафедру биотехнологии, созданную

при его непосредственном участии. В этот период А. В. Кильчевский занимался разработкой нового для Беларуси научного направления — изучением эколого-генетических основ селекции сельскохозяйственных культур, заложил принципы экологической оптимизации селекционного процесса, провел комплексные исследования по генетике накопления поллютантов и энергоэффективности сортов. Им были созданы модели сортов с заданным уровнем энергозатрат, разработаны методы оценки адаптивности генотипов к различным условиям среды. Под руководством Александра Владимировича начал работу крупнейший в Беларуси биотехнологический центр, на базе которого было организовано широко-масштабное производство оздоровленного посадочного материала растений для хозяйств и населения страны. Полученные результаты были обобщены в его докторской диссертации (1994 г.), а также отмечены Премией НАН Беларуси (1999 г.). В 2003 г. А. В. Кильчевский был избран членом-корреспондентом НАН Беларуси по специальности «биотехнология».

Многогранный талант Александра Владимировича проявился в 2004 г., когда он возглавил Институт генетики и цитологии НАН Беларуси. За 10 лет под его руководством институт стал ведущим научным центром Республики Беларусь в области молекулярной генетики и геномных биотехнологий, получили развитие такие новые направления исследований, как структурно-функциональная организация геномов растений, животных, микроорганизмов, человека, генетическая инженерия, биобезопасность, маркер-сопутствующая се-

лекция и паспортизация, онкогеномика, этногеномика, спортивная генетика. По его инициативе и при активном участии для обеспечения внедрения геномных разработок в практику при институте были созданы ключевые инновационные структуры: Республиканский центр геномных биотехнологий, Республиканский банк ДНК (получивший статус национального достояния), Центр коллективного пользования «Геном».

Этот период и последующие годы оказались плодотворными и в научной деятельности Александра Владимировича. Совместно с сотрудниками института проводились оригинальные исследования по сравнительной геномике пасленовых культур (томат, перец, баклажан, физалис) и других культур. При участии А. В. Кильчевского создано 78 сортов овощных культур, включенных в Государственный реестр. Последние годы под его руководством начаты работы по таким актуальным направлениям, как генетика долголетия и нутригеномика, ведется разработка геномных технологий для целей криминалистики и обеспечения безопасности. А. В. Кильчевский выступил инициатором создания в Институте генетики и цитологии НАН Беларуси Республиканского центра изучения микробиома, где проводятся междисциплинарные исследования микробиома человека, животных и растений для нужд медицины, сельского хозяйства, пищевой промышленности и экологии.

Александром Владимировичем создана научная школа, насчитывающая 4 доктора и 22 кандидата наук. Результаты исследований опубликованы в более чем 600 научных работах, среди которых — 4 учебника и 13 монографий, в т. ч. уникальная для СНГ 4-х томная монография «Генетические основы селекции растений», инициатором, соредактором и соавтором которой он является. Издание удостоено Премии НАН Беларуси (2015 г.). Научные заслуги Кильчевского А. В. стали основанием для избрания его академиком НАН Беларуси по специальности «геномика растений» (2017 г.).

В 2014 г. расширился масштаб научно-организационной деятельности Александра Владимировича Кильчевского. Он вошел в состав Президиума НАН Беларуси в должности главного ученого секретаря, а с 2019 г. назначен на должность заместителя Председателя Прези-

диума. Назначение способствовало тому, что Александр Владимирович, осуществляя координацию между различными учреждениями НАН Беларуси, поднял генетические исследования на междисциплинарный уровень.

Александр Владимирович активно участвует в жизни мирового научного сообщества. С 2007 г. по настоящее время он является председателем Общественного Объединения «Общество генетиков и селекционеров» Беларуси, национальным координатором Европейской биотехнологической ассоциации, членом европейского общества селекционеров и генетиков EUCARPIA, почетным членом Европейской биотехнологической ассоциации EBVTNA, членом редакционного совета ряда профильных научных журналов (BioDiscovery, EuroBiotech Journal, Proceedings of the Latvian Academy of Sciences, Генетика, Экологическая генетика, Вавиловский журнал генетики и селекции, Весці НАН Беларусі, серия биол. наук и др.).

За свою многолетнюю плодотворную деятельность Александр Владимирович Кильчевский награжден десятками премий и медалей, среди которых — медали Франциска Скорины, им. Н. И. Вавилова и С. И. Жегалова, серебряная медаль НАН Беларуси «За достижения в науке». Он является лауреатом премий НАН Беларуси, Евразийской международной премии «За огромный вклад в развитие генетики, геномики и биотехнологии растений в Республике Беларусь», Премии Союзного государства в области науки и техники. В 2024 г. ему присвоено высокое звание «Заслуженный деятель науки Республики Беларусь».

Все, кто работает рядом с Александром Владимировичем, знают его не только как талантливого ученого, организатора, педагога, но и как прекрасного человека, заряжающего окружающих своей позитивной деятельностью энергией, доброжелательностью и чувством юмора.

Отделение биологических наук НАН Беларуси, коллектив Института генетики и цитологии, коллеги, друзья, единомышленники сердечно поздравляют Александра Владимировича с юбилеем! Желают крепкого здоровья, успехов во всех начинаниях и новых свершений на благо науки, общества, государства.

ИНСТРУКЦИЯ ДЛЯ АВТОРОВ ПО ОФОРМЛЕНИЮ СТАТЕЙ

Редакция сборника научных трудов «Молекулярная и прикладная генетика» принимает к рассмотрению обзорные и экспериментальные статьи по результатам исследований, выполненных с использованием современных методов в области молекулярной и прикладной генетики, отражающие изучение фундаментальных генетических процессов на молекулярном, клеточном, организменном и популяционном уровнях. Особое внимание уделяется актуальным проблемам биоинформатики и системной биологии, биоинженерии, вопросам сохранения и рационального использования генетических ресурсов, а также исследованиям, основанным на междисциплинарном подходе.

Статьи публикуются на русском, белорусском и английском языках.

Сборник входит в Перечень научных изданий, рекомендованных ВАК для опубликования результатов диссертационных работ (для биологических, медицинских (медико-биологические аспекты), сельскохозяйственных наук), и реферируется и индексируется в РИНЦ (Российский индекс научного цитирования).

Решение о публикации принимается редакционной коллегией после рецензирования, учитывая научную значимость и актуальность представленных материалов.

ДЛЯ ОПУБЛИКОВАНИЯ РАССМАТРИВАЮТСЯ СЛЕДУЮЩИЕ ТИПЫ СТАТЕЙ

Экспериментальная статья

Основной тип статей, размещаемых в сборнике. Этот формат предназначен для публикации полученных авторами новых данных, представляющих существенный интерес для научного сообщества. Объем экспериментальной статьи не должен превышать 12 страниц. В список использованных источников следует включать не более 30 источников.

Обзорная статья

Авторы могут предложить для публикации обзорную статью по теме, представляющей интерес для широкого круга читателей. Объем обзорной статьи не должен превышать 15 страниц. В список использованных источников следует включать не более 50 источников.

Краткое сообщение

В исключительных случаях по решению редколлегии в сборнике публикуются краткие сообщения объемом не более 5 страниц, список литературы в этом случае не должен включать более 15 источников.

СТРУКТУРА СТАТЬИ

Статья набирается в текстовом редакторе Microsoft Word. Поля: верхнее и нижнее — 2 см, левое — 3 см, правое — 1,5 см. Текст набирается шрифтом Times New Roman, 12 пт, в одну колонку с одинарным межстрочным интервалом. Выравнивание — по ширине. Абзацный отступ — 1,25 см. Не допускается использование табуляции или пробелов для обозначения первой строки абзаца. *Автоматическая расстановка переносов отключена*. Все страницы должны быть пронумерованы. Для создания формул следует использовать только MathType.

1. УДК

Располагается в верхнем левом углу страницы.

2. Инициалы и фамилии авторов

Располагаются по центру, выделяются полужирным начертанием. Между инициалами и перед фамилией ставятся неразрывные пробелы (Ctrl + Shift + пробел) (И. И. Иванов).

3. Название статьи

Располагается по центру, набирается буквами как в предложении (первые буквы предложений будут заглавными, а остальные — строчными), полужирным начертанием. Переносы в заголовках не допускаются. Название статьи должно четко отражать суть работы, не содержать лишних слов. По возможности следует избегать узкой региональности, а данные о географии исследования давать в резюме.

4. Место работы авторов с почтовым адресом

Место работы авторов (аббревиатура *не допускается*; пример: Государственное научное учреждение «Институт генетики и цитологии Национальной академии наук Беларуси») располагается по центру, набирается прямым начертанием. На следующей строке указывается почтовый адрес учреждения (в следующем порядке: страна, индекс, город, улица, дом).

Если авторы работают в разных учреждениях, то после каждой фамилии и перед соответствующим местом работы надстрочными знаками проставляется цифровой индекс (1, 2 и т. д.).

5. Авторское резюме (аннотация)

Располагается после места работы авторов, набирается прямым начертанием.

Авторское резюме должно полно и понятно излагать основное содержание и результаты проведенного исследования. Если из названия статьи очевидны предмет, тема, цель, методы исследования, их не следует повторно излагать в резюме. Объем текста аннотации определяется содержанием публикации (объемом сведений, их научной ценностью и/или практическим значением) и должен быть в пределах 800–1000 знаков с пробелами.

6. Ключевые слова (5–10 слов)

Перечисляются после слов «**Ключевые слова:**». В конце ставится точка.

7. Основной текст статьи

Экспериментальная статья должна включать следующие разделы: «**Введение**», «**Материалы и методы**», «**Результаты и обсуждение**», «**Заключение**», «**Список использованных источников**». Обзорная статья и краткое сообщение в обязательном порядке должны содержать разделы: «**Введение**», «**Заключение**» и «**Список использованных источников**». Название остальных разделов выбирается на усмотрение авторов.

Названия разделов располагаются на отдельной строке по центру и выделяются полужирным начертанием. Изложение статьи должно быть ясным, без длинных исторических введений и повторений. Помимо общепринятых сокращений единиц измерения, физических, химических и математических величин и терминов (например, ДНК), допускаются аббревиатуры словосочетаний, часто повторяющихся в тексте. Все вводимые автором буквенные обозначения и аббревиатуры должны быть расшифрованы в тексте при их первом упоминании.

Вставка в текст символов (например, β , ϵ) производится только через опцию *Вставка/Символ*. Для набора надстрочных и подстрочных символов (C^2 , C_4) использовать меню *Шрифт/Надстрочный знак/Подстрочный знак*.

Общие правила набора текста

Не допускается использование буквы ё. За исключением написания в собственных именах (фамилиях, географических названиях) и для предупреждения неправильного опознания слова (всё, нёбо, совершённый). В русских текстах разделитель целой и дробной части числа — запятая.

Прямое (*не курсивное*) начертание должны иметь:

- 1) греческие символы (α , β , δ , φ , λ и др.) и знаки ($\%$, $\leq$, $\pm$, $\times$, $\neq$, ∞ , $\rightarrow$, $^\circ$, $\notin$ и др.);
- 2) единицы измерения (Вт, В, Дж, кг, м и пр.);
- 3) кириллические (т. е. русские) буквы (П, Л, Ц);
- 4) сокращения от русских или белорусских слов ($q_{\text{ср}}$);
- 5) обозначения химических элементов и соединений (Р, H_2O);
- 6) скобки () и запятые (,) в формулах и после них ($x^2 - 0,5b(1,5 + ca)$).

Курсивное начертание должны иметь названия генов, латинские (т. е. английские) буквы, обозначающие переменные (K , y , z , x , V , i , j), латинские названия семейств, родовые и видовые названия.

Без пробелов следует проставлять:

- 1) дефис (-) в сложных словах (минерал-индикатор, К-пространство) и при наращении падежных окончаний к цифрам и буквам (1-го, j -го);

- 2) короткое тире (–) (Ctrl + «–» на нумерационной клавиатуре) между цифрами, обозначающими пределы какой-либо величины (20–30 чел., 1980–1999 гг.);
- 3) знаки –, +, ≤, > в значении положительной или отрицательной величины, степени увеличения или уменьшения (+30, составляет >80);
- 4) знак % (80%);
- 5) скобки () и кавычки «» от заключенных в них слов (кавычки в в тексте на русском и белорусском языках должны быть одного рисунка — «елочки», в английском языке — “лапки”).

С пробелами проставлять:

- 1) тире (—) (Ctrl + Alt + «–» на нумерационной клавиатуре) с обеих сторон пробелом как знак препинания между словами;
- 2) математические знаки (–, +, ≤, ±, ·, ×, =, ≠ и др.) в формулах, уравнениях, когда они являются знаками арифметического действия ($R_1 + R_2 = 0,75$);
- 3) знаки №, § от следующих за ними чисел (№ 4);
- 4) многозначные числа, содержащие 5 и более знаков, разбиваются по три цифры справа на лево (10 000, 1 000 000, 1 888,88, но 1000);
- 5) сокращенные слова от имен собственных, к которым они относятся (ул. Академическая);
- 6) цифры от полных или сокращенных наименований (100 м, 2005 г.).

Требования к рисункам

Рисунки должны размещаться только после их первого упоминания в тексте статьи. **Не допускается** размещение рисунков в конце статьи (в последнем абзаце перед заключением). Если ссылка на рисунок включена в предложение, используется полное написание — рисунок 1; если слова заключаются в скобки, используется сокращение — (рис. 1).

Слово Рис. 1. и подрисовочная подпись располагаются на следующей строке после самого рисунка (**Рис. 1.** Название). Точка после названия не ставится.

Рисунки должны быть представлены в электронном виде отдельными файлами в следующих форматах: JPEG, TIFF, PNG. Название файла должно соответствовать номеру рисунка (Рис. 1, Рис. 2 и т. д.). Для отправки все файлы рисунков объединяются в одну архивную папку zip или rar. Все надписи на рис. **должны быть набраны** Times New Roman 10 пт.

Оформление таблиц

Таблицы должны быть реализованы средствами работы с таблицами редактора Microsoft Word. Не допускается вложение таблиц, созданных в других программах.

Таблицы должны быть пронумерованы и иметь названия, размещаться только после их первого упоминания в тексте статьи. Если ссылка на таблицу включена в предложение, используется полное написание — таблица 1; если слова заключаются в скобки, используется сокращение — (табл. 1).

Слово **Таблица 1** выравняется по правому краю и выделяется полужирным начертанием.

На следующей строке после «**Таблица 1**» следует название таблицы, которое набирается прямым начертанием и выравняется по центру (без абзацного отступа).

Примечание к таблице располагается на следующей строке после таблицы. Слово **Примечание** выделяется полужирным начертанием, после него ставится точка. Текст примечания следует за словом Примечание и начинается с прописной буквы. Точка после текста примечания не ставится.

Не допускается размещение таблиц в конце статьи (в последнем абзаце перед заключением).

Информация о соблюдении требований Нагойского протокола

Если работа связана с исследованиями, в которых в качестве объекта используются генетические ресурсы (растения, животные, микроорганизмы), полученные от зарубежных партнеров, необходимо дать ссылку на Международно признанный сертификат о соблюдении (Internationally Recognized Certificate of Compliance) Нагойского протокола к Конвенции о биологическом разнообразии, подтверждающий легальное получение данных ресурсов.

Информация о соблюдении биоэтических стандартов

Если работа связана с исследованиями, в которых в качестве объекта используются лабораторные животные, то необходимо указать, соблюдались ли международные, национальные и/или институциональные принципы ухода и использования животных.

Если в качестве объектов исследования выступает человек, то необходимо указать:

- соответствуют ли процедуры, выполненные в исследовании с участием людей, этическим стандартам институционального и/или национального комитета по исследовательской этике и Хельсинкской декларации и ее последующим изменениям или сопоставимым нормам этики;
- получено ли от каждого из включенных в исследование участников информированное добровольное согласие.

В разделе «Заключение» по желанию авторов приводится следующая справочная информация: Общая информация о помощи в проведении работы и подготовке статьи: сообщения о предоставлении материалов, данных, компьютерного обеспечения, приборов во временное пользование; информация о проведении исследований в центрах коллективного пользования; помощь в технической подготовке текста; а также все прочее, что оценивается как полезная помощь. Информация о грантах и другой финансовой поддержке исследований. Авторы не должны использовать сокращенные названия институтов и спонсирующих организаций.

8. Список использованных источников

Оформляется в соответствии с приказом Высшей аттестационной комиссии Республики Беларусь от 25 июня 2014 г. № 159 (в редакции приказа ВАК Республики Беларусь 01.10.2024 № 230) (ВАК–Нормативные правовые акты–Приказы ВАК–Образцы оформления БО...). Источники должны быть представлены в виде списка в порядке упоминания ссылок в тексте статьи. Нумерация — автоматическая.

Не допускаются ссылки на работы, которых нет в списке литературы, а также на неопубликованные материалы. В список использованных источников преимущественно следует включать работы не старше 10 лет (исключение составляют редкие высокоинформативные материалы).

9. На английском языке повторяются п. 2–6 (в конце статьи)

Автоматические переводы не всегда способны точно передать нюансы и смысл оригинала. Чтобы сохранить содержание и тональность текста, может потребоваться использование совершенно других формулировок.

10. Дата поступления статьи

К статье прилагаются:

1. **Акт экспертизы** (в печатном виде) о возможности опубликования в открытой печати.
2. **Сопроводительное письмо** (в печатном виде), в структуре которого должны содержаться сведения о том, что:
 - рукопись не находится на рассмотрении в другом издании;
 - не была ранее опубликована.

Авторы несут ответственность за достоверность представленных в рукописи материалов.

3. **Краткие сведения о каждом из авторов** (в электронном виде), включающие ФИО (полностью), год рождения, служебный адрес, адрес электронной почты, ученую степень, ученое звание, должность. Сведения об авторах предоставляются отдельным документом Word.

Два экземпляра статьи (один — с указанием в «шапке» авторов и мест работы (на рус. и англ. яз.), другой — без) и сопроводительные документы представляются лично либо по почте на адрес: Республика Беларусь, 220072, г. Минск, ул. Академическая, 27. Электронная копия материалов направляется на электронный адрес: redactor@igc.by, в письме необходимо указать вид статьи (экспериментальная, обзорная, краткое сообщение).

Редакция оставляет за собой право в процессе редакционной подготовки сокращать и исправлять рукопись по согласованию с автором. **Обращаем Ваше внимание!** Редакция может выборочно проверить статьи системой антиплагиат (статьи направляются на проверку (антиплагиат) в Национальную библиотеку Беларуси).

Научное издание

МОЛЕКУЛЯРНАЯ И ПРИКЛАДНАЯ ГЕНЕТИКА

СБОРНИК НАУЧНЫХ ТРУДОВ

Том 39

Ответственный за выпуск *Е. М. Гришкова*
Переводчик *М. Г. Пыжова*

Подписано в печать 21.11.2025 г. Формат 60×84^{1/8}. Бумага офсетная. Гарнитура Times.
Печать цифровая. Усл. печ. л. 18,83. Уч.-изд. л. 11,4. Тираж 60 экз. Зак. 972.

Отпечатано на государственном предприятии «СтройМедиаПроект»
Свидетельство о государственной регистрации издателя, изготовителя,
распространителя печатных изданий № 2/42 от 13.02.2014.
220123, г. Минск, ул. В. Хоружей, 13/61.

Оригинал-макет подготовлен в Государственном научном учреждении
«Институт генетики и цитологии Национальной академии наук Беларуси».
Свидетельство о государственной регистрации издателя, изготовителя,
распространителя печатных изданий № 1/51 от 08.10.2013.
220072, г. Минск, ул. Академическая, 27.