

ISSN 1999-9127

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

МОЛЕКУЛЯРНАЯ И ПРИКЛАДНАЯ ГЕНЕТИКА

**СБОРНИК НАУЧНЫХ ТРУДОВ
ТОМ 25**

Издается с 2005 года
Выходит два раза в год

Минск
2018

Молекулярная и прикладная генетика: сб. науч. тр. / Институт генетики и цитологии НАН Беларуси; редкол.: А.В. Кильчевский (гл. ред.) [и др.]. — Минск: Институт генетики и цитологии НАН Беларуси, 2018. — Т. 25. — 130 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.В. Хотылёва — зам. главного редактора;
К.У. Вильчук, С.И. Гриб, О.Г. Давыденко, А.Н. Евтушенков, А.П. Ермишин,
А.И. Ковалевич, Ф.И. Привалов, А.В. Сукало, В.А. Лемеш, С.А. Лихачёв,
Н.П. Максимова, С.Б. Мельнов, М.Е. Михайлова, И.Б. Моссэ, М.Е. Никифоров,
В.Е. Падутов, В.Н. Решетников, Е.А. Сычёва, Н.И. Дубовец, В.В. Титок, И.П. Шейко,
О.Н. Харкевич — члены редколлегии;
А.Л. Богданова — ответственный секретарь.

СОДЕРЖАНИЕ

<i>Е.А. Фомина, Т.М. Дмитриева, С.В. Малышев, О.Ю. Урбанович</i> Молекулярно-генетическая характеристика коллекции сортов пшеницы (<i>Triticum aestivum</i> L.) по аллельному составу гена чувствительности к фотопериоду <i>Ppd-D1</i> и генов, кодирующих <i>Cbf</i> -факторы <i>Fr-B2</i> локуса.....	7
<i>Ю.А. Липихина, Е.В. Евтушенко, О.М. Люсигов, И.С. Гордей, И.А. Гордей, А.В. Вершинин</i> Ржано-пшеничные амфидиплоиды секалотритикум — новая модель для изучения активности генов центрального белка CENH3.....	15
<i>А.Н. Радченко, С.И. Михальская, М.О. Дыкун, Л.В. Сирант</i> Влияние аллелей гена <i>Vp-1B</i> на устойчивость к предуборочному прорастанию у мягкой пшеницы	24
<i>Т.В. Семашко, А.И. Горина, Е.А. Мартынова, Л.Н. Валентович, Л.А. Жуковская, А.Г. Лобанок</i> Получение векторной конструкции для экспрессии в клетках мицелиальных грибов гена изофермента C2 пероксидазы из корней хрена	32
<i>В.В. Гринев</i> Сохранение интронов в транскриптом лейкозных и нормальных клеток крови человека... 44	44
<i>А.О. Рябцева, И.С. Цыбовский, С.А. Котова</i> Микросателлитные маркеры в исследовании полиморфизма дикого кабана (<i>Sus scrofa</i>) и свиньи домашней (<i>Sus scrofa domestica</i>), обитающих на территории Республики Беларусь.....	56
<i>М.Е. Михайлова, Е.Л. Романишко, А.И. Киреева</i> Детекция полиморфизма rs456206907 гена <i>SMC2</i> , детерминирующего гаплотип фертильности HH3 у крупного рогатого скота.....	67
<i>Н.В. Воронова, Ю.В. Бондаренко, С.С. Левыкина, Р.С. Шулинский</i> Митохондриальный геном <i>Aphis fabae mordvilkoii</i> Börner & Janisch, 1992	73
<i>Л.В. Можаровская</i> Структурно-функциональный анализ локусов, кодирующих антимикробные пептиды сосны обыкновенной.....	83
<i>А.В. Падутов</i> Изучение генетической структуры полусибсового потомства клонов сосны обыкновенной на лесосеменных плантациях II порядка (Краткое сообщение)	92

<i>М.В. Спринджук, А.П. Кончиц, В.В. Слизень, Л.П. Титов</i>	
Алгоритмы и программное обеспечение для обработки данных геномов растений (Обзорная статья).....	99
<i>И.Ю. Романчук, В.С. Анохина</i>	
Алкалоиды люпина: строение, биосинтез, генетика (Обзорная статья)	108
Правила для авторов.....	124

CONTENTS

<i>E.A. Fomina, T.M. Dmitrieva, S.V. Malyshev, O.Yu. Urbanovich</i> Molecular and genetic characteristics of the collection of wheat varieties (<i>Triticum aestivum</i> L.) by the allelic composition of the photoperiod sensitivity gene <i>Ppd-D1</i> and genes encoding the <i>Cbf</i> factors of the <i>Fr-B2</i> locus.....	5
<i>Y.A. Lipikhina, E.V. Evtushenko, O.M. Lyusikov, I.S. Gordei, I.A. Gordei, A.V. Vershinin</i> Rye-wheat amphidiploids secalotriticum is a new model for studying the activity of genes of centromere protein CENH3	15
<i>A.N. Radchenko, S.I. Mykhalska, M.O. Dykun, L.V. Sirant</i> Influence of <i>Vp-1B</i> gene alleles on the pre-harvest sprouting of soft wheat.....	24
<i>T.V. Semashko, A.I. Gorina, E.A. Martynava, L.N. Valentovich, L.A. Zhukovskaya, A.G. Lobanok</i> Obtaining a vector construction for expression of the peroxidase isoenzyme C2 gene from horseradish roots in filamentous fungal cells.....	32
<i>V.V. Grinev</i> Intron retention in the transcriptome of leukemic and normal human blood cells.....	44
<i>A.A. Rabtsava, I.S. Tsybovsky, S.A. Kotova</i> Microsatellite markers in the study of polymorphism of the wild boar (<i>Sus scrofa</i>) and domestic pig (<i>Sus scrofa domestica</i>), inhabiting the territory of the Republic of Belarus	56
<i>M.E. Mikhailova, E.L. Romanishko, A.I. Kireeva</i> Detection of rs456206907 polymorphism of <i>SMC2</i> gene determining HH3 fertility haplotype in cattle	67
<i>N.V. Voronova, Y.V. Bandarenka, S.S. Levykina, R.S. Shulinski</i> Mitochondrial genome <i>Aphis fabae mordvilko</i> Borner & Janisch, 1992	73
<i>L.V. Mozharovskaya</i> Structural and functional analysis of loci encoding antimicrobial peptides of scots pine	83
<i>A.V. Padutov</i> Study of the genetic half-sib progeny structure of scots pine clones on forest-seed order II orchards (Short report).....	92
<i>M.V. Sprindzhuk, A.P. Konchits, V.V. Slizen', L.P. Titov</i> Algorithms and software designed for the plant genome data processing (Review article)	99

I.Yu. Romanchuk, V.S. Anokhina

Lupine alkaloids: structure, biosynthesis, genetics (Review article)..... 108

Instructions to Authors..... 124

Е.А. Фомина, Т.М. Дмитриева, С.В. Малышев, О.Ю. Урбанович

МОЛЕКУЛЯРНО-ГЕНЕТИЧЕСКАЯ ХАРАКТЕРИСТИКА КОЛЛЕКЦИИ СОРТОВ ПШЕНИЦЫ (*TRITICUM AESTIVUM* L.) ПО АЛЛЕЛЬНОМУ СОСТАВУ ГЕНА ЧУВСТВИТЕЛЬНОСТИ К ФОТОПЕРИОДУ *PPD-D1* И ГЕНОВ, КОДИРУЮЩИХ *CBF*-ФАКТОРЫ *FR-B2* ЛОКУСА

Институт генетики и цитологии НАН Беларуси
Республика Беларусь, 220072, г. Минск, ул. Академическая, 27
e-mail: e.fomina@igc.by

Исследован аллельный состав гена *Ppd-D1*, оказывающего влияние на чувствительность к фотопериоду, а также присутствие/отсутствие делеций в локусе *Fr-B2*, влияющих на морозоустойчивость в коллекции из 98 сортов яровой и озимой пшеницы. Показано, что в сортах пшеницы белорусской селекции получили распространение аллели генов, ассоциированные с устойчивостью к холоду.

Ключевые слова: озимая пшеница, нечувствительность к фотопериоду, *Ppd-D1* ген, морозоустойчивость, *Cbf*-гены.

Введение

Пшеница является одним из важнейших злаков, культивируемых в мире и обеспечивающим около пятой части потребляемых калорий на Земле. Гексаплоидная пшеница (*Triticum aestivum* L.), посевные площади под которой составляют 95% посевных площадей, занятых под пшеницу в целом, была адаптирована к произрастанию в климатических зонах от 65° с. ш. до 45° ю. ш. с большими вариациями температурного режима в течение вегетационного периода [1]. Вследствие более длительного периода роста, высеваемая осенью озимая пшеница обладает более высоким потенциалом урожайности по сравнению с высеваемой весной яровой пшеницей. Однако для реализации данного потенциала сорта озимой пшеницы должны быть адаптированы к воздействию низких отрицательных температур в течение зимнего вегетационного периода [2].

Морозоустойчивость тесно связана с таким понятием, как закаливание — процессом, при котором растения подвергаются длительному воздействию постепенно понижающихся низких положительных температур. В результате происходит повышение холодоустойчивости и увеличение процента выживших растений при дальнейшей их экспозиции при отрица-

тельных температурах [3]. В процессе закаливания в растительных клетках происходят огромные изменения на уровне транскриптома, что, в свою очередь, приводит к синтезу белков, обеспечивающих выживание растений в условиях холодового стресса [4, 5].

Несмотря на значительные различия в уровне экспрессии генов, кодирующих такие белки среди растений, было показано, что в обеспечении холодоустойчивости участвует относительно небольшое количество локусов [6]. В частности, в чувствительном к холоду сорте Chinese Spring наблюдалось повышение морозоустойчивости в ходе проведения экспериментов по замещению в его геноме хромосом 5A, 5D и 5B на хромосомы сверхустойчивого сорта Cheyenne [7–9]. Также было выявлено, что замещение хромосомы 5A приводило к увеличению уровня экспрессии в ответ на воздействие холода гена *COR14b*, который выполняет основную роль в защите хлоропластов во время действия низких отрицательных температур [6]. Указанный ген, как и множество других генов, экспрессирующихся в ответ на воздействие холода, содержит в своей промоторной области CRT-DRE бокс — связывающий домен для транскрипционных факторов семейства *CBF* (*C-repeat Binding Factor*). Один

из таких генов, *TmCBF3*, был картирован в области хромосомы, где расположены гены, оказывающие влияние на дифференциальную экспрессию *COR14b* гена, и названной *Fr-2* (*Frost resistance-2*) локусом. Предположительно, основная роль данного локуса в обеспечении холодоустойчивости принадлежит *CBF* генам [10]. Следует отметить, что в растениях с высоким уровнем холодоустойчивости был выявлен более высокий уровень экспрессии еще целого ряда *CBF* генов, картированных в *Fr-2* локусе, по сравнению с растениями с низким уровнем холодоустойчивости [11].

Детальное изучение семейства генов *CBF* у *T. monosocum* позволило выявить 15 *CBF* генов, 11 из которых были расположены в виде кластера в локусе *Fr-2* на хромосоме 5A^m [12]. Последующее изучение данного локуса при помощи картирующей популяции *T. monosocum* показало, что вероятнее всего гены *CBF12*, *CBF14* и *CBF15* играют основную роль в наблюдаемых различиях по уровню холодоустойчивости среди растений указанной популяции [13]. *CBF* гены принадлежат к AP2/ERF семейству транскрипционных факторов и лучше всего изучены у *Arabidopsis*. Они играют важную роль в координации ответа на воздействие холода и их экспрессия индуцируется в течение 15 минут после начала воздействия низких температур, после чего, примерно через два часа, происходит активация других генов, принимающих участие в ответе на воздействие низких температур [3]. Этот так называемый *CBF* регулон содержит основную часть холодоиндуцибельных генов. Например, в трансгенном *Arabidopsis* с повышенным уровнем экспрессии *CBF2* гена происходила активация 28% всех холодоиндуцибельных генов [14]. Сюда входят гены, кодирующие первичную структуру ферментов, принимающих участие в синтезе криопротекторных и других мембран-стабилизирующих белков, которые помогают предотвратить повреждение растительных клеток при дегидратации, которая наблюдается в ходе холодового стресса [2].

Если в геноме *Arabidopsis* имеется три *CBF* гена, организованных в кластер на хромосоме 4 [15–17], то в геноме пшеницы и других злаков их значительно больше. Так у гексаплоидной пшеницы в каждом из трех геномов находится как минимум по 15 *CBF* генов, 11 из которых

располагаются в каждом из трех *Fr-2* локусов [18]. У ячменя число копий *CBF* генов в *Fr-H2* локусе варьирует, что объясняет различия в уровне морозоустойчивости [19]. У пшеницы различная степень влияния трех гомеологичных *Fr-2* локусов на уровень морозоустойчивости была продемонстрирована в эксперименте по созданию замещенных линий по каждой из трех гомеологичных хромосом 5 группы, где замещение хромосомы 5A приводило к наибольшему возрастанию уровня морозоустойчивости, а 5B — к наименьшему [7, 8].

Следует также отметить, что на адаптивность растений к осенне-зимнему и весенне-летнему периоду оказывает влияние и их чувствительность к продолжительности светового дня [20]. За нечувствительность к продолжительности светового дня отвечает неполно-доминантный аллель *Photoperiod-D1a* (*Ppd-D1a*), расположенный на хромосоме 2D. Мутации в этом аллеле позволяют растениям цвести как при коротких световых днях, так и при длинных. Основной мутацией в этом аллеле является делеция длиной 2089 пн возле кодирующей области. По данным Файт, указанный аллель способствует увеличению урожайности и в то же время снижению морозоустойчивости растений [21]. Изменения в аналогичных аллелях *Ppd-A1* и *Ppd-B1*, расположенных соответственно на хромосомах 2A и 2B, не влияют на формирование нечувствительности к продолжительности светового дня [22].

В связи с важностью генов, влияющих на чувствительность к продолжительности светового дня и кодирующих первичную структуру *Cbf*-факторов для селекционного процесса, целью проведенного исследования являлось определение аллельного состава данных генов у сортов и линий озимой пшеницы.

Материалы и методы

Объектом исследования служила коллекция, состоящая из 98 образцов сортов пшеницы, предоставленных ГУ «Государственная инспекция по испытанию и охране сортов растений», а также лабораториями тритикале и озимой пшеницы РУП «Научно-практический центр НАН Беларуси по земледелию» (г. Жодино).

ДНК из зерен выделяли по методу, предложенному Plaschke и др. [23]. Выделение проводили из двух зерновок для каждого сорта.

Анализ аллельного состава гена *Ppd-D1* проводили согласно методике, предложенной Beales и др. [22] с изменениями.

Для амплификации была использована следующая программа: 1 цикл продолжительностью 5 мин при 94 °С; 35 циклов, включающих в себя: 40 с при 94 °С, 30 с при 54 °С, 1 мин при 72 °С, заключительная элонгация 10 мин при 72 °С.

Идентификацию аллельного состава генов, кодирующих первичную структуру *Cbf*-факторов, проводили по методике, предложенной Pearce и др. [2] с изменениями.

Для амплификации была использована следующая программа: 1 цикл продолжительностью 5 мин при 94 °С; 45 циклов, включающих в себя: 1 мин при 94 °С, 1 мин при 60 °С, 1 мин при 72 °С, заключительная элонгация 10 мин при 72 °С.

Состав реакционной смеси для амплификации объемом 12,5 мкл был следующий: 1×буфер для *Taq* полимеразы «А» без $MgCl_2$; 1,5 мМ $MgCl_2$; 0,2 мМ НТФ; 0,25 нМ праймеры; 0,5 ЕА *Taq*-полимераза; 50 нг ДНК. Для анализа использовались праймеры и ПЦР реактивы производства компании «Праймтех» (Минск).

Таблица

Распространение аллелей гена *Ppd-D1* в коллекции сортов пшеницы

Вид аллеля	Название сорта	Страна происхождения	Количество сортов и линий (%)
<i>Ppd-D1a</i>	Альбатрос одесский, Багира, Барвина, Благодарна, Борвий (образец 1), Борвий (образец 2), Бунчук, Видрада, Вильшана, Герта, Годувальныця одеська, Доброчын, Заграва одесская (образец 1), Заграва одесская (образец 2), Заможність, Истина одесская, Калита, Козачий атаман, Лорд, Одесская 200, Подольнка, Полевик, Почаївка, Придеснянська напівкарликова, Роксолана, Сагайдак, Селянка одесская, Славна, Турунчук, Ужинок, Уникум, Утес, Хвест, Хоревица (образец 1), Хоревица (образец 2), Эпоха одесская, Яворина, Ярославна	Украина	59 (60,2%)
	Ариадна, Аскет, Донская полукарликовая, Донской сюрприз, Дон 95, Ермак, Ершовская 11, Жемчужина Поволжья, Корганка, Левобережная 1, Новорешовская, Памяти Калиненко, Ростовчанка 3, Синтетик, Юнона F.594	Россия	
	Balaton	Австрия	
	Co 207	Франция	
	Faur, Miranda	Румыния	
	Капылянка	Беларусь	
<i>Ppd-D1b</i>	Акорд, Кармен, Миронівська сторична, Элик	Украина	36 (36,7%)
	Богданка, Проза	Россия	
	Lupus, Saturnus	Австрия	
	Acratos, Catalus, Cubus, Dromos, Emmitt, Perfect, Samurai, Skagen	Германия	
	FT Wonder	Канада	
	Былина, Виза*, Гармония, Дарья*, Завет, Каравай, Легенда, Прэм'ера, Рассвет*, Ростань*, Сабина*, Саната, Спектр, Сюїта, Тома*, Узлет, Фантазия, Элегія (образец 1), Элегія (образец 2)	Беларусь	
<i>Ppd-D1a</i> + <i>Ppd-D1b</i>	Наусель	Украина	3 (3,1%)
	Зарница	Россия	
	Щара	Беларусь	

Примечание. Знаком «*» отмечены сорта ярового типа, все остальные сорта являются озимыми

Визуализацию фрагментов амплификации проводили после разделения методом электрофореза в трис-ацетатном буфере при помощи системы документирования гелей GelDoc 2000.

Результаты и обсуждение

Для анализа распространения аллелей гена чувствительности к фотопериоду *Ppd-D1* среди сортов пшеницы использовали праймеры Ppd-D1_F, Ppd-D1_R1 и Ppd-D1_R2, созданные для выявления в данном гене делеции размером 2089 пн в гене *Ppd-D1*. Место посадки праймера Ppd-D1_R1 находится в области делеции, поэтому ПЦР продукт размером 414 пн образуется в образцах, не несущих данную делецию, что соответствует дикому аллелю *Ppd-D1b* и, следовательно, обладающих чувствительностью к фотопериоду. Место посадки праймера Ppd-D1_R2 находится в области первого интрона, расположенного за областью делеции, и ПЦР продукт размером 288 пн образуется только в образцах, несущих в своих геномах 2089 пн делецию, т. е. мутантный аллель *Ppd-D1a*, что говорит об их нечувствительности к фотопериоду [22].

Результаты исследования коллекции сортов пшеницы по аллельному составу гена *Ppd-D1* представлены в таблице.

Как видно из таблицы, среди представленных сортов и линий 59 (60,2%) имеют в своем генотипе аллель *Ppd-D1a*, 36 (36,7%) — аллель *Ppd-D1b*. Некоторые сорта и линии (Зарница, Наусель и Щара) являются гетерозиготными по своему составу и несут в своем генотипе как *Ppd-D1a*, так и *Ppd-D1b* аллель.

Сравнительный анализ сортов зарубежной и белорусской селекции по аллельному составу указанного гена показал, что среди сортов зарубежной селекции преобладают нечувствительные к продолжительности светового дня образцы. Их доля составляет 75,3% (рисунок). Среди сортов белорусской селекции 90,5% образцов, в том числе все образцы яровой пшеницы, оказались чувствительными к продолжительности светового дня. Результаты показывают, что селекционный процесс в Беларуси велся в сторону сортов, несущих в своих генотипах *Ppd-D1b* аллель. Очевидно, что на распространение аллеля *Ppd-D1b* оказали влияние преимущества, которые он дает растениям, выращиваемым в более холодных климатических зонах. В частности, в климатических условиях Беларуси преждевременное цветение может способствовать снижению урожайности. Кроме того, присутствие аллеля *Ppd-D1b* не снижает морозоустойчивость несущих его

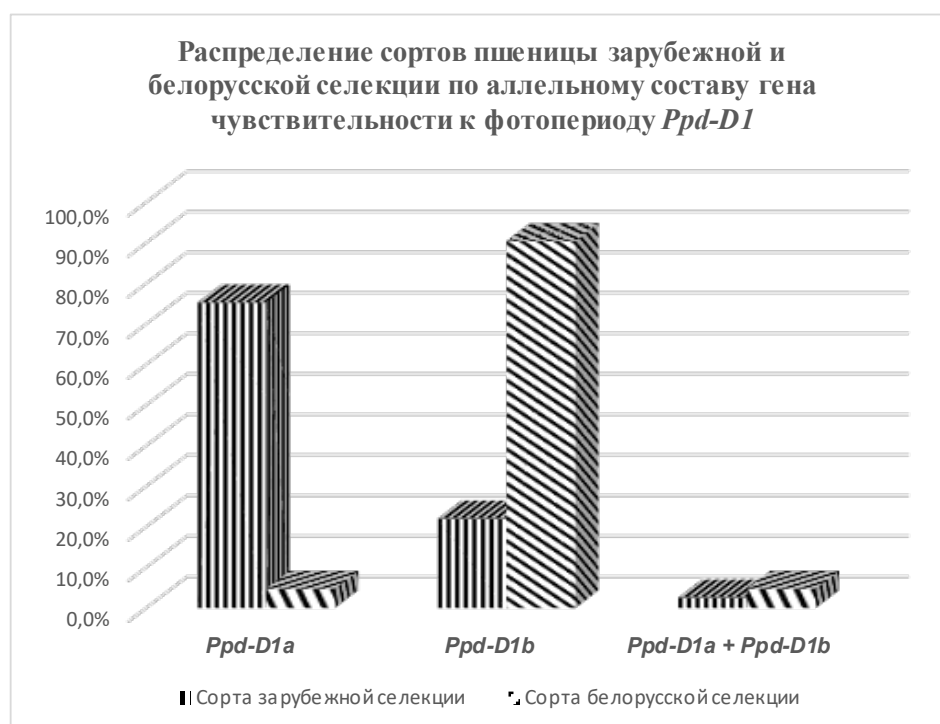


Рис. Распространение аллелей гена *Ppd-D1* среди сортов зарубежной и белорусской селекции

сортов. *Ppd-D1b* аллель преобладал у сортов стран Германии и Австрии (таблица), где климатические условия схожи с Беларусью. Для сортов стран с более теплым климатом, таких как Украина и Румыния, преобладающим оказался аллель *Ppd-D1a*, что подтверждает литературные данные о том, что указанный аллель чаще используется в селекционном процессе в странах с теплым климатом [22]. Так, например, по данным Beales и др. в Англии *Ppd-D1a* аллель способствовал раннему цветению и как следствие снижению урожайности на 1,8%, в Германии — увеличению урожайности на 7,7%, и на территории бывшей Югославии — увеличению урожайности на 33%.

Согласно Dusk et al., слабая реакция на фотопериод, обусловленная присутствием в генотипе аллеля *Ppd-D1a*, является одним из необходимых условий для реализации потенциала урожайности озимых сортов, которые возделывают в более южных районах с относительно мягкой зимой и укороченным естественным днем. Яровую пшеницу возделывают, как правило, в северных районах с продолжительным естественным летним днем. Сорта, несущие в своих генотипах аллель *Ppd-D1a*, в данных условиях не обладают преимуществом по урожаю зерна [24].

Для анализа гена *Fr-B2*, контролирующего толерантность к холоду, были использованы праймеры CBF-B4, CBF-B9-pair 1, CBF-B10 и CBF-B12-pair 2. При амплификации с ними синтезируются фрагменты размером 159, 235, 108 и 329 пн у сортов и линий, которые не имеют делеций в генах *Cbf-B4* и *Cbf-B9*, расположенных в проксимальном кластере, *Cbf-B10*, расположенном в дистальном кластере, и *Cbf-B12*, расположенном в центральном кластере указанного локуса. Pearce и др. в исследуемых сестринских линиях, полученных на основе скрещивания сортов твердой пшеницы (*Triticum durum* Desf.) Durelle и Kronos, выявили три вида делеции: из шести генов, включающих в себя гены *Cbf-B12*, *Cbf-B14* и *Cbf-B15* центрального кластера, а также гены *Cbf-B2* и *Cbf-B4* проксимального кластера и *Cbf-B16* дистального кластера, из девяти генов, включающих в себя, помимо упомянутых генов, гены *Cbf-B3* и *Cbf-B13* дистального кластера, а также из 11 генов, в которую, помимо генов шести- и девятигеновой делеции, входят

еще и гены *Cbf-B9* и *Cbf-B17* проксимального кластера [2]. Таким образом, выбранные праймеры позволяют идентифицировать все виды указанных делеций.

Проведенный анализ показал, что, несмотря на отличия в устойчивости к фотопериоду, у всех исследуемых сортов пшеницы в локусе *Fr-B2* представлено 11 *Cbf* генов. Сорта не несут известных делеций 6-го, 9-го или 11-го генов данного локуса, которые ассоциированы со снижением устойчивости к низким отрицательным температурам.

Снижение уровня морозоустойчивости у сортов, несущих в своих геномах указанные делеции, было показано Pearce и др. [2]. При этом делеции чаще встречались у сортов яровой пшеницы, чем озимой. Так, в частности, делеция девяти генов была выявлена у 50,7% культивируемых сортов яровой и 38,9% озимой твердой пшеницы (*T. turgidum* spp. *durum*) и делеция одиннадцати генов — у 22% сортов яровой и 2% озимой гексаплоидной пшеницы [2]. В исследуемых нами образцах делеции *Fr-B2* локуса отсутствовали как в озимых, так и яровых сортах.

Pearce и др. предположили, что делеции отдельных *Cbf* генов в локусе *Fr-B2* дают преимущества только в случаях произрастания растений в регионах с теплым климатом [2].

Считается, что активация под воздействием холода *Cbf* регулона приводит к активации множества генов [25] и синтезу разнообразных криопротекторных белков для защиты клеток от повреждения в результате воздействия низких отрицательных температур. Для растений поддержание *Cbf* регулона в активном состоянии является энергетически затратным [2]. Это предположение нашло подтверждение в экспериментах по созданию трансгенных линий *Arabidopsis* с повышенным уровнем экспрессии гена *CBF2*. Выращивание указанных растений при температуре 22 °C приводило к снижению скорости роста и появлению карликовых растений [26]. Вероятно, делеция ряда генов *Cbf* у сортов, выращиваемых в теплом климате, может дать им преимущества. Для более точного выяснения роли делеций *Cbf* генов *Fr-B2* локуса в адаптации растений к произрастанию в условиях теплого климата представляет интерес изучение изогенных линий яровой пшеницы,

в генотипах которых присутствуют/отсутствуют такие делеции [2].

Заключение

Исследован аллельный состав гена *Ppd-D1*, оказывающего влияние на чувствительность к фотопериоду, а также присутствие/отсутствие делеций в локусе *Fr-B2*, влияющих на морозоустойчивость в коллекции из 98 сортов яровой и озимой пшеницы. Среди представленных сортов и линий 59 (60,2%) имеют в своем генотипе аллель *Ppd-D1a* и являются нечувствительными к фотопериоду, 36 (36,7%) — аллель *Ppd-D1b* и являются чувствительными к фотопериоду. Некоторые сорта и линии (Зарница, Наусель и Щара) являются гетерозиготными по своему составу и несут в своем генотипе как *Ppd-D1a*, так и *Ppd-D1b* аллель. Анализ локуса *Fr-B2* показал, что, несмотря на отличия в устойчивости к фотопериоду, у всех исследуемых сортов пшеницы в указанном локусе представлено 11 *Cbf* генов. Показано, что в сортах пшеницы белорусской селекции получили распространение аллели генов, ассоциированные с устойчивостью к холоду.

Список использованных источников

1. Dubcovsky, J. Genome plasticity a key factor in the success of polyploid wheat under domestication / J. Dubcovsky, J. Dvorak // *Science*. – 2007. – Vol. 316. – P. 1862–1866.
2. Large deletions in the *CBF* gene cluster at the *Fr-B2* locus are associated with reduced frost tolerance in wheat / S. Pearce [et al.] // *Theor Appl Genet*. – 2013. – Vol. 126, № 11. – P. 2683–2697.
3. Thomashow, M.F. Molecular basis of plant cold acclimation: insights gained from studying the *CBF* cold response pathway / M.F. Thomashow // *Plant Physiol*. – 2010. – Vol. 154. – P. 571–577.
4. Regulation of gene expression by chromosome 5A during cold hardening in wheat / G. Kocsy [et al.] // *Mol Genet Genom*. – 2010. – Vol. 283. – P. 351–363.
5. Plant responses to cold: transcriptome analysis of wheat / M.O. Winfield [et al.] // *Plant Biotechnol J*. – 2010. – Vol. 8. – P. 749–771.
6. Two loci on wheat chromosome 5A regulate the differential cold-dependent expression of the *cor14b* gene in frost-tolerant and frost-sensitive genotypes / A. Vágújfalvi [et al.] // *Mol Genet Genet*. – 2000. – Vol. 263. – P. 194–200.
7. Sutka, J. Genetic studies of frost resistance in wheat / J. Sutka // *Theor Appl Genet*. – 1981. – Vol. 59. – P. 145–152.
8. Veisz, O. The relationships of hardening period and the expression of frost resistance in chromosome substitution lines of wheat / O. Veisz, J. Sutka // *Euphytica*. – 1989. – Vol. 43. – P. 41–45.
9. Mapping genes affecting flowering time and frost resistance on chromosome 5B of wheat / B. Toth [et al.] // *Theor Appl Genet*. – 2003. – Vol. 107. – P. 509–514.
10. The cold-regulated transcriptional activator *CBF3* is linked to the frost tolerance locus *Fr-A2* on wheat chromosome 5A / A. Vágújfalvi [et al.] // *Mol Genet Genom*. – 2003. – Vol. 269. – P. 60–67.
11. The expression of several *CBF* genes at the *Fr-A2* locus is linked to frost resistance in wheat / A. Vágújfalvi [et al.] // *Mol Genet Genom*. – 2005. – Vol. 274. – P. 506–514.
12. 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 Genom*. – 2006. – Vol. 275. – P. 193–203.
13. Identification of candidate *CBF* genes for the frost tolerance locus *Fr-A^m2* in *Triticum monococcum* / A.K. Knox [et al.] // *Plant Mol Biol*. – 2008. – Vol. 67. – P. 257–270.
14. Roles of the *CBF2* and *ZAT12* transcription factors in configuring the low temperature transcriptome of *Arabidopsis* / J.T. Vogel [et al.] // *Plant J*. – 2005. – Vol. 41. – P. 195–211.
15. The *Arabidopsis CBF* gene family is composed of three genes encoding AP2 domain-containing proteins whose expression is regulated by low temperature but not by abscisic acid or dehydration / J. Medina [et al.] // *Plant Physiol*. – 1999. – Vol. 119. – P. 463–469.
16. Low temperature regulation of the *Arabidopsis CBF* family of AP2 transcriptional activators as an early step in cold-induced *COR* gene expression / S.J. Gilmour [et al.] // *Plant J*. – 1998. – Vol. 16. – P. 433–442.
17. Stockinger, E.J. *Arabidopsis thaliana CBF1* encodes an AP2 domain-containing transcriptional activator that binds to the C-repeat/Dre, a cis-acting DnA regulatory element that

stimulates transcription in response to low temperature and water deficit / E.J. Stockinger, S.J. Gilmour, M.F. Thomashow // *Proc natl Acad Sci USA*. – 1997 – Vol. 94. – P. 1035–1040.

18. The *CBF* gene family in hexaploid wheat and its relationship to the phylogenetic complexity of cereal *CBFs* / M. Badawi [et al.] // *Mol Genet Genom*. – 2007. – Vol. 277. – P. 533–554.

19. *CBF* gene copy number variation at Frost Resistance-2 is associated with levels of freezing tolerance in temperate-climate cereals / A.K. Knox [et al.] // *Theor Appl Genet*. – 2010. – Vol. 121. – P. 21–35.

20. The influence of photoperiod genes on the adaptability of European winter wheats / A.J. Worland [et al.] // *Euphytica*. – 1998. – Vol. 100. – P. 385–394.

21. Файт, В.І. Морозостійкість і урожайність окремих сортів озимої м'якої пшениці / В.І. Файт // *Вісн. аграр. науки*. – 2005. – Vol. 11. – P. 25–29.

22. A Pseudo-Response Regulator is misexpressed in the photoperiod insensitive *Ppd-D1a*

mutant of wheat (*Triticum aestivum* L.) / J. Beales [et al.] // *Theor Appl Genet*. – 2007. – Vol. 115. – P. 721–733.

23. Plaschke, J. Detection of genetic diversity in closely related bread wheat using microsatellite markers / J. Plaschke, M.W. Ganal, M.S. Röder // *Theor. Appl. Genet*. – 1995. – Vol. 91. – P. 1001–1007.

24. Agronomic performance of hard red spring wheat isolines sensitive and insensitive to photoperiod / J.A. Dyck [et al.] // *Crop Sci*. – 2004. – Vol. 44. – P. 1976–1981.

25. Fowler, S. *Arabidopsis* transcriptome profiling indicates that multiple regulatory pathways are activated during cold acclimation in addition to the *CBF* cold response pathway / S. Fowler, M.F. Thomashow // *Plant Cell*. – 2002. – Vol. 14. – P. 1675–1690.

26. The cold-inducible *CBF1* factor-dependent signaling pathway modulates the accumulation of the growth-repressing *DellA* proteins via its effect on gibberellin metabolism / P. Achard [et al.] // *Plant Cell*. – 2008. – Vol. 20. – P. 2117–2129.

E.A. Fomina, T.M. Dmitrieva, S.V. Malyshev, O.Yu. Urbanovich

**MOLECULAR AND GENETIC CHARACTERISTICS OF THE
COLLECTION OF WHEAT VARIETIES (*TRITICUM AESTIVUM* L.) BY
THE ALLELIC COMPOSITION OF THE PHOTOPERIOD SENSITIVITY
GENE *PPD-D1* AND GENES ENCODING THE *CBF* FACTORS OF THE
FR-B2 LOCUS**

Institute of Genetics and Cytology, NASB
Minsk, 220072, the Republic of Belarus

The allelic composition of the *Ppd-D1* gene influencing the photoperiod sensitivity, as well as the presence/absence of deletions in the *Fr-B2* locus, affecting frost tolerance in the collection of 98 varieties of spring and winter wheat, was investigated. It is shown that the alleles of genes associated with cold tolerance have spread in Belarusian wheat cultivars.

Key words: winter wheat, photoperiod insensitivity, *Ppd-D1* gene, frost tolerance, *Cbf*-genes.

Дата поступления статьи: 20 августа 2018 г.

Ю.А. Липихина¹, Е.В. Евтушенко¹, О.М. Люсиков², И.С. Гордей², И.А. Гордей², А.В. Вершинин¹

РЖАНО-ПШЕНИЧНЫЕ АМФИДИПЛОИДЫ СЕКАЛОТРИТИКУМ — НОВАЯ МОДЕЛЬ ДЛЯ ИЗУЧЕНИЯ АКТИВНОСТИ ГЕНОВ ЦЕНТРОМЕРНОГО БЕЛКА CENH3

¹Институт молекулярной и клеточной биологии СО РАН
Россия, 630090, г. Новосибирск, пр-т академика Лаврентьева, 8/2
e-mail: avershin@mcb.nsc.ru

²Институт генетики и цитологии НАН Беларуси
Республика Беларусь, 220072, г. Минск, ул. Академическая, 27
e-mail: I.Gordej@igc.by

Отдаленная гибридизация является стрессовым фактором, который вызывает реорганизацию родительских геномов в гибридном потомстве («геномный шок») и сопровождается потерей генов, изменениями в экспрессии генов, молчанием генов, делециями, инверсиями и транслокациями участков хромосом. Созданные аллополиплоидные гибриды секалотритикум являются новой интересной моделью для изучения влияния отдаленной гибридизации на участки генома, ответственные за правильное прохождение процессов деления и распределения хромосом в гибридных клетках. Центромера осуществляет интегральный контроль этих процессов. Основным признаком активной центромеры является присутствие в нуклеосомах центромер специфической модификации гистона H3, обозначаемой у растений CENH3. В данной работе приведены результаты цитогенетического анализа кариотипов гибридов секалотритикум и сравнение структуры N-концевого домена CENH3 между родительскими и гибридными формами. Показано, что кариотипы созданных секалотритикум являются стабильными сбалансированными гексаплоидами и не содержат мини-хромосом с делетированными плечами или участками плеч. Оба родительских генома участвуют в формировании структуры CENH3 у гибридов секалотритикум. Высокая идентичность последовательностей CENH3 ржи и пшеницы обеспечивает высокий уровень экспрессии обеих родительских форм в гибридных геномах.

Ключевые слова: рожь, пшеница, отдаленная гибридизация, ржано-пшеничные амфидиплоиды, тритикале, секалотритикум, аллополиплоиды, геном, цитогенетика, центромера, центромерный гистон H3 (CENH3), экспрессия генов.

Введение

Аллополиплоидизация, возникающая вследствие отдаленной (межвидовой, межродовой) гибридизации, является одним из основных факторов в эволюции растений, ведущим к появлению новых видов [1, 2]. В селекционном процессе цель отдаленной гибридизации состоит в достижении у гибридов более полного проявления природного генетического потенциала или генетического разнообразия, присутствующего в каждой из родительских форм. При этом отдаленная гибридизация является стрессовым фактором, который вызывает быструю реорганизацию родительских геномов в гибридном потомстве («геномный шок» [3]), сопровождающуюся потерей генов, изменениями в экспрессии генов, молчанием

генов, делециями, инверсиями и транслокациями участков хромосом [4–6].

Принципиальное значение для расширения генофонда и создания селекционно ценных гибридов может иметь синтез ржано-пшеничных амфидиплоидов, при котором материнской формой служит рожь, а опылителем — пшеница. Такие скрещивания, как правило, затруднены вследствие проявления реакции односторонней несовместимости и ржано-пшеничные амфидиплоиды до сих пор изучены недостаточно. Преодоление этого барьера связано с большими перестройками в геномах родительских форм, что является наиболее ярким проявлением «геномного шока» и сопровождается различными хромосомными нарушениями, в том числе затрагивающими

структуру центромер. Исследованиями сотрудников лаборатории цитогеномики растений Института генетики и цитологии НАН Беларуси установлено, что использование в скрещиваниях с рожью вида-посредника тритикале в качестве источника геномов пшеницы оказалось эффективным для преодоления барьера односторонней прогамной несовместимости исходных видов [7]. Новые гибриды, у которых в качестве материнского растения использовали рожь, а в качестве отцовского источника пшеничного генома — тритикале, получили название секалотритикум (*×Secalotriticum*, син. *×Secalotriticum* = *Secale* L. *× Triticum* L.). У секалотритикум создаются более благоприятные условия для усиления экспрессии генетических систем ржи и проявления ее ценных адаптивных признаков.

Процесс образования гибридного генома и последующая его стабилизация напрямую связаны с нормализацией процесса мейоза и правильной сегрегацией хромосом. Именно центромера является тем участком хромосомы, который осуществляет контроль фундаментального процесса деления клеток, включающего спаривание и разделение сестринских хроматид, присоединение веретена деления, расхождение хромосом к полюсам клетки. Несовместимость центромер разных видов, по-видимому, является основной причиной элиминации хромосом одного из родительских геномов у гибридов [8, 9]. С молекулярной точки зрения, центромера представляет многокомпонентную структуру, состоящую из разнообразных классов последовательностей ДНК и белков. Центромерная ДНК показывает значительную вариабельность структуры между видами и даже между хромосомами одного кариотипа [10]. Среди центромерных белков, согласно исследованиям последнего десятилетия, особое место отводится центромерной модификации гистона H3, у растений обозначаемой как CENH3 [10, 11]. Это особое положение обусловлено тем, что на молекулярном уровне наиболее специализированной и универсальной характеристикой активной центромеры является присутствие CENH3 вместо канонического гистона H3 в нуклеосомах центромерного хроматина. Это свойство отмечено у хромо-

сом, имеющих разнообразные формы центромер: моноцентрические, дицентрические, голоцентрические, неоцентромеры [11]. CENH3 вместе с другими окружающими белками хроматина непосредственно ответственен за формирование в данном участке хромосомы кинетохора, который взаимодействует с микротрубочками веретена деления, по которым хромосомы двигаются к полюсам клетки при ее делении [12]. Таким образом, CENH3 можно рассматривать как эпигенетическую «метку» центромеры [13]. Как показано на некоторых видах млекопитающих и дрозофиле, в случае его потери по каким-то причинам не происходит формирования кинетохора и правильного расхождения хромосом в ходе деления клеток [14].

В связи с особой ролью CENH3 в формировании и функционировании центромер представляет значительный интерес изучение влияния отдаленной гибридизации на структуру этого белка у гибридов относительно исходных родительских форм. В настоящей работе такое исследование проведено на примере альфа-вариантов CENH3 ржано-пшеничных амфидиплоидов секалотритикум. Кроме того, используя С-бэндинг, необходимо было сравнить кариотипы родительских и гибридных форм на предмет возможных видимых хромосомных перестроек вследствие отдаленной гибридизации.

Материалы и методы

Две формы гибридов секалотритикум STr BM и STr ВД были созданы на основе гибридизации тетраплоидной ржи Верасень (^SRRRR, 2n=4x=28) с гексаплоидными тритикале сортов Михась и Дубрава (^TAABBRR, 2n=6x=42). Тритикале выступали в качестве вида-посредника — источника геномов пшеницы и ингибитора S-РНКаз ржи, что позволило преодолеть прогамную несовместимость ржи с пшеницей. Создание исходных форм тетраплоидной ржи описано в статье Гордей и др. [15].

Кариотипирование секалотритикум и молекулярно-цитогенетическую идентификацию хромосом исходных видов проводили на давленных препаратах апикальных меристем корня с помощью варианта метода дифференциального окрашивания по Гимза (C-banding) [16].

Молекулярный анализ проводили на растениях родительских форм и гибридов, выращенных в теплице. Суммарную РНК выделяли из листьев 10–12-дневных проростков из 4–5 растений родительских форм. У гибридных форм РНК выделяли из каждого растения отдельно. Выделение РНК проводили с использованием TRI Reagent (MRC, Ink., США) [17]. Чтобы избежать контаминации геномной ДНК, тотальную РНК обрабатывали RQ-RNase-Free DNAase (Promega, Madison, США). Для синтеза кДНК использовали набор реактивов RevertAid™ H Minus First Strand cDNA Synthesis Kit (Thermo Scientific, США). Полученную кДНК использовали в серии реакций ПЦР в качестве матрицы с праймерами, синтезированными специально для амплификации N-концевого участка (N-Terminal Tail, NTT) CENH3. Последовательности праймеров: 5' – ATGGCCCGCACCAAGC (F), 5' – GAAACTCGACCGACTTCTG (R). Размер продукта составлял 268 п. н. Продукты ПЦР клонировали в плазмиду pTZ57R/T (InsTAclone PCR cloning kit, Thermo Scientific, США) и секвенировали по Сэнгеру с использованием BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, США). Продукты реакции разделяли на приборе 3500 Genetic Analyzer (Applied Biosystems, США). Последовательности индивидуальных клонов, полученные для каждого образца, анализировали при помощи программы FinchTV 1.4 (Geospiza) и пакета программ FASTA [18]. Поиск идентичности нуклеотидных последовательностей осуществляли с помощью алгоритма BLAST в базе данных NCBI. Графические изображения подготовлены с использованием программы Jalview.

Результаты и обсуждение

Одним из проявлений геномного шока, возникающего вследствие объединения двух родительских геномов в гибридной клетке при отдаленной гибридизации, являются различные хромосомные изменения. Делеции и транслокации отдельных хромосомных районов и хромосомных плеч относятся к наиболее распространенным нарушениям и были обнаружены в цитогенетическом анализе как пшенично-ржаных замещенных и дополненных линий [19, 20], так и в тритикале и потомках

от скрещиваний тритикале × пшеница [21, 22]. Нарушения при этом отмечены в хромосомах обеих родительских форм.

Кариотипы секалотритикум STr BM (Верасень × Михась) и STr ВД (Верасень × Дубрава) (^SRRAABB, 2n=6x= 42) были проанализированы методом дифференциального окрашивания (С-бэндинг). Этот метод выявляет паттерн локализации гетерохроматиновых районов, который является специфичным для каждой хромосомы ржи и пшеницы и, соответственно, позволяет провести идентификацию каждой хромосомы в гибридном кариотипе. На рис. 1 показан кариотип секалотритикум STr BM. Кариотипы созданных секалотритикум являются стабильными сбалансированными гексаплоидами и не содержат мини-хромосом с делетированными плечами или участками плеч. Рисунок С-бэндинга хромосом гибридных геномов и исходных форм ржи и тритикале совпадает. Стабилизации кариотипов спо-

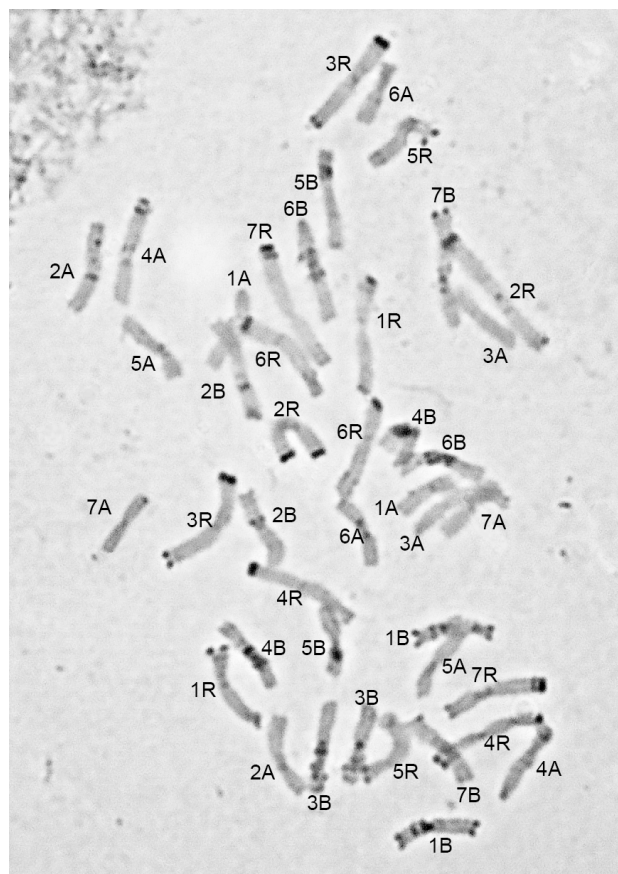


Рис. 1. Кариотип секалотритикум STr BM (Верасень × Михась) после С-бэндинга представлен полными наборами хромосом R-, A-, B-геномов

собствовали однократный беккросс ржано-тритикальных пентаплоидных гибридов F_1 (sRRABR , $5x=35$) на тритикале с последующим самоопылением в течение 15 поколений и постоянный отбор на цитологическую стабильность и фенотипическую однородность.

Диплоидный RR -геном ржи у ржано-тритикальных гибридов F_1 является фактором стабилизации мейоза и обеспечивает функциональность гамет различного хромосомного состава. Формирование генома секалотритикум происходит в F_1BC_1 на основе частично нередуцированных 21-хромосомных гамет пентаплоидов F_1 , сбалансированных по наборам хромосом гаплогеномов исходных видов ($7(R) 7(A) 7(B)$). Стабильность генома и генетическое разнообразие секалотритикум определяются типом цитоплазмы и следующими наследуемыми от генотипов ржано-тритикальных гибридов F_1 цитогенетическими факторами, в основном, связанными со структурно-функциональным состоянием центромер в их геноме:

- формирование (псевдо)унивалентов в результате десинапсиса;
- сохранение ими униполярной ориентации центромер и редукционное I деление в мейозе;
- эквационное II деление мейоза с регулярной полярной сегрегацией и низким уровнем элиминации хромосом отдаленных гибридов.

Созданные секалотритикум при отборе на продуктивность претерпевали быструю стабилизацию в мейозе. Стабилизация происходила за счет сужения спектра и быстрого снижения в F_1-F_3 частоты нарушений во втором мейотическом делении (на 15–50%), в том числе уровня элиминации хромосом (с ~30% до ~6% тетрад с микроядрами). В F_{3-5} и последующих поколениях у секалотритикум более выражена нормализация первого деления мейоза: средний уровень аномалий в MI снижался с 57,3% у секалотритикум F_3 до 16,9% в F_{5-9} , а также в среднем с 25–30% до 8–10% в AI-AII, и достигал менее 5% на стадии тетрад. Цитологическая стабилизация их генома практически завершалась в F_5 [7].

Хромосомные изменения, о которых говорилось выше, как правило, происходят быстро и наиболее интенсивно в первых

поколениях, до 5-го включительно [7, 21], и особенно в F_1 до удвоения хромосом [21]. Таким образом, продолжительные циклы скрещиваний при получении секалотритикум сыграли значительную роль в стабилизации гибридного кариотипа и восполнении возможных потерь в первых поколениях после гибридизации участков родительских геномов. В настоящее время созданный нами генофонд секалотритикум включает более 50 уникальных линий. Стабильные формы по продуктивности сравнимы или превосходят исходные тритикале, характеризуются сравнительно более широким диапазоном изменчивости, по некоторым признакам морфотипа ближе к ржи. Линии секалотритикум идентифицированы с использованием дифференциального окрашивания хромосом (C-banding) и рестрикционного анализа видоспецифических последовательностей цитоплазматической ДНК хлоропластов и митохондрий.

Различия в структуре CENH3 между родительскими формами позволяют судить о регуляции уровня экспрессии родительских форм белка в новом геномном окружении, возникающем в гибридной клетке при отдаленной гибридизации. Первая работа по изучению возможной связи между различиями в структуре CENH3 у родительских форм и процессами расхождения хромосом родительских геномов при делении гибридных клеток была проведена на гибридах, полученных от скрещивания культурного ячменя *H. vulgare* и его ближайшего дикого родственника *H. bulbosum* [23]. Молекулы CENH3 не включались в центромеры хромосом *H. bulbosum*, которые при этом инактивировались и элиминировались из гибридных эмбрионов. Возможно, отсутствие CENH3 в центромерах хромосом *H. bulbosum* обусловлено заметными различиями в структуре этого белка между видами ячменя, в особенности в структуре N-концевой (NTT) последовательности [23]. В отличие от видов ячменя последовательности ДНК α и β форм CENH3 различных видов ржи и пшеницы имеют очень высокую (95–99%) идентичность [24], что значительно затрудняет поиск межвидовых различий и, соответственно, определение характера их наследования в гибридных геномах.

Консенсусные нуклеотидные последовательности NTT α CENH3 родительских форм Верасень и Михась, приведенные на рис. 2, отличаются всего одной заменой в позиции 73, ведущей к несинонимичной замене аминокислоты серин на пролин. Она обнаружена у сорта Верасень в 73,3% проанализированных клонов (таблица), тогда как для отцовской формы тритикале Михась в этой позиции замена наблюдается только у 7,7% клонов. Полученный от скрещивания этих родительских форм гибрид секалотритикум STr BM наследует высокое содержание пролина от материнского родителя Верасень (67,4%, таблица). Однако эта тенденция не поддерживается в другом скрещивании, где гибрид STr ВД показывает столь же низкий процент замен, как и отцовский родитель — сорт тритикале Дубрава (рис. 2, таблица). С другой стороны, в этой комбинации скрещивания в нескольких позициях (122, 123, 161, 166, 198) у Дубравы наблюдается значительно высокий процент однонуклеотидных замен по сравнению с материнским родителем Верасень (рис. 2, таблица), что ведет к несинонимичным заменам аминокислот, а именно: аланин заменяется на валин, аланин на глицин, глутаминовая кислота на глутамин и лизин на аспарагин. Интересный результат заключается в том, что у гибрида секалотритикум STr ВД в этих позициях процент нуклеотидных замен (уровень полиморфизма) значительно превышает даже показатели тритикале Дубрава, демонстрируя своеобразный «гетерозисный эффект». Таким образом, полученные

результаты указывают, что оба родительских генома участвуют в формировании структуры CENH3 у гибридов секалотритикум. Вместе с тем, сравнение нуклеотидных последовательностей наиболее варибельной N-концевой части молекулы CENH3 у родительских форм и гибридов не выявляет определенной тенденции в наследовании.

Литературные данные о регуляции экспрессии различных генов и других классов последовательностей ДНК в гибридных геномах, полученных при скрещивании различных видов или контрастных популяций одного вида, указывают на сложный и неоднородный характер наследования. Явление ядрышкового доминирования, характерного для экспрессии генов рибосомальной РНК у тритикале, как пример геномного доминирования, заключается в репрессии генов рРНК ржи посредством метилирования цитозина [25]. Однако, такой тип наследования, а именно геномное доминирование одного из родителей, не характерно для большинства других примеров. Результаты изучения экспрессии генов у гибридов, полученных от различных комбинаций экотипов арабидопсиса, противоречивы. В одном случае идентифицировано общее материнское доминирование [26], тогда как в другом найдено, что материнский и отцовский геномы транскрипционно эквивалентны [27]. В следующей работе показано, что различные гибридные комбинации показывают значительные вариации в активации родительских аллелей [28]. У некоторых гибридов между экотипами арабидопсиса зафиксирован гетеро-



Рис. 2. Консенсусные нуклеотидные последовательности NTT-домена α CENH3, полученных из родительских форм ржи (Верасень), тритикале (Михась, Дубрава) и амфидиплоидов секалотритикум (STr BM и STr ВД)

Позиции нуклеотидного полиморфизма (SNP, single nucleotide polymorphism) показаны: с частотой ниже 50% — белым цветом, с частотой выше 50% — светло-серым цветом. Точная количественная оценка SNP приведена в табл. Звездочки указывают на позиции SNP, ведущие к несинонимичным заменам аминокислот, черные круги — на SNP, ведущие к синонимичным заменам. Позиции 122 и 123 расположены в одном кодоне и ведут к замене аминокислоты аланин на валин.

Таблица

Количественное определение нуклеотидного полиморфизма (SNP) в NTT-домене α CENH3, ведущего к заменам аминокислот (в % от числа секвенированных клонов)

Образец	Позиция в нуклеотидной последовательности				
	73	122, 123	161	166	198
Рожь Верасень (S RRRR, 2n = 28)	73,3	6,7	-	-	-
Тритикале Михась (T AABRR, 2n = 42)	7,7	-	-	-	-
Секалотритикум STr BM Верасень $\times$ Михась (S RRAABB, 2n = 42)	67,4	4,7	-	-	2,3
Тритикале Дубрава (T AABRR, 2n = 42)	6,7	6,7	13,3	13,3	20,0
Секалотритикум STr ВД Верасень $\times$ Дубрава (S RRAABB, 2n = 42)	2,6	44,7	39,5	39,5	36,8

Примечание. Для каждого образца родительских форм было секвенировано по 15 клонов. Для гибридов секалотритикум в среднем по 15 клонов для каждого из 5 растений. Консенсусные нуклеотидные последовательности приведены на рис. 2, несинонимичные замены аминокислот, являющиеся следствием нуклеотидного полиморфизма, приведены в тексте

зис по биомассе и урожаю семян [29]. В них 95% экспрессирующихся генов имели промежуточный уровень экспрессии. Среди большинства остальных неаддитивно экспрессирующихся генов сдвиг экспрессии был в сторону материнского родителя [29].

При изучении транскрипционной и эпигенетической адаптации хромосом кукурузы в дополненных линиях овес-кукуруза, содержащих полный геном овса и отдельные хромосомы кукурузы, большинство генов кукурузы показало специфическую для кукурузы транскрипцию, однако подавление генной активности было более распространенной тенденцией, чем активация [30].

Наибольший интерес для нас представляют исследования экспрессии родительских генов в аллополиплоидных гибридах, полученных от скрещивания видов пшеницы и ржи. Такие исследования на секалотритикум нам не известны, однако на аллогексаплоидных тритикале от скрещивания *T. turgidum* $\times$ *S. cereale* был проведен обширный анализ экспрессии генов ржи в кДНК, выделенных из различных тканей гибридных растений [31]. Были определены классы отсутствующих или молчащих

генов ржи. Сравнение между диплоидной рожью и гексаплоидным тритикале выявило 112 контигов кДНК ржи (~0,5% от общего количества), которые не определялись анализом экспрессии ни в одной из тканей тритикале, хотя их экспрессия была относительно высока в тканях ржи. Особенно интересен факт, что неэкспрессирующиеся в тритикале гены ржи имели значительно меньшую гомологию к соответствующим гомеологам в геноме пшеницы или других видов *Triticum*, чем 200 случайно отобранных генов ржи. Таким образом, гены ржи с низким подобием к их гомеологам в геномах *Triticum* имеют более высокую вероятность быть репрессированными или отсутствовать вследствие делеций в геномах аллополиплоидов. Это заключение хорошо согласуется с нашими результатами. Высокая идентичность последовательностей CENH3 между рожью и пшеницей обеспечивает высокий уровень экспрессии обеих родительских форм в гибридных геномах секалотритикум. При этом у некоторых гибридов (STr ВД) наблюдается резкое возрастание полиморфизма по нескольким позициям в последовательностях CENH3.

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (проект № 18-54-00013), Российской программы фундаментальных научных исследований (проект 0310-2018-0010) и Белорусского республиканского фонда фундаментальных исследований (проект Б18Р-170).

Заключение

В данной работе приведены результаты цитогенетического анализа кариотипов гибридов секалотритикум и сравнение структуры N-концевого домена CENH3 между родительскими и гибридными формами. Показано, что кариотипы созданных секалотритикум являются стабильными сбалансированными гексаплоидами и не содержат мини-хромосом с делетированными плечами или участками плеч. Оба родительских генома участвуют в формировании структуры CENH3 у гибридов секалотритикум. Высокая идентичность последовательностей CENH3 ржи и пшеницы обеспечивает высокий уровень экспрессии обеих родительских форм в гибридных геномах.

Список использованных источников

1. Wendel J.F. Genome evolution in polyploids // *Plant Mol. Biol.* – 2000. – Vol. 42. – P. 225–249.
2. Feldman M., Levy A.A. Genome evolution due to allopolyploidization in wheat // *Genetics.* – 2012. – Vol. 36. – P. 763–774.
3. McClintock B. Mechanisms that rapidly reorganize the genome // *Stadler Genet. Symp.* 10. – 1978. – P. 25–48.
4. Lukaszewski A.J., Gustafson J.P. Translocations and modifications of chromosomes in triticale x wheat hybrids // *Theor. Appl. Genet.* – 1983. – Vol. 64. – P. 239–248.
5. Gustafson J.P., Lukaszewski A.J., Bennett M.D. Somatic deletion and redistribution of telomeric heterochromatin in the genus *Secale* and in *Triticale* // *Chromosoma.* – 1983. – Vol. 88. – P. 82–96.
6. Kashkush K., Feldman M., Levy A.A. Gene loss, silencing and activation in a newly synthesized wheat allotetraploid // *Genetics.* – 2002. – Vol. 160. – P. 1651–1659.
7. Гордей И.А., Белько Н.Б., Люсиков О.М. Секалотритикум (*xSecalotriticum*): генетические основы создания и формирования генома. Минск: Беларуская навука. – 2011. – 214 с.
8. Sanci M., Pickering R., Kumke K. et al. Loss of centromeric histone H3 (CENH3) from centromeres precedes uniparental chromosome elimination in interspecific barley hybrids // *Proc. Natl. Acad. Sci. USA.* – 2011. – Vol. 108. P. 498–505.
9. Ishii T., Karimi-Ashtiyani R., Houben A. Haploidization via chromosome elimination: means and mechanisms // *Annu. Rev. Plant Biol.* – 2016. – Vol. 67. – P. 421–438.
10. Amor D.J., Kalitsis P., Sumer H., Choo K. Building the centromere: from foundation proteins to 3D organization // *Trends Cell Biol.* – 2004. – Vol. 14. – P. 359–368.
11. Comai L., Maheshwari S., Marimuthu P.A. Plant centromeres // *Curr. Opin. Plant Biol.* – 2017. – Vol. 36. – P. 156–167.
12. Cheeseman I.M., Desai A. Molecular architecture of the kinetochore-microtubule interface // *Nat. Rev. Mol. Cell Biol.* – 2008. – Vol. 9. – P. 33–46.
13. Allshire R.C., Karpen G.H. Epigenetic regulation of centromeric chromatin: old dogs, new tricks? // *Nat. Rev. Genet.* – 2008. – Vol. 9. – P. 923–937.
14. Talbert P.B., Bryson T.D., Henikoff S. Adaptive evolution of centromere proteins in plants and animals // *J. Biol.* – 2004. – Vol. 3. – P. 18.
15. Гордей И.А., Люсиков О.М., Гордей И.С. и др. Молекулярно-генетические основы создания нового генофонда ржи и ржано-пшеничных амфидиплоидов секалотритикум // Молекулярная и прикладная генетика. – 2017. – Т. 23. – С. 26–39.
16. Badaeva E.D., Sozinova L.F., Badaev N.S. et al. “Chromosomal passport” of *Triticum aestivum* L. em Thell. cv. Chinese Spring and standardization of chromosomal analysis of cereals. // *Cereal Res Commun.* – 1990. – Vol. 18. – P. 273–281.
17. Chomczynski P., Sacchi N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction // *Anal. Biochem.* – 1987. – Vol. 162. – P. 156–159.
18. Pearson W. R., Lipman D. J. Improved tools for biological sequence comparison // *Proc. Natl. Acad. Sci. USA.* – 1988. – Vol. 85. – P. 2444–2448.

19. Alkhimova A.G., Heslop-Harrison J.S., Shchapova A.I., Vershinin A.V. Rye chromosome variability in wheat-rye addition and substitution lines // *Chromosome Res.* – 1999. – Vol. 7. – P. 205–212.
20. Fu S., Lv Z., Guo X., Han F. Alternation of terminal heterochromatin and chromosome rearrangements in derivatives of wheat-rye hybrids // *J. Genet. Genomics.* – 2013. – Vol. 40. – P. 413–420.
21. Appels R., Gustafson J.P., May C.E. Structural variation in the heterochromatin of rye chromosomes in triticale // *Theor. Appl. Genet.* – 1982. – Vol. 63. – P. 235–244.
22. Ma X., Gustafson J.P. Timing and rate of genome variation in triticale following allopoloidization // *Genome.* – 2006. – Vol. 49. – P. 950–958.
23. Sanei M., Pickering R., Kumke K. et al. Loss of centromeric histone H3 (CENH3) from centromeres precedes uniparental chromosome elimination in interspecific barley hybrids // *Proc. Natl. Acad. Sci. USA.* – 2011. – Vol. 108. – P. 498–505.
24. Evtushenko E.V., Elisafenko E.A., Gatzkaya S.S. et al. Conserved molecular structure of centromeric histone CENH3 in *Secale* and its phylogenetic relationships // *Sci. Reports* 7. – 2017. – P. 17628.
25. Lima-Brito J., Guedes-Pinto H., Heslop-Harrison J.S. The activity of nucleolar organizing chromosomes in multigeneric F1 hybrids involving wheat, triticale and tritordeum // *Genome.* – 1998. – Vol. 41. – P. 763–768.
26. Autran D., Baroux C., Raising M.T. et al. Maternal epigenetic pathways control parental contributions to *Arabidopsis* early embryogenesis // *Cell.* – 2011. – Vol. 145. – P. 707–719.
27. Nodine M.D., Bartel D.P. Maternal and paternal genomes contribute equally to the transcriptome of early plant embryos // *Nature.* – 2012. – Vol. 482. – P. 94–97.
28. Del Toro-De Leon G., Garcia-Aguilar M., Gillmor C.S. Non-equivalent contributions of maternal and paternal genomes to early plant embryogenesis // *Nature.* – 2014. – Vol. 514. – P. 624–627.
29. Alonso-Peral M.M., Trigueros M., Sherman B. et al. Patterns of gene expression in developing embryos of *Arabidopsis* hybrids // *Plant J.* – 2017. – Vol. 89. – P. 927–939.
30. Dong Z., Yu J., Li H. et al. Transcriptional and epigenetic adaptation of maize chromosomes in Oat-Maize addition lines // *Nucl. Acids Res.* – 2018. – Vol. 46. – P. 5012–5024.
31. Khalil H.B., Ehdaei M.-R., Xu Y. et al. Identification and characterization of rye genes not expressed in allohexaploid triticale // *BMC Genomics.* – 2015. – Vol. 16. – P. 281.

Y.A. Lipikhina¹, E.V. Evtushenko¹, O.M. Lyusikov², I.S. Gordei², I.A. Gordei², A.V. Vershinin¹

RYE-WHEAT AMPHIDIPLOIDS SECALOTRITICUM IS A NEW MODEL FOR STUDYING THE ACTIVITY OF GENES OF CENTROMERE PROTEIN CENH3

¹Institute of Molecular and Cellular Biology, SB RAS
Novosibirsk, 630090, Russia

²Institute of Genetics and Cytology, NASB
Minsk, 220072, the Republic of Belarus

Remote hybridization is stressful to plants. It causes rearrangements to parental genomes in hybrid progeny ("genome shock"), including loss of genes, changes in gene expression, silencing, deletions, inversions, and translocations of chromosome regions. The developed synthetic allopolyploids secalotriticum represent a new interesting model for studying the impact of remote hybridization on the genomic regions responsible for correct cell division and chromosome distribution in hybrid cells. These processes are orchestrated by the centromere. The hallmark of an active centromere is the presence of histone H3 in centromeric nucleosomes in the form of its centromere-specific modification denoted in plants as CENH3. This work provides results of a cytogenetic analysis of secalotriticum karyotypes and compares the N-terminal domains of CENH3 between the parental and hybrid forms. It has been demonstrated that secalotriticum plants are stable balanced hexaploids devoid of mini-chromosomes with deleted arms, either in full or in part. Both parental genomes contribute to CENH3 formation in secalotriticum. A high level of nucleotide identity between rye and wheat CENH3 provides for a high level of expression of both parental forms in hybrid genomes.

Key words: rye, wheat, remote hybridization, rye-wheat amphidiploids, tritcale, secalotriticum, allopolyploids, genome, cytogenetics, centromere, centromere-specific histone 3 (CENH3), gene expression.

Дата поступления статьи: 13 августа 2018 г.

А.Н. Радченко, С.И. Михальская, М.О. Дыкун, Л.В. Сирант

ВЛИЯНИЕ АЛЛЕЛЕЙ ГЕНА *Vp-1B* НА УСТОЙЧИВОСТЬ К ПРЕДУБОРОЧНОМУ ПРОРАСТАНИЮ У МЯГКОЙ ПШЕНИЦЫ

Институт физиологии растений и генетики НАН Украины
Украина, 03022, г. Киев, ул. Васильковская, 31/17
e-mail: ales20009@ukr.net

Предуборочное прорастание зерна наносит значительный ущерб сельскому хозяйству, снижая урожай зерна и ухудшая его качество. Для изучения корреляции между аллельным состоянием гена *Vp-1B* и устойчивостью к предуборочному прорастанию были проанализированы 35 сортов пшеницы. Всего было обнаружено три аллеля гена *Vp-1B*: *Vp-1Ba*, *Vp-1Bc*, *Vp-1Bd*. Анализ коллекции образцов пшеницы показал, что 51,5% образцов имеют аллель *Vp-1Bc*, 37,1% образцов — *Vp-1Ba*, 11,4% образцов имеют аллель *Vp-1Bd*. Был определен индекс прорастания в исследованных образцах, с использованием которого все образцы были разделены на 3 группы: устойчивые, средние по устойчивости и неустойчивые к предуборочному прорастанию. Аллель *Vp-1Bc* характерен для устойчивых образцов, аллель *Vp-1Ba* — для неустойчивых образцов.

Ключевые слова: озимая мягкая пшеница, ПЦР-анализ, предуборочное прорастание, ген *Vp-1B*.

Введение

Мировое производство пшеницы превышает 650 млн т в год, и потребность в зерне постоянно растет [1]. Во многих странах, а также в Украине, в посевах озимой и яровой пшеницы преобладают краснозерные сорта, хотя наиболее высоким спросом пользуется белое зерно, которое дает максимально высокий выход муки и позволяет получать различные виды макаронных изделий высокого качества. Несмотря на это, из-за увеличения риска предуборочного прорастания, вместо белозерных в Украине преимущественно выращивают краснозерные сорта яровой и озимой пшеницы.

В годы с неблагоприятной погодой (дожди, холод) в период налива зерна и уборки урожая страдают также краснозерные сорта, при этом достаточно небольшого количества осадков (около 10–15 мм) [2, 3]. В процессе прорастания разрушается крахмал, деградируют белки и другие вещества, ухудшаются физические и технологические свойства зерна [4]. Разрушение крахмала в эндосперме в результате активности фермента альфа-амилазы может происходить и без видимых признаков прорастания зерна, что важно учитывать как при хранении зерна, так и при использовании его в хлебопекарном производстве [5].

В ряде стран проводится учет потерь от предуборочного прорастания. Например, в Канаде они составляют около 100 млн долларов в год [6]. Давно замечено, что у краснозерных сортов пшеницы устойчивость к предуборочному прорастанию, как правило, выше, чем у белозерных. Некоторые исследователи считают, что это связано с генами, которые контролируют окраску зерна [7–9]. Краснозерность контролируют доминантные аллели, а белозерность — рецессивные аллели R-генов, локализованных на длинных плечах хромосом третьей группы: *R-A1* (старое название *R2*) — на 3AL, *R-B1* (*R3*) — на 3BL и *R-D1* (*R1*) — на 3DL [10]. Доминантные R-аллели представляют собой Myb-подобные факторы транскрипции, которые регулируют экспрессию генов биосинтеза флавоноидов *ChS*, *Chi*, *F3h* и *Dfr* [11]. Положительная роль окраски зерна в устойчивости к предуборочному прорастанию подтверждается исследованиями на почти изогенных линиях [12, 13], но молекулярные механизмы этого явления не известны.

В ряде исследований говорится о том, что рядом с R-генами локализованы три *Vp-1* гена — *Vp-1A*, *Vp-1B* и *Vp-1D* (или *TaVp-A1*, *TaVp-B1* и *TaVp-D1*) [14, 15, 16], которые, с одной стороны, ускоряют созревание зерна, с другой — репрессируют его прорастание

у разных видов растений. У мягкой пшеницы установлен множественный аллелизм *Vp-1* генов [17]. Периоды покоя семян и устойчивость к предуборочному прорастанию наиболее тесно ассоциированы с рядом аллелей *Vp-1B* локуса [18], а также с *Vp-1A* локусом [19]. Устойчивость к предуборочному прорастанию возможно улучшить путем встраивания в хромосому пшеницы *Vp-1* генов кукурузы (*Zea mays*) [20].

Исходя из анализа последовательностей ДНК китайскими исследователями разработан маркер *Vp1B3*, который был использован для анализа пшеницы с различным уровнем устойчивости к предуборочному прорастанию [21]. Было четко выявлено три ампликона 845 п. н., 569 п. н., 652 п. н., которым соответствуют три аллеля гена *Vp-1B*: *Vp-1Bb*, *Vp-1Bc* и *Vp-1Ba*. Устойчивые образцы имели ампликоны 845 п. н., 569 п. н., а неустойчивые — ампликон 652 п. н. Анализ нулитетрасомных и дителесомных линий подтвердил, что этот маркер находится на хромосоме 3В. Необходимо отметить, что данные закономерности были установлены на белозерной пшенице.

При оценке аллельного состава гена *Vp-1B* на сортах мягкой пшеницы Центральной и Северной Европы были обнаружены четыре аллеля, три из которых ранее идентифицированы в китайских сортах *Vp-1Ba*, *Vp-1Bb* и *Vp-1Bc*, четвертым был новый аллель *Vp-1Bd*. Частоты различных аллелей в этом наборе европейских сортов пшеницы были следующими: *Vp-1Ba* (54%) > *Vp-1Bc* (21%) > *Vp-1Bd* (20%) > *Vp-1Ba* + *c* (4%) > *Vp-1Bb* (1 %), при этом аллель *Vp-1Bb* присутствовал только в двух французских сортах «Altria» и «Recital». Аллель *Vp-1Bd* присутствовал только в 20% европейских сортах пшеницы и отсутствовал среди китайских сортов [25, 26].

В связи с тем, что в украинском генофонде пшеницы преобладают краснозерные сорта, для определения связи между аллельным составом гена *Vp-1B* и их устойчивостью к предуборочному прорастанию актуальным было продолжение исследований в данном направлении.

Целью работы было изучение корреляции между устойчивостью к предуборочному прорастанию украинских сортов мягкой пшеницы и аллелями гена *Vp-1B*.

Материалы и методы

В работе исследованы 35 сортов мягкой пшеницы. Выделение растительной ДНК осуществляли модифицированным ЦТАБ методом. Выявление аллелей гена *Vp-1B* осуществляли путем ПЦР с использованием специфических праймеров. Последовательность праймеров к гену *Vp-1B*: праймер *Vp-1BF*: TGCTCCTTTCCCAATTGG и праймер *Vp-1BR*: ACCCTCCTGCAGCTCATTG [21].

Реакционные смеси включали: специфические праймеры, по 2 мкл буфера для ПЦР 10 × DreamTaq™ Green Buffer (Thermo Scientific), по 0,2 ммоль каждого дезоксирибонуклеотидов-3-фосфата (Thermo Scientific), 0,5–0,75 единицы полимеразы DreamTaq™ DNA Polymerase (Thermo Scientific), 50–100 нг суммарной ДНК, деионизированную воду Milli-Q до конечного объема 20 мкл. Реакции амплификации проводили в термоциклере Arctic Thermal Cycler (Thermo Scientific), Mastercycler gradient (Eppendorf) и Techne TC-312.

Программа амплификации ПЦР для выявления гена *Vp-1B* включала следующие этапы: начальная денатурация 3 мин при 94 °С; 35 циклов денатурации 30 с при 94 °С; ренатурация 30 с при 59 °С; элонгация 60 с при 72 °С и финальная элонгация 5 мин при 72 °С.

Полученные продукты ПЦР разделяли путем горизонтального электрофореза в агарозном геле в SB буфере при напряжении 3–7 В/см в зависимости от размеров ожидаемых ампликонов. Для визуализации использовали этидиум бромид в концентрации 0,5 мкг/мл. Полученные результаты документировались под УФ-светом фотоаппаратом Canon EOS 600D со светофильтром. Обработку электрофореграмм осуществляли с помощью графического редактора GIMP.

Индекс прорастания (GI) рассчитывали по Walker-Simmons: $GI = (7x_{n1} + 6x_{n2} + \dots + 1x_{n7}) / 7 \times N$, где $n1, n2, \dots, n7$ — число зерен, которые проросли на первый, второй и последующие дни до седьмого дня, соответственно, N — общее число зерен. Индекс определяли в трех повторностях для каждого образца, анализировали по 50 зерен в повторности.

Число падения определяли по методу Хагберга-Пертена на приборе Falling Number 1700 [22, 23]. Высокое число падения является показателем низкой активности альфа-амилазы

и, следовательно, хорошего содержания белка при выпекании мучных изделий.

Статистические показатели рассчитывали по Доспехову [24].

Результаты и обсуждение

Нами был проанализирован набор из 35 сортов мягкой пшеницы по аллельному составу гена *Vp-1B*, который контролирует устойчивость к предуборочному прорастанию, с помощью STS-маркера *Vp1B3*.

В исследованных образцах был определен индекс прорастания, по которому все образцы были условно разделены на 3 группы: устойчивые, средние по устойчивости и неустойчивые. Неустойчивые образцы имели среднее значение индекса прорастания 78,6, в группах устойчивых и средних по устойчивости значение индекса прорастания было 19,1 и 44,9, соответственно. Анализ индекса прорастания показал, что у исследованных генотипов по этому показателю были значительные различия (табл. 1).

Таблица 1

Характеристика сортов озимой мягкой пшеницы

Сорт	Алель	Индекс прорастания. ИП × 100	Число падения
Смуглянка	<i>Vp-1Bc</i>	9,0	444,5
Дарунок Подилля	<i>Vp-1Bc</i>	10,0	375
Ласуня	<i>Vp-1Bc</i>	17,0	_*
Украинка 0246	<i>Vp-1Bc</i>	17,5	_*
Солоха	<i>Vp-1Ba</i>	21,5	401,5
Ятрань 60	<i>Vp-1Bc</i>	23,0	204
Наталка	<i>Vp-1Bc</i>	24,0	248
Почаивка	<i>Vp-1Bc</i>	24,0	360,5
Панна	<i>Vp-1Bc</i>	26,0	338
Злука	<i>Vp-1Ba</i>	27,0	365
Нива одесская	<i>Vp-1Bc</i>	27,0	428
Каланча	<i>Vp-1Bc</i>	27,0	316
Мироновская 808	<i>Vp-1Bc</i>	31,0	_*
Крыжинка	<i>Vp-1Ba</i>	31,0	83
Достаток	<i>Vp-1Bc</i>	32,0	406,5
Веснянка	<i>Vp-1Ba</i>	33,0	_*
Снигурка	<i>Vp-1Ba</i>	34,0	365
Астарта	<i>Vp-1Ba</i>	35,0	399
Славна	<i>Vp-1Bc</i>	37,5	444,5
Киевская остистая	<i>Vp-1Bc</i>	39,0	473
Деметра	<i>Vp-1Bd</i>	40,5	_*
Лазурна	<i>Vp-1Ba</i>	42,5	209
Мироновская раннестыгла	<i>Vp-1Bc</i>	47,0	206
Малиновка	<i>Vp-1Bc</i>	48,0	441

Продолжение табл. 1

Сорт	Аллель	Индекс прорастания. ИП × 100	Число падения
Селянка	<i>Vp-1Bc</i>	50,0	— *
Куяльник	<i>Vp-1Bc</i>	58,0	385
Сонечко	<i>Vp-1Ba</i>	53,0	124
Хуртовина	<i>Vp-1Bd</i>	58,0	175,5
Фаворитка	<i>Vp-1Ba</i>	58,0	400
Переяславка	<i>Vp-1Bd</i>	60,0	215
Богдана	<i>Vp-1Bd</i>	68,0	255
Зимоярка	<i>Vp-1Ba</i>	70,0	309,5
Чигиринка	<i>Vp-1Ba</i>	73,0	191
Володарка	<i>Vp-1Ba</i>	81,0	93,5
Орийка	<i>Vp-1Ba</i>	86,0	181

Примечание. Показатели, отмеченные звездочкой, в данных образцах не определялись; $F_{\text{факт}} = 7,8713 > F_{\text{теор}} = 3,3276$; $p = 0,001859$

В данной выборке сортов нами были выявлены три аллеля гена *Vp-1B*, которые были обозначены как *Vp-1Bd*, *Vp-1Bc* и *Vp-1Ba*. Наиболее распространенным был аллель *Vp-1Bc*, который имел 18 сортов: Нива одесская, Панна, Дарунок Подилля, Почаивка, Смуглянка, Ятрань 60, Наталка, Каланча, Мироновская 808, Мироновская раннестыгла, Селянка, Достаток, Малиновка, Славна, Киевская остистая, Куяльник, Украинка 0246, Ласуня. Вторым по распространенности аллелем был *Vp-1Ba* — присутствует в 13 образцах: Астарта, Зимоярка, Лазурна, Сонечко, Фаворитка, Злука, Солоха, Снигурка, Веснянка, Крыжинка, Володарка, Чигиринка, Орийка. Аллель *Vp-1Bd* был обнаружен у четырех сортов — Деметра, Хуртовина, Переяславка, Богдана. Аллель *Vp-1Bd* имеет 25 п. н. делецию в третьей области интрона по сравнению с нуклеотидной последовательностью аллеля *Vp-1Ba* [25, 26].

Анализ коллекции образцов пшеницы показал, что 51,5% образцов имеют в своем геноме аллель *Vp-1Bc* (569 п. н.), 37,1% образцов — аллель *Vp-1Ba* (652 п. н.), 11,4% образцов имеют аллель *Vp-1Bd* (627 п. н.) (рис.1–3). Наши результаты выявили значительное аллельное многообразие гена *Vp-1B* в украинских сортах пшеницы.

В этом исследовании мы уделяли наибольшее внимание поиску ассоциаций между аллелями

гена *Vp-1B* и устойчивостью к предуборочному прорастанию (табл. 2). К устойчивым относятся образцы, которые имеют индекс прорастания на уровне 0–27, — Наталка, Каланча, Ласуня, Смуглянка, Ятрань 60, Нива одесская, Панна, Украинка 0246, Злука, Почаивка, Солоха, Дарунок Подолля. Для образцов, имеющих среднюю устойчивость, характерен индекс прорастания на уровне 27–69; это сорта — Деметра, Киевская остистая, Фаворитка, Крыжинка, Веснянка, Мироновская 808, Малиновка, Селянка, Мироновская раннестыгла, Славна, Переяславка, Сонечко, Достаток, Лазурна, Хуртовина, Астарта, Снигурка, Куяльник, Богдана. К неустойчивым относятся образцы, которые имеют индекс прорастания 70–100, — Володарка, Орийка, Чигиринка, Зимоярка.

Было показано, что для 13 образцов, несущих аллель *Vp-1Ba*, характерен более высокий индекс прорастания. Среди 13 сортов с этим аллелем семь были средними по устойчивости — Астарта, Лазурна, Сонечко, Фаворитка, Веснянка, Снигурка, Крыжинка. Два сорта были устойчивыми к предуборочному прорастанию — Злука, Солоха. Четыре сорта — неустойчивыми к предуборочному прорастанию — Володарка, Чигиринка, Орийка, Зимоярка.

Аллель *Vp-1Bc* присутствовал в 18 образцах, из них 10 устойчивых: Нива одесская, Панна,

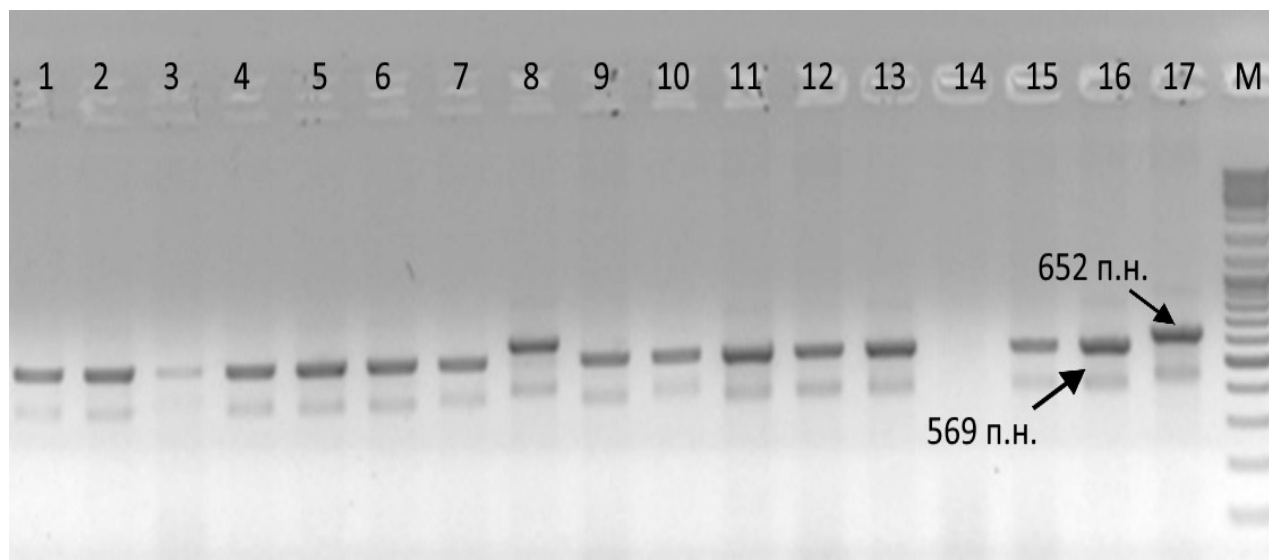


Рис. 1. Электрофореграмма детекции аллелей гена *Vp-1B*: 1 — Панна, 2 — Славна, 3 — Селянка, 4, 5 — Достаток, 6 — Почаївка, 7 — Куяльник, 8 — Сонечко, 9 — Мироновская 808, 10, 11 — Каланча, 12, 13 — Дарунок Подилля, 14 — Киевская остистая, 15 — Мироновская раннестыгла, 16 — Малиновка, 17 — Сонечко, М — маркер Ladder Mix

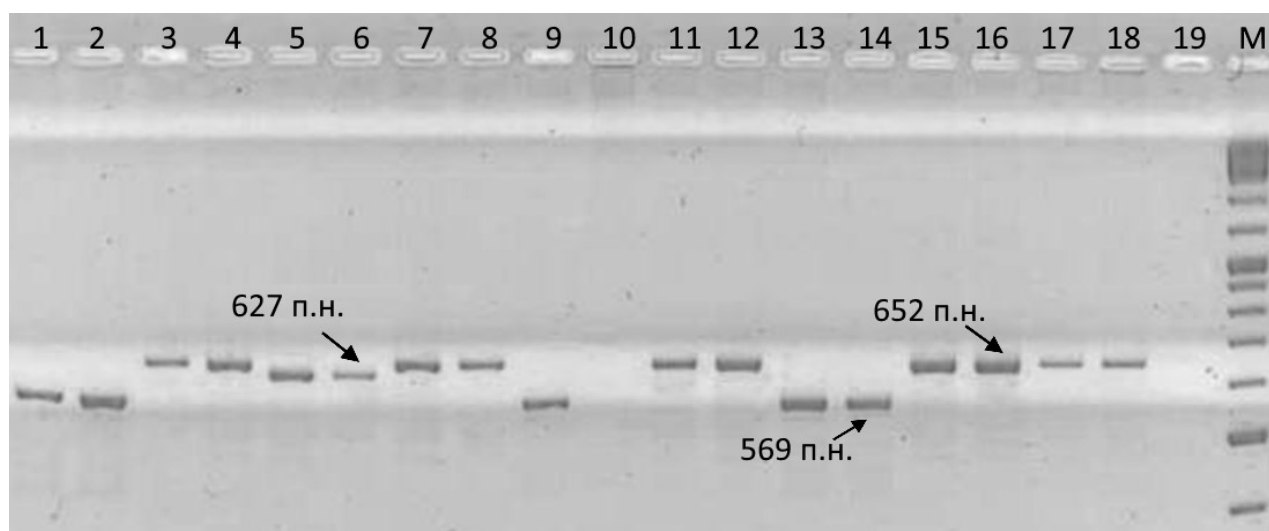


Рис. 2. Электрофореграмма детекции аллелей гена *Vp-1B*: 1, 2 — Смуглянка, 3, 4 — Злука, 5, 6 — Деметра, 7, 8 — Зимоярка, 9, 10 — Нива одесская, 11, 12 — Лазурна, 13, 14 — Ятрань 60, 15, 16, 17, 18 — Снигурка, 19 — контроль H_2O , М — маркер Ladder Mix

Дарунок Подолля, Почаївка, Смуглянка, Ятрань 60, Наталка, Каланча, Ласуня, Украинка 0246 и 8 средних по устойчивости: Мироновская 808, Мироновская раннестыгла, Селянка, Куяльник, Достаток, Малиновка, Славна, Киевская остистая.

Аллель *Vp-1Bd* присутствовал в 4 средних по устойчивости образцах: Деметра, Хуртовина, Переяславка, Богдана.

На основании полученных нами данных было установлено небольшое статистическое разли-

чие по индексу прорастания у образцов с разным аллельным составом. Показано, что в образцах с аллелем *Vp-1Bc* средний индекс прорастания — 30,3; стандартное отклонение — 7,4; стандартная ошибка — 1,8. В образцах с аллелем *Vp-1Ba* средний индекс прорастания — 49,6; стандартное отклонение — 9,1; стандартная ошибка — 2,4. В образцах с аллелем *Vp-1Bd*: средний индекс прорастания — 56,6; стандартное отклонение — 12,4; стандартная ошибка — 6,1. Было выявлено, что

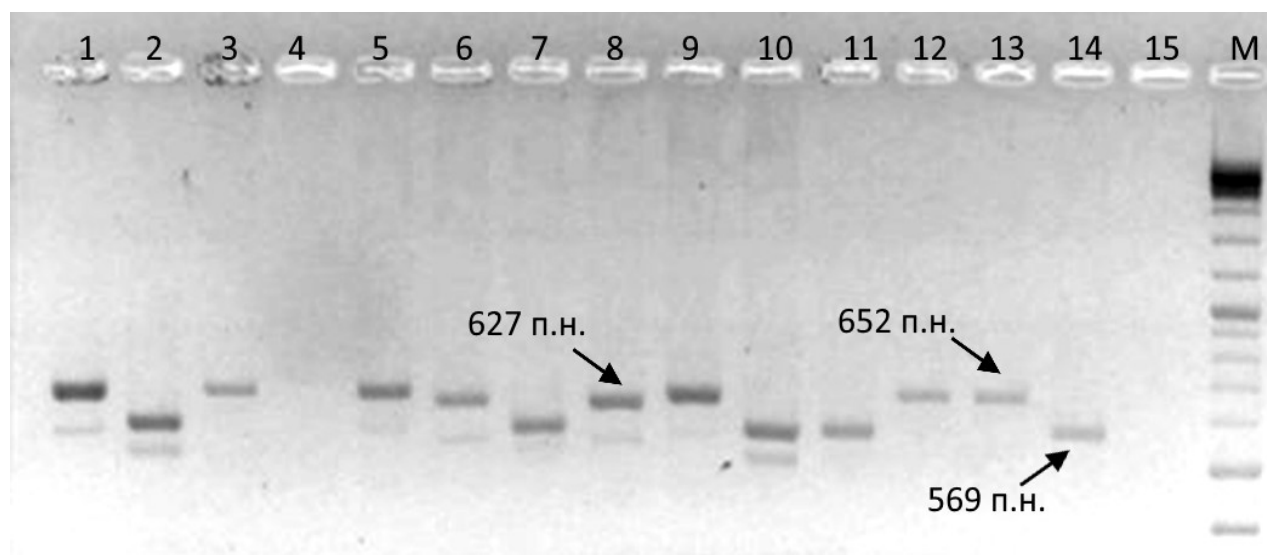


Рис. 3. Электрофореграмма детекции аллелей гена *Vp-1B*: 1 — Чигиринка, 2 — Куяльник, 3,4,5 — Орийка, 6 — Переяславка, 7 — Панна, 8 — Богдана, 9 — Веснянка, 10, 11 — Ласуня, 12 — Веснянка, 13 — Крыжинка, 14 — Киевская остистая, 15 — контроль H_2O , М — маркер Ladder Mix

Таблица 2

Встречаемость аллелей *Vp-1B* у 35 сортов озимой мягкой пшеницы

Сорта	Количество сортов	Индекс прорастания ИП $\times 100$	Диапазон индекса прорастания	<i>Vp-1Bd</i> 627 п. н.	<i>Vp-1Bc</i> 569 п. н.	<i>Vp-1Ba</i> 652 п. н.
Устойчивые	12	19,1	0–27	0 (0%)	10 (55,5%)	2 (15,3%)
Средние по устойчивости	19	44,9	27–69	4 (100%)	8 (44,5%)	7 (53,9%)
Неустойчивые	4	78,6	70–100	0 (0%)	0 (0%)	4 (30,8%)

во всей совокупности образцов существует слабая связь между устойчивостью к прорастанию и аллельным состоянием гена *Vp-1B*.

Склонность мягкой пшеницы к предуборочному прорастанию зерна связывают с высокой активностью амилазы, которая непосредственно коррелирует с числом падения Хагберга. В результате анализа нами обнаружены образцы озимой мягкой пшеницы, которые имеют число падения в пределах 83–473 с. При этом было показано, что неустойчивые образцы, которые имели число падения менее 200, характеризовались наличием аллеля *Vp-1Ba* — 5 образцов, и аллеля *Vp-1Bd* — 1 образец. Среди образцов, которые имели число падения более 200, преобладал аллель *Vp-1Bc* — 14 образцов, при этом аллель *Vp-1Ba* выявлен у 8 вариантов, у 2 вариантов — аллель *Vp-1Bd*, что не дает

четкой зависимости числа падения от аллелей гена *Vp-1B* и свидетельствует о комплексной характеристике данного показателя.

Заключение

Установлено, что исследованные образцы украинской пшеницы имеют три аллеля гена *Vp-1B*. Анализ коллекции образцов пшеницы показал, что 51,5% образцов имеют аллель *Vp-1Bc*, 37,1% образцов *Vp-1Ba*, 11,4% образцов имеют аллель *Vp-1Bd*. С наибольшей частотой выявляются аллель *Vp-1Bc* и аллель *Vp-1Ba*. В образцах краснозерной пшеницы показана слабая взаимосвязь аллелей гена *Vp-1B* и устойчивости к предуборочному прорастанию. У образцов, которые устойчивы к предуборочному прорастанию преобладает аллель *Vp-1Bc*, у неустойчивых — *Vp-1Ba*.

Список использованных источников

- [FAO] Food and Agriculture Organization, 2010: FAO Statistical Year Book. – 2010. – Rome, Italy.
- Mares D.J., Mrva K. Wheat grain preharvest sprouting and late maturity alphaamylase // *Planta*. – 2014. – Vol. 240. – P. 1167–1178.
- Mrva K., Mares D.J. Induction of late maturity alphaamylase in wheat by cool temperature // *Aust. J. Agric. Res.* – 2001. – Vol. 52. – P. 477–484.
- Fu B.X., Hatcher D.W., Schlichting L. Effects of sprout damage on durum wheat milling and pasta processing quality // *Can. J. Plant Sci.* – 2014. – Vol. 94. – P. 545–553.
- Trethowan R.M., Rajaram S., Ellison F.W. Pre-harvest sprouting tolerance in wheat in the field and under rain simulation // *Australian Journal of Agricultural Research*. – 1996. – Vol. 47. – P. 705–716.
- De Pauw R.M., Knox R.E., Singh A.K. Developing standardized methods for breeding preharvest sprouting resistant wheat, challenges and successes in Canadian wheat // *Euphytica*. – 2012. – Vol. 188. – P. 714.
- Bi H., Sun Y., Xiao Y. Characterization of *DFR* allelic variations and their associations with pre-harvest sprouting resistance in a set of red-grained Chinese wheat germplasm // *Euphytica*. – 2014. – Vol. 195. – P. 197–207.
- Cabral A.L., Jordan M.C., McCartney C.A. Identification of candidate genes, regions and markers for pre-harvest sprouting resistance in wheat (*Triticum aestivum* L.) // *BMC Plant Biology*. – 2014. – Vol. 14, № 340. – P. 1–12.
- Lin M., Zhang D., Liu S. Genome-wide association analysis on pre-harvest sprouting resistance and grain color in U.S. winter wheat // *BMC Genomics*. – 2016. – Vol. 17, № 794. – P. 1–16.
- McIntosh R.A., Yamazaki Y., Dubcovsky J. Catalogue of gene symbols for wheat. – 2008.
- Хлесткина Е.К. Гены, детерминирующие окраску различных органов пшеницы // *Вавиловский журнал генетики и селекции*. – 2012. – Т. 16. – С. 202–216.
- Liu S., Sehgal S.K., Li J. Cloning and characterization of a critical regulator for preharvest sprouting in wheat // *Genetics*. – 2013. – Vol. 195. – P. 263–273.
- Wang Y., Wang X.L., Meng J.Y. Characterization of *Tamylb10* allelic variants and development of STS marker for pre-harvest sprouting resistance in Chinese bread wheat // *Molecular Breeding*. – 2016. – Vol. 36, № 148. – P. 1–12.
- Bailey P.C., McKibbin R.S., Lenton J.R. Genetic map locations for orthologous *Vp-1* genes in wheat and rice // *Theor. Appl. Genet.* – 1999. – Vol. 98. – P. 281–284.
- Yang Y., Zhang L., Liu S.H. Characterization of the rich haplotypes of *Viviparous-1A* in Chinese wheats and development of a novel sequence-tagged site marker for pre-harvest sprouting resistance // *Molecular Breeding*. – 2014. – Vol. 33. – P. 75–88.
- Zhang Y., Xia X., He Z. The seed dormancy allele *TaSad-1A* associated with pre-harvest sprouting tolerance is mainly present in Chinese wheat landraces // *Theoretical and Applied Genetics*. – 2017. – Vol. 130. – P. 81–89.
- Yang J., Liu Y., Pu Z. Molecular characterization of high pI alphaamylase and its expression QTL analysis in synthetic wheat RILs // *Mol. Breeding*. – 2014. – Vol. 34. – P. 1075–1085.
- Chang C., Zhang H.P., Feng J.M., Yin B., Si H.Q., Ma S.X. Identifying alleles of *Viviparous1* associated with preharvest sprouting in micro-core collections of Chinese wheat germplasm // *Mol. Breeding*. – 2010. – Vol. 25. – P. 481–490.
- Chang C., Zhang H.P., Zhao Q.X. Rich allelic variations of *Viviparous1A* and their associations with seed dormancy preharvest sprouting of common wheat // *Euphytica*. – 2011. – Vol. 179. – P. 343–353.
- Huang T., Qu B., Li H.P., Zuo D.Y. A maize viviparous 1 gene increases seed dormancy and preharvest sprouting tolerance in transgenic wheat // *J. Cereal Sci.* – 2012. – Vol. 55. – P. 166–173.
- Yang Y., Zhao X.L., Xia L.Q. Development and validation of a *Viviparous-1* STS marker for pre-harvest sprouting tolerance in Chinese wheats // *Theor. Appl. Genet.* – 2007. – Vol. 115. – P. 971–980.
- Hagberg S. A rapid method for determining alphaamylase activity // *Cereal Chem.* – 1960. – Vol. 37. – P. 218–222.
- Perten H. Application of the falling number method for evaluating alphaamylase activity // *Cereal Chem.* – Vol. 41. – P. 127–139.
- Доспехов Б.А. Методика полевого опыта (с основами статистической обработки результатов исследований) // М: Агропромизд. – 1985. – 351 с.

25. Xia L.Q., Ganai M.W., Shewry P.R., He Z.H., Yang Y., Röder M.S. Exploiting the diversity of *Viviparous-1* gene associated with pre-harvest sprouting tolerance in European wheat varieties // *Euphytica*. – 2008. – Vol. 159. – P. 411–417.
26. Xia L.Q., Yang Y., Ma Y.Z. What can the *Viviparous-1* gene tell us about wheat pre-harvest sprouting // *Euphytica*. – 2009. – Vol. 168. – P. 385–394.

A.N. Radchenko, S.I. Mykhalska, M.O. Dykun, L.V. Sirant

THE INFLUENCE ALLELS GENE *VP-1B* TO THE PRE-HARVEST SPROUTING OF SOFT WHEAT

Institute of Plant Physiology and Genetics, NAS of Ukraine
Kyiv, 03022, Ukraine

Pre-harvest sprouting of grain causes considerable damage agriculture, reduces yields, and impaires the quality of grain. To detect the correlation between the allelic state of the *Vp-1B* and the resistance to pre-harvest sprouting, 35 wheat varieties were analyzed. In general, three alleles of the *Vp-1B* gene: *Vp-1Ba*, *Vp-1Bc*, *Vp-1Bd* were identified. The wheat collection on analysis showed that 51,5% of the samples had the *Vp-1Bc* allele, 37,1% of the samples had the *Vp-1Ba* allele, 11,4% of the samples had the *Vp-1Bd* allele. The germination index in the those samples was determined. According to those parameters all samples were divided into 3 groups: stable, medium-stable and unstable to pre-harvest sprouting. Allele *Vp-1Bc* was typical for resistant species, allele *Vp-1Ba* – for unstable species.

Key words: winter soft wheat, PCR-analysis, pre-harvest sprouting, *Vp-1B*.

Дата поступления статьи: 19 июня 2018 г.

Т.В. Семашко, А.И. Горина, Е.А. Мартынова, Л.Н. Валентович, Л.А. Жуковская, А.Г. Лобанок

ПОЛУЧЕНИЕ ВЕКТОРНОЙ КОНСТРУКЦИИ ДЛЯ ЭКСПРЕССИИ В КЛЕТКАХ МИЦЕЛИАЛЬНЫХ ГРИБОВ ГЕНА ИЗОФЕРМЕНТА С2 ПЕРОКСИДАЗЫ ИЗ КОРНЕЙ ХРЕНА

Институт микробиологии НАН Беларуси
Республика Беларусь, 220141 г. Минск, ул. Купревича, 2
e-mail: tsemashko@mbio.bas-net.by

Проведен анализ нуклеотидных последовательностей генов изоферментов группы С пероксидазы из корней хрена, представленных в базе данных GenBank. Сконструирован вектор для экспрессии гена изофермента С2 пероксидазы из корней хрена (*hrpC2*) в клетках мицелиальных грибов на основе синтезируемой последовательности вышеуказанного гена и плазмид: pK18-cat, несущей фрагмент гена каталазы *Penicillium adametzii* для обеспечения встройки плазмиды в геномную ДНК, и pAN7-1, имеющей в своем составе промотор и терминатор гена глицеральдегидфосфатдегидрогеназы *Aspergillus nidulans* для экспрессии *hrpC2*.

Ключевые слова: ген пероксидазы хрена, изофермент С2, экспрессионные векторы, грибы рода *Penicillium*.

Введение

Ферменты — эффективные природные биокатализаторы. Они используются в химической и пищевой промышленности, фармацевтической индустрии, бытовой химии, сельском хозяйстве и т.д. Развитие биотехнологии, генетической и белковой инженерии позволило получать ферменты с новыми свойствами для катализа самых разнообразных химических реакций [1, 2].

Особый интерес представляют ферменты класса оксидоредуктаз, в частности пероксидазы. Они способствуют осуществлению окислительных процессов, выполняют синтетическую и метаболическую, а также защитную функции.

Пероксидазы применяются в аналитической химии, медицинской диагностике и биотехнологии (в фотометрических, электрохимических, хемилюминесцентных, иммуноферментных и ДНК-гибридизационных методах анализа). Так, использование пероксидазы в качестве метки в иммуноферментном анализе позволяет определять различные вещества, такие как гормоны, онкомаркеры, аутоантитела, иммуноглобулины, лекарственные соединения и т.д. [3].

В настоящее время основным источником коммерчески доступной пероксидазы являются корни хрена (*Armoracia rusticana*) (HRP —

Horse Radish Peroxidase, КФ 1.11.1.7). Однако сезонность и нестандартность растительного сырья, сложность очистки фермента ставят производителей перед необходимостью поиска других источников пероксидазы. Кроме того, препараты HRP содержат несколько изоферментов, а их получение в гомогенном состоянии достаточно трудоемкая задача [4].

В последние годы прогресс, достигнутый в генной инженерии при клонировании эукариотических генов в гетерологических системах, позволяет решать задачи более детального анализа свойств отдельных изоформ фермента, кодируемых определенным геном, а также их получения [5]. В литературе имеются сведения о конструировании векторов для экспрессии генов HRP в клетках *Escherichia coli*, *Saccharomyces cerevisiae*, *Pichia pastoris*, *Nicotiana tabacum*, *N. benthamiana* и культурах клеток личинок насекомых [2]. Первая аминокислотная последовательность изофермента HRP C1 была определена в 1976 году Велиндером. Молекулярная масса полипептидной цепи составляла ~34 кДа [2]. В 1987–1988 года были установлены три кодирующих пероксидазу гена — *prxC1*, *prxC2* и *prxC3*. К 2014 году было опубликовано в общей сложности 28 нуклеотидных последовательностей различных изоферментов пероксидаз [5]. Установлено,

что ферменты, кодируемые этими последовательностями, различались по субстратной специфичности [3].

Система экспрессии *E. coli* наиболее популярна при исследовании кристаллической структуры HRP, поскольку позволяет довольно быстро получить дегликозилированные формы фермента. Однако отсутствие гликозилирования в 2–3 раза снижает стабильность фермента. Согласно литературным данным, все кристаллические структуры пероксидаз растений, за исключением пероксидазы арахиса, были получены в *E. coli* [3]. Способы получения рекомбинантной HRP в *E. coli* изучались многими исследователями на примере различных векторных систем. Существует несколько стратегий получения рекомбинантной HRP в *E. coli*, а именно: получение в качестве тел включения и последующий рефолдинг или получение активной пероксидазы хрена в периплазме [3, 6–8]. Недостатком подобных способов получения фермента является низкий выход целевого продукта. Так, при изучении возможности экспрессии пероксидазы в *E. coli* в работах Bartonek-Roxa с соавторами [7], было показано, что рекомбинантная HRP C1 может синтезироваться в цитоплазме клеток *E. coli* в тельцах включения в нерастворимой неактивной форме. Образование телец включения является единственным возможным способом защиты от токсического действия, вызванным HRP C1. Исследователи наблюдали ингибирование роста трансформантов *E. coli*. Ингибирующая активность обусловлена синтезом 192 нуклеотидов, кодирующих первые 64 N-концевых аминокислоты в зрелой HRP.

Ortlepp с соавторами [6] сконструировали три вектора на основе pUC18, позволяющих контролировать экспрессию с помощью плазмидного *lac* репрессора, находящегося под контролем *tac* промотора. Ими установлено, что на уровень экспрессии белка влияют изменения на C-конце. Кроме того, ими был создан вектор для периплазматической экспрессии HRP C1 в клетках *E. coli*. Успешные результаты были получены путем слияния белка DsbA (disulfide bond, тиолдисульфид-оксидоредуктазы) с N-концевым пептидом HRP. Однако такой гибридный белок HRP-DsbA демонстрировал пониженную катали-

тическую активность и нестабильность, эту проблему, как предполагают авторы, можно решить путем посттрансляционного удаления DsbA. Несмотря на вполне успешные попытки секретиции белка в периплазме клеток, большое количество рекомбинантного HRP C1 все равно обнаруживалось в тельцах включения [8].

Ряд исследований по продукции рекомбинантной HRP C1 был выполнен в эукариотических системах экспрессии. Для трансформации в *P. pastoris* использовали векторы pPICZα, содержащие последовательность, кодирующую α-фактор *S. cerevisiae* — сигнал секреции метанол-индуцируемого промотора P_{AOX1}; ген устойчивости к антибиотикам зеоцину для селекции как в *E. coli*, так и в *Pichia*, последовательность, кодирующую C-концевой пептид, содержащий полигистидиновую метку для упрощения процесса очистки рекомбинантного белка [8]. Установлено, что мутация Asn175Ser приводила к увеличению термической стабильности, благодаря дополнительным водородным связям, вызывающим стабилизацию гемовой части фермента. В *P. pastoris* также была продемонстрирована возможность продукции HRP с Fab-фрагментами (fatty acid binding protein), обеспечивающими устойчивость к атразину, что придавало HRP антигенсвязывающие свойства. Система экспрессии для получения рекомбинантной HRP с Fab-фрагментами антител также была разработана на основе вектора pPICZα [10].

Кроме того, ген *hrpC1* был клонирован в дрожжевой секретирующий вектор pYEX-S. Экспрессия гена находилась под контролем конструктивного промотора гена фосфоглицераткиназы. Вектор также содержал ген ампициллин-резистентности из *E. coli* и дрожжевые селективные маркеры. Перенос гена *hrpC1* осуществляли в *S. cerevisiae*. Активность HRP C1 составила 25 ед/л [8].

В 1992 году в клеточных культурах клеток насекомых Sf9 была выделена активно синтезирующаяся HRP C1 [11, 12]. Вектор для экспрессии конструировали на основе плазмид pVL1392 и pGEMT, в которые ген *hrp* встраивали под контроль вирусного промотора полиэдрина.

При экспрессии в растениях томата и табака гена *hrpC1*, как правило, получается

активный фермент, хотя степень его гликозилирования может отличаться от нативного белка. Известно, что степень гликозилирования фермента, синтезируемого трансгенными томатами значительно ниже, чем HRP. Уменьшение степени гликозилирования снижает стабильность фермента. Кроме того, данный белок является каталитически активным и вызывает увядание растений (как полагают, за счет окисления гормона роста). При выделении рекомбинантной HRP основной проблемой является процесс очистки фермента, а именно, его отделение от полифенольных пигментов, синтезируемых в томатах [3, 13, 14]. Преимуществами использования растительных систем для получения гетерологичных белков являются: правильный фолдинг и процессинг белка, сопоставимый с эукариотической системой экспрессии; незначительные материальные затраты на выращивание растений, что обеспечивает упрощенность производственного процесса и низкую себестоимость получаемого фермента [3, 14].

В работе Matsui с соавторами [13] сообщалось о секрети HRPС1 в клетках трансгенного табака (*Nicotiana tabacum*). Во-первых, была подтверждена роль СТРР (С-terminal propeptide; С-концевой пропептид) в клетках табака. Если в трансгенных растениях табака ген *hrpC1* экспрессируется вместе с сигнальным пептидом и СТРР, то наблюдается накопление HRPС1 в вакуолях. Когда *hrpC1* экспрессируется без СТРР, наблюдается накопление активного фермента в апопластическом пространстве. Во-вторых, было показано увеличение экспрессии *hrpC1* с помощью энхансеров — 5'-нетранслирующей области (UTR) гена алкогольдегидрогеназы *N. tabacum* (NtADH).

Как видно из литературных данных, каждая из описанных в работе систем экспрессии имеет свои преимущества и недостатки. Следует отметить, что высокий уровень экспрессии изофермента С HRP наблюдался в клетках дрожжей [3].

Цель данной работы — сконструировать вектор для экспрессии гена, кодирующего изофермент С2 HRP в клетках мицелиальных грибов. Мицелиальные грибы, синтезирующие гем-содержащие ферменты, как извест-

но, образуют гем в избытке, что позволяет не проводить дополнительной активации HRP, добавляя кофактор или вещества его индуцирующие.

Материалы и методы

Ген, кодирующий изофермент С2 HRP, синтезирован искусственно и предоставлен для работы Eurofins Genomics в составе плазмиды pEX-A258-hrp. Для анализа на наличие фрагмента гена каталазы (*cat*) в геномной ДНК использовали мицелиальные грибы *Penicillium adametzii*, *Penicillium jensenii* БИМ F-134, *Penicillium sp.* 17, *Penicillium sp.* 37, *Penicillium kapuscinski*, *Penicillium funiculosum*, *Penicillium varians* 545, *Penicillium chrysogenum* БИМ F-112, *Penicillium chrysogenum* БИМ F-32. В работе для клонирования рекомбинантных плазмид использовали *E. coli* XL-1 Blue ((F⁻::Tn10(Tet^r), *proA*⁺*B*⁺, *lacI*^q, Δ(*lacZ*) *M15* / *recA1*, *endA1*, *gyrA96*(Nal^r), *thi-1*, *hsdR17* (*r*_k⁻ *m*_k⁻), *glnV44*, *relA1*, *lac*)). Информация об использованных плаزمидах представлена в табл. 1.

Для постановки полимеразной цепной реакции (ПЦР) использовались праймеры («Праймтех», Беларусь) (табл. 2, 4). Праймеры были подобраны для выделения гена *cat* из мицелиальных грибов рода *Penicillium*, экзонных последовательностей фрагментов генов *hrpC* и для получения линейаризованного гена *hrpC2*, а также плазмид pK18-cat-pAN7-1 и pAN7-1.

ПЦР проводили с помощью *Pfu*-полимеразы или *Taq HF*-полимеразы («Праймтех», Беларусь), применяя необходимые для их работы буферные растворы того же производителя. Конечный объем смеси составил 50 мкл при использовании *Pfu*-полимеразы и 20 мкл — для *Taq HF*-полимеразы (табл. 3).

Для выделения геномной ДНК из мицелия грибов и корней *A. rusticana* использовали набор «Нуклеосорб» (комплектация «С») фирмы «Праймтех» (Беларусь) [15]. Выделение проводили по методике производителя. Предварительно грибы рода *Penicillium* выращивали в колбах Эрленмейера на качалках в течение 48 часов при температуре 25–27 °С, при постоянном вращении колб со скоростью 180 об/мин в жидкой питательной среде [16].

Таблица 1

Характеристика основных плазмид, используемых в работе

Плазмида	Характеристики	Источник получения
pJET1.2	Amp ^R ; T7-промотор; <i>PlacUV₅</i> -промотор; полилинкер; репликон плазмиды pMB1; летальный ген <i>eco47IR</i> ; 2974 п.н., используется для клонирования продуктов амплификации с тупыми концами	Плазмида получена из коллекции лаборатории «ЦАГИИ» Института микробиологии НАН Беларуси
pAN7-1	Нуг ^R ; Amp ^R ; <i>lac</i> -оператор; промотор и терминатор гена глицеральдегидфосфатдегидрогеназы <i>A. nidulans</i>	Любезно предоставлена доктором Р. Punt (Wageningen Center for Food Sciences (WCFS), Wageningen, The Netherlands)
pK18	Kan ^R ; <i>lac</i> -промотор; последовательность, кодирующая ген <i>lacZ</i> ; полилинкер в составе <i>lacZ</i> гена; ColE1-репликон	Плазмида получена из коллекции лаборатории «ЦАГИИ» Института микробиологии НАН Беларуси
pEX-A258-hrp	Amp ^R ; pUC ori; <i>lac</i> -промотор; <i>lac</i> -оператор; последовательность, кодирующая изофермент С2 пероксидазы; 3504 п.о.	Плазмида предоставлена компанией Eurofins Genomics, Люксембург

Таблица 2

Характеристика праймеров, использованных в работе

ПЦР-продукт	Название праймера	Нуклеотидная последовательность праймера в направлении 5'→3'	Размер ПЦР-продукта, п. о.
Ген фермента пероксидазы из хрена <i>Armoracia rusticana</i>	<i>P1 FHRP</i>	TGAGCAGACATCACCATGCATTCTCTTCC	1078
	<i>P2 RHRP</i>	GGTCGGCATCTACTATTCGCTAGCTCTAG	
Ген фермента каталазы мицелиальных грибов рода <i>Penicillium</i>	<i>fcat2</i>	GATTTTCATCTTCCGCCAGAAGATCCA	545
	<i>rcat3</i>	CCGTGACCGGCCATTGCCCA	
Плазмида pK18-pAN7-1	<i>P3 FpNOM</i>	GGAAGAGGAATGCATGGTGATGTCTGCTCA	10035
	<i>P4 RpNOM</i>	CTAGAGCTAGCGAATAGTAGATGCCGACC	

Таблица 3

Состав реакционной смеси для проведения ПЦР

Компонент	Объем, мкл	Концентрация	Объем, мкл
	Для <i>Pfu</i> -полимеразы		Для <i>Taq HF</i> –полимеразы
ДНК-матрица	1	10 нг/мкл	1-3
Прямой праймер	1,5	20 пМ	1,5
Обратный праймер	1,5	20 пМ	1,5
Буфер (2,5X)	20	1X	2
Нуклеотиды (2мМ)	- ¹	-	2
ДНК — полимеразы <i>Pfu</i> (2 ед/мкл)	1	1 ед/мкл	3
Вода дистиллированная стерильная	25	-	7–9

Примечание. 1 — нуклеотиды входят в состав буфера

Биомассу гриба отбирали фильтрованием, удаляли культуральную жидкость (супернатант). Затем обрабатывали мицелий ацетоном, просушивали, промывали дистиллированной водой и обрабатывали ультразвуком на установке Sonics Vibra-cell VCX 130 PB (США). Для выделения ДНК бактерий применяли набор фирмы «АртБиотех», Беларусь.

Определение концентрации ДНК осуществляли с помощью флуориметра («Invitrogen», США), используя краситель ZUBR Green-1 («Праймтех», Беларусь) согласно инструкции фирмы-производителя.

Амплификацию проводили с использованием амплификатора T100 Thermal Cycler (производство «Bio-Rad», США). Условия амплификации варьировали в зависимости от используемых полимераз, праймеров и размера получаемого фрагмента ДНК.

Электрофоретический анализ плазмидной ДНК и продуктов амплификации проводили по общепринятым методам, изложенным в руководстве Т. Маниатис и др. [17]. Электрофорез проводили при комнатной температуре в течение 40–80 мин. Для визуализации результатов использовали систему гель-документации ChemiDoc™ MP Imaging System.

Размеры фрагментов ДНК устанавливали на основании их электрофоретической подвиж-

ности в агарозном геле. Для выделения ПЦР-продуктов использовали набор «Нуклеосорб» (комплектация «G») фирмы «Праймтех» (Беларусь) [18].

Выделение плазмидной ДНК из бактериальных клеток *E. coli* XL-1 Blue проводили методом щелочного лизиса по методике Маниатиса с соавторами [17], адаптированной сотрудниками лаборатории ЦАГИИ Института микробиологии НАН Беларуси.

Эндонуклеазные реакции проводили в объеме 20 мкл с использованием 2 мкл (10X) соответствующего рестрикционного буфера, 2 мкл ДНК и 2 ед рестрицирующего фермента. Для рестрикционного анализа были использованы ферменты «ThermoScientific».

Реакции лигирования проводили в объеме 20 мкл с использованием 10X лигазного буфера и T4 ДНК-лигазы. Реакционную смесь инкубировали в течение 1 ч при температуре 22 °С. После окончания инкубирования проводили очистку лигазной смеси хлороформом. Смесь хранили при температуре -20 °С. При постановке реакций были соблюдены все условия, рекомендуемые производителем в соответствии с общепринятыми экспериментальными протоколами [19].

Сиквентс-анализ проводили по методу Сенгера с помощью автоматического секвенато-

ра 4300 DNA Analyzer (Li-COR Biosciences). Матрицей для секвенирующей реакции служила плазмидная ДНК pJET1.2-cat. Для проведения секвенирования был использован набор реактивов «DNA Cycle Sequencing Kit» («Jena Bioscience», PCR-401S) и стандартные секвенирующие праймеры к вектору pJET1.2 (pJET-F и pJET-R), меченые флуоресцентной меткой Cy5.5 (Праймтех, Беларусь).

Полученные данные анализировали с помощью компьютерных программ eSeq™ DNA Sequencing Software Version 3.1.10 (Li-COR Biosciences), BLASTN2.2.1 (NCBI сайт: <http://www.ncbi.nlm.nih.gov>), SnapGene Viewer Version 3.1.2.

Изготовленные векторные конструкции трансформировали в клетки *E. coli* XL-1 Blue и протопласты *P. adametzii* методом электропорации с помощью прибора MicroPulser BIO-RAD (США). Трансформанты отбирались по устойчивости к селективному агенту и по результатам ПЦР-анализа.

Результаты и их обсуждение

Проведен анализ 5 последовательностей генов, кодирующих изофермент С HRP, представленных в базе данных GenBank. Установлено, что все гены *hrpC* имеют единую

структуру: 4 экзона, 3 интрона. Определена гомология экзонных областей генов *hrp*, для последующего выделения участков гена *hrpC*. Показано, что сходство экзонных последовательностей 3' и 5' концов составляет 95–100%. Установлено, что анализируемые последовательности экзонов 1, 2 и 3 различных изоформ обладали наиболее высокой степенью идентичности (99–100%), существенные отличия наблюдались между 5' концевыми последовательностями экзона 4 изоформ C1 и C2, C3. На основании полученных результатов были подобраны праймеры (табл. 4).

Для амплификации использовали геномную ДНК *A. rusticana*. Нами были получены 4 экзонных области данного гена с использованием подобранных праймеров (табл. 4). Результаты проведения полимеразной цепной реакции представлены на рис. 1. Размеры полученных фрагментов соответствовали ожидаемым результатам, а именно, для экзона 1 ~ 230 п. о., для экзона 2 ~ 192 п. о., для экзона 3 ~ 169 п. о., для экзона 4 ~ 470 п. о. Наличие дополнительных фрагментов отмечено только при получении продукта экзона 3, что свидетельствует о специфичности подобранных праймеров, однако не позволяет выделить гены, кодирующие отдельные изоформы HRP.

Таблица 4

Праймеры подобранные на основании анализа нуклеотидных последовательностей экзонов

Изоформа пероксидазы	Праймеры	Последовательность	Температура плавления, °C
Экзон 1			
C1, C2, C3	<i>Fex1/3 HRP</i> прямой	ATGCATTCCCCTTCTTCTACTTC	53
C1, C2, C3	<i>Rex1/3 HRP</i> обратный	ATTAACAAAGCAGTCGTGGAAGTG	54
Экзон 2			
C1, C2, C3	<i>Fex2/3 HRP</i> прямой	GGTTGTGACGCATCGATCTTG	54
C1, C2, C3	<i>Rex2/3 HRP</i> обратный	CAAATTGACAGATTGTTGAGCTGCAA	55
Экзон 3			
C1, C2, C3	<i>Fex3/3 HRP</i> прямой	GCGGGAGGTCCTTCTTGGA	55
C1, C2, C3	<i>Rex3/3 HRP</i> обратный	CGGAGAGAGCAACGAGATCAG	56

Продолжение табл. 4

Изоформа пероксидазы	Праймеры	Последовательность	Температура, °C
Экзон 4			
C1	прямой	GTGGTCACACATTTGGAAAAAATCA	53
C2, C3	<i>Rex4/3 HRP</i> прямой	GGGGCCACACATTTGGTAAAAAT	53
C2, C3	<i>Rex4/3 HRP</i> обратный	TCACATAGAGCTAACAAAGTCAAC	52

Для дальнейшей работы ген, кодирующий изоформу C2 HRP, был синтезирован химически Eurofins Genomics и получен в составе плазмиды pEX-A258-hrp (рис. 2).

Как сказано выше, наиболее изучена изоформа C1 HRP, клонирование кодирующего ее гена было осуществлено в различные штаммы про- и эукариот [3]. Нами для работы был выбран ген *hrpC2*, поскольку в литературе мало сведений о клонировании данного гена. При этом HRP C2 в отличие от HRP C1 характеризуется более высокой скоростью окисления различных фенольных соединений [2, 3], а сходство экзонных участков генов *hrpC2* и *hrpC1* составляет 86–91%.

Хотя грибы являются одним из главных объектов, используемых для получения биологически активных метаболитов, векторов для экспрессии генов в клетках мицелиальных грибов сконструировано небольшое количество. Нами была использована конструкция для экспрессии генов в клетках грибов рода *Aspergillus* pAN7-1, содержащая сильный конститутивный промотор гена глицеральдегидфосфатдегидрогеназы *Aspergillus nidulans* (промотор *gpd*) и любезно предоставленная доктором Р. Пунт (Wageningen Center for Food Sciences (WCFS), Wageningen, The Netherlands) [20]. Ряд исследователей отмечали, что, поскольку синтезируемые дрожжами и растениями пероксидазы являются каталитически активными, то они могут влиять на рост реципиентов, вызывая снижение количества биомассы дрожжей и увядание растений [9, 21].

Для встраивания гена *hrpC2* в ДНК мицелиальных грибов было решено создать конструкцию, несущую фрагмент гена *cat*.

Предполагалось, что полученный вектор обеспечит не только встраивание *hrpC2* в хромосому, но и позволит инактивировать ген, кодирующий гем-содержащую каталазу. Каталаза и пероксидаза являются ферментами антиоксидантной системы и выполняют одну функцию — способствуют распаду пероксида водорода, образованного при окислении различных соединений.

Для выделения фрагмента гена *cat* *P. adametzii* использовались праймеры, созданные на основе анализа нуклеотидных последовательностей *cat* грибов рода *Penicillium*. В результате про-

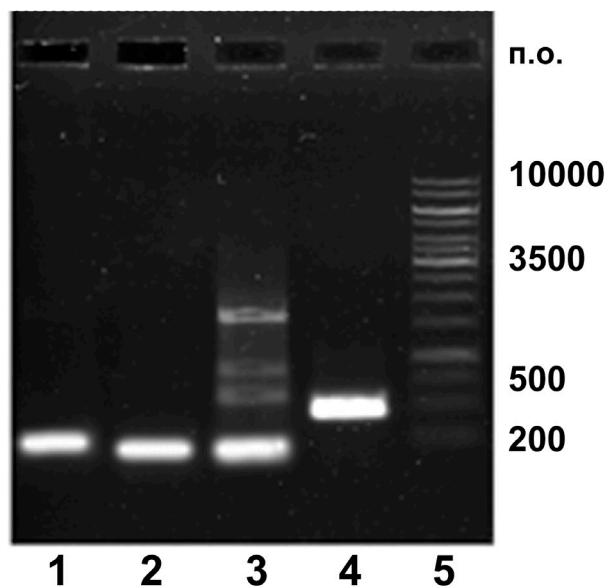
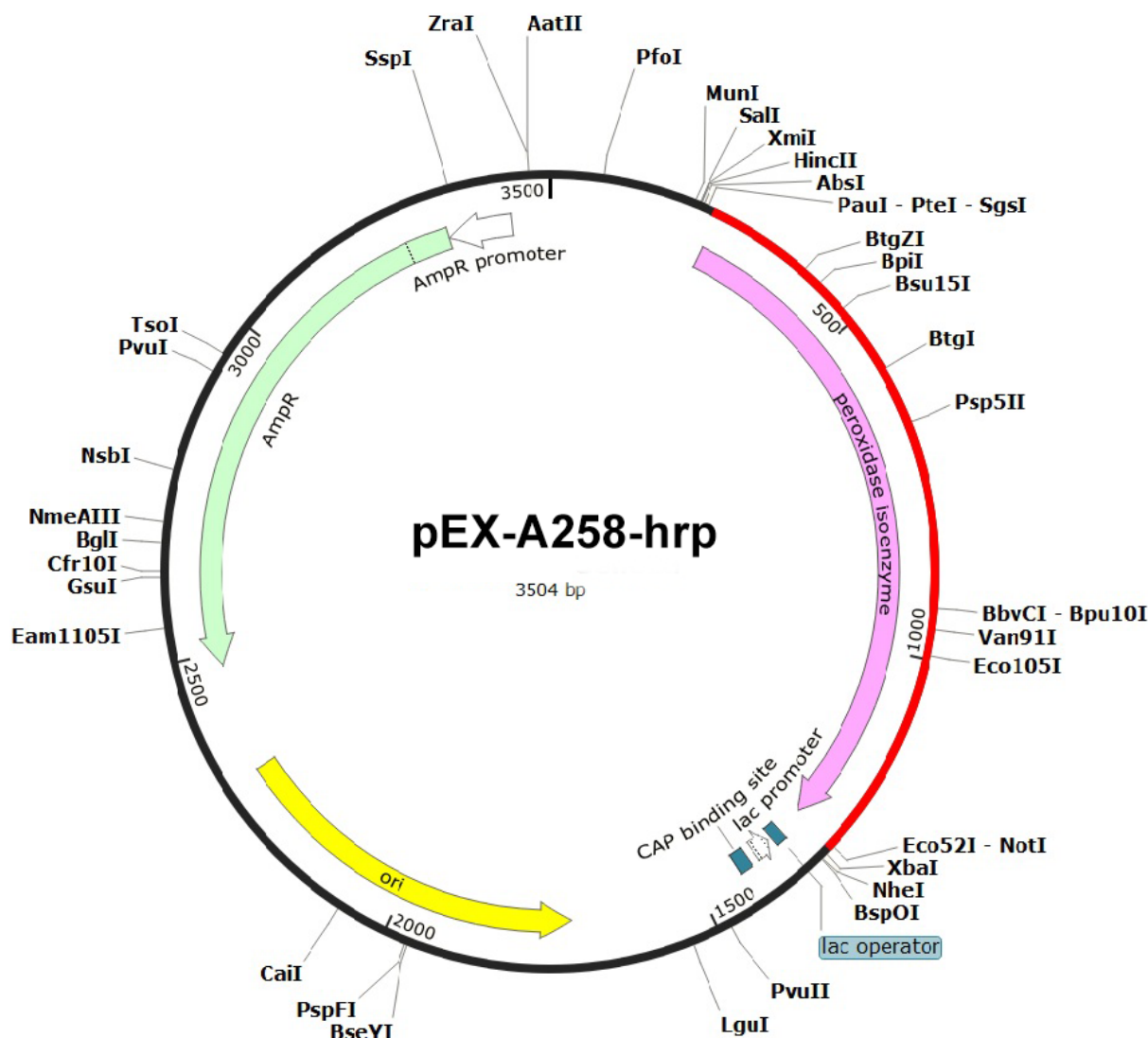


Рис. 1. Электрофореграмма фрагментов ДНК, амплифицированных на матрице тотальной ДНК, выделенной из *A. rusticana*

Номера дорожек соответствуют продуктам амплификации экзонов 1–4 гена *hrpC* согласно порядковому номеру дорожки, 5 — маркер молекулярного веса (GeneRuler™ 1 kb DNA Ladder, «ThermoScientific»)

Рис. 2. Карта плазмиды pEX-A258-hrp, несущей ген *hrpC2*

веденных экспериментов был выделен фрагмент, размер которого составил около 600 п. о. (рис. 3, дорожка 2), что полностью соответствует ожидаемому результату.

ПЦР-продукт был встроен в плазмидный вектор pJET1.2 для последующего определения нуклеотидной последовательности фрагмента *cat* *P. adametzii*. На основании сиквенс-анализа установлено, что фрагмент *cat* *P. adametzii*, имеющий размер 568 п. о., на 81–97% идентичен участкам *cat* грибов *P. chrysogenum* Wisconsin 54-1255, *P. digitatum* Pd1, *P. citrinum*, *P. marneffeii* ATCC 18224 (депонированных в GenBank NCBI (AM920431.1, XM_014675429.1, DQ288844.1, XM_002153565.1, соответственно).

В дальнейшем фрагмент с целевой последовательностью ДНК *cat* *P. adametzii* (размер вырезанного по *Vam*NI участка из вектора pJET1.2-*cat* составлял 615 п. о.) клонировали в плазмиду pK18 (предварительно также обработанную рестрицирующим ферментом *Vam*NI) (рис. 4, дорожка 2). Размер полученной плазмиды pK18-*cat* составил 2661 п. о.

Для выбора штаммов потенциальных реципиентов был проведен ПЦР-скрининг на наличие фрагмента гена *cat* для последующей его инактивации путем встройки экспрессируемого гена пероксидазы. Для этого использованы 9 штаммов мицелиальных грибов *Penicillium*, синтезирующих гем-содержащие фермен-

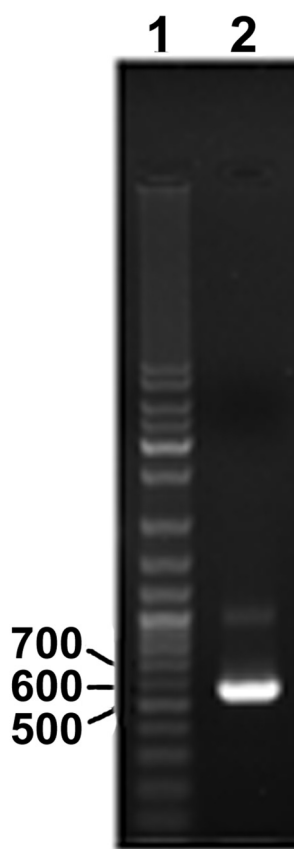


Рис. 3. Электрофореграмма амплифицированного фрагмента ДНК гена *cat* *P. adametzii*

1 — маркер молекулярного веса Long Range DNA Ladder (100bp-10kb) производства «Jena bioscience»; 2 — продукт амплификации *cat* *P. adametzii*

ты (*P. adametzii*, *P. jensenii* БИМ F-134, *Penicillium* sp. 17, *Penicillium* sp. 37, *P. kapuscinski*, *P. funiculosum*, *P. varians* 545, *P. chrysogenum* БИМ F-112, *P. chrysogenum* БИМ F-32). Согласно полученным ранее сотрудниками лаборатории ферментов данным эти штаммы продуцировали марганец-пероксидазы и каталазы [21]. Установлено, что из 9 штаммов грибов рода *Penicillium*, характеризующихся повышенным уровнем гем-содержащих ферментов, 4 (*P. adametzii*, *P. kapuscinski*, *P. chrysogenum* БИМ F-112, *P. chrysogenum* БИМ F-32) имеют аналогичный фрагмент и в дальнейшем могут использоваться в качестве реципиентных штаммов для трансформации векторной конструкции. В дальнейшем плазмиды pK18-cat и pAN7-1 были лианеризованы по HindIII и лигированы друг с другом. Таким образом, была получена плазида pK18-cat-AN7-1 10031 п. о.

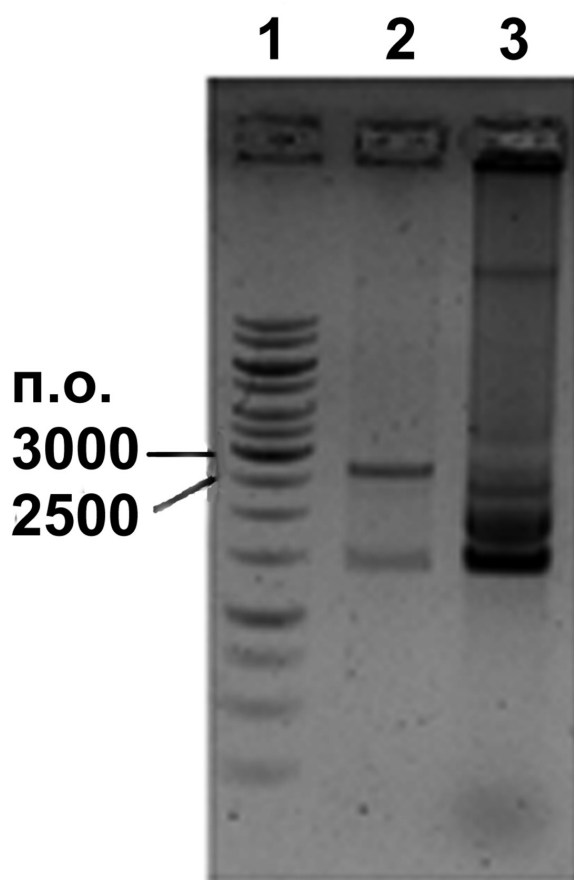


Рис. 4. Электрофореграмма плазмиды pK18

1 — маркер молекулярного веса GeneRuler™ 1 kb DNA Ladder («ThermoScientific»); 2 — плазида pK18, разрезанная по сайту рестрикции BamHI; 3 — исходная плазида pK18

Ген *hrpC2* предполагалось встраивать вместо гена устойчивости к гигромицину.

Наличие в гене пероксидазы хрена и в экспрессионных плаزمиде сайтов, подверженных воздействию наиболее часто встречающихся рестрицирующих ферментов не позволило использовать данный метод для конструирования экспрессионной плазмиды. Для этого применялся метод продолжительной перекрывающейся полимеразной цепной реакции (пп-ПЦР) [22]. На первом этапе ген *hrpC2* был выделен из pEX-A258-hrp методом ПЦР с использованием *Pfu*-полимеразы и олигонуклеотидных праймеров *P1 FHRP* и *P2 RHRP*, имеющих на концах последовательности, комплементарные плазмиде pK18-cat-AN7-1. В результате был получен фрагмент, несущий *hrpC2*, размером ~1100 п. о. (рис. 5А).

На втором этапе лианеризовали вектор pK18-cat-AN7-1 методом ПЦР с использова-

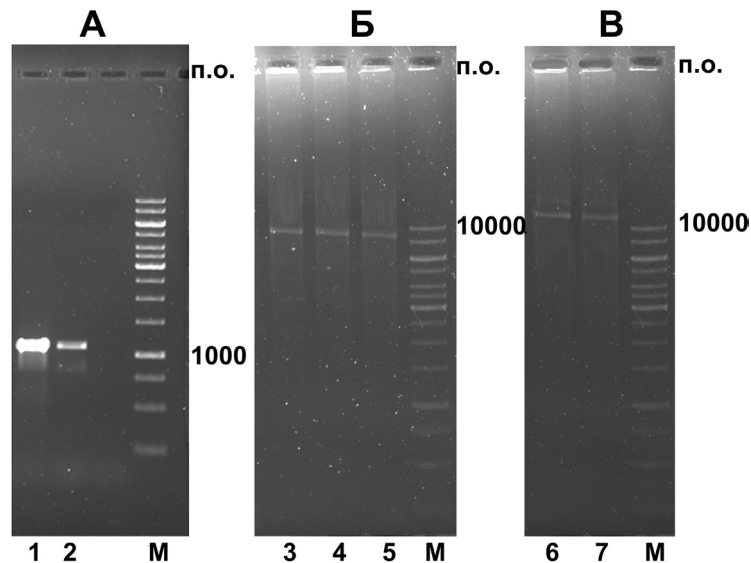


Рис. 5. Электрофореграммы продуктов амплификации при проведении клонирования гена *hrpC2* методом пп-ПЦР

А — первый этап клонирования гена *hrpC2* (температуры отжига 52 (1), 58 °C (2)); Б — амплификат линейаризованного вектора рK18-cat-AN7-1 (температура отжига: 45 (3), 50 (4), 55 °C (5)); В — амплификат вектора рK18-cat-AN7-1 с геном *hrpC2* (температура отжига: 50 (6), 55 °C (7)), М — маркер молекулярного веса (GeneRuler™ 1 kb DNA Ladder, «ThermoScientific»)

нием *Pfu*-полимеразы и олигонуклеотидных праймеров *P3 FpNOM* и *P4 RpNOM*, содержащих фрагменты, комплементарные последовательности гена *hrpC2*. В результате был получен фрагмент, несущий вектор без гена устойчивости к гигромицину, размером ~9000 п. о. (рис. 5Б). Каждый этап контролировали при помощи агарозного геля-электрофореза. Продукты амплификации очищали из реакционной смеси переосаждением ацетатом аммония и изопропиловым спиртом. На третьем этапе собирали линейаризованный вектор и целевой ген методом пп-ПЦР по следующей программе: цикл этапов предденатурации 30 с при 98 °C; элонгации 3 мин при 72 °C и последующая стандартная процедура амплификации. В результате была получена плазмида, обозначенная как рK18-cat-AN7-1-HRP размером более 10000 п. о. (рис. 5В).

В дальнейшем была проведена котрансформация рK18-cat-AN7-1-HRP и рAN7-1 в клетки *P. adametzii*. Трансформанты отбирали по устойчивости к гигромицину, ген устойчивости к данному антибиотику находился на плазмиде рAN7-1. Было получено 16 трансформантов, у 2 из которых вставка гена *hrpC2* подтверждена с использованием ПЦР-анализа (рис. 6).

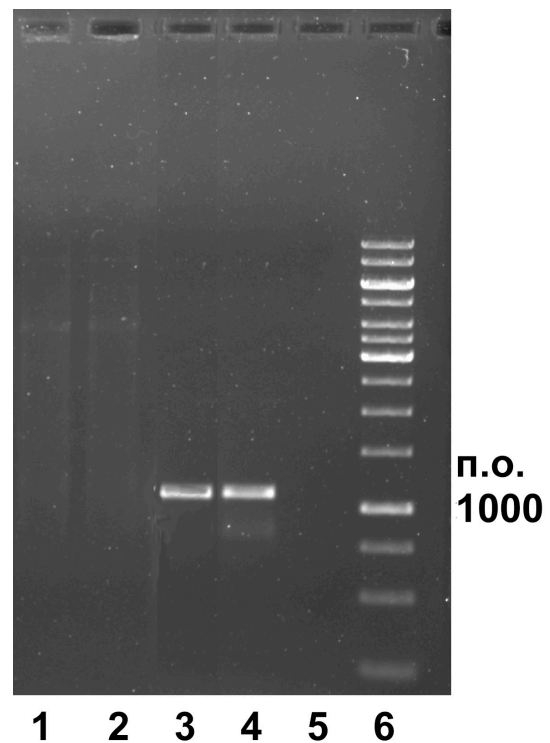


Рис. 6. Электрофореграмма ПЦР-продуктов ДНК трансформантов *P. adametzii*

1–5 — трансформанты *P. adametzii*, 6 — маркер молекулярного веса (GeneRuler™ 1 kb DNA Ladder, «ThermoScientific»)

Заключение

Таким образом, проведен анализ нуклеотидных последовательностей генов, кодирующих изофермент С HRP, представленных в базе данных GenBank. Установлено, все гены *hrpC* имеют единую структуру: 4 экзона, 3 интрона. Показано, что сходство экзонных последовательностей 3' и 5' концов составляет 95–100 %. Проведен анализ геномной ДНК грибов рода *Penicillium*. Отобраны 4 штамма мицелиальных грибов (*P. adametzii*, *P. kapuscinski*, *P. chrysogenum* БИМ F-112, *P. chrysogenum* БИМ F-32), как потенциальные реципиенты для векторных конструкций с генами пероксидазы и фрагментом гена *cat*, необходимым для встраивания конструкции в геномную ДНК. На основе синтезированной последовательности *hrpC2* и плазмид: pEX-A258-*hrp*, несущей ген *hrpC2*, pK18-*cat*, несущей фрагмент гена *cat* *P. adametzii* для обеспечения встройки плазмиды в геномную ДНК, и pAN7-1, имевшей в своем составе промотор и терминатор гена глицеральдегидфосфатдегидрогеназы *A. nidulans* для экспрессии *hrpC2*, сконструирован вектор для экспрессии гена пероксидазы хрена изофермента С2 в клетках мицелиальных грибов. Проведена его трансформация в *P. adametzii*. Наличие вставки гена подтверждено ПЦР-анализом.

Список использованных источников

1. Варфоломеев, С.Д. Химическая энзимология / С.Д. Варфоломеев. – Высшее профессиональное образование: учебник. – М., 2005. – С. 214–456.
2. Horseradish Peroxidase / N. C. Veitch [et al.] // *Advances in Inorganic Chemistry*. – 2000. Vol. 51. – P. 1–56.
3. An updated view on horseradish peroxidases: recombinant production and biotechnological applications / F.W. Krainer [et al.] // *Appl. Microbiol. Biotechnol.* – 2015. – Vol. 99, № 4. – P. 1611–1625.
4. Вяткина, О.В. Проблемы выделения и очистки растительных пероксидаз / О.В. Вяткина // Уч. записки Таврич. нац. ун-та. Сер. Биология. Химия. – 2012. – №3. – С. 271–273.
5. Peroxidase gene discovery from the horseradish transcriptome / L. Näätäsaari [et al.] // *BMC Genomics*. – 2014. – Vol. 15. – P. 1–16.
6. Expression and characterisation of a protein specified by a synthetic horseradish peroxidase gene in *Escherichia coli* / S. A. Ortlepp [et al.] // *Journal of Biotechnology*. – 1989. – Vol. 11. – P. 353–364.
7. Expression of a neutral horseradish peroxidase in *Escherichia coli* / E. Bartonek-Roxa [et al.] // *Journal of Biotechnology*. – 1994. – Vol. 37. – P. 133–142.
8. Gundinger, T. A comparative approach to recombinantly produce the plant enzyme horseradish peroxidase in *Escherichia coli* / T. Gundinger, O. Spadiut // *Journal of Biotechnology*. – 2017. – Vol. 248. – P. 15–24.
9. Functional expression of horseradish peroxidase in *Saccharomyces cerevisiae* and *Pichia pastoris* / B. Morawski [et al.] // *Protein Eng.* – 2000. – Vol. 13, № 5. – P. 377–384.
10. Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure / R.D. Gietz [et al.] // *Yeast*. – 1995. – Vol. 11, № 4. – P. 355–360.
11. Baculovirus Expression and Characterization of Catalytically Active Horseradish Peroxidase / C. Hartmann [et al.] // *Archives of Biochemistry and Biophysics*. – 1992. – Vol. 297. – P. 61–72.
12. Genetically engineered horseradish peroxidase for facilitated purification from baculovirus cultures by cation-exchange chromatography / G. Levin [et al.] // *J. Biotechnol.* – 2005. – Vol. 118, № 4. – P. 363–369.
13. High-efficiency secretory production of peroxidase C1a using vesicular transport engineering in transgenic tobacco / T. Matsui [et al.] // *J. Biosci. Bioeng.* – 2006. – Vol. 102, № 2. – P. 102–109.
14. Transient Expression and Purification of Horseradish Peroxidase C in *Nicotiana benthamiana* / Suzanne M. Huddy [et al.] // *International Journal of Molecular Sciences*. – 2018. – Vol. 19, № 115. – P. 1–13.
15. Инструкция на набор по выделению ДНК «Нуклеосорб». – Минск. – С. 9–10.
16. Михайлова, Р.В., Жуковская Л.А., Лобанок А.Г. Изучение спонтанной изменчивости *Penicillium adametzii* ЛФ F-2044 – продуцента глюкозооксидазы // *Прикладная биохимия и микробиология*. – 2007. – Т. 43, № 2. – С. 209–211.
17. Инструкция на набор по выделению ДНК «Нуклеосорб». – Минск. – С. 11.
18. Маниатис, Т. Методы генетической инженерии: молекулярное клонирование /

- Т. Маниатис, Э. Фрич, Дж. Сэмбрук. – М. : Мир, 1984. – 480 с.
19. Великов, В.А. Молекулярная биология. Практическое руководство: учеб. пособие / В.А. Великов. – Саратов: Саратовский источник, 2013. – 84 с.
20. Verdoes, J.C. Molecular genetic strain improvement for the overproduction of fungal proteins by filamentous fungi / J.C. Verdoes, P.J. Punt, C.A. van den Hondel // Appl. Microbiol. and Biotechnol. – 1995. – Vol. 43. – P. 195–205.
21. Поиск штаммов мицелиальных грибов – продуцентов гем-содержащих ферментов / Т.В. Семашко [и др.] // Микробные биотехнологии: фундаментальные и прикладные аспекты : Сборник научных трудов. – Минск: Беларуская навука, 2017. – Т. 9. – С. 115–130.
22. Рымко, А.Н. Создание штамма *Escherichia coli* – продуцента гомо-логичной нуклеозиддифосфаткиназы / А.Н. Рымко, С.В. Квач, А.И. Зинченко // Молодежь в науке – 2013: приложение к журналу «Весці НАН Беларусі», сер. біял. навук. – 2013. – № 5. – С. 135–138.

T.V. Semashko, A.I. Gorina, E.A. Martynava, L.N. Valentovich, L.A. Zhukovskaya, A.G. Lobanok

OBTAINING A VECTOR CONSTRUCTION FOR EXPRESSION OF THE PEROXIDASE ISOENZYME C2 GENE FROM HORSERADISH ROOTS IN FILAMENTOUS FUNGAL CELLS

Institute of Microbiology, NASB
Minsk, 220141, the Republic of Belarus

The nucleotide sequences of isoenzymes C horseradish peroxidase genes of group presented in the GenBank database were analyzed. The vector for the expression of the horseradish peroxidase gene of the C2 isoenzyme was constructed for mycelial fungi cells using synthesized gene sequence and some plasmids: pK18-cat, carrying the fragment of the *Penicillium adametzii* catalase gene for plasmid insertion into genomic DNA, and pAN7-1, which contains the promoter and terminator of the *Aspergillus nidulans* glyceraldehyde phosphate dehydrogenase gene for *hrpC2* expression.

Key words: horseradish peroxidase gene, isozyme C2, expression vectors, *Penicillium* fungi.

Дата поступления статьи: 10 сентября 2018 г.

В.В. Гринев

СОХРАНЕНИЕ ИНТРОНОВ В ТРАНСКРИПТОМЕ ЛЕЙКОЗНЫХ И НОРМАЛЬНЫХ КЛЕТОК КРОВИ ЧЕЛОВЕКА

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

Сохранение интронов в молекулах зрелых РНК является плохо изученной стороной активности системы сплайсинга в клетках человека. Мы разработали многоэтапный алгоритм поиска *in silico* интронных последовательностей, сохраняющихся в молекулах зрелых РНК, секвенированных по методу *RNA-Seq*. Результаты применения данного алгоритма хорошо согласуются с данными по реконструкции полноразмерных молекул РНК, проведенной для разных типов клеток человека, и говорят о функциональной важности генов, РНК продукты которых содержат интроны.

Ключевые слова: сохранение интронов, сплайсинг РНК, *RNA-Seq*, транскриптом.

Введение

Сплайсинг РНК является многоступенчатым скоординированным молекулярным процессом, ведущим к образованию молекул зрелых РНК [1]. Во время этого процесса происходит определение границ интронов (или экзонов согласно альтернативной модели) и их удаление с последующим объединением экзонов в единую молекулу. Однако порой удаления некоторых интронов не происходит, и они могут быть обнаружены в составе полиаденилированных молекул РНК. На сегодняшний момент открыто три класса таких интронов: 1) экзитроны (exitrons), которые входят в состав кодирующих белок экзонов [2]; 2) удержанные интроны (retained introns), присутствующие в цитоплазматических РНК [3]; 3) задержанные интроны (detained introns), удаление которых идет с существенным запаздыванием по сравнению с другими интронами, из-за чего молекула РНК не может покинуть ядро [4, 5].

Во всех этих случаях существование самого по себе феномена сохранения интронов не вызывает сомнений, однако остаются открытыми вопросы о причинах их сохранения, тканевой, клеточной и генной специфичности, а также функциональной роли при разных состояниях клеток [3]. Поскольку ключевым моментом в изучении таких интронов является их обнаружение, особенно по данным полнотранскриптомного секвенирования, то мы задались

целью разработать алгоритм поиска *in silico* интронных последовательностей, сохраняющихся в молекулах зрелых РНК, сосредоточившись при этом только на транскриптоме лейкозных и нормальных клеток крови человека.

Материалы и методы

Представленное исследование основано на двух группах данных полнотранскриптомного секвенирования полиаденилированной клеточной РНК, проведенного по методу парноконцевого *RNA-Seq*: оригинальных, полученных на базе Детского онкологического центра Вульфсона при Ньюкаслском университете (г. Ньюкасл, Великобритания), и публичных, находящихся в свободном доступе в базе *ENA*. К первой группе относятся данные, полученные при секвенировании РНК, выделенной из двух образцов клеток линии *Kasumi-1*, и такого же количества образцов клеток линии *SEM*. Клеточная линия *Kasumi-1* (*ATCC*[®] *CRL-2724*TM) является модельной клеточной линией положительной по транслокации *t*(8;21) (*q22;q22*) формы острого миелоидного лейкоза человека, а линия *SEM* (*DSMZ* № *ACC 546*) — положительной по транслокации *t*(4;11) (*q21;q23*) формы острого лимфобластного лейкоза. Вторая группа была представлена данными секвенирования РНК, выделенной из лейкозных клеток пациентов, несущих транслокацию *t*(8;21)(*q22;q22*) (образцы с кодами

доступа GSM1521606 и GSM1521614) или инверсию *inv*(16) (образцы с кодами доступа GSM1521580 и GSM1636713). Кроме того, в эту же группу входили данные по двум образцам *CD34⁺* гемопоэтических стволовых/прогениторных клеток, полученных из пуповинной крови (образцы с кодами доступа GSM1843222 и GSM1843223), и двум образцам (с кодами доступа GSM1552801 и GSM1552802) таких же клеток, но выделенных из нормального костного мозга человека.

FASTQ файлы, полученные при секвенировании транскриптома целевых клеток, подвергались стандартной процедуре предпроцессинга, после чего риды (короткие секвенированные фрагменты РНК) картировались по эталонной сборке *GRCh38.p7* генома человека с помощью программы *subjunc* [6]. Суммаризация интронных ридов осуществлялась с помощью алгоритма *featureCounts* [7]. В дальнейшем количественные данные подвергались многоступенчатой фильтрации, нормализации и трансформации, более подробно описанной в контексте изложения полученных результатов.

Сборка полноразмерных РНК осуществлялась с помощью алгоритма *Cufflinks* v.2.2.1 [8]. Сборка проводилась с опорой на эталонные модели генов человека, представленные в 85-ом выпуске базы данных *Ensembl* [9]. Для корректной оценки уровня экспрессии собранных транскриптов на алгоритм подавался сиквенс эталонной сборки *GRCh38.p7* генома человека, а также игнорировались риды, картированные по генам рибосомальных РНК и генам митохондриального генома. Значения остальных параметров сборки использовались по умолчанию. Собранные транскрипты были подвергнуты многоуровневой фильтрации с помощью оригинального *R* кода *filtrTrans* (<https://github.com/VGrinev/transcriptome-analysis/tree/master/TranscriptsFeatures/filtrTrans.r>).

Функциональная аннотация генов проводилась с помощью оригинального *R* кода *enrichmentTest* (<https://github.com/VGrinev/transcriptome-analysis/tree/master/TranscriptsFeatures/enrichmentTest.r>) с использованием двустороннего точного теста Фишера и наборов функциональных групп генов, загруженных в виде *OBO* файлов с репозитория *Gene Ontology Consortium* [10]. Карта обогаще-

ний построена с помощью плагина *Enrichment Map* v.2.2.1 для пакета *Cytoscape* v.3.6.0 [11].

Статистические тесты и другие рутинные вычислительные процедуры проводились в соответствии с общепринятыми стандартами. Вся аналитическая работа, результаты которой представлены в настоящем исследовании, проведена в среде программирования *R* с использованием как уже готовых решений, доступных через репозиторий *Bioconductor* или хранилище *CRAN*, так и с помощью оригинальных программ, написанных автором исследования.

Результаты и обсуждение

Обнаружение интронов, сохранившихся в молекулах зрелых РНК, не является тривиальной задачей. Для ее решения мы разработали многоэтапный алгоритм поиска *in silico* таких последовательностей в структуре молекул зрелых РНК, секвенированных по технологии *RNA-Seq*.

На первом этапе был получен список геномных координат всех не сохраняющихся в молекулах зрелых РНК интронов эталонного генома *GRCh38.p7* человека, аннотированного в 85-ом выпуске базы данных *Ensembl* [9]. Из этих координат были исключены интервалы, соответствующие одноэкзонным генам и генам малых РНК (*miRNA*, *miscRNA*, *rRNA*, *tRNA*, *ribozyme*, *vaultRNA*, *sRNA*, *snRNA*, *scaRNA*, *scRNA* и *snoRNA*), а оставшиеся интервалы были разбиты на интронные бины длиной не менее 50 нуклеотидов и сгруппированы по их генной принадлежности. В данном случае под интронным бином подразумевается уникальный фрагмент интрона (либо весь интрон), не перекрывающийся с другими такими же уникальными интронными фрагментами. К полученному списку координат интронных бинов были добавлены геномные координаты, которые в базе данных *Ensembl* уже аннотированы как сохраняющиеся интроны. Эти координаты были процессированы так же, как и координаты не сохраняющихся интронов и, кроме того, разделены на четыре группы в зависимости от локализации в гене (ассоциированы ли они с первым экзоном, внутренними экзонами или последним экзоном гена) или принадлежности к одноэкзонному(-ым) транскрипту(-там)

гена. Наш окончательный список содержал координаты 314 168 бинов не сохраняющихся в молекулах зрелых РНК интронов, а также 5089, 3858, 6655 и 836 бинов сохраняющихся интронов, ассоциированных, соответственно, с первыми, внутренними и последними экзонами генов или одноэкзонными транскриптами. Распределение длин этих бинов показано на рис. 1А. Обращает на себя внимание, что бины интронов, аннотированных в базе данных *Ensembl* как сохраняющиеся в молекулах зрелых РНК, более короткие, чем бины несохраняющихся интронов: основной пик распределения размеров таких бинов совпадает с пиком минимальных интронов, длина которых обычно лежит в диапазоне от 50 до 150 нуклеотидов [12].

На втором этапе с помощью алгоритма *subjunc* [6] проводилось картирование *RNA-Seq* ридов на эталонный геном *GRCh38.p7* человека с последующей суммаризацией ридов, перекрывающихся с координатами интронных бинов. Суммаризация ридов осуществлялась с помощью алгоритма *featureCounts* [7], причем учитывались только уникально картированные риды и только те, которые перекрывались с интронным бином не менее чем на 5 нуклеотидов. При использовании такого подхода в работе с *RNA-Seq* данными по транскриптому клеток линии *Kasumi-1* нами обнаружено, что $44,9 \pm 0,5\%$ бинов, произошедших из несохраняющихся интронов, получает не менее одного рида. Доля таких же бинов, но принадлежащих сохраняющимся интронам, получилась несколько больше: $58,5 \pm 0,8\%$ в случае с бинами, ассоциированными с одноэкзонными транскриптами, до $69,9 \pm 0,6\%$ бинов, связанных с последними экзонами генов. Наблюдаемые процентные различия согласуются с различиями и по уровню представленности в транскриптом клеток линии *Kasumi-1*: распределение уровней экспрессии бинов несохраняющихся интронов смещено влево по сравнению с распределениями для других групп бинов, что проиллюстрировано на рис. 1Б. Схожие результаты наблюдались и при работе с другими наборами *RNA-Seq* данных.

Третий этап нашего анализа был нацелен на расчет эффективной длины интронных бинов. Под эффективной длиной в данном

случае подразумевается та часть последовательности, которая является уникальной в геноме. Необходимость такой корректировки обусловлена присутствием в интронных областях наших генов множества повторов [13]. В результате риды, которые картируются по таким областям генома, не учитываются при суммаризации, что искажает реальный уровень экспрессии интрона. Для корректировки длины интронных бинов мы воспользовались данными по картируемости генома человека [14] и рассчитали эффективную длину каждого бина с помощью оригинального *R* кода. Как видно на рис. 1В, при использовании такого подхода существенной корректировке подвергается длина только части интронных бинов, большая же часть корректируется минимально, что и обуславливает высокий уровень корреляции между исходной и скорректированной длиной.

На четвертом этапе рассчитывались нулевые распределения ридов по интронам генов. Под нулевым, или теоретически ожидаемым, распределением подразумевается гипотетическая ситуация, при которой все интронные риды гена случайным образом и с равной вероятностью распределяются по его интронам. Нулевые распределения строились нами на основе экспериментальных данных, или эмпирических распределений, описанных выше, по следующей формуле:

$$NDR_g^i = \frac{R_s \sqrt{Intr_{EL}^i - R_L}}{\sum_{i=1}^N \sqrt{Intr_{EL}^i - R_L}},$$

где NDR_g^i — это математическое ожидание для i -го интронного бина гена g (другими словами, это то количество интронных ридов гена g , которое пришлось бы на i -тый интронный бин этого же гена если бы риды распределялись по интронам случайным образом и с равной вероятностью);

R_s — общее количество уникально картированных интронных ридов гена g ;

$Intr_{EL}^i$ — эффективная длина i -го интронного бина гена g ;

R_L — длина ридов в используемой *RNA-Seq* библиотеке;

N — количество интронных бинов, принадлежащих гену g .

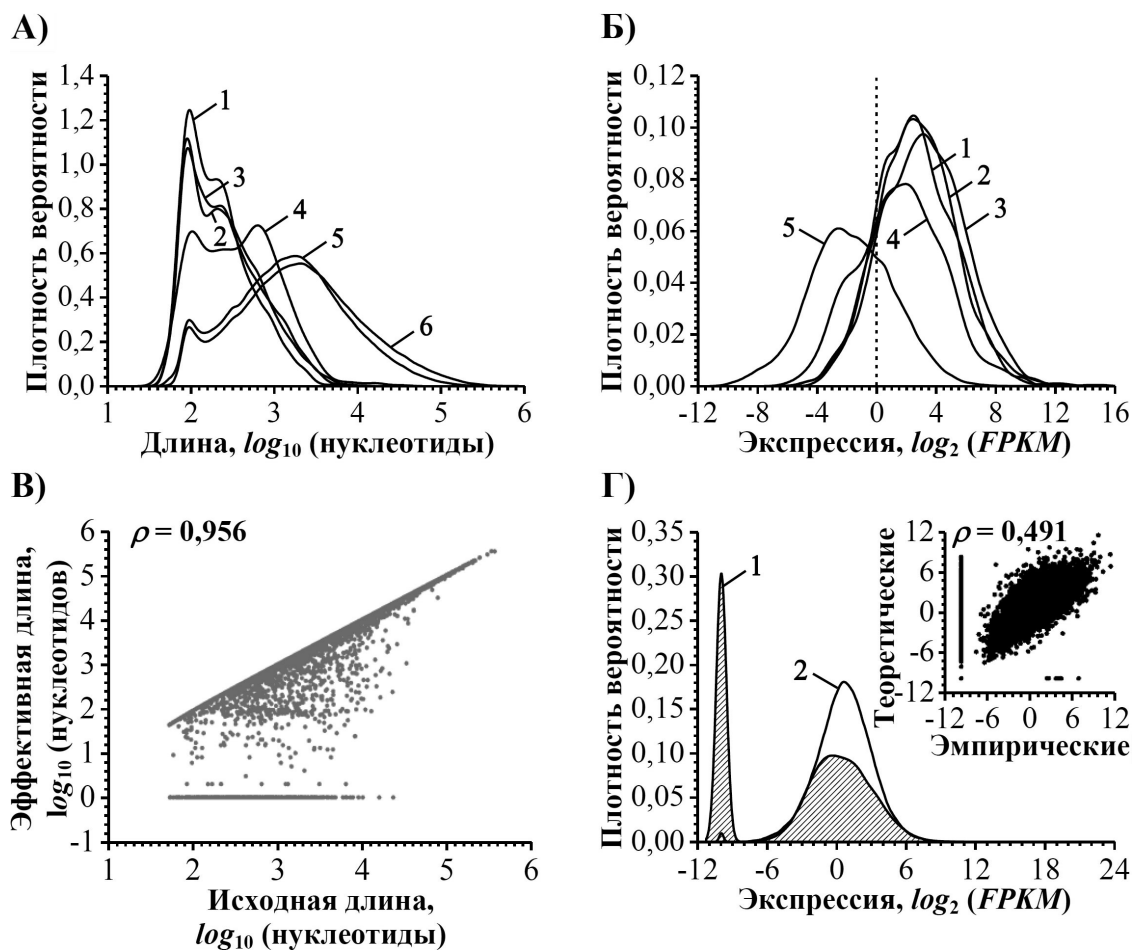


Рис. 1. Описательная статистика интронных бинов, вовлеченных в исследование

Экспериментальные данные, представленные на этом рисунке, были получены на клетках линии *Kasumi-1*, однако схожие результаты наблюдались и при работе с другими типами клеток.

А) Распределение длин бинов сохраняющихся интронов, ассоциированных с первыми, внутренними и последними экзонами генов или одноэкзонными транскриптами генов (кривые 1, 2, 3 и 4, соответственно) и бинов несохраняющихся интронов (кривая 5). Кривая 6 демонстрирует распределение длин всех исходных интронов эталонного генома *GRCh38.p7* человека, которые использовались при расчете геномных координат интронных бинов.

Б) Экспрессия интронных бинов в клетках линии *Kasumi-1*. На этом графике показаны распределения бинов сохраняющихся интронов, ассоциированных с первыми, внутренними и последними экзонами генов или одноэкзонными транскриптами генов (кривые 1, 2, 3 и 4, соответственно) и бинов несохраняющихся интронов (кривая 5), экспрессия которых подтверждается хотя бы одним ридом. Экспрессия выражена в *FPKM* (от англ. fragments per kilobase per million) — количестве секвенированных фрагментов, приходящихся на 1 кб длины интронных бинов и на 1 млн фрагментов в *RNA-Seq* библиотеке. Штриховая вертикальная линия соответствует одному *FPKM*, что в среднем эквивалентно ~30 ридам на один интронный бин.

В) Соотношение между исходной и эффективной длиной интронных бинов.

Г) Эмпирическое (кривая 1) и нулевое (кривая 2) распределение ридов по интронным бином генов, экспрессирующихся в клетках линии *Kasumi-1*. На дополнительной диаграмме рассеивания показано соотношение между эмпирическими и теоретически ожидаемыми уровнями экспрессии интронных бинов генов, активных в указанных клетках.

Наш подход в расчете нулевого распределения схож с предложением Braun C. J. с соавт. [5], в частности, мы также стабилизировали дисперсию через извлечение корня квадратного из эффективной длины интронных бинов, скорректированной по длине ридов. Однако наш подход предполагает использование

исходных данных по суммаризации ридов, а не нормализованных, как в работе Braun C. J. с соавт. Кроме того, дискретизация значений нулевого распределения в нашем подходе не является обязательной, а зависит от того, какая статистическая модель будет использоваться в дальнейшем при идентификации сохраненных

интронов. По нашему мнению, эти два отличия сводят к минимуму деформацию значений нулевого распределения. Тем не менее, это распределение ожидаемо отличается от эмпирического почти полным отсутствием значений равных нулю (рис. 1Г) и умеренно с ним коррелирует (рис. 1Г, вставка).

Пятый этап нашего анализа включал поиск статистически значимых различий между эмпирическими и нулевыми распределениями интронных ридов. Для этого и эмпирические, и теоретически ожидаемые значения объединялись в одну матрицу данных, которая подвергалась стандартным процедурам предпроцессинга и линейного моделирования. На этом этапе мы использовали два метода линейного моделирования: простой и обобщенный. Первый метод основан на допущении, что эмпирические данные могут быть описаны с помощью распределения Гаусса, а второй — распределения Пуассона. В нашей работе практическая реализация этих методов целиком базировалась на функциональных возможностях библиотек *R*-функций *edgeR* и *limma* [15, 16] и *DESeq2* [17], соответственно. На выходе любой из указанных методов выдает логарифмы (по основанию 2) кратности различий, или $\log_2 FC$, между эмпирическим и теоретически ожидаемым количеством интронных ридов, подтверждающих сохранение в молекулах зрелых РНК каждого из интронных бинов. Кроме того, каждый коэффициент $\log_2 FC$ сопровождается как оригинальным *p*-значением, так и скорректированным *q*-значением — *p*-значением с поправкой на множественность проверок нулевых гипотез, проведенной по методу Беньямини-Хохберга [18].

Как видно на рис. 2А, описанный выше подход позволяет идентифицировать множество интронных бинов, представленность которых в молекулах РНК статистически значимо выше, чем теоретически ожидалась. Так, в случае с клетками линии *Kasumi-1* таковых обнаруживается 7963 варианта.

Однако, как показано на рис. 2Б, сохранение таких интронных бинов в молекулах РНК подтверждается разным количеством ридов, в том числе лишь единичными ридами, и это несмотря на то, что длина бинов порой может достигать до нескольких тысяч пар оснований.

Очевидно, что в такой ситуации необходимо выбрать тот минимальный уровень подтверждаемости события, выше которого мы будем рассматривать его как значимое, а ниже — проигнорируем. Выбор такого порога и последующая финальная фильтрация полученных результатов составили суть шестого этапа алгоритма поиска сохраненных интронов.

При выборе такого порога мы ориентировались на два критерия: 1) возможность получить независимое подтверждение экспрессии с помощью количественной полимеразной цепной реакции и 2) схожесть распределения длин интронных бинов, претендующих на роль сохраняющихся в молекулах РНК, с таковым для интронов, которые уже аннотированы в базе данных *Ensembl* как сохраняющиеся в молекулах зрелых РНК человека (см. рис. 1А). Основываясь на собственном опыте, а также опыте других исследователей, мы пришли к выводу, что для того, чтобы экспрессию интрона, обнаруженную по данным *RNA-Seq*, можно было независимо подтвердить и другими методами, например, с помощью стандартной или количественной полимеразной цепной реакции, необходимо, чтобы по нему картировалось не менее 5 ридов. Что касается второго критерия, то, как видно на рис. 2В, чем существеннее экспрессия интронных бинов подтверждается ридами, тем ближе распределение их длин к длинам сохраняющихся интронов из базы данных *Ensembl*. В конечном итоге мы остановились на компромиссном подходе фильтрации, выбрав пороговый уровень экспрессии 40 *FPKM*.

Воспользовавшись выше описанным алгоритмом действий и проведя соответствующую фильтрацию полученных данных, нам удалось идентифицировать множество интронов (в форме интронных бинов), сохраняющихся в молекулах зрелых РНК в транскриптом лейкозных и нормальных клеток крови человека. Минимальное количество сохраненных интронов было обнаружено в транскриптом *CD34⁺* гемопоэтических стволовых/прогениторных клеток нормального костного мозга человека: 425 уникальных интрона, принадлежащих 394 генам. Несколько большее количество таких интронов было идентифицировано в транскриптом *CD34⁺* клеток пуповинной крови человека: 1106 уникальных

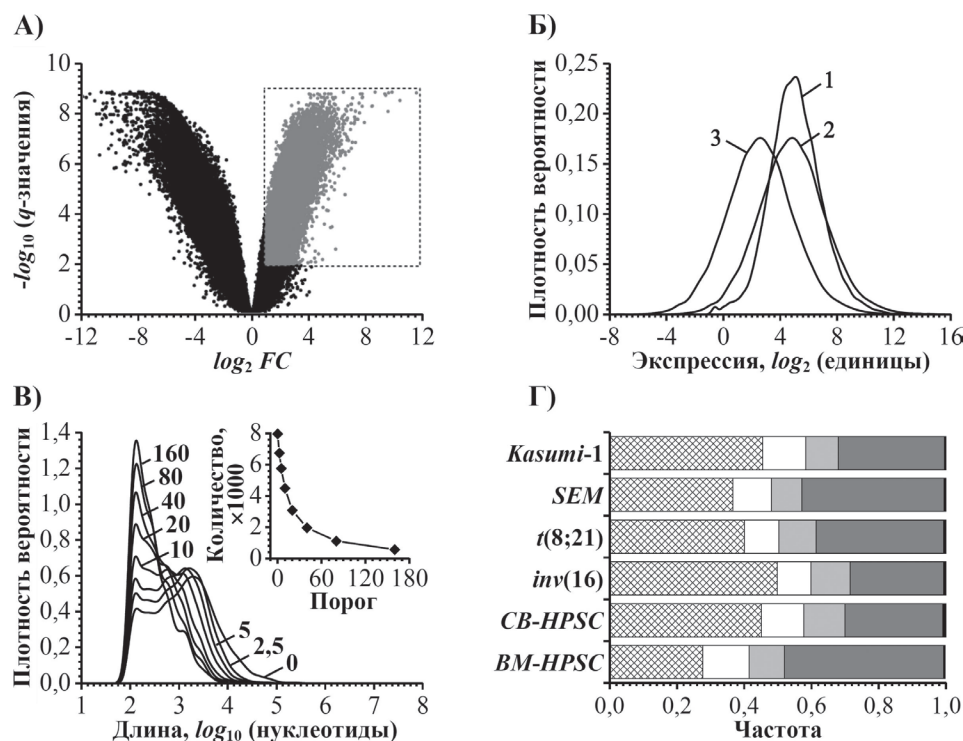


Рис. 2. Идентификация сохраненных интронов в молекулах РНК

Экспериментальные данные, представленные на этом рисунке, были получены при изучении тотальной РНК, выделенной из клеток линии *Kasumi-1*, и секвенированной методом *RNA-Seq*. Линейное моделирование было проведено с помощью функциональных возможностей библиотек *R*-функций *edgeR* и *limma*. Схожие результаты были получены при изучении транскриптомов других типов клеток человека, а также при использовании функциональных возможностей библиотеки *R*-функций *DESeq2*.

А) График-вулкан, демонстрирующий статистические различия между эмпирическим и теоретически ожидаемым количеством ридов, подтверждающих сохранение интронных бинов в молекулах РНК. Ось абсцисс соответствует логарифмам (по основанию 2) кратности различий между эмпирическими и теоретически ожидаемыми значениями. Положительные величины коэффициента $\log_2 FC$ указывают на то, что реальное количество ридов, подтверждающих сохранение интронных бинов в молекулах РНК, выше, чем теоретически ожидаемое. По оси ординат отложены отрицательные логарифмы (по основанию 10) соответствующих q -значений. Прямоугольник ограничивает область статистически значимых ($p < 0,001$, $q < 0,01$) событий с не менее чем двукратной положительной разницей между эмпирическими и теоретически ожидаемыми значениями.

Б) Представленность интронных бинов, выделенных на рис. 2А, в транскриптоме клеток линии *Kasumi-1*. Уровень экспрессии этих бинов выражен с помощью трех разных единиц: 1) количество ридов, приходящихся на один интронный бин без учета его размера (кривая 1), 2) количество ридов, приходящихся на один интронный бин с учетом его размера (кривая 2) и 3) количество ридов, приходящихся на один интронный бин с учетом его размера и размера *RNA-Seq* библиотеки (кривая 3).

В) Взаимосвязь между длиной интронных бинов и уровнем их представленности в транскриптоме клеток линии *Kasumi-1*. Экспрессия интронных бинов выражена в *FPKM*. Цифры на графике указывают на пороговые значения *FPKM*, использованные при фильтрации интронных бинов, выделенных на рис. 2А. На дополнительной вставке показано количество интронных бинов, прошедших фильтрацию при том или ином пороговом значении *FPKM*.

Г) Распределение интронных бинов, сохраняющихся в молекулах зрелых РНК лейкозных и нормальных клеток крови человека, по разным группам: бины интронов, аннотированных в базе данных *Ensembl* как не сохраняющиеся в молекулах зрелых РНК человека (▨), а также интронов, аннотированных как сохраняющиеся в ассоциации с первыми (□), внутренними (■) или последними (■) экзонами генов либо одноэкзонными (■) транскриптами генов. Аббревиатуры *CB-HPSC* и *BM-HPSC* обозначают $CD34^+$ гемопоэтические стволовые/прогениторные клетки, полученные из пуповинной крови или нормального костного мозга человека, соответственно.

интрона из 892 генов. Однако максимальное разнообразие сохраненных интронов наблюдается в транскриптоме лейкозных клеток, причем как клеток перевиваемых клеточных линий, так и клеток, выделенных из организ-

ма пациентов. Так, в транскриптоме клеток линии *Kasumi-1* при созревании молекул РНК 1528 генов произошло 1903 события сохранения интронов, а в клетках больных лейкозом с инверсией *inv(16)* количество таких событий

достигло 2609 и затронуло сплайсинг молекул РНК, кодируемых 1898 генами. Эти различия являются интригующим результатом, однако, обусловлен ли он биологическими причинами или же это технический артефакт, остается вопросом открытым и требующим дальнейших тщательных исследований. Интересно также отметить, что, несмотря на различия в общем количестве, соотношение разных групп сохраненных интронов остается схожим между разными типами клеток, что проиллюстрировано на рис. 2Г.

Для последующей валидации полученных результатов мы провели реконструкцию транскриптомов изучаемых клеток с помощью алгоритма *Cufflinks* [8]. Далее, воспользовавшись инфраструктурой *GenomicRanges* [19], мы идентифицировали те реконструи-

рованные транскрипты, которые содержат обнаруженные *in silico* интроны. Фактически на данном этапе исследования мы проверяли соответствие между покрытием интронной последовательности ридами и включением ее при сборке полноразмерных молекул РНК, как это показано на рис. 3А для гена *NABP1* человека.

Как видно на рис. 3Б, около 90% интронных бинов, классифицированных согласно *in silico* анализу как сохраняющиеся в молекулах зрелых РНК, присутствует в транскриптах, реконструированных с помощью *Cufflinks*. При этом экспрессия транскриптов, содержащих интроны, не отличается от таковой для транскриптов без интронов (рис. 3В). Более того, такие транскрипты получаются более длинными (очевидно, из-за вставки интрона), чем

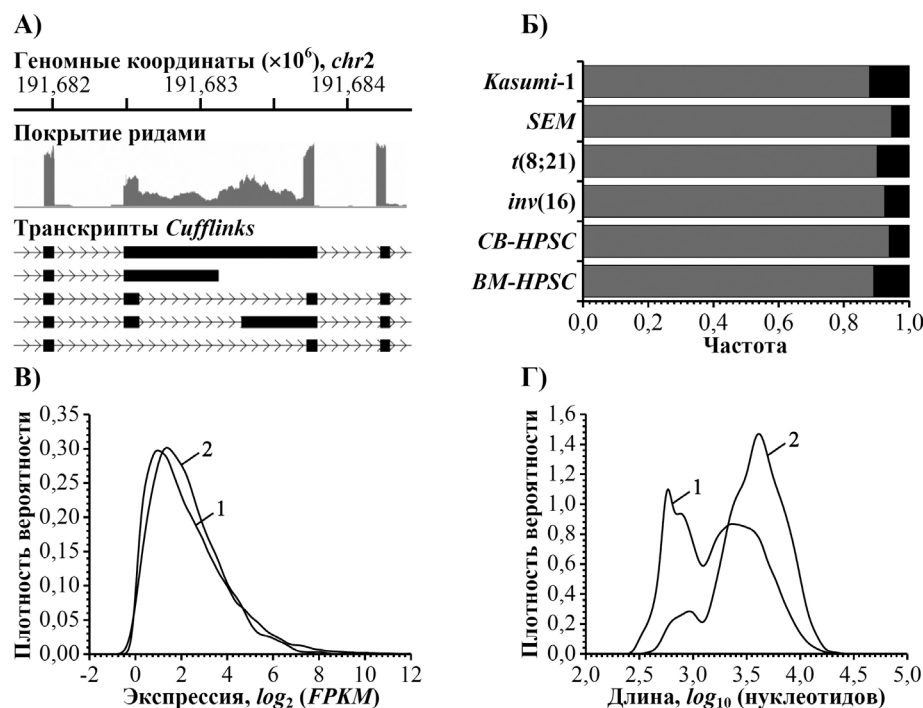


Рис. 3. Присутствие интронов в молекулах РНК, реконструированных с помощью *Cufflinks*

А) Сохранение интрона в молекулах зрелых РНК гена *NABP1* человека, активного в клетках линии *Kasumi-1*. На рисунке показана только та часть гена, где идентифицирован сохраняющийся интрон. Сохранение этого интрона подтверждается как на уровне «сырых» ридов, картированных по гену *NABP1*, так и на уровне реконструированных транскриптов.

Б) Частота встречаемости интронов, определенных *in silico* как сохраняющиеся, в транскриптах, реконструированных с помощью *Cufflinks*. Область интронов, присутствующих в транскриптах, залита серым цветом, не присутствующих — черным цветом.

В) Экспрессия транскриптов, не содержащих интроны (кривая 1), и транскриптов, в которых присутствуют интроны (кривая 2), в клетках линии *Kasumi-1*. Схожие результаты были получены для других типов клеток, вовлеченных в исследование.

Г) Распределение длины транскриптов, не содержащих интроны (кривая 1), и транскриптов, в которых присутствуют интроны (кривая 2). Рисунок основан на данных, полученных для клеток линии *Kasumi-1*, однако схожие результаты наблюдались и при работе с другими типами клеток, вовлеченных в исследование.

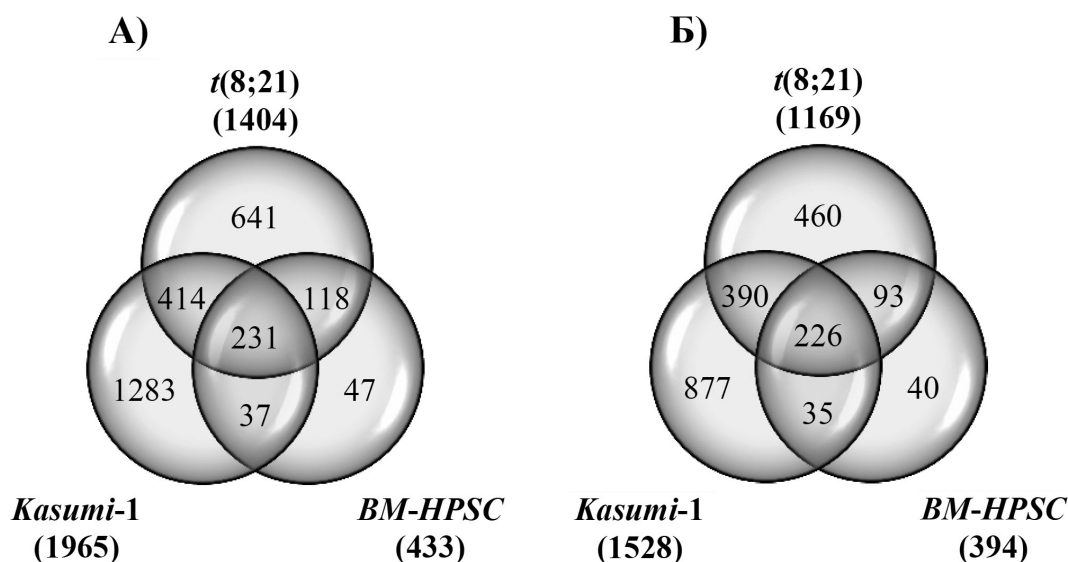


Рис. 4. Разнообразие сохраненных интронов и генов, их кодирующих, в лейкозных и нормальных клетках

Этот рисунок основан на результатах, полученных при проведении линейного моделирования с помощью функциональных возможностей библиотек *R*-функций *edgeR* и *limma*. Схожие результаты были получены и при использовании библиотеки *R*-функций *DESeq2*.

А) Разнообразие интронов, сохраняющихся в зрелых молекулах РНК в транскриптоме клеток линии *Kasumi-1*, лейкозных клеток пациентов с транслокацией *t(8;21)(q22;q22)*, а также *CD34⁺* гемопоэтических стволовых/прогениторных клеток, полученных из нормального костного мозга человека. Общее количество сохраненных интронов, обнаруженных в транскриптоме тех или иных целевых клеток, указано в круглых скобках.

Б) Разнообразие генов, в зрелых молекулах РНК которых были обнаружены интроны. Типы клеток, показанные на этом рисунке, те же, что и на рис. 3А. Общее количество генов, для которых характерно сохранение интронов, указано в круглых скобках.

транскрипты без сохраняющихся интронов (рис. 3Г). Все это указывает на то, что транскрипты, содержащие интроны, скорее всего не являются шумом, появляющимся при работе системы сплайсинга, а, более вероятно, играют какую-то биологическую роль.

В рамках представленной работы мы приняли только базовый вариант функционального анализа наблюдаемого феномена — провели функциональное аннотирование генов, для которых были показаны сохраняющиеся интроны, с помощью генных онтологий. Хотя очевидно, что такой подход не может быть исчерпывающим и что нужен дальнейший более глубокий функциональный анализ. Поскольку спектр сохраненных интронов и генов, для РНК продуктов которых характерно такое явление, в разных клетках различается (рис. 4), то мы провели функциональную аннотацию для каждого из исследуемых типов клеток по отдельности, а затем интегрировали полученные результаты в единую карту обогатчений.

В качестве примера такого рода функционального анализа мы взяли результаты, полученные для клеток линии *Kasumi-1*, лейкозных клеток пациентов, содержащих такую же хромосомную перестройку, как и клетки модельной линии, а также *CD34⁺* гемопоэтических стволовых/прогениторных клеток, полученных из нормального костного мозга человека. Обобщенный результат показан на рис. 5. Видно, что списки генов, для РНК продуктов которых характерно сохранение интронов, не случайны, а статистически значимо обогащены генами из нескольких функциональных групп. Большая часть этих групп так или иначе связана с функционированием самих молекул РНК — от их образования при транскрипции через процессинг (включая этап сплайсинга) до трансляции и внутриклеточной локализации белковых продуктов. Отдельно можно также выделить группы генов, контролирующих состояние хромосом и клеточный цикл, т. е. влияющих на пролиферацию клеток.

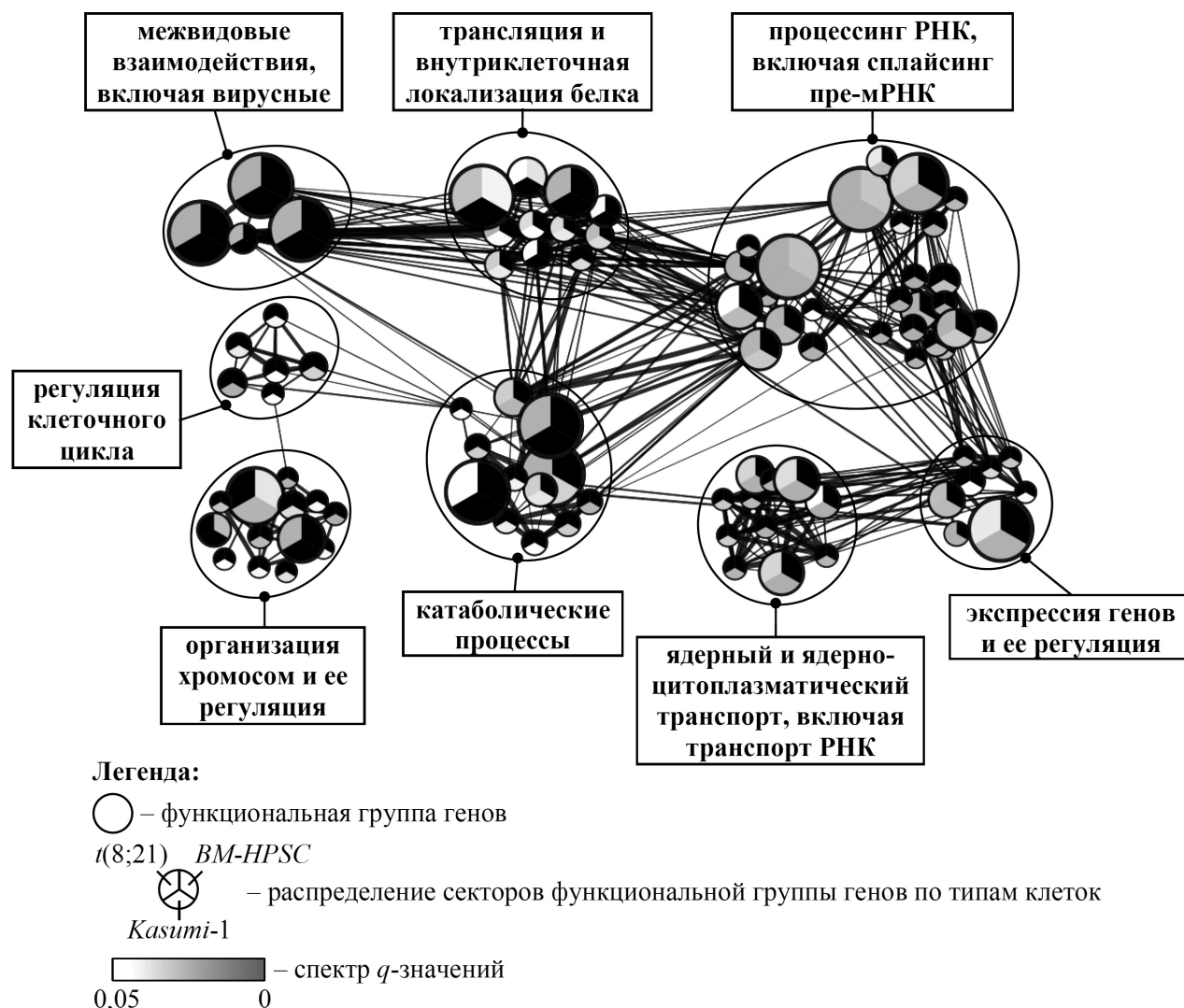


Рис. 5. Функциональная аннотация генов, в зрелых молекулах РНК которых были обнаружены сохраненные интроны

Представленная функциональная аннотация основана на генных онтологиях и аннотациях, разработанных консорциумом *Gene Ontology Consortium*. Карта обогащений построена с помощью плагина *Enrichment Map v.2.2.1* для пакета *Cytoscape v.3.6.0*. Функционально схожие группы генов объединены в метагруппы, названия которых даны в прямоугольных сноках. Диаметр круга функциональной группы генов пропорционален количеству ее членов, найденных среди генов с интронами, сохраняющимися в зрелых молекулах РНК в транскриптоме изучаемых клеток. Черная заливка сектора обозначает отсутствие статистически значимого обогащения генами данной функциональной группы списка генов с сохраняющимися интронами в изучаемом типе клеток. Интенсивность серой заливки сектора зависит от уровня статистической значимости (выраженной в виде скорректированных q -значений) обогащения списка генов с сохраняющимися интронами членами данной функциональной группы: чем меньше q -значение, тем интенсивнее заливка. Толщина линий, соединяющих функциональные группы, пропорциональна степени перекрытия списков генов этих групп.

Заключение

Таким образом, нами разработан многоэтапный алгоритм поиска *in silico* интронных последовательностей, сохраняющихся в молекулах зрелых РНК. Разработанный алгоритм эффективен как при работе с оригинальными, так и с публичными *RNA-Seq* данными. Кроме того, он позволяет идентифицировать сохранившиеся интроны как в транскрипто-

ме нормальных клеток крови человека, так и лейкозных.

Результаты поиска *in silico* интронных последовательностей, сохраняющихся в молекулах зрелых РНК, хорошо согласуются с результатами независимой реконструкции полноразмерных молекул РНК с помощью алгоритма *Cufflinks*, проведенной с использованием тех же первичных *RNA-Seq* данных.

Более того, функциональный анализ, основанный на генных онтологиях, показывает, что выбор генов, для РНК продуктов которых характерно сохранение интронов, не является случайным, а затрагивает преимущественно те гены, которые участвуют в метаболизме самих РНК, трансляции и локализации белков, а также контролируют клеточное деление.

Вместе с тем, полученные результаты поднимают целый ряд вопросов, связанных с феноменом сохранения интронов в молекулах зрелых РНК. В частности, требуют разрешения вопросы о детерминантах этого явления, механизмах избирательности в действиях системы сплайсинга, клеточной специфичности, а также влиянии на кодирующий потенциал молекул РНК и функциональных последствий для белков и клетки в целом. Все эти вопросы могут быть предметом дальнейших исследований.

Автор выражает искреннюю признательность профессору молекулярной гематологии Олафу Хайденрайху из Ньюкаслского университета (Prof. Olaf Heidenreich, Wolfson Childhood Cancer Centre, Northern Institute for Cancer Research, Newcastle University, Newcastle upon Tyne, United Kingdom) за всестороннюю помощь в генерации первичных данных полнотранскриптомного секвенирования клеток линий Kasumi-1 и SEM. Кроме того, автор благодарен доктору Полу Боцу из Массачусетского технологического института (Dr. Paul L. Boutz, Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, USA) за совместное детальное, критическое и продуктивное обсуждение методологии обнаружения интронов, сохраняющихся в зрелых молекулах РНК.

Список использованных источников

1. Ramanouskaya, T.V., Grinev, V.V. The determinants of alternative RNA splicing in human cells // *Mol. Genet. Genomics.* – 2017. – Vol. 292. – P. 1175–1195.
2. Yamile, M., Markus, H., Zahra, A., Andrea, B., Maria, K. Unmasking alternative splicing inside protein-coding exons defines exitrons and their role in proteome plasticity // *Genome Res.* – 2015. – Vol. 25. – P. 995–1007.
3. Jacob, A.G., Smith, C.W.J. Intron retention as a component of regulated gene expression programs // *Hum. Genet.* – 2017. – Vol. 136. – P. 1043–1057.
4. Boutz, P.L., Bhutkar, A., Sharp, P.A. Detained introns are a novel, widespread class of post-transcriptionally spliced introns // *Genes Dev.* – 2015. – Vol. 29. – P. 63–80.
5. Braun, C.J., Stanciu, M., Boutz, P.L., Patterson, J.C., Calligaris, D., Higuchi, F., Neupane, R., Fenoglio, S., Cahill, D.P., Wakimoto, H., Agar, N.Y.R., Yaffe, M.B., Sharp, P.A., Hemann, M.T., Lees, J.A. Coordinated splicing of regulatory detained introns within oncogenic transcripts creates an exploitable vulnerability in malignant glioma // *Cancer Cell.* – 2017. – Vol. 32. – P. 411–426.
6. Liao, Y., Smyth, G.K., Shi, W. The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote // *Nucleic Acids Res.* – 2013. – Vol. 41. – P. e108.
7. Liao, Y., Smyth, G.K., Shi, W. featureCounts: an efficient general-purpose program for assigning sequence reads to genomic features // *Bioinformatics.* – 2014. – Vol. 30. – P. 923–930.
8. Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., Pachter, L. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation // *Nat. Biotechnol.* – 2010. – Vol. 28. – P. 511–515.
9. Zerbino, D.R., Achuthan, P., Akanni, W., Amode, M.R., Barrell, D., Bhai, J., Billis, K., Cummins, C., Gall, A., Girón, C.G., Gil, L., Gordon, L., Haggerty, L., Haskell, E., Hourlier, T., Izuogu, O.G., Janacek, S.H., Juettemann, T., To, J.K., Laird, M.R., Lavidas, I., Liu, Z., Loveland, J.E., Maurel, T., McLaren, W., Moore, B., Mudge, J., Murphy, D.N., Newman, V., Nuhn, M., Ogeh, D., Ong, C.K., Parker, A., Patricio, M., Riat, H.S., Schuilenburg, H., Sheppard, D., Sparrow, H., Taylor, K., Thormann, A., Vullo, A., Walts, B., Zadissa, A., Frankish, A., Hunt, S.E., Kostadima, M., Langridge, N., Martin, F.J., Muffato, M., Perry, E., Ruffier, M., Staines, D.M., Trevanion, S.J., Aken, B.L., Cunningham, F., Yates, A., Flicek, P. Ensembl 2018 // *Nucleic Acids Res.* – 2018. – Vol. 46. – P. D754–D761.
10. The Gene Ontology Consortium. Expansion of the Gene Ontology knowledgebase and

resources // *Nucleic Acids Res.* – 2017. – Vol. 45. – P. D331–D338.

11. Merico, D., Isserlin, R., Stueker, O., Emili, A., Bader, G.D. Enrichment map: a network-based method for gene-set enrichment visualization and interpretation // *PLoS One.* – 2010. – Vol. 5. – P. e13984. – doi: 10.1371/journal.pone.0013984

12. Yu, J., Yang, Z., Kibukawa, M., Paddock, M., Passey, D.A., Wong, G.K. Minimal introns are not “junk” // *Genome Res.* – 2002. – Vol. 12. – P. 1185–1189.

13. Li, W., Freudenberg, J. Mappability and read length // *Front. Genet.* – 2014. – Vol. 5. – doi: 10.3389/fgene.2014.00381.

14. Derrien, T., Estelle, J., Marco Sola, S., Knowles, D.G., Raineri, E., Guigo, R., Ribeca, P. Fast computation and applications of genome mappability // *PLoS One.* – 2012. – Vol. 7. – P. e30377. – doi: 10.1371/journal.pone.0030377.

15. Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., Smyth, G.K. limma pow-

ers differential expression analyses for RNA-sequencing and microarray studies // *Nucleic Acids Res.* – 2015. – Vol. 43. – P. e47.

16. Robinson, M.D., McCarthy, D.J., Smyth, G.K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data // *Bioinformatics.* – 2010. – Vol. 26. – P. 139–140.

17. Love, M.I., Huber, W., Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2 // *Genome Biol.* – 2014. – Vol. 15. – P. 550.

18. Benjamini, Y., Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing // *J. R. Stat. Soc. Ser. B.* – 1995. – Vol. 57. – P. 289–300.

19. Lawrence, M., Huber, W., Pages, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan, M.T., Carey, V.J. Software for computing and annotating genomic ranges // *PLoS Comput. Biol.* – 2013. – Vol. 9. – P. e1003118. – doi: 10.1371/journal.pcbi.1003118

V.V. Grinev

INTRON RETENTION IN THE TRANSCRIPTOME OF LEUKEMIC AND NORMAL HUMAN BLOOD CELLS

Belarusian State University
Minsk, 220030, the Republic of Belarus

Retention of introns in mature *RNA* molecules is a dark side of the splicing machinery activity in human cells. We have developed a multi-step algorithm for *in silico* identification of introns retained in mature *RNA* molecules using the *RNA-Seq* technique. Algorithm application results are in good agreement with the reconstruction results of full-length *RNA* molecules carried out for different types of human cells and confirm the functional significance of genes whose *RNA* products contain introns.

Key words: intron retention, *RNA* splicing, *RNA-Seq*, transcriptome.

Дата поступления статьи: 30 марта 2018 г.

А.О. Рябцева¹, И.С. Цыбовский², С.А. Котова²

МИКРОСАТЕЛЛИТНЫЕ МАРКЕРЫ В ИССЛЕДОВАНИИ ПОЛИМОРФИЗМА ДИКОГО КАБАНА (*SUS SCROFA*) И СВИНЬИ ДОМАШНЕЙ (*SUS SCROFA DOMESTICUS*), ОБИТАЮЩИХ НА ТЕРРИТОРИИ РЕСПУБЛИКИ БЕЛАРУСЬ

¹Государственный комитет судебных экспертиз Республики Беларусь
Республика Беларусь, 220073, г. Минск, ул. Кальварийская, 43; e-mail: alunchik_90@mail.ru

²Государственное учреждение «Научно-практический центр Государственного комитета судебных экспертиз Республики Беларусь»
Республика Беларусь, 220114, г. Минск, ул. Филимонова, 25

В судебно-экспертной идентификации объектов животного происхождения важнейшее значение имеют методы молекулярной биологии, из которых наиболее распространенными являются анализ микросателлитных локусов (STR) и митохондриальной ДНК. Генотипирование образцов и последующая статистическая оценка результатов предоставляет следственным органам высоко доказательную информацию, для чего необходимы справочные базы частот встречаемости исследуемых ДНК-маркеров. С целью подбора панели информативных маркеров для ДНК-идентификации биологических образцов вида кабан (*Sus scrofa*), обитающего на территории Республики Беларусь, проведено исследование 26 видоспецифичных STR-маркеров. На основе полученных данных о генетическом полиморфизме диких (n=719) и домашних (n=304) представителей вида произведен отбор 16 STR-локусов, которые позволяют установить происхождение животного от конкретной особи.

Ключевые слова: полиморфизм, микросателлиты, дикий кабан, свинья домашняя, идентификация.

Введение

На сегодняшний день широкий спектр преступлений связан с объектами животного происхождения (незаконная охота, торговля находящимися под угрозой исчезновения видами, трансграничное перемещение трофеев и изделий (дериватов) животного происхождения, мошенничество с составом и истинным происхождением мяса и полуфабрикатов, кражи и жестокое обращение с животными и др.). Одним из наиболее распространенных объектов, в отношении которого фиксируются факты незаконной охоты на территории Республики Беларусь, является кабан (*Sus scrofa*) [1]. В то же время в стране активно ведется разведение высококачественных пород свиньи домашней, в отношении которых совершаются преступления имущественного характера, связанные с кражами животных или мясопродуктов.

Для получения доказательной информации при расследовании данных преступлений широко применяются молекулярные методы

исследований. В судебных лабораториях по всему миру проводится молекулярно-генетическая идентификация видов, основанная на анализе митохондриальной или ядерной ДНК, полученной из тканей животных (костей, чешуй и других биологических объектов с мест преступлений). Кроме того, генетические маркеры используются для определения пола, оценки биологического происхождения (родства), установления принадлежности особи к конкретной популяции, распознавания гибридов.

Современные тенденции в подборе генетических маркеров направлены на повышение разрешающей способности, специфичности, эффективности, доступности и технологичности анализа. Согласно основным требованиям, ДНК маркер должен отвечать комплексу характеристик: быть высоко полиморфным для выявления генетического разнообразия; иметь кодоминантный тип наследования для возможности дифференциации гетерозиготных орга-

низмов. Микросателлиты (STR, от англ. Short Tandem Repeats) имеют ряд преимуществ перед другими маркирующими системами: они множественны, высокополиморфны, широко представлены по всему геному, легко выявляются и идентифицируются, в связи с чем, стали наиболее надежным и эффективным методом анализа индивидуальных особенностей организмов. Кроме того, относительно высокие темпы мутирования этих локусов приводят к накоплению популяционно-специфических мутаций, что позволяет проводить детальный анализ популяционной структуры. В судебной экспертизе ДНК-анализ на основе STR-маркеров имеет особое значение, поскольку позволяет получить информацию доказательного уровня при исследовании биологических следов, которые не идентифицируются морфологически: следы крови, фрагменты внутренних органов, видоизмененные со временем останки и т. п.

Цель данной работы — подбор ДНК-маркеров для решения комплекса экспертных задач, возникающих в процессе расследования уголовных и административных преступлений, объектами которых становились дикий кабан/свинья домашняя или продукты их технологической переработки.

Материалы и методы

Полиморфизм микросателлитных маркеров кабана (*Sus scrofa*) был исследован в массиве из 719 образцов дикого кабана из различных районов Республики Беларусь и массиве из 304 образцов 6 пород свиньи домашней. В массиве образцов свиньи домашней представлены следующие породы: белорусская мясная (n=53), белорусская крупная белая (n=50), белорусская черная пестрая (n=20), йоркширская (n=55), ландрас (n=54) и дюрок (n=72). Биологические образцы представляли собой мышечную и хрящевую ткань.

Для экстракции ДНК образцы инкубировали при постоянном встряхивании и температуре 56 °С в 10 мМ Трис-НСl буфере pH 7.5, содержащем 2% додецилсульфата натрия (SDS), 100 мМ NaCl, 20 мМ EDTA с добавлением 5 мкл протеиназы К (20 мг/мл) в течение 24 часов. Лизат биологических образцов далее подвергали процедуре очистки на силикагеле [2].

Амплификацию образцов проводили при следующих условиях: 0,2 мМ dNTP, 10 мМ трис-НСl буфер pH 8.6; 50 мМ KCl; 1,5 мМ MgCl₂ и 0,75 U Taq-полимеразы. Концентрации праймеров варьировали в диапазоне 0,032–0,052 μМ.

Анализ нуклеотидной последовательности проводили с использованием набора BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, США) с последующей очисткой продуктов секвенирования на колонках NucleoSEQ (Macherey-Nagel, Германия). Анализ нуклеотидной последовательности исследуемых микросателлитных локусов проводили при помощи программных пакетов: «SeqScape Software 3» и «Sequencing Analysis Software 6».

Синтез праймеров для генотипирования и секвенирования произведен ОДО «Синтол», Россия.

Характерное для каждого из образцов сочетание аллелей выявляли путем электрофоретического разделения продуктов ПЦР в генетическом анализаторе «3500 Genetic Analyzer» (Applied Biosystems, США). Определение размеров выявленных аллелей (в п. н.) и соответствующих генотипов ДНК в исследуемых локусах проводили с использованием внутреннего стандарта размера «GeneScan-600 LIZ Size Standard v.2.0» (Applied Biosystems, США) на основе специализированной программы «GeneMapper ID-X» (Applied Biosystems, США) и электрофореграмм продуктов ПЦР образцов контрольной ДНК. Контрольная ДНК была представлена образцами из популяций дикого кабана и свиньи домашней с известными генотипами.

Статистическую обработку результатов генотипирования проводили с использованием программного пакета «Arlequin v.3.5.1.3» (Population Genetics CPMG Lab, Швейцария) и программного средства «PowerStats V12 v.2» (Promega, США).

Результаты и обсуждение

В работе были использованы 26 STR-маркеров. Подбор STR-локусов проводили с учетом рекомендаций FAO (от англ. Food and Agriculture Organization), представленных в библиотеке DAD-IS[3] (от англ. Domestic Animal Diversity Information System) и панели ISAG (от англ. International Society for Animals Genetics) [4, 5].

В судебно-экспертном ДНК-анализе подбор локусов имеет особое значение, поскольку достоверность полученных экспертом результатов может быть проверена назначением повторного исследования образцов в другой лаборатории с оценкой воспроизводимости результатов. Генотипирование многих динуклеотидных STR-маркеров характеризуется высоким уровнем статтер-пиков и нередко выраженным аденилированием (+/- А-пики), которые могут приводить к несопоставимости результатов, полученных при проведении первичных и повторных экспертиз, особенно при исчислении аллелей локуса по молекулярному размеру фрагмента (пары нуклеотидов).

Поэтому основу идентификационной панели составили локусы с тетрануклеотидной структурой тандема: FH2148, FH1727, FH1701, FH2709, FH2478, FH1733, FH3637, FH1900, FH1696, FH1589, FH4219, NLRIP0001, S0766, D0768, S0719, FH3558, S0663. В то же время, локусы с динуклеотидной структурой тандема, специфичные для свиньи домашней, изучены в мировой научной литературе значительно подробнее, что позволяет оценить их возможности также и в дифференциации биологических образцов генетически близкородственного происхождения (дикого кабана и свиньи домашней). В настоящей работе задействованы следующие локусы с динуклеотидным повтором тандема: SW240, SW857, S0101, S0005, S0355, S0710, SW951, S0227, SW403 [6–11].

Оценка пригодности панели локусов для криминалистической идентификации проводилась на основе скрининга образцов, анализа генетических и криминалистических параметров массивов генотипов, результатов исследования первичной структуры аллелей на предмет выявления природы полиморфизма, наличия инсерций/ делеций и т. п.

Отдельным и немаловажным фактором является дальнейшая необходимость мультиплексирования локусов для создания наборов, технологичных в работе. Объединение праймеров различных локусов в один набор, во-первых, усредняет условия ПЦР для всех локусов, и, во-вторых, создает дополнительные требования, такие как недопустимость перекрывания аллельных диапазонов различных локусов, отсутствие взаимного влияния праймеров друг на друга в ходе ПЦР и др.

В ходе исследований при высоком уровне полиморфизма, установленном в предварительных экспериментах (рис. 1), в дальнейшем не использовались в мультиплексах локусы FH1657, FH3558, S0719, S0663. Так, для локусов FH1657 и FH3558 показан широкий диапазон молекулярных размеров аллелей, что создает риск перекрывания аллелей с соседними локусами — в дальнейшем мультиплексировании использовался 1 локус (FH3637) с широким диапазоном аллелей. Аллели локуса S0719 не оптимальны по размеру (460–520 п. н.), поскольку увеличивается время электрофоретического разделения продуктов ПЦР. При генотипировании образцов в мультиплексе для локуса S0663 установлен значительный дисбаланс интенсивности наработки «коротких» и «длинных» аллелей, который может приводить к «ложной» гомозиготности.

В локусах S0355, S0710, SW951, SW403 и FH4219 доминирующие аллели встречаются с частотой 0,865–0,997 в группе диких, а у домашних животных доминирующий аллель локуса S0227 встречается с частотой 0,816. Такой характер распределения аллелей делает эти локусы малоинформативными при ДНК-идентификации образцов. Однако, ввиду установленных различий в распределении частот аллелей у диких и домашних животных, эти локусы вносят существенный вклад при дифференциации неизвестного образца по типу «дикий»/«домашний» (на примере локуса FH4219 — рис. 2).

Таким образом, в дальнейшем подробно изучался полиморфизм локусов FH2148, FH1727, FH1701, FH2709, FH2478, FH1733, FH3637, FH1900, FH1696, FH1589, NLRIP0001, S0766, SW240, SW857, S0101, S0005. Для локусов с тетрануклеотидными повторами также было проведено изучение первичных последовательностей выявленных аллелей. В общей сложности была установлена нуклеотидная структура 71 аллеля из 255, суммарно выявленных в ходе генотипирования выборок диких и домашних животных. В результате анализа первичной структуры аллелей подтверждено, что аллели локусов FH1589, FH1733, FH1727, FH2478, FH1696, FH1701, FH1900, FH2709, FH2148, FH3637, NLRIP0001 имеют тетрануклеотидную структуру тандемной области, что при выявленном

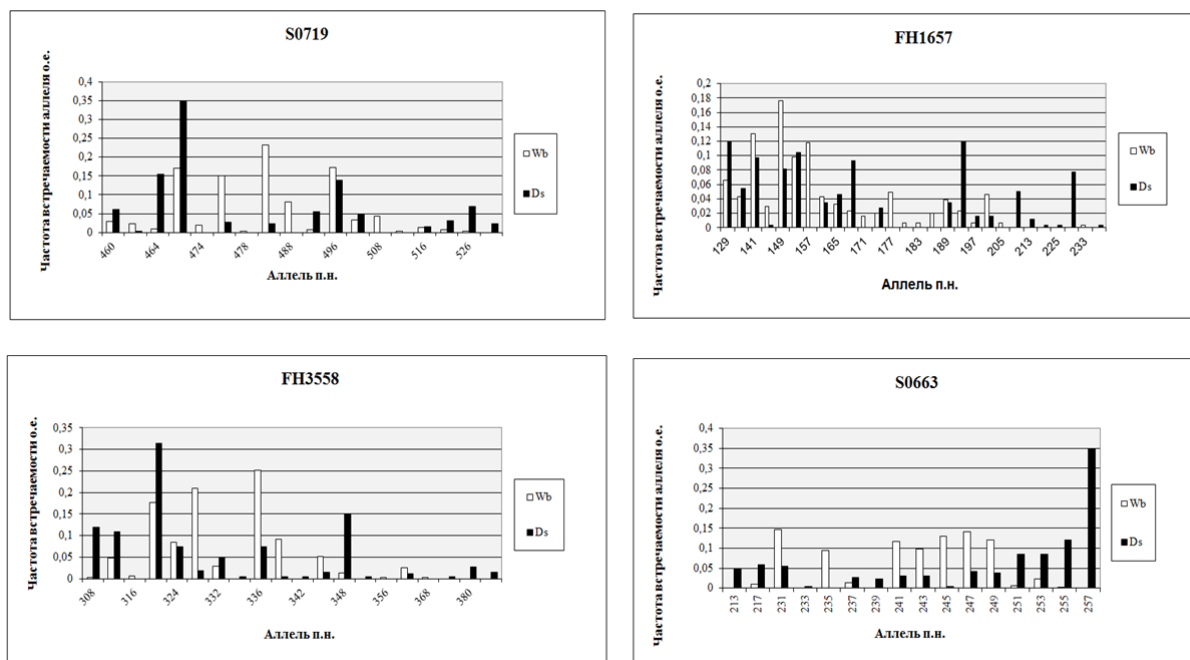


Рис. 1. Диаграмма распределения частот встречаемости аллелей локусов FH1657, FH3558, S0719, S0663 у дикого кабана (Wb, *Sus scrofa*) и свиньи домашней (Ds, *Sus scrofa domestica*)

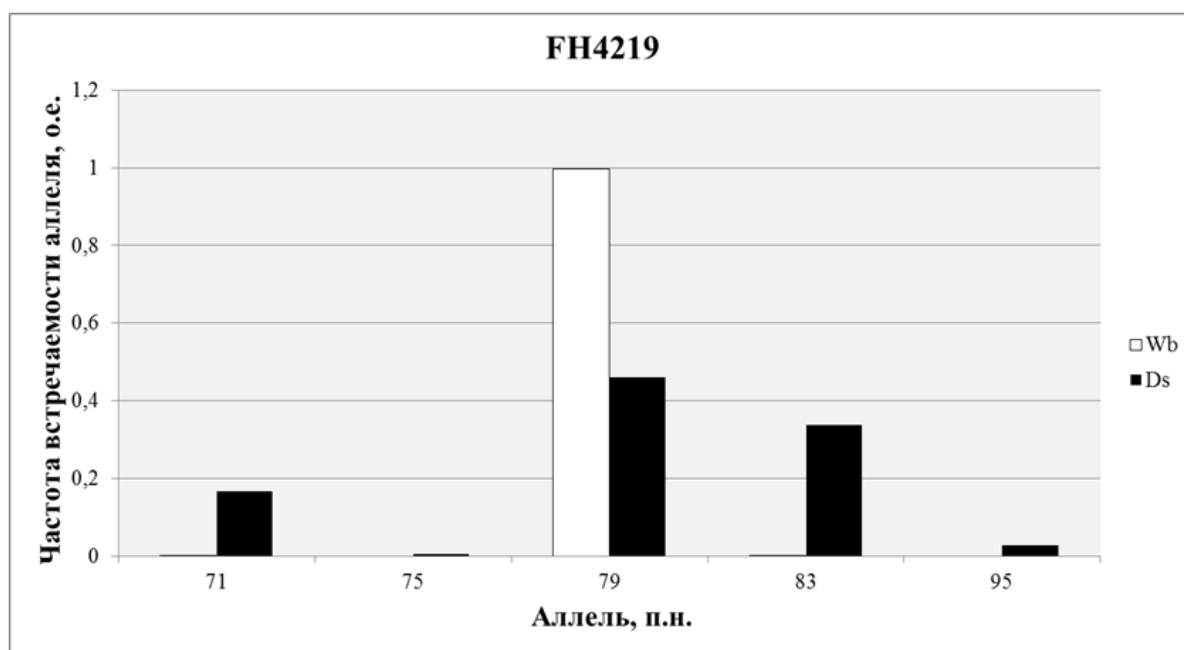


Рис. 2. Диаграмма распределения частот встречаемости аллелей локуса FH4219 у дикого кабана (Wb, *Sus scrofa*) и свиньи домашней (Ds, *Sus scrofa domestica*)

уровне полиморфизма позволяет использовать данные локусы для решения идентификационных задач.

По данным секвенирования 16 аллелей локуса D0768 установлено, что наряду с повторяющимся тандемным мотивом $-(GAAA)_n$ - существует участок, содержащий динуклеотидные

повторы $-(GA)_5-CA-(GA)_n$ -, где для n экспериментально установлены значения 11–14, 16 и 22. Наличие двух повторяющихся тандемов дает различную структуру первичных последовательностей для двух пар изоаллелей (в 288 п. н. и 309 п. н.), т. е. аллелей, совпадающих по молекулярным размерам, но различных по

первичной структуре. По причине наличия изоаллельных вариантов данный локус в дальнейшей работе не использовался.

Локус S0766 также содержит два варьирующих tandemных мотива, тетрауклеотидный (GAAA) и динуклеотидный (CA), что потенциально определяет возможность существования изоаллелей. Поскольку экспериментально установлено, что локус S0766 вносит существенный вклад при молекулярно-генетической дифференциации образцов дикого и домашнего происхождения, локус включен в разработанную панель, однако генотипы локуса S0766 не рекомендуется использовать при идентификации образцов.

В выборках дикого кабана и свиньи домашней суммарно выявлено 255 аллельных вариантов (212 в группе диких и 201 в группе домашних животных). Наиболее полиморфными в обеих группах животных оказались локус S0005 с суммарным количеством аллельных вариантов 32, локус FH2148 с суммарным количеством аллельных вариантов 22, и локусы FH1589 и FH3637, имеющие по 21 варианту аллелей, а наименее полиморфными — локусы FH1900, FH2478 и FH1733 (по 10 аллельных вариантов).

Аллельное разнообразие в целом выше у диких кабанов.

Результаты исследования полиморфизма отобранных 16 локусов приведены в табл. 1.

Таблица 1

Аллельные диапазоны и количество выявленных аллелей в выборках дикого кабана и свиньи домашней

Локус	Аллельные диапазоны у дикого кабана, п. н.	Аллельные диапазоны у свиньи домашней, п. н.	Количество аллелей у дикого кабана	Количество аллелей у свиньи домашней	Общее количество аллелей
FH1589	136–196 (аллели 9–24)*	136–204	16	19	21
FH1696	349–401 (аллели 10–23)	345–393	11	12	15
FH1701	175–231 (аллели 6–20)	175–231	12	15	15
FH1727	206–254 (аллели 16–28)	210–246	13	8	13
FH1733	282–314 (аллели 8–16)	274–310	9	9	10
FH1900	231–263 (аллели 3–11)	239–279	9	7	10
FH2148	232–340 (аллели 20–47)	244–376	17	18	22
FH2478	272–308 (аллели 5–14)	276–304	10	8	10
FH2709	114–182 (аллели 14–30.2)	114–186	18	13	19
FH3637	133–291 (аллели 2–41)	137–323	16	14	21
NLRIP0001	333–369 (аллели 10–19)	337–381	10	10	13
S0005	193–267	203–269	27	26	32
S0101	195–219	193–217	13	8	14
S0766	432–468	434–464	9	9	11

Продолжение табл. 1

Локус	Аллельные диапазоны у дикого кабана, п. н.	Аллельные диапазоны у свиньи домашней, п. н.	Количество аллелей у дикого кабана	Количество аллелей у свиньи домашней	Общее количество аллелей
SW240	95–127	93–127	15	14	18
SW857	144–158	136–160	7	11	11
В расчете на 1 локус	—	—	13,3	12,7	15,9

Примечание. В скобках приведена нумерация аллелей соответственно числу повторяющегося тандема

Сравнительный анализ показал, что размерные диапазоны аллелей для обеих групп животных в целом совпадают, однако для всех исследуемых локусов показано наличие аллельных вариантов, характерных только для одной из групп животных.

Для локуса FH2709 установлено наличие 6 аллелей с неполными тандемными повторами — минорных аллелей, первичная структура 4 из них подтверждена секвенированием. Идентификация минорных аллелей может представлять определенные сложности. Локус

FH3637 имеет широкий размерный диапазон аллелей (190 п. н.) с нерегулярным распределением аллелей и одной распространенной доминантой как у диких, так у домашних животных — аллелем 19 (0,4).

В табл. 2 приведено частотное распределение аллелей тетрауклеотидных локусов в выборке дикого кабана. Нумерация аллелей соответствует числу тандемных повторов, жирным шрифтом выделены аллели, для которых секвенированием установлена первичная последовательность.

Таблица 2

Частотное распределение аллелей в выборке дикого кабана

№	FH1589	FH1696	FH1701	FH1727	FH1733	FH1900	FH2148	FH2478	FH2709	FH3637	NLRP 0001
2										0,066	
3						0,005				0,078	
4						0,002				0,013	
5						0,028		0,002			
6			0,001			0,051		0,001		0,002	
7						0,110		0,308			
8			0,035		0,003	0,154		0,139			
9	0,001		0,072		0,072	0,412		0,051			
10	0,004	0,013	0,175		0,159	0,228		0,040			0,033
11	0,097		0,135		0,337	0,010		0,110			0,033
13	0,043	0,027	0,261		0,141			0,161		0,001	0,276
14	0,026	0,169	0,188		0,043			0,001	0,024	0,006	0,305
15	0,146	0,267	0,090		0,028						0,164
16	0,092	0,220	0,001	0,004	0,003				0,014		0,005

Продолжение табл. 2

№	FH1589	FH1696	FH1701	FH1727	FH1733	FH1900	FH2148	FH2478	FH2709	FH3637	NLRP0001
17	0,098	0,156	0,001	0,083					0,038	0,052	0,079
17.2									0,374		
18	0,242	0,056		0,037					0,007	0,329	0,047
19	0,102	0,010		0,084					0,226	0,403	0,004
20	0,121	0,019	0,009	0,103			0,001		0,065	0,001	
21	0,015			0,182			0,051		0,149		
22	0,006			0,001					0,019	0,001	
23	0,003	0,001		0,013			0,054		0,020	0,039	
24	0,001			0,063			0,070		0,017		
25				0,232			0,055		0,001		
26				0,120			0,051		0,033		
26.2									0,002		
27				0,068			0,062		0,002		
28				0,011			0,113				
28.2									0,001		
29							0,108				
29.2									0,008		
30							0,117				
30.2									0,001		
31							0,201				
32							0,089				
33							0,021			0,004	
34							0,004				
35							0,002			0,002	
37							0,001				
39										0,001	
41										0,002	
47							0,001				

Сравнительный анализ частот встречаемости аллелей исследуемых локусов в анализируемых группах (дикого кабана и свиньи домашней) показал наличие статистически значимой генетической дифференциации между ними ($F_{st} = 0,07072$; $P < 0,000001$).

Анализ соответствия частот распределения аллелей в популяции закону Харди-Вайнберга

показал, что в суммарной выборке дикого кабана отклонение от равновесия Харди-Вайнберга было выявлено в 7 локусах, а в выборке домашних свиней не соответствовали равновесию 14 локусов. После применения поправки Bonferroni [12] количество локусов, не находящихся в равновесии, составило 3 и 11 для выборок диких кабанов и домашних свиней, соответственно.

Выборки диких кабанов и домашних свиней в целом показали высокие значения ожидаемой и наблюдаемой гетерозиготности (табл. 3) в исследованных локусах, за исключением локусов FH1900 и S0101 в группе домашних животных.

Криминалистическую применимость локусов для идентификационных целей оценивали по следующим критериям: вероятность исключения исследуемого образца по заданному генотипу (power of exclusion, PE); вероятность различения генотипов двух неродственных индивидуумов (power of discrimination, PD); информационное содержание полиморфизма (polymorphism information content, PIC). Значения PIC для исследованных локусов приведены в табл. 4. Критерий PIC выявляет дискриминационную способность маркера, зависит от числа известных аллелей и распределения их частот и эквивалентен генному разнообра-

зию. Чем выше данное значение, тем более информативным является исследуемый локус.

В результате показано, что параметры PIC различаются как между локусами, так и для одних и тех же локусов в различных выборках образцов.

В целом оценка генетико-популяционных параметров локусов подтверждает применимость исследованной панели локусов для решения идентификационных задач в судебном ДНК-анализе образцов животного происхождения.

Способность отобранных локусов идентифицировать исследуемые образцы (cumulative power of discrimination, CPD) [13] составляет 99,99999999999999% для популяции дикого кабана и 99,99999999999999% для свиньи домашней.

Возможность использования панели локусов для установления родства была доказана

Таблица 3

Значения наблюдаемой и ожидаемой гетерозиготности в выборках дикого кабана и свиньи домашней

Локус	He для диких	Ho для диких	He для домашних	Ho для домашних
FH1589	0,86467	0,86926	0,87196	0,80592
FH1696	0,81918	0,80389	0,86697	0,73684
FH1701	0,83249	0,80529	0,85613	0,71711
FH1727	0,86412	0,85953	0,73407	0,69408
FH1733	0,78742	0,78164	0,85635	0,78289
FH1900	0,73898	0,72184	0,55924	0,46711
FH2148	0,89299	0,84562	0,81103	0,71053
FH2478	0,80871	0,79277	0,76675	0,70066
FH2709	0,77786	0,77330	0,85505	0,81250
FH3637	0,71477	0,68567	0,72456	0,60526
NLRIP0001	0,79008	0,76356	0,71582	0,66118
S0005	0,87375	0,84006	0,92425	0,90132
S0101	0,79599	0,80250	0,65251	0,59539
S0766	0,78571	0,78720	0,78464	0,66776
SW240	0,81150	0,78581	0,81535	0,80592
SW857	0,66528	0,67177	0,84313	0,80592
Среднее	0,801469	0,786857	0,789863	0,7169

Таблица 4

Значения генетико-популяционного параметра PIC для исследованных микросателлитных локусов особей рода кабанов (*Sus*)

№ п/п	Локус	PIC для дикого кабана	PIC для свиньи домашней
1	FH1589	0,85082	0,85920
2	FH1696	0,79531	0,85240
3	FH1701	0,81165	0,83985
4	FH1727	0,85010	0,69517
5	FH1733	0,75834	0,84089
6	FH1900	0,70316	0,46171
7	FH2148	0,88381	0,79829
8	FH2478	0,78353	0,72784
9	FH2709	0,75126	0,83973
10	FH3637	0,67147	0,68434
11	NLRIP0001	0,76129	0,68005
12	S0005	0,86305	0,91942
13	S0101	0,76920	0,61116
14	S0766	0,75213	0,75550
15	SW240	0,78799	0,79543
16	SW857	0,60301	0,82489

исследованием семейных групп свиньи домашней. Уровень различения потомков (сисбсов) в первом поколении с помощью 16 STR-локусов исследовали в 6 семейных группах. Из них 2 группы включали обоих родителей и по 13 сисбсов в каждом помете. Остальные семейные группы были неполными: группа № 1 — свиноматка и 9 сисбсов одного помета; группа № 2 — свиноматка и 7 сисбсов; группа № 4 — свиноматка и 11 сисбсов; группа № 5 — свиноматка и 8 сисбсов.

В результате исследований панель из 16 STR-локусов позволила индивидуализировать каждую особь в первом поколении во всех семейных группах.

Заключение

Для идентификации кабана (*Sus scrofa*) в литературе описано большое количество микросателлитных локусов, которые потенциально

могут быть эффективным инструментом для решения задач судебной экспертизы. При этом в судебно-экспертном приложении к STR-маркерам помимо информативности, зависящей от степени их полиморфизма, предъявляются дополнительные требования, такие как способность к мультиплексированию, к полиморфизму нуклеотидной последовательности, устойчивости амплификации и др.

В результате экспериментального исследования 26 микросателлитных локусов было отобрано 16 видоспецифичных STR-маркеров, с использованием которых проведено скрининговое исследование 719 образцов дикого кабана из различных регионов Беларуси и 304 образцов 6 пород свиньи домашней, разводимой на территории Республики. Исследование первичной структуры аллелей с тетрануклеотидным tandemным повтором позволило вести общепринятое универсаль-

ное исчисление аллелей по числу tandemных повторов.

Получены приоритетные знания фундаментального уровня о генетическом полиморфизме микросателлитных локусов животных вида кабан (диких и домашних), обитающих на территории Республики Беларусь. Экспериментально доказано, что отобранные 16 STR-локусы пригодны для достоверной идентификации образцов дикого кабана и свиньи домашней, что существенно в формировании объективной доказательственной информации в судебно-следственном процессе при расследовании правонарушений в отношении животных данного вида. Полученные сведения о полиморфизме микросателлитных локусов рода кабанов могут использоваться в мониторинге развития диких популяций, охотоведении, в генетике животных, научных исследованиях и селекции в свиноводстве.

Авторы выражают благодарность Гандже А.И. и Журиной Н.В., сотрудникам РУП «Научно-практический центр НАН Беларуси по животноводству», за предоставление для исследования коллекционных образцов пород свиньи домашней.

Список использованных источников

1. Добыча (изъятие) основных видов охотничьих животных // Национальный статистический комитет Республики Беларусь [Электронный ресурс]. – 1998. – Режим доступа: <http://belstat.gov.by/homep/ru/indicators/envir.php>. – Дата доступа: 08.02.2011.
2. Boom, R. et al. Rapid and simple method for purification of nucleic acids / R. Boom [et al.] // Journal of Clinical Microbiology. – 1990. – Vol. 28, № 3. – P. 495–503.
3. Domestic Animal Diversity Information System [Электронный ресурс]. – Режим доступа: <http://www.fao.org/dad-is/en/>. – Дата доступа: 17.10.2017.
4. Laval, G. et al. Genetic diversity of eleven European pig breeds / G. Laval et al. // Genetics Selection Evolution. – 2000. – Vol. 32, № 2. – P. 187–203.
5. Fan, B. et al. Population genetic variability and origin of Auckland Island feral pigs / B. Fan et al. // Journal of the Royal Society of New Zealand. – 2005. – Vol. 35, № 3. – P. 279–285.
6. Cherel, P., Glinisson, J., Pires, J. Tetranucleotide microsatellites contribute to a highly discriminating parentage test panel in pig: Tetranucleotide microsatellites / P. Cherel, J. Glinisson, J. Pires // Animal Genetics. – 2011. – Vol. 42, № 6. – P. 659–661.
7. Rohrer, G.A. et al. A microsatellite linkage map of the porcine genome / G.A. Rohrer et al. // Genetics. – 1994. – Vol. 136, № 1. – P. 231–245.
8. Yaemmeklin, W. et al. Efficacy of Microsatellite Markers in Parentage Control in Swine / W. Yaemmeklin et al. // Thai J. Vet. Med. – 2009. – Vol. 39, № 3. – P. 259–265.
9. Chen, K. et al. Targeted oligonucleotide-mediated microsatellite identification (TOMMI) from large-insert library clones / K. Chen et al. // BMC genetics. – 2005. – Vol. 6, – P. 54.
10. Robic, A. et al. Isolation of 28 new porcine microsatellites revealing polymorphism / A. Robic et al. // Mamm Genome. – 1994. – Vol. 5, – P. 580–583.
11. Alexander, L.J. et al. Physical assignments of 68 porcine cosmid and lambda clones containing polymorphic microsatellites / L.J. Alexander et al. // Mamm. Genome. – 1996. – Vol. 7, – P. 368–372.
12. Watsh B. Multiple Comparisons: Bonferoni Corrections and False Discovery Rates. Lecture Notes for EEB 581/ B. Walsh // – 2004. – 17 p.
13. Olowofeso, O. et al. Estimation of the Cumulative Power of Discrimination in Haimen Chicken Populations with Ten Microsatellite Markers / O. Olowofeso et al. // Animal Breeding and Genetics. – 2005. – Vol. 8, – P. 1066–1070.

A.A. Rabtsava¹, I.S. Tsybovsky², S.A. Kotova²

MICROSATELLITE MARKERS IN THE STUDY OF POLYMORPHISM OF THE WILD BOAR (*SUS SCROFA*) AND DOMESTIC PIG (*SUS SCROFA DOMESTICUS*), INHABITING THE TERRITORY OF THE REPUBLIC OF BELARUS

¹State Forensic Examination Committee of the Republic of Belarus
Minsk, 220073, the Republic of Belarus

²State Institution “Scientific and Practical Centre of the State Forensic Examination Committee of the Republic of Belarus”
Minsk, 220114, the Republic of Belarus

Molecular biology methods have a high value in the forensic examination of animal origin objects. Analysis of STR loci and of mitochondrial DNA are the most common methods. Genotyping of samples and subsequent statistical evaluation of results provide investigative authorities with highly evidentiary information. This requires reference databases of the studied DNA marker frequencies. Analysis of 26 species-specific microsatellite markers was carried out to establish the most informative markers for the DNA identification of biological samples of the European wild boar (*Sus scrofa*), inhabiting the territory of the Republic of Belarus. Using the obtained data on the genetic polymorphism of wild (n=719) and domestic (n=304) species representatives, the selection of 16 STR loci was performed. The selected loci allow to verify the animal origin from a concrete individual.

Key words: polymorphism, microsatellites, wild boar, domestic pig, identification.

Дата поступления статьи: 28 марта 2018 г.

М.Е. Михайлова, Е.Л. Романишко, А.И. Киреева

ДЕТЕКЦИЯ ПОЛИМОРФИЗМА RS456206907 ГЕНА *SMC2*, ДЕТЕРМИНИРУЮЩЕГО ГАПЛОТИП ФЕРТИЛЬНОСТИ ННЗ У КРУПНОГО РОГАТОГО СКОТА

Институт генетики и цитологии НАН Беларуси
Республика Беларусь, 220072, г. Минск, ул. Академическая, 27
e-mail: M.Mikhailova@igc.by

Интенсивная селекция, направленная на увеличение молочной продуктивности, привела к появлению у коров голштинской породы проблем, связанных со снижением их репродуктивной способности. Изучение генетической структуры популяции крупного рогатого скота голштинизированной породы по гену *SMC2*, детерминирующему фертильность крупного рогатого скота, актуально для развития племенного животноводства Республики Беларусь. Негативное влияние данного рецессивного генетического дефекта приводит к эмбриональной смертности или гибели теленка в ранний постэмбриональный период. Данный наследственный дефект относится к гаплотипу фертильности ННЗ. Частота животных-носителей гаплотипа фертильности ННЗ (ННЗС), обусловленного мутацией в гене *SMC2*, составляет 3,46%. Во избежание распространения мутации необходимо исключать из селекционного процесса животных-носителей этого дефекта, а также строго контролировать качество ввозимой племенной продукции (материала) в Республику Беларусь.

Ключевые слова: крупный рогатый скот, полиморфизм гена *SMC2*, гаплотип фертильности ННЗ.

Введение

В настоящее время голштинская порода крупного рогатого скота является самой распространенной породой молочного направления в мире, которую, согласно данным FAO (Food and Agriculture Organization), разводят по крайней мере в 161 стране [1]. В Республике Беларусь на долю молочного скота голштинской и голштинизированной черно-пестрой пород приходится основная часть поголовья. В связи с интенсивной селекцией, направленной на увеличение молочной продуктивности, у коров голштинской породы появились проблемы, связанные со снижением их репродуктивной способности [2]. До недавнего времени снижение воспроизводительной способности объясняли, главным образом, послеродовыми проблемами клинического характера и развитием метаболического стресса, возникающего во время лактации. В настоящее время считается, что, по крайней мере, половина такого снижения детерминирована генетическими факторами [3].

Исследование геномов, с использованием NGS-технологий (new generation sequencing), показало, что усредненный геном несет 250-300 вариантов последовательностей с наруше-

нием физиологических функций (LoF, loss-of-function). Открытие целого ряда LoF-мутаций стало возможным благодаря разработке и применению нового подхода, так называемого картирования гомозиготности. Возрастание негативного влияния LoF-мутаций на фертильность коров связывают с поступательным ростом гомозиготности в породах, поскольку в гомозиготном состоянии такие мутации могут быть летальными, приводя к эмбриональной гибели. Так, средний коэффициент инбридинга у голштинского скота в США вырос с 0,22% в 1965 году до 6,53% в 2015 году [4]. Причинами возрастания степени инбридинга стали относительно низкое исходное генетическое разнообразие в большинстве молочных пород как следствие их происхождения от ограниченного числа родоначальников; интенсивное использование для искусственного осеменения относительно небольшого числа выдающихся быков-производителей, практикуемое в течение более 50 лет; жесткая селекция по ограниченному числу признаков. Некоторые голштинские быки стали отцами нескольких тысяч сыновей, оцененных по качеству потомства, и нескольких сотен тысяч и даже более миллиона

лактующих дочерей. Кроме того, из приблизительно 5000 быков, оцениваемых ежегодно по качеству потомства, почти 50% — потомки 10 наиболее популярных быков. В этой связи диагностика LoF-мутаций, ассоциированных с летальными наследственными заболеваниями, становится одним из ключевых элементов в системе генетического мониторинга популяций сельскохозяйственных животных.

Число LoF-мутаций, обуславливающих наследственные аномалии и вызывающих эмбриональную и раннюю постэмбриональную смертность, в молочных породах крупного рогатого скота до недавнего времени ограничивалось единицами. Ранее в голштинской породе КРС были идентифицированы LoF-мутации, приводящие к наследственным генетическим дефектам: дефицит уридин-монофосфатсинтазы (DUMPS), дефицит лейкоцитарной адгезии (BLAD), комплексный порок позвоночника (CVM), брахиспинальный синдром (BY). В настоящее время, наряду с традиционной аббревиатурой, они известны как гаплотипы фертильности (HND, HNB, HNC, HH0, соответственно). Использование метода полногеномного секвенирования позволило выявить новые мутации, ассоциированные с гаплотипами фертильности голштинского скота: HH1 (*APAF1*, C→T, Q579X), HH3 (*SMC2*, T→C, F1135S), HH4 (*GART*, A→C, N290T), HH5 (*TFB1M*, 138kb Del), HCD (*APOB*, 1,3 kb Ins), являющиеся причиной эмбриональной смертности.

На данный момент в голштинской породе регистрируется 10 гаплотипов фертильности (HCD, HH0, HH1, HH2, HH3, HH4, HH5, HNB, HNC, HND), оказывающих влияние на процент успешных осеменений (с наступлением стельности) и (или) ассоциированных с эмбриональной и ранней постэмбриональной смертностью на различных стадиях и встречающихся с частотой от 0,01 до 2,95%. Негативное влияние этих гаплотипов, оцененное в популяции североамериканских голштинов, проявляется в снижении частоты осеменений, завершившихся стельностью на 1,0–9,9%. Распространению гаплотипов фертильности способствует то, что их носители — выдающиеся быки-производители, интенсивно используемые для искусственного осеменения в странах, занимающихся молочным скотоводством, в том числе и в Республике Беларусь.

Материалы и методы

В качестве биологического материала для выделения ДНК использовали сперму племенных быков-производителей и образцы цельной крови или хрящевой ткани (ушной выщип) коров. Выделение и очистку ДНК из образцов биологического материала проводили с использованием коммерческих наборов реагентов для выделения ДНК «Нуклеосорб В», «Нуклеосорб С» («Праймтех», Беларусь) и NucleoSpin Tissue («Macherey-Nagel», Германия) согласно прилагаемым инструкциям по применению.

Концентрацию выделенной ДНК определяли с помощью флуориметра Qubit 2.0 («Invitrogen», Австрия) с использованием набора реагентов Molecular probes Qubit dsDNA BR Assay kit («Life technologies», США). Разведение ДНК до рабочей концентрации проводили стерильной деионизированной водой (MQ-H₂O) в зависимости от используемого молекулярно-генетического метода.

Полиморфные варианты гена *SMC2* выявляли методом ПЦР в режиме «реального времени» (ПЦР-РВ). Для постановки ПЦР-РВ использовали термоциклер CFX 96 Real-Time PCR Detection System («Bio-Rad», США). Анализ результатов проводили в программе Bio-Rad CFX Manager (Version 3.1). Выявление полиморфизма гена *SMC2* проводили с помощью аллель-специфических праймеров. Реакционная смесь для ПЦР-РВ содержала 2 x Maxima Sybr Green PCR Mix («Thermo Scientific», EU), по 300 нМ каждого праймера, деионизированную воду и 5–30 нг геномной ДНК исследуемого образца. Программа амплификации: 95 °C — 10 мин; 95 °C — 10 с; 58–60 °C — 20 с; 62 °C — 30 с (50 циклов). Учет результатов проводили при снятии показаний флуоресценции на стадии отжига при 58–60 °C.

Секвенирование по Сэнгеру исследуемых локусов гена *SMC2* (аллели дикого типа и мутантные аллели) осуществляли на приборе 3500 Genetic Analyzer («Applied Biosystems», США). ПЦР-фрагменты экстрагировали из агарозного геля, очищали с помощью набора DNA Gel Extraction Kit («Thermo scientific», EU). Для постановки секвенирующей ПЦР использовали Big Dye Terminator v3.1 Cycle Sequencing Kit («Applied Biosystems», США).

и обратные праймеры для амплификации фрагментов гена *SMC2*. ПЦР проводили по следующей программе: 95 °C 5 мин; 50 циклов — 95 °C 30 с, 55 °C 10 с, 60 °C 40 мин; 12 °C 5 мин. Полученные амплификаты очищали от непрореагировавших ddNTP. Анализ результатов проводили с использованием программы Sequencing Analysis («Applied Biosystems», США).

Результаты и обсуждение

Гаплотип фертильности HH3 (OMIA: 001824-9913) локализован на хромосоме 8 в позиции с 92,485,682 до 96,594,716 (74 маркера) и рассчитан на основе координат генома (UMD 3.1.) [5].

Установлено, что причиной эмбриональной смертности телят является несинонимная однонуклеотидная замена — SNP (Т / С) в экзоне 24 гена структурной поддержки хромосом (*SMC2*) в положении 95 410 507 (UMD3.1), (rs456206907). Данный полиморфизм является причиной замен как аминокислоты 1135 от фенилаланина до серина (Phe1135Ser) (GenBank #: XP_002689921.2), так и кодируемого белка в домене Р-loop-нуклеозидтри-фосфатгидролазы (НТФазы). Белок *SMC2* играет важную роль в процессах репарации ДНК, конденсации хромосом и их сегрегации в процессе клеточного деления. [5, 6–10].

Изучение происхождения выявленных скрытых животных-носителей мутации гена *SMC2* показало, что наиболее ранними известными родоначальниками мутации являются известные быки-производители американской селекции Glendell ARLINDA CHIEF (1968 г. р.), USA 1556373 и Gay View SKYLINER (1954 г. р.) USA 1244845. Анализ результатов, проведенный в североамериканской и французской популяциях голштинов у осемененных коров, отцы которых являются скрытыми носителями гаплотипа фертильности HH3, с быками — скрытыми носителями аналогичного гаплотипа, показал снижение степени стельности, соответственно, на 3,2 и 5,4–5,5% [11, 12]. Показано, что частота встречаемости скрытых носителей HH3 среди быков-производителей составляет 3,0–4,7% в США и 5,0% во Франции [13, 14]. Проведенный российскими учеными анализ родословных быков-произво-

дителей (n=560) в племенных предприятиях России показал, что у тридцати быков-производителей в поколениях F1-F2 предков встречаются носители мутантного аллеля в гене *SMC2*. Частота встречаемости животных-носителей гаплотипа фертильности HH3 российских племпредприятий варьирует от 2,7 до 8,4%, в среднем составляет 5,4%. Скрытые носители принадлежат двум широко распространенным генеалогическим линиям — Рефлекшн Соверинга и Уэс Идеала [1, 15].

В результате проведенного по племенным картам анализа родословных быков, используемых для осеменения в сельхозпредприятиях Республики Беларусь, и согласно базе данных Holstein Association USA [32, 33], были определены вероятные носители мутантного аллеля гена *SMC2*, детерминирующего гаплотип фертильности HH3. Секвенированием был выявлен носитель мутантного аллеля — бык Джудо, №100552. Образец ДНК, выделенный из спермы этого быка, использовали в качестве положительного контрольного образца для детекции полиморфизма rs456206907 в гене *SMC2* с помощью метода аллель-специфической ПЦР в режиме реального времени. Каждый образец амплифицировался с 2-мя парами праймеров. В каждой паре один из праймеров был общим, а второй комплементарен мутантному/дикому аллелям гена *SMC2*. Визуализация накопления ПЦР-продуктов происходила в режиме реального времени в присутствии интеркалирующего красителя Sybr Green I. Для интерпретации результатов ПЦР-РВ значение ΔC_t тестируемых образцов сравнивали со значением ΔC_t контрольного образца, содержащего два типа аллелей в гетерозиготе и отрицательного контрольного образца.

Специфичность полученных ПЦР-продуктов была подтверждена секвенированием по Сэнгеру (рисунок).

Нами проведено тестирование быков-производителей и маточное поголовье в отобранной выборке (n=433) для выявления животных-носителей мутантного аллеля гена *SMC2*, ассоциированного с гаплотипом фертильности HH3.

Выявлено 15 особей-носителей мутации *SMC2*, из них 4 быка производителя: Юрген/600411, Джудо/100552, Мотто/400595, Блик/200459; особей маточного поголовья — 11 голов.

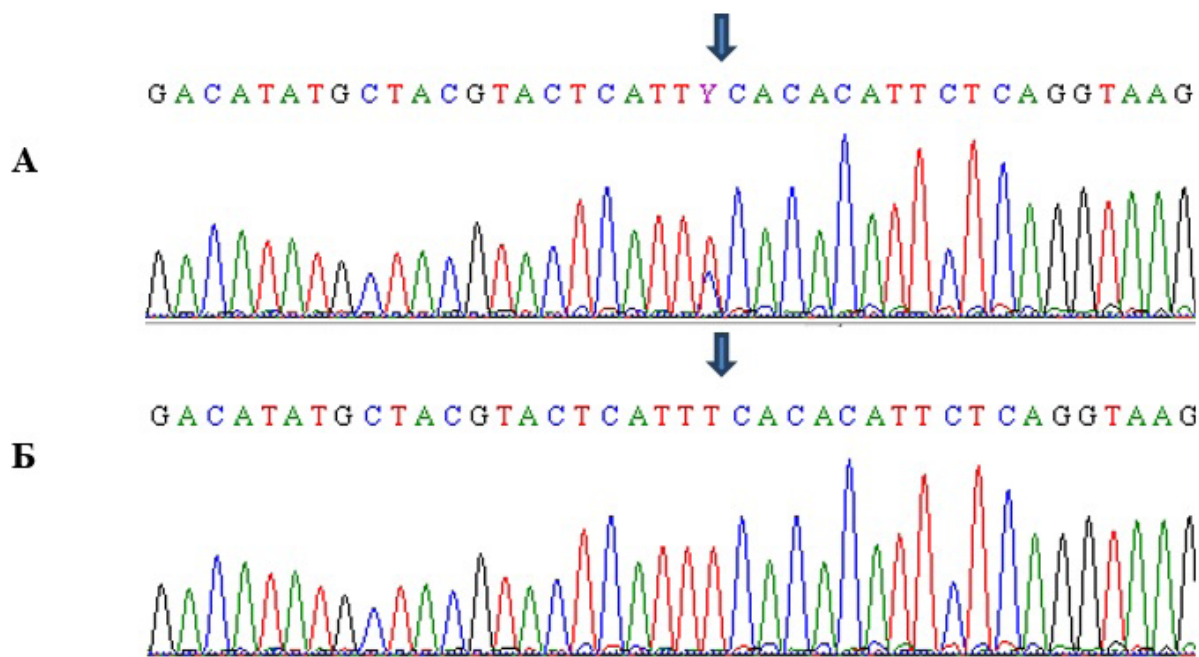


Рис. Секвенирование по Сенгеру фрагмента гена *SMC2*: А — генотип ТС (Y), животное-носитель мутантного аллеля; Б — генотип ТТ, здоровое животное

Таким образом, в результате проведенного скрининга популяции крупного рогатого скота выявлены животные-носители мутации в гене *SMC2*, ассоциированной с пониженной плодовитостью КРС (гаплотип фертильности ННЗ). Частота животных-носителей мутации в гене *SMC2* составляет 3,46%.

Заключение

Носительство гаплотипа ННЗ, связанного с пониженной фертильностью у быков-производителей, сперма которых была ввезена из-за рубежа — Юрген/600411, Джудо/100552, Мотто/400595, Блик/200459, подтверждает необходимость строгого контроля качества ввозимой племенной продукции (материала).

К маточному поголовью с установленным носительством гаплотипов во избежание скрытых аборт и эмбриональной смертности гомозиготного потомства следует подбирать только проверенных на генетические аномалии производителей.

Предложено исключить из селекционного процесса быков-производителей — носителей гаплотипа фертильности ННЗ и строго контролировать качество ввозимой племенной продукции (материала) для избежания распространения в селекционном стаде мутаций, отрицательно влияющих на фертиль-

ность, учитывая, что скрытые носители гаплотипа фертильности ННЗ принадлежат двум широко распространенным генеалогическим линиям — Рефлекшн Соверинга и Уэс Идеала.

Список использованных источников

1. Зиновьева, Н.А. Гаплотипы фертильности голштинского скота // Сельскохозяйственная биология. — 2016. — Т. 51. — С. 423–435.
2. Barbat, A. Female fertility in French dairy breeds: current situation and strategies for improvement / Barbat A., Le Mezec P., Ducrocq V. // J. Reprod. Dev. — 2010. — Vol. 56. — P. 15–21.
3. Dobson, H. The high producing dairy cow and its reproductive performance / Dobson H., Smith R.F., Royal M.D. Reprod. Domest. // Anim. — 2007. — Vol. 42(Suppl. 2). — P. 17–23.
4. Council on dairy cattle breeding. Inbreeding trend for holstein or red & white cows. — Mode of access: <https://www.cdcb.us/eval/summary/inbrd.cfm>. — Date of access: 27.02.2016.
5. Van Raden, P.M. Harmful recessive effects on fertility detected by absence of homozygous haplotypes // Van Raden P.M., Olson K.M., Null D.J. // J. Dairy Sci. — 2011. — Vol. 94. — P. 6153–6161.
6. Hayes, B. 1000 Bull Genomes Consortium. The 1000 Bull Genomes project — toward genomic selection from whole genome sequence

data in dairy and beef cattle / Hayes B., Daetwyler H.D., Fries R. // Proc. Plant and Animal Genome XXI Conf. San Diego. – 2013. – P. 150.

7. Schmiesing, J.A. Identification of two distinct human SMC protein complexes involved in mitotic chromosome dynamics / Schmiesing J.A., Ball A.R. Jr., Gregson H.C. // PNAS USA. – 1998. – Vol. 95(22). – 12906–12911.

8. Pausch, H. Homozygous haplotype deficiency reveals deleterious mutations compromising reproductive and rearing success in cattle / Pausch H., Schwarzenbacher H., Burgstaller // J. BMC Genomics. – 2015. – Vol. 16. – P. 312.

9. Mc Clure, Bovine exome sequence analysis and targeted SNP genotyping of recessive fertility defects BH1, HH2 and HH3 reveal causative mutation in SMC2 for HH3/ Mc Clure, M.C., Bickhart D., Null D. // PLoS ONE. – 2014. – Vol. 9. – P. 125–130.

10. Daetwyler, H.D. Whole-genome sequencing of 234 bulls facilitates mapping of monogenic and complex traits in cattle / Daetwyler H.D., Capitan A., Pausch H. // Nature Genet. – 2014. – Vol. 46. – P. 858–865.

11. Washburn, S.P. Trends in reproductive performance in Southeastern Holstein and Jersey DHI herds / Washburn S.P., Silvia W.J., Brown C.H. // J. Dairy Sci. – 2002. – Vol. 85. – P. 244–251.

12. Lucy, M.C. Reproductive loss in high-producing dairy cattle: where will it end? // J. Dairy Sci. – 2001. – Vol. 84. – 1277–1293.

13. Oltenacu, P.A. The impact of genetic selection for increased milk yield on the welfare of dairy cows / Oltenacu P.A., Broom D.M. // Animal Welfare. – 2010. – Vol. 19. – P. 39–49.

14. Durbin, R.M. The 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing // Nature. – 2010. – Vol. 467. – P. 1061–1073.

15. Романенкова, О.С. Разработка тест-системы для диагностики гаплотипа фертильности крупного рогатого скота HH3, ассоциированного с ранней эмбриональной смертностью / О.С. Романенкова, Е.А. Гладырь, О.В. Костюнина, Н.А. Зиновьева // Достижения науки и техники АПК. – 2015. – Т. 29, № 11. – С. 91–94.

M.E. Mikhailova, E.L. Romanishko, A.I. Kireeva

DETECTION OF RS456206907 POLYMORPHISM OF *SMC2* GENE DETERMINING HH3 FERTILITY HAPLOTYPE IN CATTLE

Institute of Genetics and Cytology, NASB
Minsk, 220072, the Republic of Belarus

Intensive breeding aimed at an increase in milk productivity has led to the problems associated with a decreased reproductive capacity in Holstein cows. Study of a genetic structure of the cattle population of holsteinized breed by the *SMC2* gene that determines the cattle fertility is relevant for the development of livestock breeding in the Republic of Belarus. A negative effect of this recessive genetic defect in the homozygous state leads to embryonic death or calf death in the early postembryonic period. This hereditary anomaly refers to the HH3 fertility haplotype. Carriers of the HH3 fertility haplotype have been identified. The frequency of the HH3 fertility haplotype in animal carriers (HH3C) determined by the *SMC2* gene mutation is 3,46%. In order to avoid the spread of this mutation, it is necessary to exclude the carriers of such mutation from the breeding process and strictly control the quality of imported pedigree products (material) to the Republic of Belarus.

Key words: cattle, *SMC2* gene polymorphism, HH3 fertility haplotype.

Дата поступления статьи: 13 сентября 2018 г.

Н.В. Воронова, Ю.В. Бондаренко, С.С. Левыкина, Р.С. Шулинский

**МИТОХОНДРИАЛЬНЫЙ ГЕНОМ *APHIS FABAE MORDVILKOEI*
BORNER & JANISCH, 1992**

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

Результаты секвенирования, аннотации и анализа митохондриального генома *Aphis fabae mordvilkoii* приведены в настоящей работе. Показано, что мтДНК *A. fabae mordvilkoii* имеет типичные для тлей структуру и размер и не несет перестроек, обнаруженных в других известных митохондриальных геномах тлей рода *Aphis* L.

Ключевые слова: *Aphis fabae mordvilkoii*, тли, митохондриальный геном.

Введение

Тли (Aphididae), представляющие собой одно из семейств грудехоботных насекомых (Hemiptera: Sternorrhyncha), это сравнительно небольшая группа фитофагов, объединяющая около 5000 рецентных видов в 24 подсемействах [1–3]. Хорошо известной особенностью этого таксона являются облигатный циклический партеногенез и полифения, т. е. существование нескольких морфологически различных форм внутри одного годичного цикла. Благодаря живорождению и партеногенезу телескопического типа, тли способны к взрывному увеличению численности, что, вкупе с активным расселением и миграцией на другие растения, делает их крайне опасными вредителями сельского хозяйства [4–6].

Все тли, за исключением 35 видов относимых к полифагам, способны питаться и размножаться на строго ограниченном перечне видов растений, т. е. являются более или менее строгими специалистами [7]. Полифаги населяют растения из десятков (обычно не менее 40) ботанических семейств [8], среди которых, как правило, множество сельскохозяйственных культур, пряно-ароматических и лекарственных растений, причем способность полифагов метаболизировать колоссальное разнообразие растительных алкалоидов и иных защитных веществ привлекает все большее внимание исследователей во всем мире [9]. Однако, несмотря на тщательное изучение, феномен возникновения полифагии у тлей, который, по сути, представляет собой эволюционную реверсию,

до сих пор вызывает множество вопросов, и механизм этого явления остается не полностью понятным [10–12].

Из 35 видов, относимых к полифагам, 33 относятся к одному из наиболее молодых и прогрессивных подсемейств тлей — Aphidinae. *Aphis fabae* s.l., свекловичная или, иначе, черная бобовая тля, это один из видов-полифагов, распространенный повсеместно, питающийся на калине, жасмине и бересклете (первичное кормовое растение) и множестве как сорных травянистых, так и возделываемых растений, включая свеклу, бобовые, подсолнечник, картофель и томаты, морковь и многие другие важные сельскохозяйственные культуры [13, 14].

В последнее время *Aphis fabae* s.l. разделяют на 6 подвигов (*A. fabae cirsiacanthoidis*, *A. fabae eryngii*, *A. fabae evonymi*, *A. fabae fabae*, *A. fabae mordvilkoii* и *A. fabae solanella*), как это указано информационным ресурсом «Fauna Europaea» [15]. Все подвиды морфологически идентичны, и их дифференциация возможна только по ассоциации с конкретными кормовыми растениями. *A. fabae mordvilkoii* Börner & Janisch, 1922, в частности, использует в качестве первичного растения бересклет (*Euonymus europaeus*) и различные виды чубушника (*Philadelphus* sp.) и широко распространен в Беларуси, где эти растения часто высаживают в городах в качестве орнаментальной растительности. Как вторичное растение *A. fabae mordvilkoii* чаще всего использует лопухи, которые, как известно, относятся к лекарственным растениям, поскольку

их вегетативные части содержат, кроме прочего, много разнообразных эфирных масел [16].

A. fabae s.l., как и другие полифаги, переносит множество вирусов растений как с одних культур на другие, так и от сорных растений — на возделываемые. Несмотря на питание этих тлей на культурных растениях, где они неизбежно подвергаются обработкам, случаи появления устойчивых к инсектицидам рас *A. fabae* s.l. в научной литературе до сих пор описаны не были. С нашей точки зрения, изучение тлей этого вида как самостоятельно, так и в сравнении с полифагами, способными к формированию устойчивости к инсектицидам, может быть крайне интересно с точки зрения понимания эволюционных механизмов формирования полифагии и любой другой биохимической резистентности, в том числе резистентности к синтетическим инсектицидам. В этой работе мы представляем первые результаты в области сравнительной геномики *A. fabae* s.l. — результаты сравнительного анализа митохондриальных геномов тлей, что особенно интересно, учитывая, что митохондриальные гены не связаны с адаптацией животных к питанию и, следовательно, могут предоставить исследователям информацию об особенностях молекулярной эволюции вида в общем, позволяя обнаружить какие-либо различия в скорости мутации ДНК и/или особенностях действующего на вид отбора (разумеется, если такие различия имеются). В рамках этой работы мы поставили перед собой задачу оценить изменчивость митохондриальных геномов тлей, используя для анализа все доступные геномы митохондрий тлей рода *Aphis* L., опубликованные разными исследователями в предыдущие годы, и геном *A. fabae mordvilkoii*, коллектированной в Беларуси.

В последние годы полные митохондриальные геномы насекомых все чаще используются в работах, связанных с изучением их филогении и уточнением систематики [17]. В качестве филогенетических маркеров могут использоваться структура генома, т. е. такие признаки, как число и очередность расположения генов, размер и строение регуляторных областей, а также нуклеотидные и аминокислотные последовательности белок-кодирующих генов и вторичная структура транспортных и рибосомальных РНК [18–21]. Показано, что эти при-

знаки обладают различной степенью консервативности и могут использоваться для решения разнообразных исследовательских задач во многих таксонах насекомых.

В общем случае митохондриальный геном насекомых представлен кольцевой двуцепочечной молекулой ДНК длиной 12–20 Кб и содержит 13 белок-кодирующих генов (никогда не имеющих интронов), 22 гена тРНК и 2 гена рРНК (малой и большой субъединицы митохондриальных рибосом) [22, 23]. Кроме этого, в митохондриальном геноме насекомых присутствует богатая аденином и тиминном регуляторная область, часто содержащая многочисленные короткие повторы. Любопытной особенностью митохондриальных геномов насекомых является то, что гены в них могут частично перекрываться, т. е. одни и те же нуклеотиды могут одновременно входить в терминирующий кодон вышележащего гена и в старт-кодон гена, располагающегося в геноме ниже [24, 25].

На момент написания этой работы в мире было опубликовано только 3 митохондриальных генома тлей рода *Aphis* L.: два генома, принадлежащих *A. craccivora*, один — из Европы, а второй — из США, и геном *A. gossypii*, причем оба этих вида так же, как и *A. fabae* s.l., относятся к полифагам [26–28].

Материалы и методы

Колония *A. fabae mordvilkoii* была коллектирована в г. Минск (Беларусь) с *Philadelphus* sp. Принадлежность тлей к указанному виду и подвиду была подтверждена ДНК-штрих-кодированием (сравнение последовательности гена COI с последовательностями, депонированными в BOLD). Тотальную ДНК выделили с использованием набора Blood-Animal-Plant DNA Preparation Kit (Jena Bioscience). Полногеномное секвенирование провели по технологии IonTorrent.

Извлечение прочтений митохондриальной ДНК из общего пула прочтений проводили путем последовательной фильтрации и выравнивания данных полногеномного секвенирования *A. fabae mordvilkoii* по ядерным и митохондриальным геномам близкородственных организмов (*A. glycines* и *A. craccivora*). В частности, локальное выравнивание общего пула прочтений по ядерно-

му геному *A. glycines* проводили по алгоритму Смита-Ватермана в программе Bowtie2 с использованием параметра «very-sensitive-local» (высококочувствительное локальное выравнивание). Фильтрацию не выравненных прочтений из карты выравнивания последовательностей (SAM) проводили с использованием программного пакета Samtools версии 0.1.19 при помощи опции sort. Преобразование форматов данных провели в программе bedtools версии 2.25.0. Далее провели выравнивание полученных прочтений вышеуказанным способом по митохондриальному геному *A. craccivora* и отбор выравненных прочтений. Полученный набор последовательностей использовали в дальнейшей процедуре *de novo* сборки в программе MITObim 1.9, которая использует алгоритмы сборки ассемблера MIRA. Данный ассемблер проводит сборку вокруг высоко консервативных участков митохондриального генома, в нашем случае были использованы гены COI и *cytB*. Нуклеотидные последовательности этих генов были получены из международной базы данных нуклеотидных последовательностей GenBank NCBI (COI [FJ965710.1] и *cytB* [GU205347.1]). Параметры сборки для полученного пула прочтений оптимизировали следующим образом: значение *k*-меры варьировали от 11 до 41 с шагом в 2 единицы. Оптимальное значение *k*-меры выбрали по основным статистическим показателям, таким как максимальная длина контиги, количество контиг, N50. Перекрытие интервалов контиг провели в программе MEGA7.

Аннотацию митохондриального генома *A. fabae mordvilkoii* провели по гомологии с ре-

ференсными геномами, находящимися в свободном доступе в GenBank NCBI, информация о которых приведена в табл. 1.

Карту митохондриального генома *A. fabae mordvilkoii* построили в программе SnapGene с соблюдением пропорций размера генов по отношению друг к другу (GSL Biotech; на сайте snapgene.com). Моделирование вторичной структуры тРНК провели в онлайн сервисе ARWEN [29]. Нуклеотидную композицию рассчитали в программе MEGA7 [30].

Последовательность полного митохондриального генома *A. fabae mordvilkoii* была депонирована в GenBank NCBI (код доступа MG897128).

Результаты и обсуждение

Полный митохондриальный геном *A. fabae mordvilkoii* имеет размер 15,346 Kb. Как это часто отмечается для беспозвоночных животных, содержание аденина и тимина в митохондриальной ДНК *A. fabae mordvilkoii* в сумме значительно превышает ожидаемые 50% и достигает 83,6%. Такая диспропорция содержания нуклеотидов в митохондриальных геномах обычно считается следствием агрессивности среды в матриксе митохондрий [31, 32]. У *A. fabae mordvilkoii* пропорция (A+T)/(G+C) > 80/20 сохраняется в самых разных в функциональном отношении областях генома (табл. 2).

В геноме *A. fabae mordvilkoii* обнаруживаются 13 белок-кодирующих генов и 24 РНК-кодирующих, из которых 2 кодируют РНК большой и малой субъединиц митохондриальных рибосом и 22 представляют собой гены транспортных РНК (табл. 3).

Таблица 1

Митохондриальные геномы тлей рода *Aphis* L., использованные для аннотации

Вид тли	Размер генома, п. н.	Число генов	Длина, п. н.		Код доступа в GenBank
			D-петля	Некодирующая область	
<i>A. craccivora</i>	15 305	37	557	—	KX 447142.1
<i>A. craccivora</i>	15 308	37	557	—	NC 031387.1
<i>A. gossypii</i>	15 869	37	627	784	NC 024581.1

Таблица 2

Нуклеотидный состав митохондриального генома *Aphis fabae mordvilko*

Характеристика	Процентное содержание нуклеотидов					Длина, п. н.
	Т	С	А	Г	А+Т	
Полный геном	38,6	10,3	45,0	6,2	83,6	15346
Белок-кодирующие гены (без учета цепи)	38,1	10,8	44,9	6,2	83,0	10947
Белок-кодирующие гены, (+)-цепь	43,3	11,4	38,5	6,9	81,8	6735
Белок-кодирующие гены, (-)-цепь	29,9	9,9	55,1	5,1	85,0	4212
Гены тРНК	40,4	8,7	44,7	6,2	85,1	1383
Гены рРНК	39,0	10,4	45,8	4,9	84,8	2043
Регион, несущий повторы	38,6	7,6	49,7	4,1	88,3	197
D-петля	41,0	5,6	43,6	9,8	84,6	661

Таблица 3

Характеристика митохондриального генома *Aphis fabae mordvilko*

Название области	Продукт/функциональное назначение	Координаты в геноме	Длина, п. н.	Кодирующая цепь
<i>Cox1</i>	Субъединица I цитохром с оксидазы	1–1536	1536	+
L	тРНК лейцина	1531–1602	72	+
<i>Cox2</i>	Субъединица II цитохром с оксидазы	1603–2274	672	+
K	тРНК лизина	2289–2362	74	+
D	тРНК аспарагиновой кислоты	2362–2424	63	+
<i>Atp8</i>	Субъединица VIII АТФ-синтазы F0	2425–2583	159	+
<i>Atp6</i>	Субъединица VI АТФ-синтазы F0	2564–3217	654	+
<i>Cox3</i>	Субъединица III цитохром с оксидазы	3217–4002	786	+
G	тРНК глицина	4002–4064	63	+
Nad3	Субъединица III НАДН-дегидрогеназы	4062–4418	357	+
A	тРНК аланина	4419–4482	64	+

Продолжение табл. 3

Название области	Продукт/функциональное назначение	Координаты в геноме	Длина, п. н.	Кодирующая цепь
R	тРНК аргинина	4482–4547	66	+
N	тРНК аспарагина	4545–4610	66	+
S	тРНК серина	4610–4672	63	+
E	тРНК глутаминовой кислоты	4675–4738	64	+
<i>Rep_region</i>	Область повторов	4733–4929	197	+
F	тРНК фенилаланина	4918–4981	64	–
<i>Nad5</i>	Субъединица V НАДН-дегидрогеназы	4982–6652	1671	–
H	тРНК гистидина	6700–6766*	67	–
<i>Nad4</i>	Субъединица IV НАДН-дегидрогеназы	6764–8086	1325	–
<i>Nad4L</i>	Субъединица IV L НАДН-дегидрогеназы	8080–8370	291	–
T	тРНК треонина	8372–8432	61	+
P	тРНК пролина	8435–8502	68	–
<i>Nad6</i>	Субъединица VI НАДН-дегидрогеназы	8494–8997	504	+
<i>CytB</i>	Цитохром B	9000–10115	1116	+
S	тРНК серина	10122–10192	71	+
<i>Nad1</i>	Субъединица I НАДН-дегидрогеназы	10197–11132	936	–
L	тРНК лейцина	11133–11198	66	–
<i>RrnL</i>	16S рРНК	11199–12460	1262	–
V	тРНК валина	12459–12520	60	–
<i>RrnS</i>	12S рРНК	12532–13312	781	–
<i>D-loop</i>	Область D-петли	13329–13989	661	+
I	тРНК изолейцина	13984–14046	63	+
Q	тРНК глутамин	14043–14109	67	–

Окончание табл. 3

Название области	Продукт/функциональное назначение	Координаты в геноме	Длина, п. н.	Кодирующая цепь
М	тРНК метионина	14115–14180	66	+
<i>Nad2</i>	Субъединица II НАДН-дегидрогеназы	14181–15158	978	+
W	тРНК триптофана	15157–15218	62	+
C	тРНК цистеина	15210–15278	69	+
Y	тРНК тирозина	15279–15346	68	–

Примечание. * — перед геном располагается наиболее длинная некодирующая область (не считая региона, несущего повторы, и D-петли) длиной 49 нуклеотидов

Белок-кодирующие гены в митохондриальном геноме *A. fabae mordvilkoii* в основном имеют типичные для членистоногих старт- и стоп-кодоны (ATN и TAA соответственно). Только в гене *cytB* роль стоп-кодона выполняет кодон ATA, а в гене *Nad4* стоп-кодон отсутствует, как это часто отмечается в митохондриальных геномах других тлей. 17 генов в мтДНК *A. fabae mordvilkoii* перекрываются, при этом чаще всего они имеют 1 общий нуклеотид (5 случаев), по одному разу встреча-

ются перекрытия на 4 и 7 нуклеотидов. Наибольшее перекрытие обнаруживается между генами *Atp8* и *Atp6* и составляет 19 нуклеотидов. Между некоторыми генами есть короткие некодирующие вставки. Исключение составляют гены *Nad5* и тРНК аминокислоты гистидин, вставка между которыми составляет 48 нуклеотидов (см. табл. 3).

Порядок генов в митохондриальном геноме *A. fabae mordvilkoii* оказался в целом типичным для тлей (рис. 1).

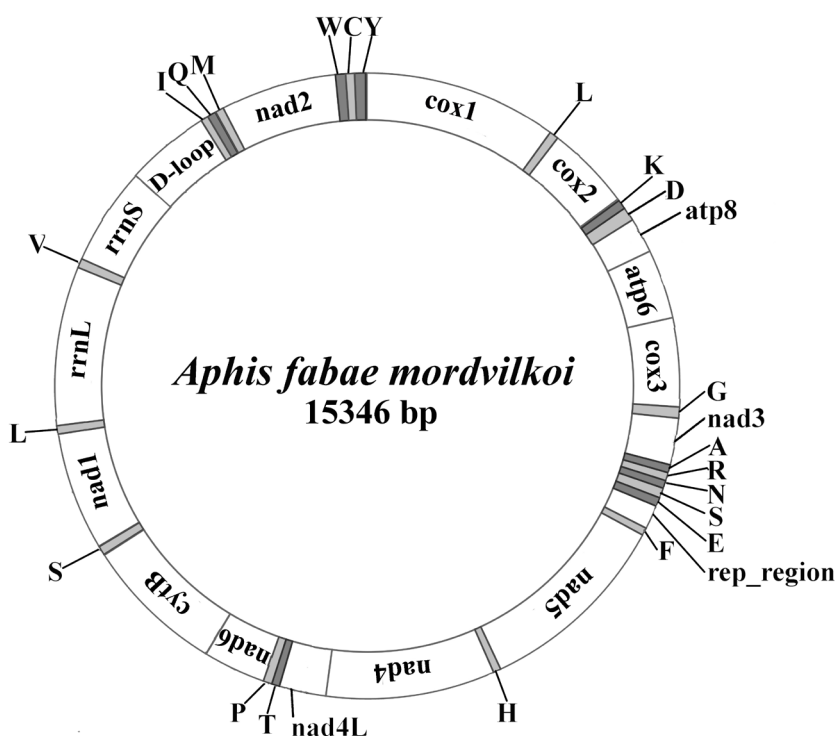


Рис. 1. Карта митохондриального генома *Aphis fabae mordvilkoii*

Митогеном *A. fabae mordvilko* несет все гены тРНК, наличие которых типично для митохондриальных геномов насекомых. Первичная и вторичная структура тРНК также типичны, в том числе обычным является отклонение структуры одной из двух тРНК серина (tRNA-Ser с антикодоном «gtc») от типичной формы «клеверного листа» (рис. 2).



Ситуация с регионом, содержащим повторы или, так называемым, «repeat-region», у трех сравниваемых видов рода *Aphis* значительно различается. В обоих митохондриальных геномах *A. craccivora* эта область отсутствует, что отмечается и у других видов тлей. В митохондриальном геноме *A. gossypii* эта область присутствует и является довольно протяженной — 784 п. н. У *A. fabae mordvilkoii* repeat-region, напротив, необычно короткий — только 197 п. н. Следует отметить, что repeat-region в митохондриальном геноме насекомых встречается сравнительно нечасто, в связи с чем многие исследователи считают его типичным признаком мтДНК тлей [33]. Тем не менее, как уже было сказано, не все виды тлей несут в своих митохондриальных геномах эту область с до сих пор не установленной функцией [34]. Интересным также является то, что в литературе этот участок обычно описывают как АТ-богатый [35, 36]. Однако содержание А+Т у *A. fabae mordvilkoii* является крайне высоким во всем митохондриальном геноме (86,3%), что не позволяет обозначить repeat-region этих тлей как область со значительным, относительно других областей митохондриального генома, повышением доли аденинов и тиминов, в сравнении с содержанием G+C (содержание А+Т в области, содержащей повторы, равно 88,3%).

Область D-петли у всех трех видов рода *Aphis* является вполне типичной для тлей и у *A. fabae mordvilkoii* состоит из 661 п. н.

Заключение

Секвенирование генома *A. fabae mordvilkoii* фауны Беларуси показало, что митохондриальный геном тлей этого вида является типичным для тлей несмотря на то, что геномы двух других представителей того же рода, которые были расшифрованы, имеют отклонения от общей для тлей структуры. Размеры митохондриального генома трех видов тлей рода *Aphis* несколько различаются за счет варьирования в длине региона, содержащего повторы. В отличие от *A. craccivora*, митохондриальный геном *A. fabae mordvilkoii* имеет область, обычно обозначаемую как «repeat-region», которая, однако, значительно короче, чем аналогичный участок генома у *A. gossypii*.

Список использованных источников

1. Introduction to aphid species file [Electronic resource] / C. Favret, D.C. Eades. — 2009. — Mode of access: <http://Aphid.Species-File.org>. — Date of access: 26.07.2018.
2. Nafria, J.M.N. Update to the Registers of family-group and genus-group taxa of Aphidoidea (Hemiptera, Sternorrhyncha) / J.M.N. Nafria, C. Favret // Boln. Asoc. esp. Ent., — 2014. — Vol. 38, № 1. — P. 1–23.
3. Aphid Species File. Version 5.0/5.0. [Electronic resource] / C. Favret. — 2014. — Mode of access: <http://Aphid.SpeciesFile.org>. — Date of access: 29.07.2018.
4. Impact of the English Grain Aphid, *Sitobion avenae* (F.) (Homoptera: Aphididae), on the Yield of Wheat Plants Subjected to Water Deficits / A. Fereres [et al.] // Environmental Entomology. — 1988. — Vol. 17, № 3. — P. 596–602.
5. Malik, Y.P., Impact of aphid (*Lipahis eyrsimi*) intensity on plant growth and seed characters of Indian mustard. / Y.P. Malik, B. Deen // Indian J. Ent. — 1998. — Vol. 24. — P. 286–287.
6. Carter, C.I. Impact of green spruce aphid on growth: Can a tree forget its past? / C.I. Carter // Forestry Commission Research and Development Paper. — 1977. — Vol. 116. — P. 1–8.
7. Blackman, R.L. Aphids on the world's crops: an identification and information guide. / R.L. Blackman, V.F. Eastop // Wiley, John & Sons, Incorporated. — 476 p.
8. Weber, G. Genetic variability in host plant adaptation of the green peach aphid, *Myzus persicae* / G. Weber // Entomologia Experimentalis et Applicata. — 1985. — Vol. 38, № 1. — P. 49–56.
9. Cardoza Y.J. Phloem Alkaloid Tolerance Allows Feeding on Resistant *Lupinus angustifolius* by the Aphid *Myzus persicae* / Y.J. Cardoza [et al.] // Journal of Chemical Ecology. — 2006. — Vol. 32, № 9. — P. 1965–1976.
10. Jermy, T. Evolution of Insect Host Plant Relationships / T. Jermy // The American Naturalist. — 1984. — Vol. 124, № 5. — P. 609–630.
11. Powell, G. Host plant selection by aphids: behavioral, evolutionary, and applied perspectives / G. Powell, C.R. Tosh, J. Hardie // Annual Review of Entomology. — 2006. — Vol. 51, № 1. — P. 309–330.
12. Evolutionary history of aphid-plant associations and their role in aphid diversification /

- J. Peccoud [et al.] // *Comptes Rendus Biologies*. – 2010. – Vol. 333, № 6–7. – P. 474–487.
13. Berim, M.N. *Aphis fabae Scopoli – Black Bean Aphid*, 2009 / M.N Berim // *Interactive Agricultural Ecological Atlas of Russia and Neighboring Countries*. / Available from: <http://www.agroatlas.ru>.
14. Black bean aphid. [Electronic resource] / Rothamsted Insect Research. Rothamsted Research. / Mode of access: <http://www.agroatlas.ru>. – Date of access: 01.03.2013.
15. Fauna Europaea – all European animal species on the web / Y. de Jong [et al.] // *Biodiversity Data Journal*. – 2014. – Vol. 2. – 4034 p.
16. A review of the pharmacological effects of *Arctium lappa* (burdock) / Y.S. Chan [et al.] // *Inflammopharmacology*. – 2011. – Vol. 19, № 5. – P. 245–254.
17. Huang, X.L. Aphids as models for ecological and evolutionary studies: Aphids as models for ecological and evolutionary studies / X.L. Huang, G.X. Qiao // *Insect Science*. – 2014. – Vol. 21, № 3. – P. 247–250.
18. The mitochondrial genome sequence of the Tasmanian tiger (*Thylacinus cynocephalus*) / W. Miller [et al.] // *Genome Res*. – 2009. – Vol. 19. – P. 213–220.
19. Broughton, R.E. The complete sequence of the zebrafish (*Danio rerio*) mitochondrial genome and evolutionary patterns in vertebrate mitochondrial DNA / R.E. Broughton, J.E. Milam, B.A. Roe // *Genome Res*. – 2001. – Vol. 11. – P. 1958–1967.
20. The complete sequence of the mitochondrial genome of the crustacean *Penaeus monodon*: are malacostracan crustaceans more closely related to insects than to branchiopods? / K. Wilson [et al.] // *Mol. Biol. Evol*. – 2000. – Vol. 17. – P. 863–874.
21. The complete mitochondrial genome of the bag-shelter moth *Ochrogaster lunifer* (Lepidoptera, Notodontidae) / P. Salvato [et al.] // *BMC Genomics*. – 2008. – Vol. 9, № 1. – P. 331.
22. Wolstenholme, D.R. Animal Mitochondrial DNA: Structure and Evolution / D.R. Wolstenholme // *International Review of Cytology*. – Elsevier, 1992. – Vol. 141. – Animal Mitochondrial DNA. – P. 173–216.
23. Boore, J.L. Animal mitochondrial genomes / J.L. Boore // *Nucleic Acids Res*. – 1999. – Vol. 27, № 8. – P. 1767–1780.
24. Comparative and phylogenomic studies on the mitochondrial genomes of Pentatomomorpha (Insecta: Hemiptera: Heteroptera) / J. Hua [et al.] // *BMC Genomics*. – 2008. – Vol. 9. – P. 610.
25. The Complete Mitochondrial Genome of the Damsel Bug *Alloeorhynchus bakeri* (Hemiptera: Nabidae) / H. Li [et al.] // *International Journal of Biological Sciences*. – 2012. – Vol. 8, № 1. – P. 93–107.
26. Complete mitochondrial genome of *Aphis gossypii* Glover (Hemiptera: Aphididae) / S. Zhang [et al.] // *Mitochondrial DNA*. – 2016. – Vol. 27, № 2. – P. 854–855.
27. All 37 Mitochondrial Genes of Aphid *Aphis craccivora* Obtained from Transcriptome Sequencing: Implications for the Evolution of Aphids / N. Song [et al.] // *Plos One*. – 2016. – Vol. 11, № 6. – P. 97–103.
28. Comparison of complete mitochondrial DNA sequences between old and new world strains of the cowpea aphid, *Aphis craccivora* (Hemiptera: Aphididae) / W. Sun [et al.] // *Agri Gene*. – 2017. – Vol. 4. – P. 23–29.
29. Laslett, D. ARWEN: a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences / D. Laslett, B. Canback // *Bioinformatics*. – 2008. – Vol. 24, № 2. – P. 172–175.
30. Kumar, S. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets / S. Kumar, G. Stecher, K. Tamura // *Molecular Biology and Evolution*. – 2016. – Vol. 33, № 7. – P. 1870–1874.
31. Miquel, J. An update on the mitochondrial-DNA mutation hypothesis of cell aging / J. Miquel // *Mutation Research/DNAging*. – 1992. – Vol. 275, № 3–6. – P. 209–216.
32. Cameron, S.L. Insect Mitochondrial Genomics: Implications for Evolution and Phylogeny / S.L. Cameron // *Annual Review of Entomology*. – 2014. – Vol. 59, № 1. – P. 95–117.
33. Wang, Y. Comparative Analysis of Mitochondrial Genomes of Five Aphid Species (Hemiptera: Aphididae) and Phylogenetic Implications / Y. Wang, X.L. Huang, G.X. Qiao // *Plos One*. – 2013. – Vol. 8, № 10. – P. 77511.
34. Ren, Z.M. Complete mitochondrial genome of the North American Rhus gall aphid *Melaphis rhois* (Hemiptera: Aphididae: Eriosomatinae) / Z.M. Ren, J. Wen // *Mitochondrial DNA Part B*. – 2017. – Vol. 2, № 1. – P. 169–170.

35. New Views on Strand Asymmetry in Insect Mitochondrial Genomes / S.J. Wei [et al.] // Plos One. – 2010. – Vol. 5, № 9. – P. 12708.
36. Complete nucleotide sequence and organization of the mitochondrial genome of *Sirthenea flavipes* (Hemiptera: Reduviidae: Peiratinae) and comparison with other assassin bugs / J. Gao [et al.] // Zootaxa. – 2013. – Vol. 3669, № 1. – P. 1.

N.V. Voronova, Y.V. Bandarenka, S.S. Levykina, R.S. Shulinski

**MITOCHONDRIAL GENOME *APHIS FABAE MORDVILKOI*
BORNER & JANISCH, 1992**

Belarusian State University
Minsk, 220030, the Republic of Belarus
e-mail: nvoronova@bsu.by

The results of sequencing, annotation and analysis of the *Aphis fabae mordvilkoii* mitochondrial genome are presented in the Article. It was found that *A. fabae mordvilkoii* mtDNA has a typical aphid's structure and size and does not contain any rearrangements, which other known mitochondrial aphid genomes of the *Aphis* L. genus have.

Key words: *Aphis fabae mordvilkoii*, aphids, mitochondrial genome.

Дата поступления статьи: 11 сентября 2018 г.

Л.В. Можаровская

СТРУКТУРНО-ФУНКЦИОНАЛЬНЫЙ АНАЛИЗ ЛОКУСОВ, КОДИРУЮЩИХ АНТИМИКРОБНЫЕ ПЕПТИДЫ СОСНЫ ОБЫКНОВЕННОЙ

Институт леса НАН Беларуси
Республика Беларусь, 246001, г. Гомель, ул. Пролетарская, 71
e-mail: milamozh@yandex.by

С использованием высокопроизводительного секвенирования транскриптомов проростков сосны обыкновенной идентифицировано четыре, ранее не описанные, экспрессируемые последовательности, детерминирующие антимикробные пептиды (АМП). На основании полученных результатов проведена аннотация и структурно-функциональный анализ выявленных EST-маркеров. Проведенный сравнительный анализ идентифицированных паралогичных АМП-генов выявил высокий уровень их сходства (92–97%).

Ключевые слова: сосна обыкновенная, высокопроизводительное секвенирование, антимикробные пептиды, инфекционное полегание.

Введение

Сосна обыкновенная является главной лесообразующей породой Беларуси — долевое участие сосновой формации в лесном фонде республики составляет порядка 50%. Сосновые насаждения имеют важное хозяйственное и средообразующее значение [1]. Ежегодный объем проводимых мероприятий по лесовосстановлению сосняков превышает 16 тыс. га с использованием более 60 млн штук посадочного материала.

Основным негативным фактором при выращивании посадочного материала в условиях лесных питомников являются инфекционные заболевания, среди которых доминируют болезни, вызываемые некротрофными патогенными грибами родов *Fusarium*, *Alternaria*, *Verticillium*, *Rhizoctonia*, *Botrytis*, поражающие корневую систему и стебель проростков — инфекционное полегание сеянцев [2]. Одной из стратегий повышения устойчивости лесного посадочного материала является использование при выращивании генотипов растений, характеризующихся наследственно обусловленной резистентностью к фитопатогенным микроорганизмам.

За последние десятилетия, с развитием молекулярно-генетических методов у древесных растений, в том числе и хвойных видов, было описано большое число генов, ассоциированных с устойчивостью к патогенам

(R-белки) [3–8]. Одной из основных групп защитных белков, являются антимикробные полипептиды (АМП). АМП растений характеризуются небольшим размером (<100 аминокислот) и большинство из них относятся к цистеин-насыщенным пептидам, в молекулах которых содержится четное число остатков цистеина, образующих дисульфидные связи, что придает их молекулам высокую структурную стабильность и предотвращает их денатурацию вне клетки. На основе гомологии аминокислотных последовательностей, а также пространственной структуры, АМП классифицируются на следующие семейства: тионины, дефензины, хитин-подобные белки, гевеино- и нотиноподобные белки, неспецифические липидпереносящие белки, макроциклические пептиды, Snakin/GASA-семейство и др. [9–12].

Экспрессия антимикробных пептидов в растительном организме происходит как конститутивно, так и в ответ на микробное заражение [8]. Механизм физиологического воздействия антимикробных пептидов растений на патогены объясняется моделью Шая-Мацудзаки-Хуанга (SMH) [13–15], в основе которой показано взаимодействие АМП с мембраной патогена и последующим смещением липидов мембраны, изменением ее структуры и, в некоторых случаях, проникновением пептида внутрь микробных клеток.

Основываясь на данной модели, различными авторами предложено несколько гипотез причин гибели патогенного микроорганизма: вследствие деполяризации мембраны; создания механических повреждений; активации процессов разрушающих клеточную стенку, включая индукцию гидролитических ферментов; нарушения распределения липидов мембраны, приводящих к ее функциональному дисбалансу; повреждения органоидов клетки [16–18].

В настоящее время в базе нуклеотидных последовательностей NCBI содержатся в основном фрагментарные, полученные различными авторами и не систематизированные данные структурной организации антимикробных пептидов. Для хвойных видов растений количество аннотированных последовательностей антимикробных пептидов является ограниченным: *Pinus monticola* (AMP1: JQ771157 – JQ771173; AMP1: AY596276); *Pinus pinaster* (AMP1: HM210085; AMP2: HM210086); *Pinus sylvestris* (AMP1: AF410952; AMP2: AF410953; AMP3: AF410954; AMP4: AF410955); *Pseudotsuga menziesii* (antimicrobial peptide 2-like: KC155997) [4, 5, 7]. В связи с этим исследования, направленные на молекулярно-генетическую идентификацию антимикробных пептидов хвойных видов, представляют особый интерес как в аспекте оценки эволюционного развития иммунной системы растений, так и для выявления устойчивых к инфекционным заболеваниям генотипов.

Исходя из вышесказанного, целью работы являлся поиск наиболее представленных антимикробных пептидов в транскриптоме проростков сосны обыкновенной с последующим их структурно-функциональным анализом.

Материалы и методы

Проростки *P. sylvestris* (n=14), выращивали в лабораторных условиях при температуре воздуха 22 °C и фотопериоде 10 ч в течение 14 суток. Выделение тотальной РНК проводили из корня и гипокоты проростков набором GeneJET Plant RNA Purification Mini Kit (Thermo Scientific, США) по методике фирмы-производителя. Очистку РНК от примесей геномной ДНК выполняли набором DNase I,

RNase-free (Thermo Scientific, США), а для ингибирования рибонуклеаз — RNase Inhibitor (Thermo Scientific, США). Для реакции обратной транскрипции матричной РНК с получением двухцепочечной кДНК использовали набор Maxima H Minus Double-Stranded cDNA Synthesis Kit (Thermo Scientific, США).

Библиотеки кДНК (размер фрагмента ≈200 п. н.) готовили с использованием набора Ion Plus Fragment Library Kit (Thermo Scientific, США). Эмульсионную ПЦР проводили с использованием набора Ion PGM Template OT2 200 Kit (Thermo Scientific, США) согласно инструкции компании-производителя, в планшетном амплификаторе Ion One Touch 2 System (Thermo Scientific, США). Этап обогащения микросфер производился с использованием автоматической пробоподготовки Ion OneTouch ES с применением наборов Ion PGM Template OT2 Solutions 200 Kit и Ion PGM Enrichment Beads (Thermo Scientific, США). Реакцию секвенирования производили на базе геномного анализатора Ion PGM System (Thermo Scientific, США) с использованием набора Ion PGM Sequencing 200 Kit v2 и полупроводникового микрочипа Ion 314 Chip v2 (Thermo Scientific, США). Первоначальную обработку данных, поступающих от геномного анализатора, осуществляли в автоматическом режиме при помощи программного обеспечения Ion Torrent Suite v. 4.0 (Thermo Scientific, США). Окончательную обработку информации и сборку транскриптов генов производили с использованием программного обеспечения SeqMan NGen v. 11 (DNASTAR, Израиль).

Для идентификации кодирующих последовательностей антимикробных пептидов использовали ISS-подход (анализ сходства последовательностей) в базе данных нуклеотидных последовательностей и консервативных доменов GenBank NCBI [19]. Трансляцию нуклеотидных последовательностей с последующим множественным выравниванием проводили с использованием программного пакета CLC Sequence Viewer 6.3 (Qiagen, США). Для кластерного анализа, с построением дендрограмм генетического сходства, использовали программное обеспечение MEGA 6.0. Оценку воспроизводимости результатов кластеризации проводили с использованием бутстреп-анализа (1000 повторов).

Результаты и обсуждение

По результатам высокопроизводительного секвенирования библиотек кДНК исследуемых проростков *P. sylvestris* с последующей сборкой, было получено более 28 тыс. последовательностей (контигов). Для контигов (уровень покрытия ≥ 20) № 43, 44, 45, 82 выявлен высокий уровень сходства с генами семейства антимикробных пептидов, содержащих домен MiAMP1, ранее описанный для антимикробного пептида австралийского ореха *Macadamia integrifolia* [20]. С использованием ISS-подхода для нуклеотидных последовательностей контигов № 43 и 45 было показано сходство с последовательностями генов антимикробного пептида 1 (АМП1) на 99% (ген *Sp-AMP1 Pinus sylvestris*, депонент GenBank AF410952) и 97% (ген *PpAMP1 Pinus pinaster*, депонент GenBank HM210085), соответственно. Нуклеотидная последовательность контига № 44 на 99% совпадала с последовательностью гена *Sp-AMP3* антимикробного пептида 3 (АМП3) *Pinus sylvestris* (депонент GenBank AF410954). Анализ сходства последовательностей для контига № 82 показал идентичность на 99% с последовательностями генов трех антимикробных пептидов (АМП2, АМП3, АМП4) *Pinus sylvestris* (депоненты GenBank AF410953, AF410954, AF410955). Также был проведен анализ сходства между нуклеотидными последовательностями новых идентифицированных АМП (контиги № 43, 44, 45, 82), установлен высокий уровень консервативности (сходство 92–97%).

С использованием программы Protein BLAST базы данных GenBank NCBI был проведен поиск гомологов транслируемых аминокислотных последовательностей (АК-последовательностей) идентифицированных пептидов с последующей оценкой степени сходства. Для контига № 43 сходство аминокислотной последовательности транслируемого региона с АК-последовательностью кодируемой геном *Sp-AMP1* (депонент GenBank AF410952, protein_id AAL05052) составило 97%, наибольшее сходство в 98% наблюдалось с АК-последовательностью *PpAMP1* (депонент GenBank HM210085, protein_id ADJ53037) с наличием двух несинонимичных замен, выражающихся в изменении полипептидной цепи в положениях L10 $\leftrightarrow$ M10, R84 $\leftrightarrow$ G84, и одной

синонимичной, относящейся к позиции A9. Сравнение нуклеотидного и аминокислотного выравниваний показало, что несинонимичные замены вызваны миссенс-мутациями по типу трансверсии в первых позициях кодонов (TTG $\leftrightarrow$ ATG; CGT $\leftrightarrow$ GGT), синонимичная замена обеспечена сайленс-мутацией-транзицией в третьей позиции кодона (GCT $\leftrightarrow$ GCC). Аминокислотная и нуклеотидная последовательности транслируемого региона контига № 44 имела 100% гомологию с последовательностями пептида и гена *Sp-AMP3* (депонент GenBank AF410954, protein_id AAL05054). Для контига № 45 сходство АК-последовательности транслируемого региона с АК-последовательностью, кодируемой геном *PpAMP1* (депонент GenBank HM210085, protein_id ADJ53037), составило 97%. При сравнении нуклеотидного и аминокислотного выравниваний было выявлено наличие трех несинонимичных замен (V9 $\leftrightarrow$ M9; M10 $\leftrightarrow$ L10; A92 $\leftrightarrow$ D92), вызванных миссенс-мутациями: транзицией в первой позиции кодона (GTG $\leftrightarrow$ ATG), трансверсиями в первой (ATG $\leftrightarrow$ CTG) и второй (GCT $\leftrightarrow$ GAT) позициях кодонов, а также одной синонимичной заменой в 101 позиции (F101) за счет сайленс-мутации по типу транзиции в третьей позиции кодона (TTT $\leftrightarrow$ TTC). Анализ сходства аминокислотной последовательности транслируемого региона контига № 82 показал идентичность на 99% с АК-последовательностями: *Sp-AMP2* (депонент GenBank AF410953, protein_id AAL05053), *Sp-AMP3* (депонент GenBank AF410954, protein_id AAL05054), *Sp-AMP4* (депонент GenBank AF410955, protein_id AAL05055). Для транслируемых аминокислотных последовательностей, кодируемых генами *Sp-AMP2* и *Sp-AMP4*, была установлена 100% гомология (при этом выявлено различие в нетранслируемом регионе). Для нуклеотидной последовательности АМП контига № 82 по отношению к данным последовательностям наблюдалось наличие одной несинонимичной замены, выражающейся в АК-последовательности в положении Q25 $\leftrightarrow$ E25, вызванной миссенс-мутацией по типу трансверсии в первой позиции кодона (CAG $\leftrightarrow$ GAG). При сравнении последовательностей АМП контига № 82 с последовательностью *Sp-AMP3* выявлена одна несинонимичная замена, выражающаяся в аминокислотной за-

мене (S70 ↔ A70), обеспеченная двумя транс-версиями в первом и третьем положении кодо-на (TCT ↔ GCG), а также одна синонимичная по типу транзиции в третьем положении кодо-на (ACT ↔ ACC), кодирующего аминокислоту треонин в 68 положении.

Для проведения кластерного анализа ис-пользовались идентифицированные нами по-следовательности АМП *P. sylvestris*, а также ранее известные для хвойных видов, взятые из базы данных GenBank. С использованием программы ORF Finder для нуклеотидных по-следовательностей анализируемых хвойных растений была идентифицирована откры-тая рамка считывания длиной 318 нуклеоти-дов (105 аминокислот), содержащая домен пептида MiAMP1 и 26 (27 для HM210086) аминокислотных остатков сигнальной по-следовательности (короткая аминокислотная последовательность белков, которая опреде-ляет направление их транспорта в клетке после трансляции и способствует проникновению

молекул сквозь мембраны во внеклеточное пространство) [21]. Ранее в исследовании К. Казана с соавторами было показано, что сигнальная последовательность антимикроб-ного пептида MiAMP1 необходима для экзо-генной секреции АМП в межклеточные про-странства для ингибирования находящихся в них патогенов [22].

Анализ множественного выравнивания ами-нокислотных последовательностей данных антимикробных пептидов показал наличие 33 консервативных аминокислот (рис. 1), большая часть из которых (94%) была локализована в функциональном MiAMP1-домене. Харак-терной особенностью домена являлось обяза-тельное наличие не менее шести цистеиновых остатков, обеспечивающих сохранение ста-бильности третичной структуры полипептида, необходимой для сохранения функциональной активности.

С помощью алгоритма ближнего соседа (neighbor-joining, NJ) для исследуемых аминокис-

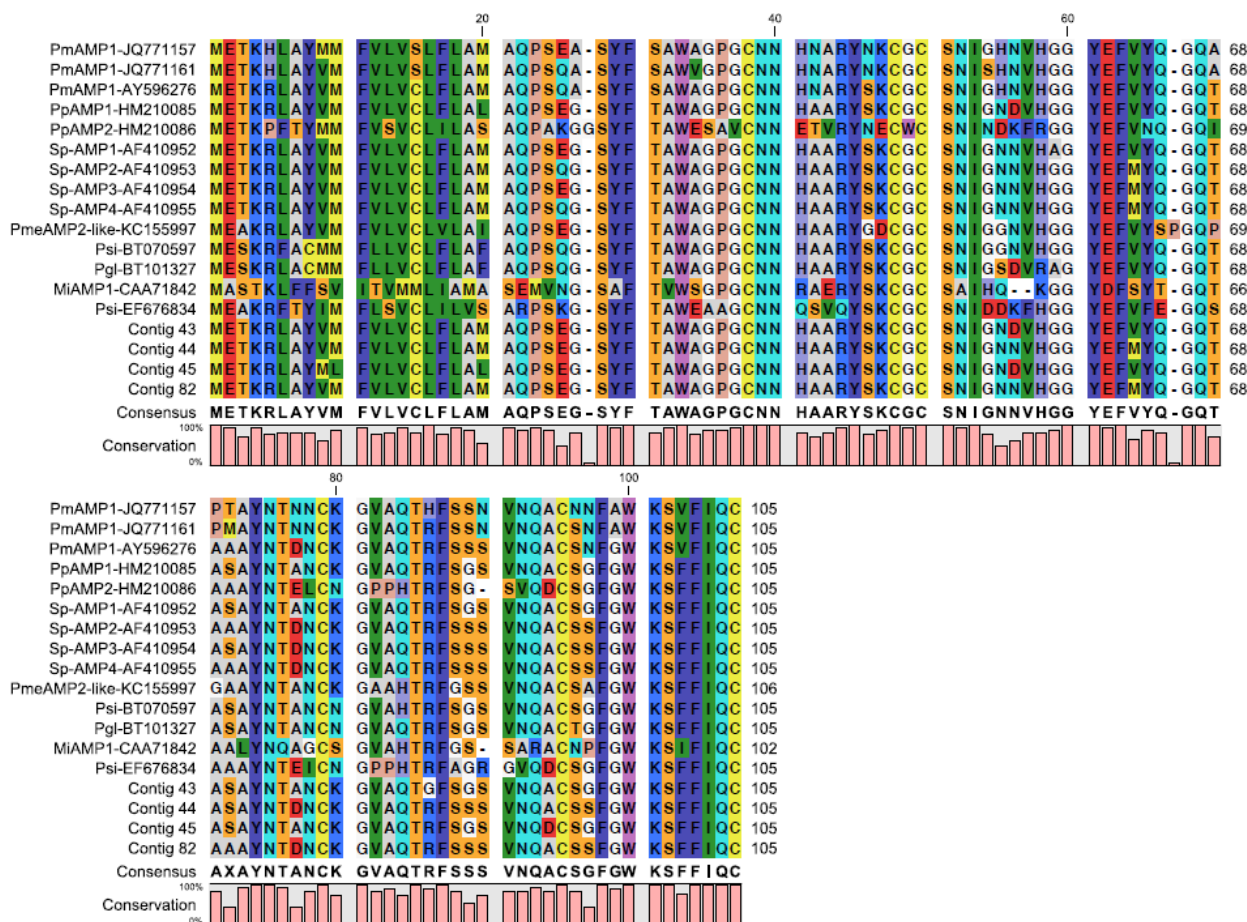


Рис. 1. Выравнивание аминокислотных последовательностей АМП хвойных видов, а также *Macadamia integrifolia*, содержащих домен MiAMP1

кислотных последовательностей АМП (включая АМП1 *Macadamia integrifolia*) проведен кластерный анализ и построена дендрограмма (рис. 2).

Как видно из рис. 2, структура дендрограммы представлена двумя основными кластерами с достоверно подтвержденными высокими уровнями значения бутстрепа (730–1000). При этом внутри кластера, сформированного антимикробным пептидом 2 *Pinus pinaster* и неаннотированным пептидом *Picea sitchensis* (депоненты GenBank HM210086 и EF676834), наблюдался высокий уровень генетической дифференциации: 0,1287 и 0,1268. Второй обширный кластер был сформирован аминокислотными последовательностями различных АМП хвойных родов *Pinus* (ранее аннотированных как: АМП1, АМП2, АМП3, АМП4, предшественник АМП), *Picea* (ранее не аннотированные) и *Pseudotsuga* (аннотированный как АМП-2-подобный пептид). Кроме того, в структуре дендрограммы АМП-2 видов одного рода *P. sylvestris* (AF410953) и *P. pinaster* (HM210086) вошли в разные кластеры и различия между их аминокислотными структурами были значительно выше, чем между АМП-2 *P. sylvestris* (AF410953) и АМП-2-подобным пептидом *Pseudotsuga*

menziesii (KC155997). Также важно отметить, что для антимикробных пептидов *P. sylvestris*, идентифицированных Ф. Асигбу с соавторами, АМП2 (AF410953) и АМП4 (AF410955) наблюдалась полная идентичность, и они сформировали общую ветвь [6]. Полученные результаты кластеризации, на наш взгляд, свидетельствуют о том, что депонированные различными авторами данные, относительно молекулярной структуры и аннотации антимикробных пептидов хвойных растений, разобщены и требуют дальнейшей систематизации.

Анализ внутривидовой представленности АМП, на примере *P. sylvestris*, показал широкое разнообразие антимикробных пептидов с низким уровнем структурных различий, что, предположительно, обусловлено общим происхождением и последующей дивергенцией генов-паралогов, кодирующих данные пептиды. Исходя из литературных данных, описывающих особенности коэволюции растений и патогенных микроорганизмов, у древесных растений в ответ на формирование широкого спектра таксонов патогенного микроорганизма наблюдается увеличение разнообразия в геномах растений семейств генов антимикробных пептидов [24, 25].

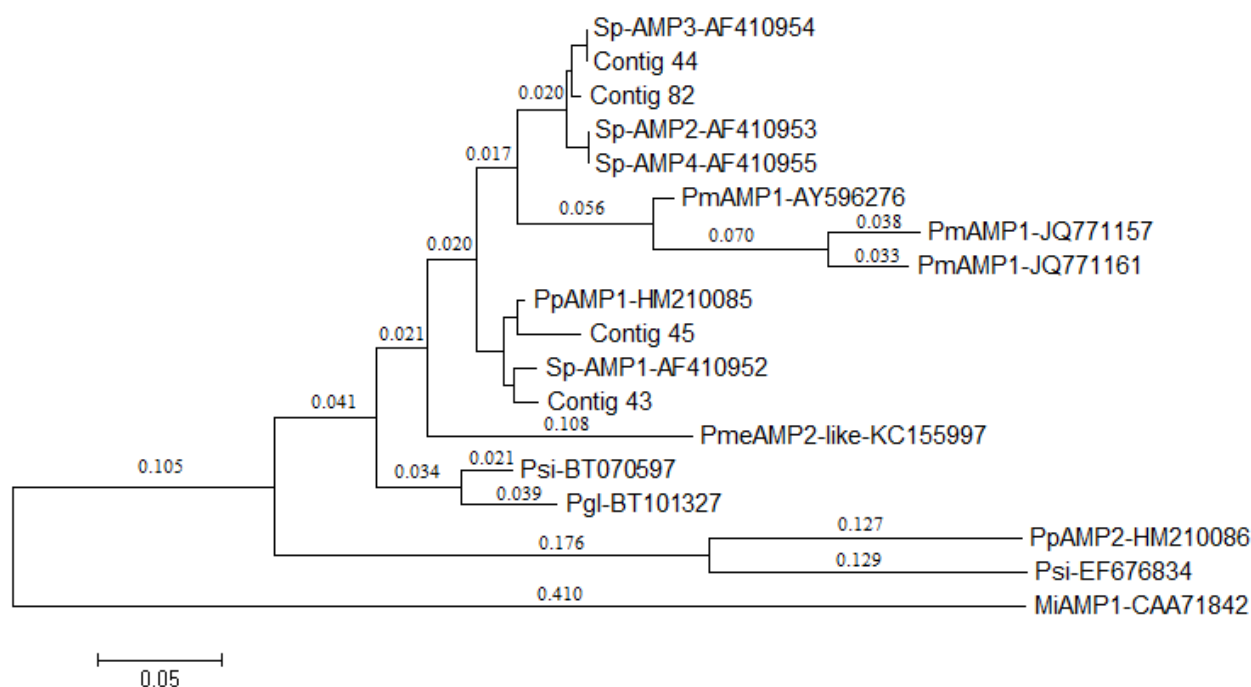


Рис. 2. Дендрограмма генетического сходства исследуемых аминокислотных последовательностей АМП. Шкала внизу — генетическое расстояние [23]

Для идентифицированных доменов анти-микробных пептидов был проведен молекулярный анализ (*in silico*) с использованием программы IPC. Прогнозируемая молекулярная масса пептидного региона составила от 8,35 кДа (для контига № 43) до 8,59 (контиг № 44), теоретическая изоэлектрическая точка (pI) также оказалась сходной: 7,25 (для контигов № 43 и 45) и 7,53 для контигов № 44 и 82), что указывает на сходство их физико-химических параметров.

Трехмерное моделирование пространственной структуры аминокислотных последовательностей структурных доменов MiAMP1 исследуемых АМП проведено на сервере программы SWISS-MODEL. В качестве референсной модели использовался домен MiAMP1 австралийского ореха *Macadamia integrifolia* с установленной ранее структурой [20], для которой характерно наличие восьми β -цепей, формирующих две β -складки с топологией «греческого ключа». Аналогичная пространственная организация показана и для доменов антимикробных пептидов, выявленных нами у

сосны обыкновенной (рис. 3). В то же время, исходя из структурного анализа установлено, что электростатический потенциал исследуемых доменов сосны отличался от референсного домена MiAMP1, являющегося амфипатической молекулой, у которой большая часть положительных зарядов сконцентрировано на одной стороне пептида. Для всех четырех доменов было характерно замещение положительно заряженного аминокислотного остатка E16 на нейтральный A16, также в положении 65 отрицательно заряженный R замещен на нейтрально заряженный Q. Пары доменов контигов 43/45 и 44/82 характеризовались сходной структурной организацией и электростатическим потенциалом, но для домена контига № 45 наблюдался дополнительный положительный заряд в положении D66. Однако, не смотря на наличие полиморфизма в доменной области, пространственная структура, с присутствием специфичных аминокислот, для всех АМП сохранялась.

Полученные данные кластерного и молекулярно-структурного анализа свидетельствуют

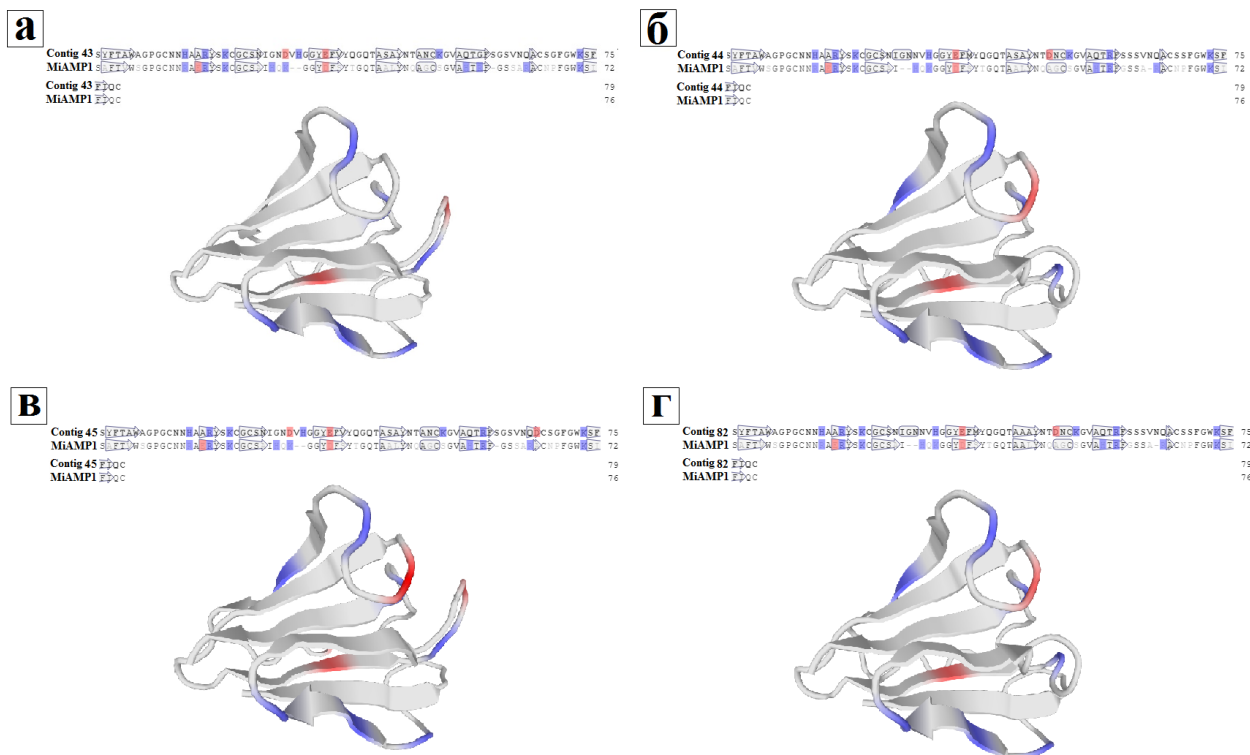


Рис. 3. Трехмерная модель доменов MiAMP1 идентифицированных АМП проростков сосны обыкновенной (а — контиг № 43, б — контиг № 44, в — контиг № 45, контиг № 82) на основе референсной модели домена MiAMP1 *Macadamia integrifolia* с отображенным электростатическим потенциалом, β -цепи обозначены стрелкой

о наличии структурного разнообразия паралогичных генов *P. sylvestris*, кодирующих функционально сходные антимикробные пептиды. На наш взгляд, это может быть обусловлено особенностью таксономического состава патогенной микрофлоры, вызывающей болезни сеянцев сосны обыкновенной. Как отмечалось выше, основными патогенами выступает группа аскомицетных некротрофных грибов со сходными патогенетическими свойствами, что обуславливает выработку у растения тождественных защитных механизмов. В то же время, эффективность действия антимикробных пептидов достигается за счет незначительных изменений в их структуре, обеспечивающих наибольшее сродство к различным видам патогенов. Аналогичные результаты относительно структурно-функционального разнообразия АМП среди хвойных растений были получены в исследованиях для *P. pinaster* и *P. monticola* [5, 7].

Как указывалось ранее и подтверждается результатами наших исследований, антимикробные пептиды являются наиболее широко представленной группой защитных белков (продуктов R-генов) в метаболоме проростков *P. sylvestris*, что определяет их значимость в качестве объекта исследования для установления механизмов резистентности сосны обыкновенной к различным инфекционным заболеваниям.

Заключение

В результате проведенных исследований для транскриптома проростков сосны обыкновенной идентифицировано четыре ранее не описанных гена, кодирующих антимикробные пептиды, проведен их структурно-функциональный анализ. Установлено, что они являются паралогами и содержат функциональный домен MiAMP1, ассоциированный с устойчивостью растений к патогенным микроорганизмам. Уровень генетического сходства экспрессируемых нуклеотидных последовательностей идентифицированных АМП-генов составил 92–97%, при этом генетические изменения в меньшей степени затрагивали функциональную область (MiAMP1-домен), что указывает на сохранение их антимикробной специфичности. Выявлена относительная консервативность домена MiAMP1, что позво-

ляет использовать его в качестве EST-маркера для идентификации устойчивых генотипов сосны обыкновенной.

Список использованных источников

1. Министерство лесного хозяйства [Электронный ресурс]. – Режим доступа: <https://www.mlh.by/>. – Дата доступа: 23.07.2018.
2. Федоров, Н. И. Лесная фитопатология: учебник для студентов специальности «Лесное хозяйство» / Федоров Н. И. // Минск: БГТУ. – 2004. – 462 с.
3. Natural roles of antimicrobial peptides in microbes, plants and animals / G. Maroti [et al.] // Research in microbiology. – 2011. – Vol. 162, № 4. – P. 363–374.
4. Anti-microbial peptide (AMP): nucleotide variation, gene expression, and host resistance in the white pine blister rust (WPBR) pathosystem / J. J. Liu, A. Zamany, R. A. Sniezko // Planta. – 2013. – Vol. 237, № 1. – P. 43–54.
5. Cloning and characterization of a putative antifungal peptide gene (*Pm-AMP1*) in *Pinus monticola* / A. K. Ekramoddoullah, J. J. Liu, A. Zamani // Phytopathology. – 2006. – Vol. 96 № 2. – P. 164–170.
6. Isolation of a novel antimicrobial peptide gene (*Sp-AMP*) homologue from *Pinus sylvestris* (Scots pine) following infection with the root rot fungus *Heterobasidion annosum* / F. O. Asiegbu [et al.] // FEMS microbiology letters. – 2003. – Vol. 228, № 1. – P. 27–31.
7. Canales, J., Avila C., Canovas F. M. A maritime pine antimicrobial peptide involved in ammonium nutrition / J. Canales, C. Avila, F. M. Canovas // Plant, cell & environment. – 2011. – Vol. 34, № 9. – P. 1443–1453.
8. Zasloff, M. Antimicrobial peptides of multicellular organisms // Nature. – 2002. – Vol. 415, № 6870. – P. 389.
9. Garcia-Olmedo, F. Plant defense peptides / F. Garcia-Olmedo, A. Molina, J.M. Alamillo // Biopolymers (Peptide Sci.). – 1998. – Vol. 47. – P. 479–491.
10. Tam, J. P. Antimicrobial peptides from plants / J. P. Tam, S. Wang, K. H. Wong // Pharmaceuticals. – 2015. – Vol. 8, № 4. – P. 711–757.
11. Silverstein, K. A. T. et al. Small cysteine-rich peptides resembling antimicrobial peptides have been under-predicted in plants /

- K. A. T. Silverstein [et al.] // *The Plant Journal*. – 2007. – Vol. 51, № 2. – P. 262–280.
12. Oliveira-Lima, M. Snakin: structure, roles and applications of a plant antimicrobial peptide / M. Oliveira-Lima // *Current Protein and Peptide Science*. – 2017. – Vol. 18, № 4. – P. 368–374.
13. Matsuzaki, K. Why and how are peptide–lipid interactions utilized for self-defense? Magainins and tachyplesins as archetypes // *Biochimica et Biophysica Acta (BBA)-Biomembranes*. – 1999. – Vol. 1462, № 1. – P. 1–10.
14. Yang, L. Crystallization of antimicrobial pores in membranes: magainin and protegrin / L. Yang [et al.] // *Biophysical Journal*. – 2000. – Vol. 79, № 4. – P. 2002–2009.
15. Shai, Y. Mechanism of the binding, insertion and destabilization of phospholipid bilayer membranes by α -helical antimicrobial and cell non-selective membrane-lytic peptides // *Biochimica et Biophysica Acta (BBA)-Biomembranes*. – 1999. – Vol. 1462, № 1. – P. 55–70.
16. Westerhoff, H. V. Magainins and the disruption of membrane-linked free-energy transduction / H. V. Westerhoff // *Proceedings of the National Academy of Sciences*. – 1989. – Vol. 86, № 17. – P. 6597–6601.
17. Bierbaum, G. Induction of autolysis of staphylococci by the basic peptide antibiotics Pep 5 and nisin and their influence on the activity of autolytic enzymes / G. Bierbaum, H. G. Sahl // *Archives of microbiology*. – 1985. – Vol. 141, № 3. – P. 249–254.
18. Kragol, G. The antibacterial peptide pyrrhocoricin inhibits the ATPase actions of DnaK and prevents chaperone-assisted protein folding / G. Kragol [et al.] // *Biochemistry*. – 2001. – Vol. 40, № 10. – P. 3016–3026.
19. Basic local alignment search tool (BLAST NCBI) [Электронный ресурс]. – Режим доступа: <https://blast.ncbi.nlm.nih.gov/Blast.cgi>. – Дата доступа: 28.07.2018
20. McManus, A. M. MiAMP1, a novel protein from *Macadamia integrifolia* adopts a greek key β -barrel fold unique amongst plant antimicrobial proteins 1 / A. M. McManus [et al.] // *Journal of molecular biology*. – 1999. – Vol. 293, № 3. – P. 629–638.
21. Тарантул В. З. Словарь биотехнологических терминов. // ИНИЦ Роспатента – 2005. – 124 с.
22. Kazan, K. Enhanced quantitative resistance to *Leptosphaeria maculans* conferred by expression of a novel antimicrobial peptide in canola (*Brassica napus* L.) / K. Kazan [et al.] // *Molecular Breeding*. – 2002. – Vol. 10, № 1–2. – P. 63–70.
23. Nei, M. Genetic distance between populations / M. Nei // *The American Naturalist*. – 1972. – Vol. 106, № 949. – P. 283–292.
24. Баранов, О. Ю. Генетические механизмы коэволюции в системе «хозяин-паразит» (на примередревесныхрастенийифитопатогенных грибовлесных ценозов Беларуси): диссертация доктора биологических наук: 03.02.07 – генетика: защищена 21.12.2017, утверждена: 16.06.2018. – Гомель, 2017. – с. 208.
25. Stahl, E. A. Plant–pathogen arms races at the molecular level / E. A., Stahl, J. G. Bishop // *Current opinion in plant biology*. – 2000. – Vol. 3, № 4. – P. 299–304.

L.V. Mozharovskaya

STRUCTURAL AND FUNCTIONAL ANALYSIS OF LOCI ENCODING ANTIMICROBIAL PEPTIDES OF SCOTS PINE

Institute of Forest, NASB
Gomel, 246001, the Republic of Belarus

Using high throughput sequencing of seedling transcriptomes of Scots pine, four expressed sequences (not previously described) determining antimicrobial peptides were identified. Based on the obtained results, the annotation and structural-functional analysis of identified EST-markers were performed. A comparative analysis of identified paralogous AMP genes revealed their high similarity level (92–97%).

Key words: Scots pine, Next-Generation Sequencing, antimicrobial peptide, infectious lodging of seedlings.

Дата поступления статьи: 29 августа 2018 г.

А.В. Падутов

ИЗУЧЕНИЕ ГЕНЕТИЧЕСКОЙ СТРУКТУРЫ ПОЛУСИБСОВОГО ПОТОМСТВА КЛОНОВ СОСНЫ ОБЫКНОВЕННОЙ НА ЛЕСОСЕМЕННЫХ ПЛАНТАЦИЯХ II ПОРЯДКА (Краткое сообщение)

Институт леса НАН Беларуси
Республика Беларусь, 246001, г. Гомель, ул. Пролетарская, 71
e-mail: apadutov@yandex.by

В статье представлены результаты молекулярно-генетического анализа потомства сосны обыкновенной, полученного из семян, собранных в разных частях кроны. Показано, что материнский генотип является основным фактором формирования генетической структуры семенного потомства. Выявлена гетерогенность генетической структуры мужских гаметофитов, участвующих в опылении мегастробил как различных клонов, так и разных частей кроны одного и того же материнского дерева. Установлено, что уровень ожидаемой гетерозиготности потомства из верхней, средней и нижней частей кроны статистически достоверно не отличается.

Ключевые слова: лесосеменная плантация, молекулярно-генетический анализ, сосна обыкновенная, система скрещивания.

Введение

Сосна обыкновенная (*Pinus sylvestris* L.) является главной лесообразующей породой Беларуси — доленое участие сосняков составляет около 50% от всей лесопокрытой площади страны [1]. Сосновые леса имеют важное экономическое значение в качестве источника сырья для деревообрабатывающей промышленности страны, а также выполняют широкий спектр экологических и социальных функций.

Одним из центральных вопросов, связанных с ведением лесного хозяйства, является обеспечение возобновления лесных ресурсов, компенсирующих их изъятие как в результате лесопромышленной деятельности (заготовка древесины и недревесной продукции), так и в результате сокращения древостоев за счет негативного воздействия различных факторов биотической (болезни и вредители) и абиотической (ураганы, паводки, аномальные значения температур и др.) природы. С этой целью в Беларуси ежегодно проводятся лесовосстановительные мероприятия на общей площади ≈ 40 тыс. га, значительная часть из которых связана с посадкой леса. Для обеспечения

потребностей лесного хозяйства в посадочном материале производится сезонная заготовка семян как в естественных древостоях, так и в специализированных искусственных насаждениях — лесосеменных плантациях (ЛСП) [2].

В настоящее время лесосеменные плантации являются одним из основных источников получения семян с высокими наследственными качествами. По структуре ЛСП разделяются на два типа — ЛСП первого и второго порядка, что связано с селекционным статусом деревьев произрастающих на плантациях. Лесосеменные плантации первого порядка закладываются прививками черенков от плюсовых деревьев — отобранных в естественных древостоях по комплексу хозяйственно ценных признаков. ЛСП второго порядка закладываются клоновым материалом элитных деревьев — плюсовых деревьев, прошедших селекционно-генетическую оценку по потомству.

Несмотря на то, что генетическая структура ЛСП является основой для формирования наследственных характеристик семенного потомства, использование только этого кри-

терия при создании объектов постоянной лесосеменной базы не позволяет в полной мере прогнозировать генетические особенности партий семян и, вследствие этого, максимально реализовать селекционный потенциал используемого клонового материала плюсовых и элитных деревьев [3]. Среди дополнительных факторов, способных оказывать влияние на формирование наследственных характеристик потомства, наибольшей значимостью характеризуется система скрещивания, которая зависит от фенологических и репродуктивных (постгамная совместимость) свойств материнских деревьев, пространственного размещения клонов по территории ЛСП и т. д. [4].

На текущий момент времени одним из наиболее информативных способов оценки наследственных характеристик партий семян, получаемых на лесосеменных плантациях, является использование методов ДНК-маркирования, которые позволяют проводить генотипирование отдельных индивидов в потомстве, оценивать их генетическое родство и разнообразие, устанавливать источник происхождения пыльцы и др.

Следует подчеркнуть, что в случае анализа семенного потомства, полученного в результате открытого опыления, молекулярно-генетические методы дают возможность не только установить эффективную численность индивидов на ЛСП (клоны, участвующие в скрещивании), но и определить репродуктивные особенности структуры плантации, включая перечень факторов, оказывающих наибольшее влияние на процессы опыления и эмбриогенеза [4]. Последующий анализ генетического разнообразия пыльцевого пула, совместимости и превалирования различных вариантов генотипических комбинаций дает возможность оптимизировать клоновую структуру ЛСП, исходя из аспектов их репродуктивного сочетания.

Исходя из всего вышесказанного, целью данного исследования являлся молекулярно-генетический анализ семенного потомства клонов сосны обыкновенной на ЛСП II порядка, включая оценку гетерогенности пыльцевого облака и особенностей опыления в различных частях кроны.

Материалы и методы

Сбор экспериментального материала для исследований проводился на ЛСП сосны обыкновенной II порядка, расположенной на территории Старо-Дятловичского лесничества Гомельского опытного лесхоза. На изученном участке для исследований были выбраны клоны, представленные в трех повторностях, расположенные на равном удалении друг от друга. Для проведения исследования с выбранных нами клонов были собраны шишки по отдельности из трех условно разделенных частей кроны (верхней, средней, нижней). Извлечение семян производилось из шишек, предварительно высушенных в контролируемых условиях при температуре 45 °С (сушильный шкаф для семян DL 600, BCC AB, Швеция).

С целью увеличения объема доступного биологического материала для выделения ДНК, перед проведением генетического анализа, семена проращивались в контролируемых условиях (в аппарате Якобсена 5101 Series RuMed, Германия) в течение 7 дней при температуре 25 °С. Для дальнейшего выделения ДНК из каждой группы проростков (относящихся к определенному дереву и части кроны) было отобрано по 10 индивидов, из которых иссекали одинаковые по размеру фрагменты (L=5 мм) в области гипокотилия и объединяли их в единый образец.

Выделение ДНК производилось на основании использования модифицированного СТАВ метода [5]. Полученные препараты ДНК разбавляли до одинаковой концентрации. В качестве ДНК-маркера был использован наиболее вариабельный из пяти рассмотренных (Ptx 4001, Ptx 3116, Psyl 17, Psyl 36, Psyl 44) микросателлитный локус, локализованный в клеточном ядре — Psyl 17. Структура праймера и условия ПЦР-амплификации микросателлитных локусов соответствовали литературным данным [6].

Предварительное электрофоретическое фракционирование проводили в 2% агарозном геле. Окончательный электрофоретический анализ ампликонов SSR-локусов выполнялся с помощью генетического анализатора ABI Prism 310 (Thermo Fisher Scientific, США) согласно протоколу, рекомендуемому компанией изготовителем. Анализ полученных резуль-

татов проводился на основании использования программного пакета GeneMapper 4.1 (Thermo Fisher Scientific, США). Генетико-статистическая обработка проводилась с помощью программного обеспечения PopGen32 и Statistica 10 [7].

Результаты и обсуждение

На первом этапе исследований для каждой изученной группы проростков на основе анализа электрофоретических профилей SSR-локуса нами была описана их генетическая структура, выраженная в частотах встречаемости аллельных вариантов (рис. 1).

Проведенная сравнительная оценка генетической структуры полусибсового потомства показала, что наибольшим уровнем сходства, в целом, характеризуются группы, происходящие от одного и того же или генотипически одинаковых (по изученному локусу) материнских деревьев (табл. 1), что указывает на ведущую роль генотипа матери при формировании генетической структуры потомства (рис. 2). В то же время наблюдается некоторая тенденция того, что семена, собранные в верхней части кроны, характеризовались большими

генетическими отличиями по сравнению с группами, относящимися к средней и нижней части кроны. Данные результаты могут быть обусловлены определенными различиями в генетической структуре пыльцевого пула, участвовавшего в формировании семенного потомства, что привело к изменениям соотношения частот аллельных вариантов. Так, частота встречаемости аллеля 217 (локус Psyl 17) в средней и нижней части кроны составляла 27,7% и 25,0%, соответственно, в то же время в верхней части кроны она равнялась 43,9%. Разнонаправленность кластеризации потомства из различных частей кроны клонов 28/423, 28/521/ 28/522, 28/513 и 28/424 объясняется сочетанием двух факторов — сходством генотипа материнских растений — Psyl 17^{220/222} и генетической гетерогенностью пыльцевого пула. В то же время потомство клонов с уникальными генотипами (28/533 — Psyl 17^{217/217} и 28/427 — Psyl 17^{220/220}) характеризовалось наибольшим генетическим сходством внутри группы, что выражалось в формировании единых кластеров на дендрограмме (рис. 2).

На следующем этапе работы проведено вычленение материнских аллелей для полу-



Рис. 1. Генетическая структура потомства сосны обыкновенной (локус Psyl 17, клон 28/533, верхняя часть кроны)

Таблица 1

Генотипы изученных клонов сосны обыкновенной (локус Psyl 17)

№ клона	28/423	28/424	28/427	28/513	28/521	28/522	28/529	28/533
Материнский генотип	220/222	220/222	220/220	220/222	220/222	220/222	217/220	217/217

чения дендрограммы, иллюстрирующей степень генотипической дифференциации групп потомств по отцовскому генотипу (рис. 3). Из полученных результатов следует, что генетическая структура потомства с различных частей кроны сосны, при кластеризации не связана друг с другом. Также следует отметить тот факт, что при вычленении материнских аллелей у клонов 28/533 и 28/427 одна из ча-

стей кроны переместилась в другой кластер. Это еще раз подтверждает, что генотип матери имеет наибольшее значение при формировании семенного материала. На разделение по различным кластерам потомства, полученного от одного дерева, также могут влиять и другие факторы, например различные сроки раскрытия шишек при созревании в различных частях кроны, расположение клонов

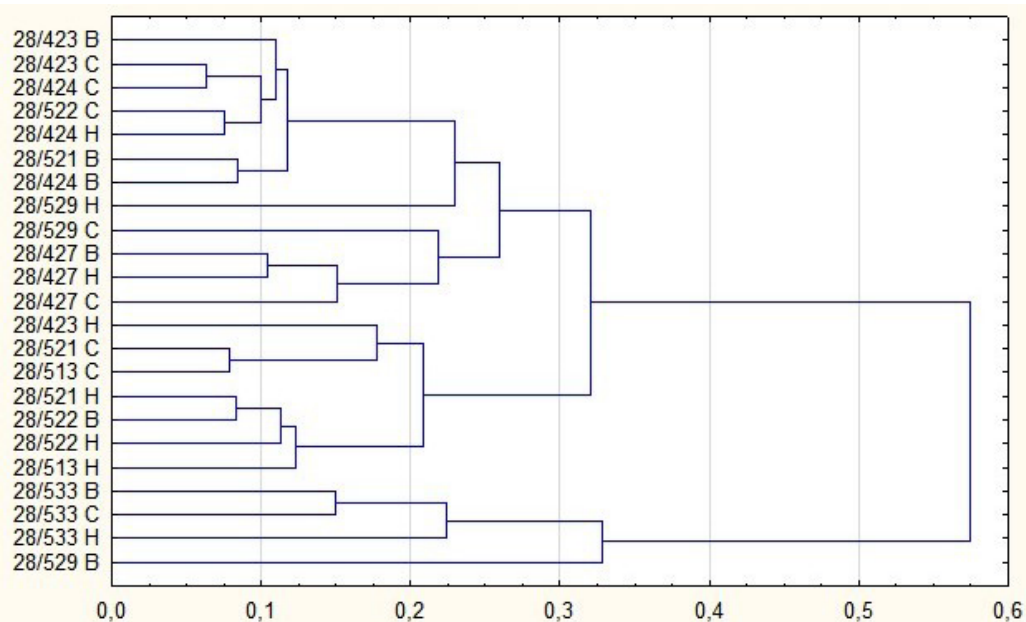


Рис. 2. Дендрограмма, иллюстрирующая степень генотипической дифференциации между семенным потомством изученных клонов сосны обыкновенной

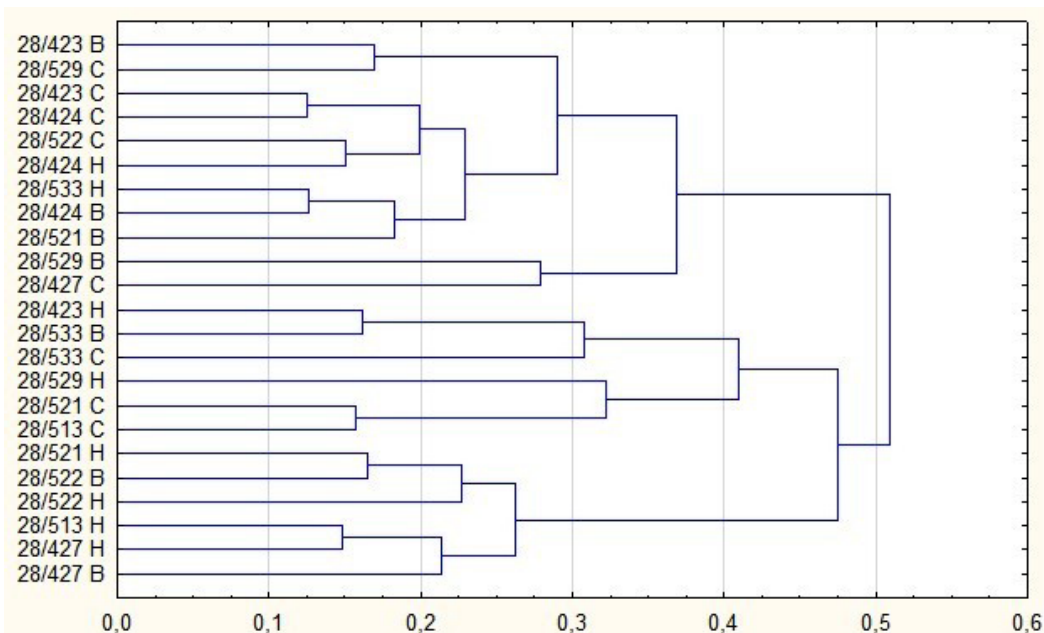


Рис. 3. Дендрограмма, иллюстрирующая степень генотипической дифференциации изученных клонов сосны обыкновенной после вычленения материнских аллелей

Таблица 2

Ожидаемая гетерозиготность клонов сосны обыкновенной

№ клона	Часть кроны	H_e
28/423	Верх	0,61
	Середина	0,66
	Низ	0,62
28/424	Верх	0,69
	Середина	0,65
	Низ	0,67
28/427	Верх	0,67
	Середина	0,58
	Низ	0,67
28/513	Верх	–
	Середина	0,64
	Низ	0,72
28/521	Верх	0,71
	Середина	0,61
	Низ	0,72
28/522	Верх	0,72
	Середина	0,66
	Низ	0,76
28/529	Верх	0,64
	Середина	0,61
	Низ	0,67
28/533	Верх	0,67
	Середина	0,62
	Низ	0,57

на плантации, распространение пыльцевого облака возле дерева, роза ветров и т. д.

Показатели ожидаемой гетерозиготности (H_e) (табл. 2), у потомства из различных ча-

стей кроны клонов 28/533 и 28/427, являющихся гомозиготными, в среднем ниже, чем у остальных деревьев. Наиболее высокий уровень гетерозиготности был отмечен

у клона под номером 28/522. При сравнении средней гетерозиготности потомства из верхней, средней и нижней части кроны, были получены следующие значения — 0,67, 0,63 и 0,68, соответственно, при этом данные различия являются статистически недостоверными.

При анализе такого показателя, как число аллелей на локус, следует выделить клон 28/533, у которого отсутствовали редкие аллели. Количество не редких аллелей у потомства остальных клонов варьировало от 2 до 7 штук.

Заключение

Анализ семенного материала, собранного с клонов на ЛСП II порядка показал, что генотип матери имеет наибольшее значение при формировании семенного материала. При кластеризации потомства на основании данных только по отцовскому вкладу не наблюдается образования групп по принадлежности к материнскому клону, поэтому генотипическая структура потомства одного клона, полученного из различных частей кроны, может значительно различаться, из-за гетерогенности отцовского пыльцевого пула. Уровень ожидаемой гетерозиготности потомства из верхней, средней и нижней частей кроны статистически не отличается между собой.

Список использованных источников

1. Министерство лесного хозяйства Республики Беларусь [Электронный ресурс]. – Режим доступа: <http://www.mlh.by/ru>. – Дата доступа: 25.07.2018.
2. Волкович, А.П. Лесное семеноводство: тексты лекций для студентов специальности 1-75 01 01 «Лесное хозяйство» специализации 1-75.01 01 06 «Лесовосстановление и питомническое хозяйство» / А. П. Волкович. – Минск: БГТУ, 2014. – 107 с.
3. Динамика структуры и продуктивности лесных формаций в Республике Беларусь / Л. Н. Рожков [и др.] // Труды БГТУ. Сер. I, Лесн. хоз-во. – 2007. – № XV. – С. 98–102.
4. Wasielewska, M. Genetic diversity and mating system of Scots pine plus trees / M. Wasielewska, M. Klemm, J. Burczyk // Dendrobiology. – 2005. – Vol. 53. – P. 57–62.
5. Селекция лесных пород / Г.И. Молотков, И.Н. Патлай и др. – М.: Лесная пром-сть. 1982. – 217 с.
6. Ефимов, Ю.П. Семенные плантации в лесной селекции семеноводства: Авто-реф. дис. на соиск. учен. степ. д. с.-х. н. – Йошкар-Ола, 1997. – 45 с.
7. Халафян, А. А. Statistica 6. Статистический анализ данных: учебник: [для вузов по специальности «Статистика» и другим экон. специальностям] / А. А. Халафян. – [3-е изд.]. – Москва: Бином, 2008. – 503 с.

A.V. Padutov

**STUDY OF THE GENETIC HALF-SIB PROGENY STRUCTURE OF
SCOTS PINE CLONES ON FOREST-SEED ORDER II ORCHARDS
(Short report)**

Institute of Forest, NASB
Gomel, 246001, the Republic of Belarus

The Article presents molecular-genetic analysis results of the Scots pine progeny obtained from the seeds collected in different parts of the crown. It was shown that the maternal genotype contribution is the main factor in the genetic structure formation of the seed progeny. The genetic structure heterogeneity of male gametophytes, participating in the megastrobile pollination of both different clones and different crown parts of the same mother tree was identified. It was established that the expected heterozygosity level in progeny from upper, middle and lower parts of the crown is not significantly different statistically.

Key words: seed orchards, molecular-genetic analysis, Scots pine, mating system.

Дата поступления статьи: 29 августа 2018 г.

М.В. Спринджук, А.П. Кончиц, В.В. Слипень, Л.П. Титов

АЛГОРИТМЫ И ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ДЛЯ ОБРАБОТКИ ДАННЫХ ГЕНОМОВ РАСТЕНИЙ (Обзорная статья)

Объединенный институт проблем информатики НАН Беларуси
Республика Беларусь, 220012, г. Минск, ул. Сурганова, 6
e-mail: stepanenkoatvei@yandex.by

В статье рассматриваются основные алгоритмы и программное обеспечение для обработки данных сборки геномов растений.

Ключевые слова: сборка геномов, биоинформатика, геномика, геном, прикладная математика, программное обеспечение, математическая биология.

Введение

За последние годы геномы более 50 видов сельскохозяйственных и нескольких лесохозяйственных растений были просеквенированы, представляя новые ресурсы для улучшения селекции и информацию для изучения генного состава растений. Секвенирование — не конечная цель, конечная цель — это улучшение качества и увеличение урожайности сельскохозяйственных культур и лесных ресурсов, а также более глубокое понимание геномной эволюции [46]. Автоматические методы расшифровки последовательностей ДНК (секвенирования) позволяют в течение короткого времени (несколько дней) и сравнительно недорого (несколько тысяч долларов) получить сотни миллиардов коротких последовательностей из четырех букв А, Т, G, С, полученных прочтением фрагментов

входного образца ДНК одного или нескольких организмов.

Существует несколько компаний, выпускающих устройства для получения коротких чтений. Самыми распространенными на рынке этих устройств являются продукты компании Illumina, однако в последнее время секвенаторы Ion Torrent стремительно набирают популярность благодаря своей дешевизне. Сравнительные характеристики секвенаторов этих двух компаний представлены в табл. 1 [48, 49].

Основные алгоритмы сборки геномов растений

De novo в переводе с латыни означает «сначала», заново. Сборка генома *de Novo* предполагает ассемблирование генома из коротких чтений без ссылочного эталона-генома.

Таблица 1

Сравнительные характеристики секвенаторов Illumina и Ion Torrent

Параметр	Ion Torrent Proton	Illumina HiSeq 2500
Цена системы, \$	243000	740000
Стоимость годового обслуживания, \$	19400	59200
Стоимость на 1 миллиард пар нуклеотидов, \$	16,67	46,00
Длительность одного запуска, ч	8	27

В противоположность *de novo* сборке существует и широко применяется сборка геномов с использованием проаннотированного ссылочного генома-эталона. Есть публикации, показывающие преимущества и недостатки обоих подходов [44].

Задачей сборки генома является восстановление последовательности ДНК (ее длина составляет от миллионов до миллиардов нуклеотидов у разных живых существ) на основании информации, полученной в результате секвенирования. Этот процесс делится, как правило, на следующие этапы:

1. Исправление ошибок в данных секвенирования.
2. Сборка квазиконтигов — фрагментов, префиксы и суффиксы которых были получены на этапе секвенирования.
3. Сборка контигов — максимальных непрерывных последовательностей нуклеотидов, которые удалось восстановить.
4. Построение скэффолдов — последовательностей контигов, разделенных промежутками, для длин которых известны верхние и нижние оценки.

Одной из наиболее часто используемых при сборке генома математических моделей является граф де Брюйна. На его использовании основаны следующие программные средства сборки генома: Velvet, ALLPATHS, ABySS, SOAPdenovo, EULER. Одним из недостатков, которым обладают перечисленные программные средства, является большой объем оперативной памяти, необходимый для сборки генома размером в миллиард нуклеотидов [48].

В алгоритмах ассемблирования (сборки геномов из коротких ридов) должны учитываться такие факторы, как неравномерность покрытия (число чтений, содержащих тот или иной нуклеотид генома) геномов ридов, которые по технологии получаются из случайных фрагментов ДНК; возможность и частота ошибок в ридов; возможность наличия химерных ридов, составленных из разных частей ДНК; наличие в геномах длинных повторов, которые могут приводить к невозможности восстановления полной последовательности даже теоретически.

Наиболее сложной является сборка генома *de novo*. Задача сборки при наличии образца, например, сборки генома индивидуального человека при наличии референсного генома

(секвенированный, собранный и проаннотированный геном организма того же вида, к которому относится анализируемый образец) более простая. Предложены десятки алгоритмов сборки *de novo*. Большинство из них основаны на построении графа де Брюйна и нахождении Эйлерова пути в нем. Однако из-за разных эвристик, заложенных на разных этапах сборки, результаты применения сборщиков существенно отличаются. Важным обстоятельством, усложняющим сравнение алгоритмов, является то, что отсутствуют универсальные метрики оценки качества сборки. Причина заключается в том, что нет количественной оценки ошибок разного типа. Например, что лучше: большее число длинных контигов (однозначно расшифрованных непрерывных последовательностей) и скэффолдов (нескольких контигов, склеенных в единую последовательность с возможными пропусками или недостоверно определенными нуклеотидами между ними) либо уменьшение числа химер: контигов и скэффолдов, ошибочно склеенных из фрагментов разных хромосом.

Большой разрыв между желаемым результатом сборки *de novo* и получаемыми результатами, высокая актуальность создания более качественных сборщиков геномов для медицины и биологии ставят эту задачу в первый ряд актуальных вычислительных задач.

Особые сложности возникают при попытках сборки метагеномных данных с различными смесями и примесями геномов организмов различного происхождения.

Нами, на основе литературы и собственного опыта обработки геномных данных, предложены алгоритмы, представленные на рис. 1 и 2.

Обзор особенностей доступных программ-сборщиков геномов

Все перечисленные сборщики геномов фигурируют в научных работах последнего времени, посвященных сравнению качества работы геномных сборщиков.

Для контроля качества коротких чтений, фильтрации и тримминга перед ассемблированием-сборкой, как правило, применяются FASTX-Toolkit и FastQC.

ABYSS — собирает геном в модели распределенной памяти, используя библиотеку MPI. Для определения участков перекрытий использует распределенную хеш-таблицу [23, 32].

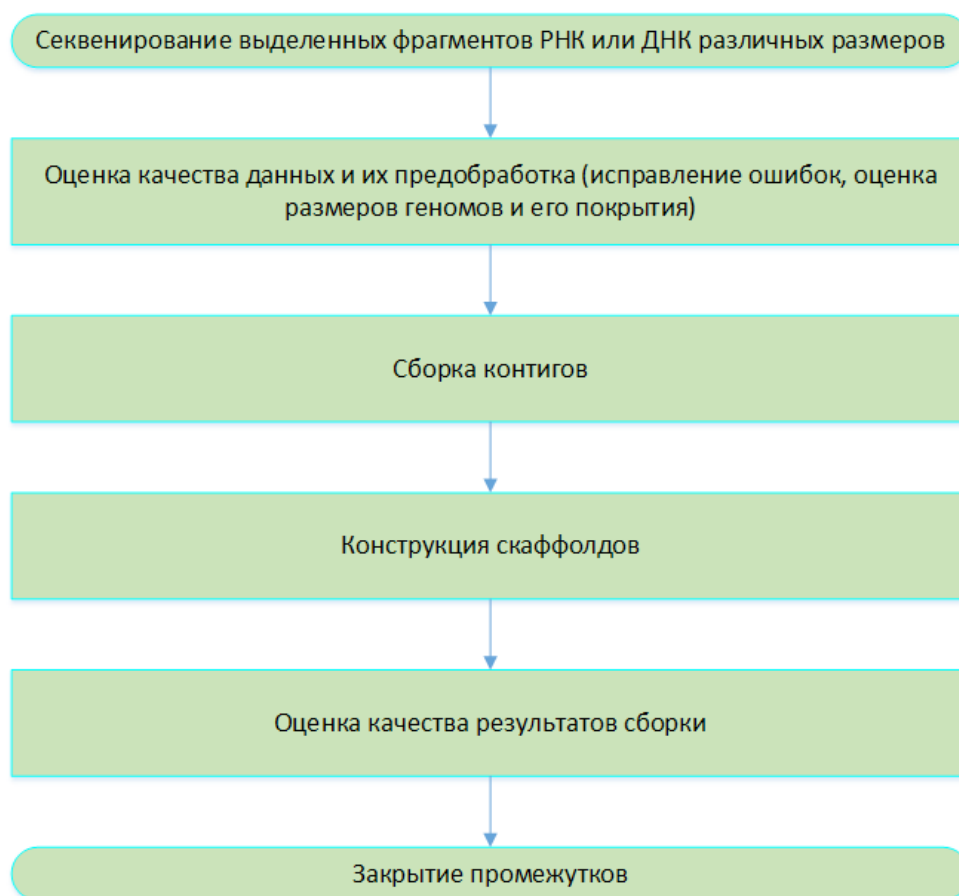


Рис. 1. Общий алгоритм сборки геномов (пояснение см. в тексте). По источнику [47]

RAY — может собирать геном в модели распределенной памяти, используя библиотеку MPI. Использует оригинальную эвристику для выявления повторов в геноме.

SOAPdenovo — использует разреженную хеш-таблицу для экономии оперативной памяти, что может приводить к ошибкам сборки. Для этого сборщика характерно относительно малое время работы [39].

SPAdes — использует различные параметры сборки для участков генома с различной глубиной покрытия [5].

Velvet — один из первых и самых распространенных сборщиков для коротких ридов. Необходимо вручную задавать большой набор параметров сборки [2, 22, 25, 40, 41].

Результаты сборки обрабатываются с целью удаления повторов и соответствующих промежутков программами REPuter [15, 16], Tandem Repeat Finder [17, 38], and RepeatMasker [4, 6, 9, 13, 33, 34] RECON (Bao and Eddy, 2002), Repeat Modeler RepeatScout и REPET. Что касается картирования/выравнивания геномов

против ссылочного генома-эталона, можно применять инструменты анализа BLAST [4] и Blat [14]. Также разработаны более продвинутые инструменты: Maq [7], Bowtie, SOAPaligner и BWA. Эти же программы широко применяются для картирования геномов человека, млекопитающих и микроорганизмов [1, 12, 19, 20]. Существуют также так называемые компаративные ассемблеры: AMOScmp [29], ABBA [28] и Celera Assembler [10] для сборки с использованием генома-эталона и для анализа сложных метагеномных данных с геновыми примесями.

Пути преодоления ошибок секвенирования, сборки и картирования геномов

Проблемой секвенирования геномов является тот факт, что только малая порция генома кодирует непосредственно сами реальные активные гены, а эти гены окружены повторяющейся ДНК, которая трудно поддается анализу и идентификации.



Рис. 2. Алгоритм обработки данных геномов лесных деревьев на примере сосны

В процессе чтения секвенаторами допускаются технические ошибки. Ошибки бывают трех типов:

1. Ошибки вставки — в основном проявляются в прочтении более длинных, чем в исходном геноме, последовательностей одинаковых нуклеотидов (например, вместо «АА» было прочитано «ААА»).

2. Ошибки удаления — в этом случае в прочитанной нуклеотидной последовательности может не хватать одного нуклеотида (например, вместо «ACGT» было прочитано «AGT»).

3. Ошибки замены — в таких случаях некоторые нуклеотиды были прочитаны неверно (например, вместо нуклеотида А был прочитан нуклеотид G).

Секвенаторы компании Ion Torrent совершают ошибки вставки и удаления значительно чаще, чем ошибки замены, в связи с чем возникает необходимость разработки методов их исправления.

Сборки генома, построенные сборщиками на основе ридов длины 100–150 нуклеотидов, фрагментированы и содержат ошибки. Обычная практика при решении задачи сборки генома заключается в запуске нескольких сборщиков с различными параметрами, а затем выбору наилучшего варианта согласно некоторым соображениям. Однако недавние исследования показывают, что достаточно распространена ситуация, при которой одни сборщики показывают лучший результат по одному из статистических критериев в сравнении с другими программами и уступают им же по другому критерию.

Одной из первых программ для объединения набора контигов от разных сборщиков стала Reconciliator. В ней производится поиск участков, являющихся уникальными как для ведомой, так и для ведущей последовательности. На следующем этапе закрываются пропуски в контигах из первого набора с использованием последовательностей из второго

набора. В случае наличия нескольких вариантов выбирается тот, который отвечает лучшему статистическому критерию. Похожий подход использует GAA, строящий граф соответствия между наборами контигов, используемый для объединения сборок, и ZORRO, который предваряет шаг объединения этапом фильтрования контигов, содержащих ошибки. Программа GAM-NGS [35] ищет в контигах блоки соответствия, которые зависят от количества ридов, картирующихся на сравниваемые последовательности — таким образом, не проводят процедуру выравнивания каждого контига с каждым. Исходя из информации, полученной на стадии картирования ридов, строится граф сборок, анализируя который можно установить участки несоответствия между наборами контигов.

Процесс объединения результатов работ геномных сборщиков является ресурсоемкой задачей. Время работы программы Reconciliator на одном из входных наборов данных составляет 24 часа, потребление оперативной памяти при этом составляет больше 100 Гб.

Методы, основанные на исправлении k -меров, работают не с чтениями, а с их подстроками длины k — k -мерами. Если в чтениях присутствуют только ошибки замены, то можно для каждого k -мера перебрать все возможные варианты ошибок, которые могли в нем произойти (их 3^k вариантов). Такой подход не работает для ошибок вставки и замены, так как при таких ошибках k -меры превращаются соответственно в $(k+1)$ -меры и $(k-1)$ -меры.

Методы, основанные на графе де Брейна, в целом менее эффективны, чем основанные на k -мерах, так как требуют хранения в памяти большого графа, из-за чего затрудняется их масштабируемость, а также из-за сложности графа, получаемого в случае больших геномов [48].

Так как парные короткие чтения имеют переклест приблизительно размером в 20 пар оснований, их можно переассемблировать в псевдо-длинные чтения (*псевдоконтиги*, *суперчтения*), используя программное обеспечение ALLPATHS-LG [45], MaSuRCA [43] и FLASH [21]. Также, особенно при использовании программ-сборщиков без встроенной функции исправления ошибок секвенирования, полезно применять программное обе-

спечение Quake, которое эффективно находит и исправляет ошибки секвенирования перед сборкой генома. Еще одна методика для преодоления ошибок и повторов в геномах — это применение гибридного подхода с использованием сборки и картирования геномов коротких и длинных чтений одновременно.

Опыт секвенирования и сборки геномов березы

Имеется практически единственная публикация [37], в которой в деталях представляется процесс секвенирования и сборки генома карликовой березы (*Betula nana*). Применялось программное обеспечение SOAPdenovo-63mer версия 2.04.3. Генерированные контиги обрабатывались программой GapCloser [3], была выбрана длина k -меров между 28 и 38. Качество результатов сборки контролировалось программным обеспечением Assemblathon statistics 2 [8, 11] и CEGMA [24]. На основе большей длины, более полноценного покрытия генома и меньшего количества контигов отбирался лучший результат сборки. Лучший результат показал k -мер=35 и параметр M3 SOAPdenovo сборщика. Авторы применяли также программное обеспечение Bowtie и Stacks для картирования и запроса вариантов с использованием аннотированного ссылочного генома-эталона.

Заключение

Сборка любого генома, в особенности растения лесохозяйственного назначения, представляет собой новую многогранную задачу. Для преодоления специфических проблем, связанных с особенностями геномов растений, нужно применять дополнительное программное обеспечение для удаления повторов как в исходных коротких чтениях, так и в контигах, необходимо экспериментировать с различным программным обеспечением, предназначенным для создания псевдоконтигов. Недавно появилось программное обеспечение, позволяющее собирать геномы полиплоидных растений и выполнять сборку геномов с использованием эталонного ссылочного генома [18, 26, 27, 30, 31, 36, 42]. Из литературных источников можно также сделать вывод об острой необходимости применять для задач геномики мощнейшие дорогостоящие процес-

соры и большую оперативную память вплоть до 1ТБ и более; также, возможно, уже имеется опыт применения суперкомпьютерных технологий. Будущие исследования геномики лесохозяйственных и сельскохозяйственных растений будут выполнимы только при наличии адекватных поставленным задачам технических, программных и аналитических технологий, средств, алгоритмов и методик. Следующей ступенью в изучении геномов растений будут исследования по изучению разнообразия геномных вариаций лесных и сельскохозяйственных растений, глубокая аннотация их функциональных элементов и создание предиктивных количественных моделей, соотносящих генотип с фенотипом.

Список использованных источников

1. Abuin, J.M., Pichel, J.C., Pena, T.F., Amigo, J. BigBWA: approaching the Burrows-Wheeler aligner to Big Data technologies // *BIOINFORMATICS*. – 2015. – Vol. 31, № 24. – P. 4003–4005.
2. Afiahayati, Sato, K., Sakakibara, Y. MetaVelvet-SL: an extension of the Velvet assembler to a de novo metagenomic assembler utilizing supervised learning // *DNA Res.* – 2015. – Vol. 22, № 1. – P. 69–77.
3. Aung, H.L., Nyunt, W.W., Fong, Y., Cook, G.M., Aung, S.T. First 2 Extensively Drug-Resistant Tuberculosis Cases From Myanmar Treated With Bedaquiline // *Clin Infect Dis.* – 2017. – Vol. 65, № 3. – P. 531532.
4. Bajpai, A., Sridhar, S., Reddy, H.M., Jesudasan, R.A. BRM-Parser: a tool for comprehensive analysis of BLAST and RepeatMasker results // *In Silico Biol.* – 2007. – Vol. 7, № 4–5. – P. 399–403.
5. Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M.и др. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing // *J Comput Biol.* – 2012. – Vol. 19, № 5. – P. 455–77.
6. Bedell, J.A., Korf, I., Gish, W. MaskerAid: a performance enhancement to RepeatMasker // *BIOINFORMATICS*. – 2000. – Vol. 16, № 11. – P. 1040–1041.
7. Bottcher, R., Amberg, R., Ruzius, F.P., Guryev, V., Verhaegh, W.F.и др. Using a priori knowledge to align sequencing reads to their exact genomic position // *Nucleic Acids Res.* – 2012. – Vol. 40, № 16. – P. e125.
8. Bradnam, K.R., Fass, J.N., Alexandrov, A., Baranay, P., Bechner, M.и др. Assemblathon 2: evaluating de novo methods of genome assembly in three vertebrate species // *Gigascience*. – 2013. – Vol. 2, № 1. – P. 10.
9. Chen, N. Using RepeatMasker to identify repetitive elements in genomic sequences // *Curr Protoc Bioinformatics*. – 2004. – Vol. 4. – P. Unit 4 10.
10. Denisov, G., Walenz, B., Halpern, A.L., Miller, J., Axelrod, N.и др. Consensus generation and variant detection by Celera Assembler // *BIOINFORMATICS*. – 2008. – Vol. 24, № 8. – P. 1035–1040.
11. Earl, D., Bradnam, K., St John, J., Darling, A., Lin, D.и др. Assemblathon 1: a competitive assessment of de novo short read assembly methods // *Genome Res.* – 2011. – Vol. 21, № 12. – P. 2224–2241.
12. Gholami, M., Arbabi, A., Sharifi-Zarchi, A., Chitsaz, H., Sadeghi, M. ARYANA: Aligning Reads by Yet Another Approach // *BMC Bioinformatics*. – 2014. – Vol. 15, № 9. – P. S12.
13. Huda, A., Jordan, I.K. Analysis of transposable element sequences using CENSOR and RepeatMasker // *Methods Mol Biol.* – 2009. – Vol. 537. – P. 323–336.
14. Kent, W.J. BLAT--the BLAST-like alignment tool // *Genome Res.* – 2002. – Vol. 12, № 4. – P. 656–664.
15. Kurtz, S., Choudhuri, J.V., Ohlebusch, E., Schleiermacher, C., Stoye, J.и др. REPuter: the manifold applications of repeat analysis on a genomic scale // *Nucleic Acids Res.* – 2001. – Vol. 29, № 22. – P. 4633–4642.
16. Kurtz, S., Schleiermacher, C. REPuter: fast computation of maximal repeats in complete genomes // *BIOINFORMATICS*. – 1999. – Vol. 15, № 5. – P. 426–427.
17. Lee, J.C., Tseng, B., Ho, B.C., Linacre, A. pSTR Finder: a rapid method to discover polymorphic short tandem repeat markers from whole-genome sequences // *Investig Genet.* – 2015. – Vol. 6. – P. 10.
18. Lischer, H.E.L., Shimizu, K.K. Reference-guided de novo assembly approach improves genome reconstruction for related species // *BMC Bioinformatics*. – 2017. – Vol. 18, № 1. – P. 474.
19. Liu, X., Han, S., Wang, Z., Gelernter, J., Yang, B.Z. Variant callers for next-generation

sequencing data: a comparison study // PLoS ONE. – 2013. – Vol. 8, № 9. – P. e75619.

20. Luo, R., Cheung, J., Wu, E., Wang, H., Chan, S.H. и др. MICA: A fast short-read aligner that takes full advantage of Many Integrated Core Architecture (MIC) // BMC Bioinformatics. – 2015. – Vol. 16, Suppl. 7. – P. S10.

21. Magoc, T., Salzberg, S.L. FLASH: fast length adjustment of short reads to improve genome assemblies // BIOINFORMATICS. – 2011. – Vol. 27, № 21. – P. 2957–2963.

22. Namiki, T., Hachiya, T., Tanaka, H., Sakakibara, Y. MetaVelvet: an extension of Velvet assembler to de novo metagenome assembly from short sequence reads // Nucleic Acids Res. – 2012. – Vol. 40, № 20. – P. e155.

23. Nielsen, C.B., Jackman, S.D., Birol, I., Jones, S.J. ABySS-Explorer: visualizing genome sequence assemblies // IEEE Trans Vis Comput Graph. – 2009. – Vol. 15, № 6. – P. 881–888.

24. Parra, G., Bradnam, K., Korf, I. CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes // BIOINFORMATICS. – 2007. – Vol. 23, № 9. – P. 1061–1067.

25. Powell, D.R., Seemann, T. VAGUE: a graphical user interface for the Velvet assembler // Bioinformatics. – 2013. – Vol. 29, № 2. – P. 264–265.

26. Rausch, T., Koren, S., Denisov, G., Weese, D., Emde, A.K. и др. A consistency-based consensus algorithm for de novo and reference-guided sequence assembly of short reads // BIOINFORMATICS. – 2009. – Vol. 25, № 9. – P. 1118–1124.

27. Ruttink, T., Sterck, L., Rohde, A., Bendixen, C., Rouze, P. и др. Orthology Guided Assembly in highly heterozygous crops: creating a reference transcriptome to uncover genetic diversity in *Lolium perenne* // Plant Biotechnol J. – 2013. – Vol. 11, № 5. – P. 605–617.

28. Salzberg, S.L., Sommer, D.D., Puiu, D., Lee, V.T. Gene-boosted assembly of a novel bacterial genome from very short reads // PLoS Comput Biol. – 2008. – Vol. 4, № 9. – P. e1000186.

29. Schatz, M.C., Phillippy, A.M., Sommer, D.D., Delcher, A.L., Puiu, D. и др. Hawkeye and AMOS: visualizing and assessing the quality of genome assemblies // Brief Bioinform. – 2013. – Vol. 14, № 2. – P. 213–224.

30. Schneeberger, K., Ossowski, S., Ott, F., Klein, J.D., Wang, X. и др. Reference-guided

assembly of four diverse *Arabidopsis thaliana* genomes // Proc Natl Acad Sci U S A. – 2011. – Vol. 108, № 25. – P. 10249–10254.

31. Silva, G.G., Dutilh, B.E., Matthews, T.D., Elkins, K., Schmieder, R. и др. Combining de novo and reference-guided assembly with scaffold_builder // Source Code Biol Med. – 2013. – Vol. 8, № 1. – P. 23.

32. Simpson, J.T., Wong, K., Jackman, S.D., Schein, J.E., Jones, S.J. и др. ABySS: a parallel assembler for short read sequence data // Genome Res. – 2009. – Vol. 19, № 6. – P. 1117–1123.

33. Tarailo-Graovac, M., Chen, N. Using RepeatMasker to identify repetitive elements in genomic sequences // Curr Protoc Bioinformatics. – 2009. – Vol. Chapter 4. – P. Unit 4 10.

34. Tempel, S. Using and understanding RepeatMasker // Methods Mol Biol. – 2012. – Vol. 859. – P. 29–51.

35. Vicedomini, R., Vezzi, F., Scalabrin, S., Arvestad, L., Policriti, A. GAM-NGS: genomic assemblies merger for next generation sequencing // BMC Bioinformatics. – 2013. – Vol. 14, Suppl. 7. – P. S6.

36. Wang, B., Ekblom, R., Bunikis, I., Siitari, H., Hoglund, J. Whole genome sequencing of the black grouse (*Tetrao tetrix*): reference guided assembly suggests faster-Z and MHC evolution // BMC Genomics. – 2014. – Vol. 15. – P. 180.

37. Wang, N., Thomson, M., Bodles, W.J., Crawford, R.M., Hunt, H.V. и др. Genome sequence of dwarf birch (*Betula nana*) and cross-species RAD markers // Mol Ecol. – 2013. – Vol. 22, № 11. – P. 3098–3111.

38. Wang, X., Huang, B., Blair, B., Eglezos, S., Bates, J. Selection of optimal combinations of loci by the Optimal Combination Finder computer program from a group of variable number tandem repeat loci for use in *Staphylococcus aureus* food poisoning case investigations // J Med Microbiol. – 2012. – Vol. 61, № 5. – P. 631–639.

39. Xie, Y., Wu, G., Tang, J., Luo, R., Patterson, J. и др. SOAPdenovo-Trans: de novo transcriptome assembly with short RNA-Seq reads // BIOINFORMATICS. – 2014. – Vol. 30, № 12. – P. 1660–1666.

40. Zerbino, D.R. Using the Velvet de novo assembler for short-read sequencing technologies // Curr Protoc Bioinformatics. – 2010. – Vol. 11. – P. 5.

41. Zerbino, D.R., McEwen, G.K., Margulies, E.H., Birney, E. Pebble and rock band: heuristic resolution of repeats and scaffolding in the velvet short-read de novo assembler // PLoS One. – 2009. – Vol. 4, № 12. – P. e8407.
42. Zhong, C., Yang, Y., Yooseph, S. GRASP: guided reference-based assembly of short peptides // Nucleic Acids Res. – 2015. – Vol. 43, № 3. – P. e18.
43. Zimin, A.V., Marcais, G., Puiu, D., Roberts, M., Salzberg, S.L. и др. The MaSuRCA genome assembler // Bioinformatics. – 2013. – Vol. 29, № 21. – P. 2669–2677.
44. Chiu, K.P. Next-generation sequencing and sequence data analysis. Bentham Science Publishers, 2015.
45. Gnerre, S., MacCallum, I., Przybylski, D., Ribeiro, F.J., Burton, J.N. и др. High-quality draft assemblies of mammalian genomes from massively parallel sequence data // Proceedings of the National Academy of Sciences. – 2011. – Vol. 108, № 4. – P. 1513–1518.
46. TÜRktaş, M., Kurtoglu, K.Y., Dorado, G., Baohong, Z., Hernandez, P. и др. Sequencing of plant genomes — a review // Turkish Journal of Agriculture & Forestry. – 2015. – Vol. 39, № 3. – P. 361–376.
47. Wang, X. Next-generation sequencing data analysis. 2016. – 246 p.
48. Александров, А., Шалыто, А. Метод исправления ошибок вставки и удаления в наборе чтений нуклеотидной последовательности // Научно-технический вестник информационных технологий, механики и оптики. – 2016. – Vol. 16, № 1. – P. 108–14.
49. Романенков, К.В., Сальников, А.Н., Алексеевский, А.В. Параллельный метод объединения результатов работы программ по сборке генома // Вестник Южно-Уральского государственного университета. Серия: Вычислительная математика и информатика. – 2016. – Vol. 5, № 1. – P. 24–34.

M.V. Sprindzhuk, A.P. Konchits, V.V. Slizen', L.P. Titov

**ALGORITHMS AND SOFTWARE DESIGNED FOR THE PLANT
GENOME DATA PROCESSING
(Review article)**

United Institute of Informatics Problems, NASB
Minsk, 220012, the Republic of Belarus

Authors describe the principal algorithms and software designed and implemented for the data processing of plant genomes.

Key words: genome assembly, bioinformatics, genomics, genome, applied mathematics, software, mathematical biology.

Дата поступления статьи: 11 апреля 2018 г.

И.Ю. Романчук, В.С. Анохина

АЛКАЛОИДЫ ЛЮПИНА: СТРОЕНИЕ, БИОСИНТЕЗ, ГЕНЕТИКА (Обзорная статья)

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

В обзоре приведены данные опубликованных работ по биосинтезу, генетике и изменению содержания алкалоидов в процессе онтогенеза растений и селекции разных видов люпина, а также фунгицидному действию отдельных алкалоидов.

Ключевые слова: алкалоиды, генетика, люпин белый, желтый, узколистый.

Введение

Использование разных методов идентификации алкалоидов позволило выявить не только их качественное и количественное наличие у разных видов люпина, но и тестировать геномы по известным генам алкалоидности. Обзор такой информации важен при разработке новой стратегии в селекции люпина разных направлений использования, а также апробации перспективных селекционных образцов на разных этапах онтогенеза растений. Знание структуры и генетического контроля разных форм алкалоидов люпина приобретает все большую необходимость при создании стабильно безалкалоидных сортов для производства кормов и кондитерской промышленности, использовании отдельных алкалоидов в фармакологической промышленности и познании их роли в формировании устойчивости растений к биотическим и абиотическим стрессорам.

Алкалоиды люпина: строение, содержание, характеристики

Алкалоиды люпина относятся к группе производных хинолизида [1], большая часть из них имеет состав $C_{15}H_xNO_2$, где «х» может быть равен 20, 22, 24 и 26. Наиболее распространены у разных видов люпина люпинин, люпанин, спартеин, гидроксилупанин, ангустифолин. Всего выявлено более 150 хинолизиновых алкалоидов [1, 2, 3], многие из которых видоспецифичны. В люпине желтом содержится также алкалоид индольной группы грамин [4].

Содержание алкалоидов в семенах дикорастущих и культивируемых горьких люпинов колеблется от 3,5 до 8% [1, 5]. Эти образцы используются для сидеральных целей. Образцы со средним содержанием алкалоидов от 0,1 до 0,3% относятся к кормовым [6]. При содержании в семенах 0,02% алкалоидов и менее образцы могут быть рекомендованы для употребления в пищу человеку, в кондитерской и другой промышленности [7]. Абсолютно безалкалоидных генотипов люпина не обнаружено.

Уровень алкалоидности растений и семян образцов люпина различного происхождения может изменяться в 2–6 раз в зависимости от внешних и внутренних условий [1, 3, 8, 9], этапа онтогенеза [1, 10]. Так, в фазу ветвления, когда фотосинтетическая деятельность растений особенно высока, наибольшее количество алкалоидов содержится в молодых листьях [1, 11]. В фазу цветения растений происходит интенсивный отток этих веществ из листьев и стеблей к местам плодообразования, где отмечается максимальное количество алкалоидов к началу созревания бобов. По мере формирования семян продолжается дальнейшее перемещение алкалоидов в них [12]. В последние годы высказывается предположение, что часть алкалоидов может быть синтезирована и в семенах [11]. В ходе онтогенеза растения изменяется также качественный состав алкалоидного комплекса. Так, в фазу начала ветвления растений гидроксилупанина содержится несколько больше, чем люпанина. В фазах цветения и образования бобов на растении в листьях, а

затем и в семенах количественно преобладает люпанин [10]. Показано селективное перемещение алкалоидов группы хинолизидина по флоэме, поскольку компонентный состав алкалоидного комплекса в листьях богаче, чем в семенах [11].

Количество алкалоидов в семенах и растениях люпина зависит от содержания отдельных компонентов комплекса. Однако связь эта может быть различной. У диких линий люпина узколистного показан высокий положительный корреляционный коэффициент между содержанием 13-гидроксилупанина и ангустифолина и общей алкалоидностью растений. Среди возделываемых сортов из реестра Польши такое же влияние на количество алкалоидов оказывал люпанин [5].

По данным [6, 8], чем меньше алкалоидов суммарно находится в образце люпина, тем меньше содержится в нем компонентов алкалоидного комплекса. Однако этот факт требует уточнения, поскольку показано различие в лимитах детекции для каждого отдельного алкалоида [13, 14] в зависимости от метода исследования. Так, нижние пределы детекции и количественного определения люпанина при использовании газой хроматографии составляют 2 мкг/мл и 3 мкг/мл, соответственно [13], лимиты по этим же показателям с использованием гель-электрофореза [14] — 7,9 мкг/мл и 8,12 мкг/мл, соответственно. Сходные результаты отмечены для спартеина, 13 α -гидроксилупанина, ангустифолина. В связи с этим возможны некоторые различия в характеристике состава алкалоидного комплекса у сортов разных видов люпина вследствие неполной детекции. Также методы экстракции и используемые системы растворителей оказывают влияние на степень и полноту выделения отдельных компонентов алкалоидного комплекса и, как следствие, общего суммарного содержания алкалоидов. Поэтому обоснована необходимость стандартизации методик определения алкалоидности растений разных видов люпина.

В современной хемосистематике алкалоиды люпина ограниченно используются в качестве хемотаксономических маркеров видов люпина из-за указанных причин [1], однако для паспортизации образцов, защиты авторских прав и выбора направления использования сорта,

несомненно, имеют значение наряду с морфологическими, биохимическими, молекулярно-генетическими и другими характеристиками.

Биосинтез алкалоидов группы хинолизидина

По современным представлениям биосинтез алкалоидов начинается с декарбоксилирования лизина до кадаверина при участии фермента лизиндекарбоксилазы, изученной у люпина узколистного [2, 11]. Особенно интенсивно происходит этот процесс в хлоропластах молодых листьев при их освещении [1]. Далее кадаверин подвергается окислительному дезаминированию медь-аминоксидазой (CuAO) с получением 5-аминопентанала, образующего в результате спонтанной циклизации основание Шиффа (рис. 1). В дальнейшем происходит образование люпинина, спартеина, мультифлорина, люпанина и, после этерификации, их производных [11]. По данным автора [2], конечными продуктами биосинтеза алкалоидов являются эфиры алкалоидов, и именно они накапливаются в семенах, что требует дополнительного изучения.

К настоящему времени этапы образования алкалоидов изучены не полностью. Известно только о некоторых ферментах процесса биосинтеза — лизиндекарбоксилазе и нескольких трансферазах [2, 11], одна из которых характерна только для хинолизидинсодержащих растений [2], клонирована и охарактеризована — это (–)-13- α -гидроксимультифлорин/(+)-13- α -гидроксилупанин-О-тиглоилтрансфераза.

При изучении экспрессии генов, кодирующих трансферазы биосинтеза алкалоидов, показано, что начальные этапы этого процесса происходят в листьях [11]. Последующая этерификация алкалоидов происходит в корнях растений и гипокотилиях проростков, где показана экспрессия генов трансфераз, ацетилирующих 13- α -гидроксилупанин и 13- α -гидроксимультифлорин [1, 11].

При изучении высоко- и низкоалкалоидных форм люпина желтого, белого, узколистного было установлено, что уровень лизина видоспецифичен [15], однако при этом заметна общая тенденция: у изученных высокоалкалоидных форм содержание лизина ниже, чем у сладких образцов [15, 16], что, видимо, свя-

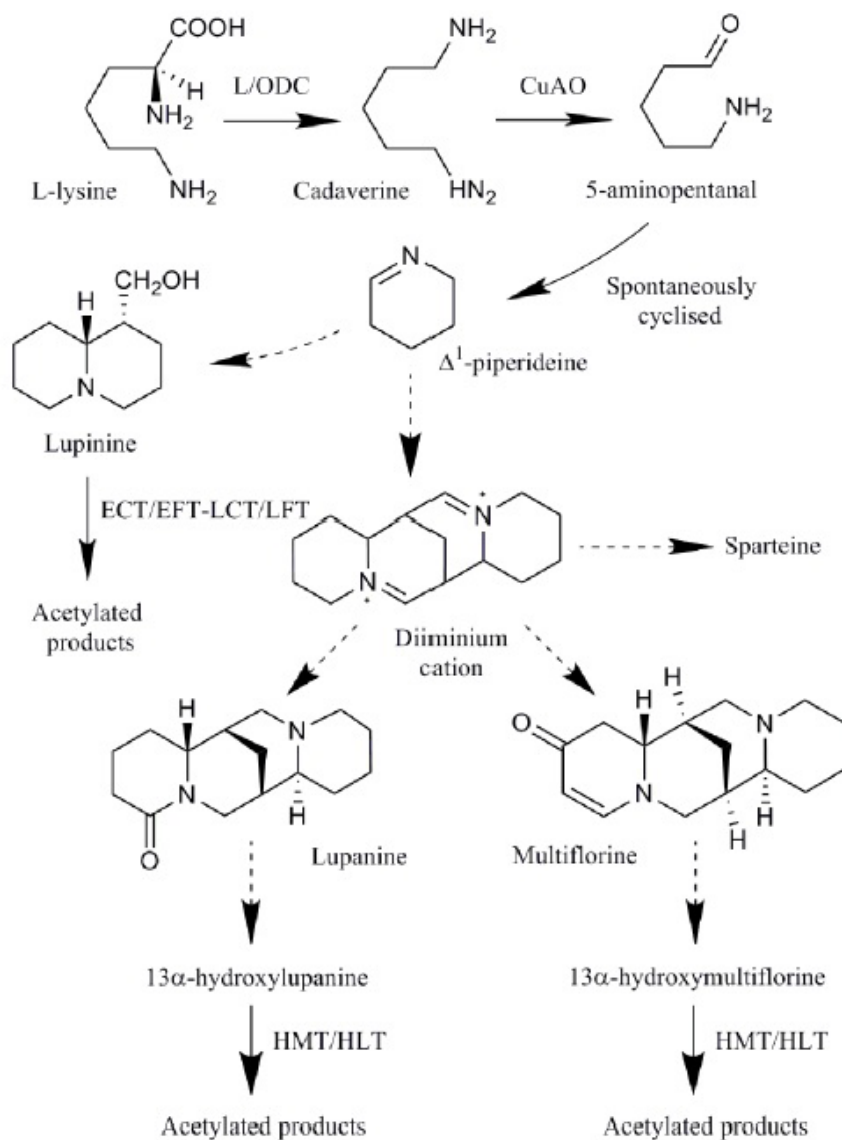


Рис. 1. Схема биосинтеза хинолизидиновых алкалоидов люпина [2]

L/ODC — лизин/орнитиндекарбоксилаза, ECT/EFT-LCT/LFT — p -кумароил-CoA/ферулоил-CoA: (+)-эпилюпинин/(–)-люпинин О-кумароил/ферулоилтрансфераза, HMT/HLT — тиглоил-CoA:(–)-13 α -гидроксимультифлорин/(+)-13 α -гидроксилупанин О-тиглоилтрансфераза, CuAO — медь-аминоксидаза

зано с блокированием синтеза алкалоидов на этапе образования кадаверина из лизина. Сходная тенденция отмечена для некоторых других аминокислот: орнитина и аргинина [15].

Зная основные этапы биосинтеза алкалоидов и гены, отвечающие за этот процесс, можно предположить, что образцы с высоким содержанием алкалоидов не несут мутантных генов этого признака. В то же время повышение количества люпанина и других основных алкалоидов по сравнению с их эфирами среди возделываемых пищевых и кормовых сортов указывает на мутации в генах, детерминирую-

щих процесс этерификации. Возможен также вариант сорта, когда мутация затрагивает ген лизиндекарбоксилазы, что приводит к резкому снижению всех компонентов алкалоидного комплекса в равной степени и повышению количества лизина как универсального предшественника всех хинолизидиновых алкалоидов.

Учитывая данные о биосинтезе алкалоидов и генетике алкалоидности в селекции люпина, можно получить образцы не только с желаемым количеством алкалоидов, но и с интересующим исследователя компонентным составом алкалоидного комплекса.

Генетика признака алкалоидности

Высокое содержание алкалоидов — доминантный признак по отношению к низкому их содержанию. Безалкалоидные и малоалкалоидные формы возникают в виде рецессивных мутаций генов, что приводит либо к частичному снижению уровня алкалоидности, либо к практически полному отсутствию признака [17]. Алкалоидность — сложный количественный признак, в его формировании принимают участие не менее 6 генов [17, 18], что было экспериментально обосновано при построении комплементационной карты — схемы при гибридизации 12 сортов люпина узколистного разного происхождения [18]. Гены, отвечающие за каждый из этапов процесса биосинтеза алкалоидов, наследуются независимо. При комплементарном эффекте взаимодействия неаллельных генов в F1 происходит реверсия к дикому генотипу (восстановление нормального биосинтеза алкалоидов) при гибридизации различающихся по генам низкой алкалоидности условно безалкалоидных сортов. В F2 возможно образование двух типов гибридов: безалкалоидных, как и родители, и алкалоидных. Пары родительских безалкалоидных сортообразцов, дающие алкалоидное потомство, были названы «несовместимыми» [17, 18]. Такое восстановление синтеза алкалоидов у гибридов при скрещивании неаллель-

ных мутантов было названо «перерождение» кормовых форм в алкалоидные.

При скрещивании образцов с минимальной алкалоидностью возможно также частичное восстановление биосинтеза алкалоидов (в 2–5 раз по сравнению с исходными сладкими формами). Это свидетельствует о том, что часть неаллельных генов алкалоидности гибрида комплементарна, за счет чего и происходит неполное восстановление биосинтеза алкалоидов. Некоторые исследователи полагают, что ген малоалкалоидности таких сортов представляет собой «блок» генов (совокупность генов) по этому признаку [8, 18, 19].

Первые шаги в современной селекции кормового люпина связаны с фактом обнаружения Зенгбушем среди высокоалкалоидных растений спонтанных мутантов с низким содержанием алкалоидов [20]. При изучении этого признака позднее было установлено, что алкалоидность контролируется у этого образца геном, который был назван по названию образца, у которого он был обнаружен, — Stamm 411 [20], а позднее был идентифицирован как *iucundus*. Полный список генов пониженного содержания алкалоидов люпина узколистного в сравнении с высокоалкалоидными формами, их фенотипическое проявление и данные об авторах представлены в табл. 1.

Образцы с геном *tantalus* были получены путем радиационного мутагенеза [22]. По

Таблица 1

Гены, детерминирующие алкалоидность люпина узколистного

Название	Символ	Фенотипическое проявление, алкалоидность, % ОСВ	Выделен впервые у образца, автор
depressus	dep	снижение уровня алкалоидности	Stamm 14 (выделил Федотов, 1936, описал Troll, 1955) [21, 8]
esculentus	esc	снижение уровня алкалоидности	Stamm 415 (Sengbusch, 1942) [21, 8]
iucundus	iuc	снижение уровня алкалоидности	Stamm 411 (Sengbusch, 1942) [21, 8]
tantalus	tant	снижение уровня алкалоидности	№6034 и №6035 [22]
angustifolius	a1, a2, a3	аллель a1 — 0,06 %, аллель a2 — до 0,02%, аллель a3 — до 0,04%	[23]

данным [24, 25], ген малоалкалоидности *iucundus* расположен в 7 группе сцепления. По другим данным [26], этот ген находится в девятой группе сцепления. Данный факт указывает на противоречивые данные, полученные при анализе этих генов алкалоидности, и требует дополнительных исследований.

Гены люпина желтого, отвечающие за снижение количества алкалоидов, представлены в табл. 2. Следует упомянуть, что гены хинолизидиновых алкалоидов и индольного алкалоида грамина наследуются независимо друг от друга [10].

У люпина белого низкая алкалоидность также определяется мутантными генами [22, 26, 27], наследуемыми независимо (табл. 3).

В ходе последних исследований люпина узколистного идентифицирован регион, охватывающий ген *iucundus* и отвечающий как за общее содержание алкалоидов, так и за уровни ангустифолина и 13-гидроксилюпанина, но не люпанина [25]. Ближайшим маркером, сцепленным с указанным регионом, является именно маркер гена *iucundus*. Следовательно, данный ген не блокирует начальные этапы биосинтеза алкалоидов люпина. Локус

Таблица 2

Гены, детерминирующие алкалоидность люпина желтого [21]

Название	Символ	Фенотипическое проявление, алкалоидность, % ОСВ	Выделен впервые у образца, автор
amoenus	am	до 0,013%	Stamm 80 (Sengbusch, 1928)
dulcis	dul	до 0,05%	Stamm 8 (Sengbusch, 1928)
liber	lib	до 0,01%	Stamm 102 (Sengbusch, 1928)
V351	V351	снижение уровня алкалоидности	Stamm V351 (Lamberts, 1955)
		наличие грамина	Jach, 1975 [4]

Таблица 3

Гены, детерминирующие алкалоидность люпина белого [21]

Название	Символ	Фенотипическое проявление
mitis	mit	максимальное уменьшение содержания всех компонентов алкалоидного комплекса в равной степени [27]
pauper	pau	снижение уровня алкалоидности
nutricius	nut	повышении концентрации дегидролюпанина и снижении концентрации дегидроангустифолина [27, 28]
exiguus	exi	снижение уровня алкалоидности
primus	prim	снижение уровня алкалоидности
quintus	q	снижение уровня алкалоидности
reductus	red	снижение уровня алкалоидности
suavis	sua	снижение уровня алкалоидности

13-гидроксильюпанина наиболее тесно связан с маркером DAWA504.280 [25], из чего может следовать разобщенность этих регионов и, как следствие, указание на действие разных генов малоалкалоидности.

Аллели гена *dulcis* люпина желтого, видимо, связаны с образованием основных структурных алкалоидов, поскольку соотношение лизина и кадаверина у сладких сортов с этим мутантным геном и у горьких диких образцов, доминантных по этому гену, не изменяется, как и активность ацилтрансфераз хинолизидиновых алкалоидов [11].

Состав алкалоидного комплекса люпина белого при наличии генов малоалкалоидности также различен [21]. Ген *pauper*, видимо, связан с регуляцией начальных этапов биосинтеза алкалоидов [27]. Повышение концентрации дегидролюпанина и снижение уровня дегидроангустифолина у образцов с геном *nutricius* указывает на иной фермент цепи биосинтеза [27].

Наличие рецессивных аллелей, изменяющих качественный и количественный состав компонентов комплекса алкалоидов, показано и для других видов люпина. Так, аллель *mutal* у люпина изменчивого (*L. mutabilis* L.), полученный в результате химического мутагенеза, в гомозиготном состоянии снижает количество алкалоидов менее 2,0%, сдвигая при этом соотношение (люпанин+спартеин):

(17-оксоспартеин+4-гидроксильюпанин) в сторону двух последних [29].

На основании ряда исследований были предложены [цит. по 8] генотипы сортов люпина узколистного, различающихся по алкалоидности, которые могут иметь следующие формулы (табл. 4).

Из данных таблицы следует, что, кроме дикого типа в геномах культурных сортообразцов, имеется хотя бы один мутантный ген, блокирующий синтез алкалоидов [8, 17]. Рекомбинация таких генотипов между собой в правильно подобранной системе скрещиваний может привести к выщеплению образцов с минимальным содержанием алкалоидов (следы) [6, 8, 17, 18], что желательно для современной селекции.

Гены алкалоидности плеiotропно действуют на формирование хозяйственно ценных признаков [21, 22, 25, 30]: локусы, отвечающие за алкалоидность люпина узколистного, тесно коррелируют с локусами, контролирующими количество семян и массу их на растении [25]. Было показано, что наличие рецессивных генов малоалкалоидности у люпина отрицательно коррелирует с жизнеспособностью и продуктивностью растений [31]. Так, ген *iucundus* в рецессивном состоянии снижает фертильность пыльцы растений люпина узколистного [22], вызывает предрасположенность к заболеваниям и растрескиваемости бобов. При этом ген *tantalus* не снижает про-

Таблица 4

Предполагаемые формулы генотипов, различающихся по алкалоидности сортов люпина узколистного [8]

Название генотипа, ген малоалкалоидности	Алкалоидность (фенотип), % ОСВ	Формула генотипа
Дикий тип	2% и более	++ ++ ++ ++ ++
Сидеральный тип 1	около 1,2%	alc1alc1 ++ ++ ++ ++
Сидеральный тип 2	около 0,6%	++ alc2alc2 ++ ++ ++
Белорусский 155	около 0,2%	++ ++ ++ alc4alc4 ++
Штамм 411 (<i>iucundus</i>)	0,01–0,06%	alc1alc1 alc2alc2 ++ ++ alc5alc5
Линия Б-аЛ1	до 0,01%	alc1alc1 alc2alc2 alc3alc3 ++ ++

Примечание. + — аллель дикого типа;

alc1-alc5 — мутантные гены, определяющие пониженное количество алкалоидов

дуктивности и жизнеспособности растений в отличие от других мутантных генов [31] и поэтому более желателен для использования в селекции [27]. Ген *dulcius* люпина желтого снижает урожайность в 1,5 раза [30]. Наличие гена *V 351* в рецессивном состоянии у этого же вида приводит к резкому снижению жизнеспособности растений, в связи с чем носителей этого гена не рекомендуется использовать в селекции [21, 30], что следует учитывать в дальнейшем подборе пар родительских форм.

Наличие разных генов алкалоидности в генотипе каждого конкретного образца можно установить как методами классической генетики при анализирующих скрещиваниях, так и в результате молекулярно-генетического анализа при применении праймеров к каждому конкретному гену. Сегодня идет интенсивный поиск маркеров и конструирование праймеров к генам алкалоидности люпина узколистного, белого, желтого. Однако не все типы маркеров могут быть применены к культуре люпина в решении данной проблемы. Так, изучение «горьких» сортов люпина узколистного с применением RAPD-анализа не выявило полиморфизма ДНК в зависимости от фенотипа сорта, что делает непригодными данный тип маркеров для анализа популяций [2]. Только для гена *iucundus* люпина узколистного осуществлен сегодня синтез ISSR-праймера *iucLi* [24], амплификоны с применением ко-

торого выявляют полиморфизм гена по этому признаку (описано не менее 6 бэндов) [24, 32]. Такое большое количество описанных аллелей, возможно, связано с тем, что по данным [6, 8] этот ген малоалкалоидности представляет собой блок из 4-х рецессивных генов с соответствующим распределением их при гибридизации. У люпина белого осуществлен синтез праймера к гену *pauper*, выявляющего генотипы с горьким (207 п. о.) и сладким (209 п. о.) аллелями [33]. При изучении алкалоидности листьев и семян у образцов стержневой коллекции люпина желтого выявлен ряд участков генома, связанных с биосинтезом люпинина, спартеина и грамина, что планируется использовать для разработки видоспецифичных праймеров к генам алкалоидности люпина желтого. В результате проведенного анализа 164 образцов люпина желтого показаны на карте хромосом регионы, связанные с морфологией корней и отвечающие за содержание алкалоидов и сахаров [34]. Ведется поиск генов биосинтеза алкалоидов, в частности гена лизиндекарбоксилазы и других ферментов, катализирующих образование алкалоидов [2, 35].

В результате анализа литературы по алкалоидности люпина нами выделены группы генотипов, у которых установлены разными авторами соответствующие гены малоалкалоидности (табл. 5).

По данным молекулярно-генетического анализа у образцов люпина узколистного Rancher,

Таблица 5

Образцы разных видов люпина, несущие гены малоалкалоидности

Вид люпина, ген, сорт	Источник
Люпин узколистный	
ген <i>iucundus</i> (<i>iuc</i>)	
<i>Современные австралийские сорта</i> : Belara, Chittick, Coromup, Danja, Geebung, Gungurru, Ilyarrie, Jenabillup, Jindalee, Kalya, Mandelup, Marri, Merrit, Moonah, Myallie, Quilinoch, Rencher, Tallerack, Tanjil, Unicrop, Uniharvest, Uniwhite, Wandoo, Warrah, Wonga, Yandee, Yorrell, 71 A07 B, 75 F261 и др.	[21, 24, 26, 32–37]
<i>Современные отечественные сорта</i> : Ашчадны, Миртан, Першацвет, ПЧБ9	[32, 36]
<i>Старые и современные сорта Австралии, Германии, Дании, Швеции, Польши, России</i> : Беньяконский 335, Брянский 1272, Вада-18, ГЛ-174-86/1, ДМ-антоциановый, Ладный, Ланедекс-1, Лаф-рбс-2, Мюнхенбергский синий II, Немчиновский 846, Северный 3, Blanko, Borre, Frost, Kazan, Myrzynek, Mirela, New Zealand Blue, Pflugs Blau Lupine, Fest, №4243, p22730 и др.	[8, 21, 37, 38]

Продолжение табл. 5

Вид люпина, ген, сорт	Источник
ген <i>tantalus</i> (tan)	
№ 6034, № 6035	[21, 24, 26]
ген <i>depressus</i> (dep)	
Stamm 14 (417)	[21, 37]
ген <i>esculentus</i> (esc)	
Stamm 415	[21, 27]
другие гены, не аллельные гену <i>iucundus</i> (iuc)	
ген сорта Белорусский 155	
Белорусский 155, Немчиновский кормовой 1, АПБ155	[6, 21, 37]
гены линии Б-аЛ1	
ДВС1, Б-аЛ1, Мужин, Муженек, Темно-зеленый, Ясны	[6, 21]
Люпин желтый	
ген <i>dulcius</i> (dul)	
Брянский 81, Брянский 27, Детер 38, Детер 122. Мотив 369, Проминь, Stamm 8 (Wt 97506)	[21]
ген <i>amoenus</i> (am)	
Stamm 80 (Wt 97507)	[21]
ген <i>liber</i> (lib)	
Stamm 8 (Wt 97506)	[21]
ген V 351	
Stamm 102 (Wt 97764)	[21]
Люпин белый	
ген <i>rauper</i> (rau)	
Andromeda, Astra, Esta 1-15, Hamburg, Ida, Kraftquell, Kiev Mutant, Luxor, Magna, Minibean, Minori, Pflugs Gela, Rosetta, Tiptop, Ultra, Vladimir, WALAB 2008, Борки, Владимир, Диета, Олежка, Пищевой, Севериновский, Синий парус	[28, 33]
ген <i>exiguus</i> (exi)	
Neuland, Start	[33]
ген <i>nutricius</i>	
Nahrquell	[33]

Брянский 1121, Сидерат 38, p22823, p22868 отсутствует низкоалкалоидный аллель *iucundus* [32]. Поскольку современный австралийский сорт Yorrell, анонсированный как носитель доместизируемого аллеля *iucundus*, является малоалкалоидным, но в молекулярно-генетическом анализе не показывает аллеля «сладкого» генотипа, этими же авторами был сделан вывод о возможном изменении последовательности ДНК у образца Yorrell в области амплификации.

На основе характеристики гибридов первого поколения, полученных при скрещивании родительских компонентов с низким содержанием алкалоидов, нами был составлен список образцов, являющихся аллельными и неаллельными мутантами по этому при-

знаку. Совокупные данные представлены в табл. 6 и 7.

Серия проведенных авторами скрещиваний позволила установить образцы с мутантным геном *iucundus*, в потомстве которых происходит частичное восстановление алкалоидности. Это комбинации Миртан (*iuc*) x Frost (*iuc*), Frost (*iuc*) x Yorrell (*iuc*), Frost (*iuc*) x Wonga (*iuc*), Frost (*iuc*) x Fest (*iuc*), Fest (*iuc*) x Frost (*iuc*), Fest (*iuc*) x Першацвет (*iuc*), Першацвет (*iuc*) x Frost (*iuc*), Першацвет (*iuc*) x Fest (*iuc*) [36]. Хотя у этих форм и есть мутантный ген, восстановление алкалоидности происходит до малоалкалоидных форм, что свидетельствует о присутствии других генов, отвечающих за биосинтез алкалоидов, либо о «блоке» генов

Таблица 6

«Совместимые» по признаку низкой алкалоидности сорта люпина узколистного и люпина желтого

Вид люпина, сорт (материнская форма)	Совместимые сорта (отцовская форма)
Люпин желтый	
Weiko	Белорусский 6, Кормовой 3 [10]
Надежный	Престиж, Демидовский [36]
МЛ	БСХА-13, БСХА-19, Фотон [36]
БСХА-13	Фотон, Престиж, Демидовский [36]
Люпин узколистный	
Денлад	Ланедекс 3, Надежда [39]
Данко	Миртан, Unicrop [39]
Ладенекс 3	Надежда, Денлад, L-300 [39]
L-300	Надежда, Денлад, Ланедекс 3 [39]
Надежда	Ланедекс 3, Денлад, L-300 [39]
Натальевский 5	Unicrop [40]
Северный 3	Стадолиценский Л-610, Беляконский розовый сладкий, Borre [10,18]
Немчиновский 846	La II, Стадолиценский Л-610 [37]
Северный 3	Unicrop [37]
Мужин	Unicrop [37]

Таблица 7

«Несовместимые» по признаку низкой алкалоидности сорта люпина узколистного и люпина желтого

Вид люпина, сорт, материнская форма	Несовместимые сорта, отцовская форма
Люпин желтый	
Престиж	МЛ [36]
Люпин узколистный	
Оборницкий	Белорусский 155, Мирзин, Мирзинец, Вогге [10, 18]
Данко	Кучный [40]
Мужин	Оборницкий [18]
Белорусский 155	Оборницкий, Вогге, Быстрорастущий сладкий, Мирзин, Мирзинец, Unicrop, Харчовы, Немчиновский кормовой 1 [10, 18, 37, 40]
Немчиновский 846	Брянский 35, Unicrop, Uniwhite, La 8 [10, 21, 36, 40]

iuscundus. Следует отметить, что во всех вариантах скрещивания сорта Fest с образцами с геном *ius* нами было получено в F1 алкалоидное потомство. Сходное формирование признака алкалоидности (малоалкалоидное потомство в первом поколении) было показано и для образца Брянский 1121 при скрещивании его с сортом Першацвет (*ius*) [36]. Это указывает на «несовместимость» этих двух указанных образцов [10, 21, 32, 36].

В связи с тем, что количественное содержание компонентов комплекса может варьировать в зависимости от генотипа, становится понятным различие в степени токсичности суммарного комплекса алкалоидов, что и было показано в работе [6, 8]. Всего выделяют по соотношению отдельных алкалоидов три типа комплекса:

а) ангиустифолиус, с максимальным содержанием наиболее токсичного люпанина (50–80%) и с меньшим количеством двух других основных алкалоидов (13-гидроксилупанина 10–20% и ангиустифолина 5–20%). Характерен для большинства изученных сортов люпина узколистного;

б) линифолиус (алкалоидов 24%, 9% и 67%, соответственно). Описан у люпина льнолистного;

в) опсиантус (6–12%, 34–72% и 16–47% этих же алкалоидов, соответственно, а также ново-

го алкалоида мультифлорина 1–13%). Данный комплекс описан у люпина поздноцветущего, отнесенного, как и льнолистный, к виду люпин узколистный [6, 8]. Сходный качественный состав имеют образцы Белорусский 155 и некоторые другие.

Алкалоидные комплексы двух последних групп (линифолиус и опсиантус) менее токсичны и представляют собой рецессивные мутации по отдельным локусам (табл. 8) [8, 40] с доминированием высокотоксичного комплекса ангиустифолиус.

Для отдельных сортов (Gungurru, Warrah) показана малая токсичность стерни при выпасе скота [41, 42], что может быть связано либо с малым количеством алкалоидов в зеленой массе люпина, либо с низкой степенью токсичности алкалоидного комплекса указанных сортов. Данный факт требует дополнительного исследования с помощью современных методов хроматографического разделения алкалоидов.

Изучив историю сортов [21, 41, 42 и др.] и приняв во внимание филогенетические исследования, проведенные с применением биохимических маркеров [43], можно предположить, что образцы, имеющие в родословной ген из образца Stamm 411, будут содержать максимальное количество генов малоалкалоидности,

Таблица 8

Формулы генотипов разных по степени токсичности алкалоидных комплексов люпина узколистного [8, 40]

Состав алкалоидного комплекса	Формула генотипа
Высокотоксичный (ангустифолиус)	aco1 ⁺ aco2 ⁺
Слаботоксичный Льюлистный (линифолиус)	aco1 aco1
Слаботоксичный Ржавый — Ltac (опсиантус)	aco2 aco2

что предпочтительно для дальнейшей селекции культуры в направлении снижения синтеза алкалоидов.

В процессе доместикации разных видов люпина в генотипах происходило накопление генов с фенотипическим проявлением желаемых признаков (устойчивость к заболеваниям, нестрессиваемость бобов, отсутствие твердокаменности, нетребовательность к яркой окраске и др.). Для ряда сортов отечественной и зарубежной селекции раскрыты родословные [21, 41, 42] с указанием накопленных в генотипе рецессивных генов. Однако для признака «низкая алкалоидность» у люпина узколистного информация о распространении гена в процессе селекции показана только для гена *iucundus* [21]. Можно предположить, что в ходе длительной селекции люпина происходило накопление разных генов малоалкалоидности в генотипе, что получило свое отражение в создании образцов со стабильно минимальными количествами этих токсичных вторичных метаболитов. Анализ данных литературы позволил предположить, что у образца Frost есть только 1 доместичированный ген малоалкалоидности — *iucundus* [9, 37], у сортов Гелена и Немчиновский 846 показано наличие 2 генов пониженного содержания алкалоидов [6]. Сорта Illyarrie, Эмир, Митан, Дикаф-14 несут по 3 гена низкого содержания алкалоидов в своих генотипах, что получает у них свое выражение в уровне алкалоидности менее 0,1% [43, 44]. У образцов Yorrell и Ладный присутствуют 4 рецессивных гена, контролирующих алкалоидность. Уровень алкалоидов у них не превышает 0,06% даже в неблагоприятных по климатическим условиям годы.

По австралийской программе селекции сорта люпина белого несут в себе только ген малоалкалоидности *pauper* с целью исключения переопыления неаллельных образцов [21].

Снижение количества алкалоидов в семенах и зеленой массе люпина в ходе длительной селекции привело к созданию кормовых и условно пищевых сортов [45]. Это дополнительно свидетельствует о накоплении рецессивных генов малоалкалоидности в генотипах таких сортов. Селекция же сидеральных форм была направлена на отбор алкалоидных образцов в связи с более высокой устойчивостью их к действию стрессорных факторов различной природы и накоплением большей биомассы для использования их на зеленое удобрение независимо от содержания алкалоидов.

Биологическое действие алкалоидов люпина

Алкалоиды люпина обладают широким биологическим действием. Так, известен их токсичный эффект [1]. Не обладая мутагенным действием [46] и не индуцируя апоптоз опухолевых клеток [1], алкалоиды тем не менее являются сильнодействующими веществами с нейротоксическим и кардиотоксичным эффектами. Отдельные из них токсичны для крупного рогатого скота [47], вызывают гибель взрослых животных. Анагирин действует тератогенно. Для других алкалоидов на проростках лука репчатого показан антимуагенный эффект [46], видимо, основанный на большом количестве сопряженных связей в их структурной формуле. Возможно, поэтому спартеин не окисляется пероксидазой из хрена, оставаясь стабильной молекулой [48]. Также показано

влияние алкалоидов группы хинолизидина на секрецию инсулина β -клетками островков Лангерганса поджелудочной железы человека [цит. по 1] и прирост массы тела у потомства крольчих, получающих корм с добавкой люпиновой муки, что может быть связано со стимуляцией лактации за счет окситоциноподобного действия пахикарпина и спартеина [49].

Степень токсичности алкалоидов люпина различается. Наиболее токсичны люпанин и спартеин, степень токсичности эфиров алкалоидов на порядок ниже, а соли алкалоидов менее токсичны, чем чистые препараты [1, 7 и др.]. Показано, что летальная доза алкалоида для животных зависит от вида животного, возраста, а также от способа введения вещества, что указывает на разрушение алкалоидов либо в желудочно-кишечном тракте, либо в воротной системе печени [1, 11].

Установлено, что состав алкалоидного комплекса также влиял на развитие личинок тли, особенно снижали количество вредителей этого вида 13-гидроксилупанин и 13-тиглоилоксилупанин. Генотипы с максимальным содержанием указанных алкалоидов могут быть источниками устойчивости к тле при селекции на безалкалоидность [1, 50]. Именно эти алкалоиды менее токсичны для остальных животных, что также представляет интерес для новой модели будущего сорта. Возможно, эволюция тли происходила на образцах с высокотоксичным комплексом алкалоидов, что позволило насекомым отдельных видов выработать устойчивость, связанную со способностью тли выделять алкалоиды во внешнюю среду секретирующими железами [1].

В отношении фитопатогенных грибов также показана различная степень токсичности алкалоидов, при этом *Fusarium oxysporum*, возбудитель трахеомикозного увядания, оказался практически нечувствителен к токсичному действию люпинина и спартеина, в то время как рост мицелия и спорообразование у возбудитель корневых гнилей *F.avenaceum* ингибировался этими алкалоидами [51] в системе *in vitro*, что следует учитывать при создании сортов, устойчивых к возбудителям разных форм фузариоза.

Действие многих алкалоидов на организм человека и животных хорошо изучено, а их роль

в жизни самих растений исследована пока недостаточно. По современным представлениям наиболее вероятна теория «защитной» роли алкалоидов [31, 44, 46, 48], что указывает на необходимость присутствия этих вторичных метаболитов в составе растений. Так, показано, что в растениях люпина в течение нескольких часов после нанесения травмы растению происходит резкое повышение уровня алкалоидов в 4 раза [1]. Установлено, что низкая алкалоидность снижает устойчивость к вредителям и переносимым ими вирусам [28]. В то же время анализ вторичных метаболитов в зеленой массе инфицированных и неинфицированных растений люпина узколистного показал, что в ответ на заражение спорами гриба *Colletothrychum lupini* не изменяется уровень алкалоидов и жирных кислот, как и при обработке метаболитами патогена. Количество же флаваноидов (в частности, фитоалексинов) значительно изменяется в сторону увеличения у пораженных растений [52]. Возможна и запасающая роль алкалоидов в качестве источников азота [1].

Вторичные метаболиты — аттрактанты для насекомых и, возможно, сигнальные соединения [1], регулирующие ряд защитных и физиологических процессов в растениях люпина, поэтому понятна необходимость присутствия алкалоидов в зеленой массе и семенах люпина. В связи с вышесказанным сегодня одна из рассматриваемых перспективных моделей сорта люпина — образцы с высоким содержанием алкалоидов в листьях и низким — в семенах, благодаря чему растения люпина смогут повысить устойчивость к неблагоприятным факторам среды и сформировать семена, которые можно будет использовать на корм скоту. Поэтому сегодня исследования направлены не только на изучение механизмов действия генов при биосинтезе, но также значительное внимание уделяется транспортным системам, благодаря которым будет возможен либо селективный перенос малотоксичных алкалоидов из листьев в генеративные органы, либо алкалоиды в семена транспортироваться не будут [1, 2, 11]. Вторая из возможных моделей будущего сорта — растения с большим содержанием низкотоксичных алкалоидов, что также сделает возможным употребление зеленой массы растений и семян в качестве кормовой культуры [21, 43, 46].

Заключение

За более чем 100 лет изучения алкалоидности люпина получены знания о составе алкалоидов, их биологическом действии, роли в растениях, генетике этого признака. Однако многие вопросы остаются открытыми, что является предметом дальнейших исследований как фундаментального, так и прикладного характера в плане реконструирования новых геномов у разных видов люпина для разнонаправленного их использования. Эволюция геномов растений разных видов люпина в культурном возделывании привела к реконструированию разных генотипов относительно отдельных генов алкалоидности и их взаимодействия в процессе онтогенеза растений. Поэтому при характеристике образцов разных типов коллекций люпина обязательно должен быть указан в их паспортах этот селекционно значимый признак.

Список использованных источников

1. Wink, M. A summary of 30 years of research in lupins and lupin alkaloids / M. Wink // *Lupin crops – an opportunity for today, a promise for the future: Proc. of the 13th International Conference*, 6–10 June 2011, Poznan, Poland. – 2011. – P. 225–228.
2. Bunsupta, S. et al. Quinolizidine alkaloid biosynthesis: recent advances and future prospects / S. Bunsupta, M. Yamazaki, K. Saito // *Frontiers in Plant Science / Plant Metabolism and Chemodiversity*. – 2012. – Vol. 3. – Art. 239.
3. Методические рекомендации по использованию зерна бобовых культур в кормлении сельскохозяйственных животных / сост.: В.М. Голушко [и др.]. – Жодио, 2003. – 24 с.
4. Козлова, Л.С. [и др.]. Изменчивость кормовых форм люпина желтого по содержанию хинолизидиновых алкалоидов и грамина / Л.С. Козлова, В.С. Анохина, А.Г. Купцова // *Изменчивость и отбор*. – Минск, 1980. – С. 165–169.
5. Quantitative and qualitative content of alkaloids in seeds of a narrow-leaved lupin / K.A. Kamel [et al.] // *Genet. Resour. Crop. Evol.* – 2015. – Vol. 3. – P. 239–245.
6. Купцов, Н.С. Люпин – генетика, селекция, гетерогенные посевы / Н.С. Купцов, И.П. Такунов. – Брянск: Клиновская городская типография, 2006. – 576 с.
7. Naganowska, B. [et al.] *Lupin Crops – an opportunity for today, a promise for the future: Proc. of the 13th International Lupin Conference*, 6–10 June 2011, Poznan, Poland / B. Naganowska, P. Kachlicki, B. Wolko (eds). – International Lupin Assotiation, Canterbury, New Zealand, 2011. – 256 p.
8. Генетические основы селекции растений. В 4 т. Т.2. Частная генетика растений / науч. ред. А.В. Кильчевский, Л. В. Хотылева. – Минск: Беларус. навука, 2010. – 579 с. – ISBN 978-985-08-1127-1.
9. Polish Lupinus Collection Database [Electronic resource]. – Mode of access: www.igr.poznan.pl/resources/Lubin. – Date of access: 28.03.2017.
10. Козлова, Л.С. [и др.]. Генетика признака алкалоидности в связи с проблемой создания кормовых сортов люпина / Л.С. Козлова, В.С. Анохина // *Проблемы и перспективы селекции зерновых, зернобобовых и кормовых культур в 12 пятилетке: сб. тез. докл. конф.* – 1985. – С. 61–63.
11. Quinolizidine Alkaloid Biosynthesis in Lupins and Prospects for Grain Quality Improvement / K. M. Frick [et al.] // *Frontiers in Plant Science*. – 2017, January. – Vol. 8, Art. 87. – Mode of Access : www.frontiersin.org. – Date of access: 11.08.2017.
12. Рыбникова, В.А. Особенности завязывания семян люпина узколистного / В.А. Рыбникова // *Сб. трудов 2-го Съезда ВОГиС, Санкт-Петербург, 1–5 февраля 2000*. – СПб., 2000. – С.64–65.
13. Quinolizidine alkaloids in seeds of lupin genotypes of different origins / G. Boschini [et al.] // *J. Agric. Food Chem.* – 2008. – Vol. 56. – P. 3657–3663.
14. Ganzera, M. [et al.]. Determination of quinolizidine alkaloids in different *Lupinus* species by NACE using UV and MS detection / M. Ganzera, A. Kruger, M. Wink // *J. Pharm. Biomed. Anal.* – 2010. – Vol. 53. – P. 1231–1235.
15. Markiewicz, M. [et al.]. Preliminary studies of correlation between amines contents in lupin seeds / M. Markiewicz, V.R. Villanueva, A. Santerre // *Proc. of 5th International Lupin Conference*, Poznan, Poland, July 5–8 1988. – P. 307–314.
16. Nowacki, E. Inheritance and biosynthesis of alkaloids in lupin / E. Nowacki // *Genetica Po-*

lonica. – 1963. – Vol. 7, № 2. – P. 34–48.

17. Турбин, Н.В. [и др.]. Генетическое изучение алкалоидности у люпина / Н.В. Турбин, В.С. Анохина // Селекция, семеноводство и приемы возделывания люпина. – Орел, 1974. – С. 109–121.

18. Турбин, Н.В. [и др.]. Генетическая комплементация у межсортных гибридов люпина / Н.В. Турбин, В.С. Анохина, Л.Е. Садыкова. – В сб. «Проблемы экспериментальной генетики». – Минск, «Наука и техника». – 1972. – С. 42–52.

19. Люпин: биология, селекция и технология возделывания: учебное пособие для студентов агрономических специальностей / Г.И. Тарануха. – Горки: БГСХА, 2001. – 112 с.

20. Sengbusch, R. Süßlupinen und Öllupinen. Die Entstehungsgeschichte einiger neuer Kulturpflanzen / R. Sengbusch // Landwirtschaftliche Jahrbücher. – 1942. – Vol. 91. – P. 719–880.

21. Анохина, В.С. [и др.]. Люпин: селекция, генетика, эволюция / В.С. Анохина, Г.А. Дебелый, П.М. Коноров. – Минск: БГУ, 2012. – 271 с.

22. Sachow, F. Ein neues Gen für Alkaloidarmut bei *Lupinus angustifolius* / F. Sachow // Zuchter. – 1967. – Vol. 37, № 1. – P. 35–38.

23. Mikolajczyk, J. Lubin / J. Mikolajczyk. – Poznan, 1966. – 256 с.

24. Development of a DNA marker tightly linked to low-alkaloid gene *iucundus* in narrow-leaved lupin (*Lupinus angustifolius* L.) for marker-assisted selection / X. Li [et al.] // Crop and Pasture Science. – 2011. – Vol. 62. – P. 218–224.

25. Kroc, M. [et al.]. QTL analysis of seed yield components and alkaloid content in *Lupinus angustifolius* / Lupin Crops – an opportunity for today, a promise for the future: Proc. of the 13th International Lupin Conference, 6–10 June 2011, Poznan, Poland. – 2011. – P. 104–107.

26. Nelson, M. [et al.]. Domestication genes in narrow-leaved lupin / M. Nelson, D. Ganesalingham, W. Cowling // Mode of access: www.agwine.adelaide.edu.au/docs/example-6-domestication-genes-lupin-1.pdf. – Date of access: 28.02.2017.

27. Gill, E. [et al.]. The control of Alkaloid by Mutant alleles in *Lupinus albus* and *Lupinus angustifolius* / E. Gill, M. Harrison, W. Williams // Zeitschrift für Pflanzenzüchtung. – 1983. – Vol. 90, № 1. – P. 32–41.

28. Головченко, О.В. [и др.]. Результаты изучения генетической природы снижения синтеза алкалоидов у белого люпина / О.В. Головченко, Н.В. Солодюк, Т.М. Левченко // Научное обеспечение люпиносеяния в России: тез. докл. Междунар. науч.-практич. конф. ВНИИ люпина, 12–14 июля 2005 г. – Брянск: Читай-город, 2005. – С. 61–66.

29. Genetical control of alkaloid production in *Lupinus mutabilis* and the effect of a mutant allele *Mutal* isolated following chemical mutagenesis / W. Williams [et al.] // Euphytica. – 1984. – Vol. 33, № 3. – P. 811–817.

30. Hackbarth, J.I. [et al.]. Die Gene für Alkaloidarmut bei *Lupinus luteus* / J.I. Hackbarth, R. Sengbusch. – 1934. – 34 p.

31. Седлецкий, М.А. [и др.]. Изучение влияния количественного содержания алкалоидов на продуктивность и устойчивость к болезням у мутантов-аналогов / М.А. Седлецкий, В.И. Головченко // Сборник : Химический мутагенез и проблемы селекции. – М. : Наука, 1991. – С. 206.

32. Генетическое типирование хозяйственно ценных признаков люпина узколистного / Е.Н. Сысолятин [и др.] // Весці НАН Беларусі. – 2015. – № 4. – С. 44–50.

33. Development of a sequence-specific PCR marker linked to the gene “pauper” conferring low-alkaloids in white lupin (*Lupinus albus* L.) for marker assisted selection / R. Lin [et al.] // Mol. Breeding. – 2009. – Vol. 23. – P. 153–161.

34. Assessing root variability, growth, morphology, and metabolite content, in a diverse germplasm collection of yellow lupin (*L. luteus* L.) / G. Del Canto [et al.] // Proceedings Book of the XIV International Lupin conference Milan, Italy, 21–26 June 2015. – 2015. – P. 47.

35. Uncovering genetic factors associated with leaf and seed alkaloid content in the yellow lupin (*Lupinus luteus* L.) / C. Osorio [et al.] // Proceedings Book of the XIV International Lupin conference Milan, Italy, 21–26 June 2015. – 2015. – P. 101.

36. Романчук, И. Ю. [и др.]. О влиянии цитоплазмы материнского растения на наследование признака алкалоидности люпина / И. Ю. Романчук, В. С. Анохина // Генетика и биотехнология XXI века: проблемы, достижения, перспективы : материалы II Междунар. науч. конф. 13–16

октября 2015 г., Минск. – Минск: Полиграфт, 2015. – С. 120.

37. Дебелый, Г.А. Зернобобовые культуры в Нечерноземной зоне РФ / Г.А. Дебелый. – Москва, Немчиновка, 2009. – 260 с.

38. Куницкая, М.П. [и др.]. Изучение аллельности мутаций малоалкалоидности у образцов люпина узколистного / М.П. Куницкая, В.С. Анохина // Научное обеспечение люпиносеяния в России : тез. докл. Междунар. науч.-практич. конф., ВНИИ люпина, 12–14 июля 2005 г. – Брянск: Читай-город. – 2005. – С. 46–48.

39. Власова, Е.В. Установление аллельности генов малоалкалоидности и детерминации ветвления у шести колосовидных детерминантов люпина узколистного / Е.В. Власова // Генетические ресурсы культурных растений в XXI веке: Состояние, проблемы, перспективы : сборник тезисов докладов 2 Вавиловской междунар. конф., 26–30 ноября 2007. – СПб.: ВИР. – С. 53–54.

40. Купцоў, М.С. Заканамернасці змянення алкалоіднага комплексу вузкалісцевага лубіну ў працесе селекцыі на кармавыя вартасці / М.С. Купцоў, Т.П. Міронава // Весці акадэміі аграрных навук Беларусі. – 1993. – № 4. – С. 63–69.

41. Gladstones, J.S. *Lupinus angustifolius* L. (narrow-leaved lupin) cv. Gungurru / J.S. Gladstones // Austral. J. exper. Agr. – 1989. – Vol. 29, № 6. – P. 913–914.

42. Hawthorne, W.A. [et al.] *Lupinus angustifolius* L. cv. Warrarrah / W.A. Hawthorne, J.S. Gladstones // Austral. J. exper. Agr. – 1989. – Vol. 29, № 6. – P. 911–912.

43. Молекулярно-генетическая и биохимическая характеристика образцов стержневой коллекции люпина желтого / И. Ю. Романчук [и др.] // Труды БГУ. Серия: Физиологические, биохимические и молекулярные основы функционирования биосистем. – 2015. – Т. 10, ч. 1. – С. 127–132.

44. Влияние содержания алкалоидов на выживаемость растений люпина узколистного в полевых условиях / И.Ю. Цибульская [и др.] // Регуляция роста, развития и продуктивности растений: сб. матер. IV Междунар. науч. конф., Минск, 26–28 октября 2005 г. – Минск: Право и

экономика, 2005. – С. 244.

45. Цибульская, И.Ю. [и др.]. Межвидовое и внутривидовое разнообразие коллекционных образцов желтого и узколистного люпина по биохимическим показателям / И.Ю. Цибульская, В.С. Анохина // Научное обеспечение люпиносеяния в России : Тез. докл. Международной науч.-практич. конф., ВНИИ люпина, 12–14 июля 2005 г. – Брянск: Читай-город. – 2005. – С. 46–48.

46. Анохина, В.С. [и др.]. Алкалоид люпина – люпанин – природный ингибитор прорастания семян / В.С. Анохина, И.Ю. Цибульская // Регуляция роста, развития и продуктивности растений: матер. V Междунар. науч. конф. Минск, 28–30 ноября 2007 г. – Мн.: Право и экономика. – 2007. – С. 13.

47. Lupine induced «Crooked Calf Disease» in Washington and Origin: Identification of the alkaloid profiles in *Lupinus sulfureus*, *Lupinus leucophyllus* and *Lupinus sericeus* / S.T. Lee [et al.] // Agric. Food Chem. – 2007. – Vol. 55, № 26. – P. 10649–10655

48. De Pinto, M.C. [et al.]. Superoxide anion scavenger properties of sparteine, a quinolizidine alkaloid from *Lupinus* / M.C. De Pinto, M.C. de Canto, A.R. Arcelo // J. Plant Physiol. – 1997. – Vol. 150, № 1–2. – P. 5–8.

49. Effect of diets containing whole white lupin seeds on rabbit doe milk yield and milk fatty acid composition as well as the growth and health of their litters / Z. Volek [et al.] // J. Anim. Sci. 2014. – Vol. 92, № 5. – P. 2041–2049.

50. Correlation of the alkaloid content and composition of narrow-leaved lupins (*Lupinus angustifolius* L.) to aphid susceptibility / J. Philippi [et al.] // J. of Pest. Science. – Original paper. – P. 1–15.

51. Анохина, В.С. [и др.]. Алкалоиды люпина: их фунгицидные эффекты / В.С. Анохина, Л.Н. Каминская, И.Ю. Цибульская // Молекулярная и прикладная генетика. – 2008. – Т. 8. – С. 138–142.

52. Changes of phenolic secondary metabolite profiles in the reaction of narrow-leaf lupin (*Lupinus angustifolius* L.) plants to infections with *Colletotrichum lupini* fungus or treatment with its toxin / Wojakowska, A. [et al.] // Metabolomics. – 2013. – Vol. 9. – P. 575–589.

I.Yu. Romanchuk, V.S. Anokhina

LUPINE ALKALOIDS: STRUCTURE, BIOSYNTHESIS, GENETICS (Review article)

Belarusian State University
Minsk, 220030, the Republic of Belarus

The review contains data on the published works on biosynthesis, genetics and changes in the alkaloid content in the process of plant ontogeny and breeding of various lupine species, and the fungicidal effect of certain alkaloids.

Key words: alkaloids, genetics, yellow lupine, white lupine, narrow-leaved lupine.

Дата поступления статьи: 4 декабря 2017 г.

ПРАВИЛА ДЛЯ АВТОРОВ

Статьи должны быть написаны в сжатой и ясной форме и содержать:

- соответствующий индекс универсальной десятичной классификации литературы (УДК);
- название на русском и английском языках;
- инициалы и фамилии авторов на русском и английском языках;
- полное название учреждений, в которых выполнялось исследование и их почтовые адреса;
- ключевые слова (3–5 слов) на русском и английском языках;
- аннотацию на русском и английском языках (рекомендуемый средний объем аннотации — 500 знаков с пробелами). Аннотация должна ясно излагать содержание статьи и быть пригодной для опубликования в аннотациях к журналам отдельно от статьи;
- текст статьи (стандартизировать, используя подзаголовки «Введение», «Материалы и методы», «Результаты и обсуждение», «Заключение»), обзорная статья может содержать подзаголовки: «Введение», «Заключение» и названия подразделов;
- список использованных источников;
- дату поступления статьи в редакцию.

Объем статьи должен составлять не менее 14 000 знаков, включая пробелы, до 10–12 страниц. После распечатки статья должна быть вычитана автором (авторами). На последней ее странице должна(ы) быть подпись(и) автора(ов). Текст статьи идентичного содержания представляется в электронном виде (по e-mail или на флеш-накопителе) и на бумажном носителе в 2 экз. В виде отдельного документа представляются краткие сведения о каждом из авторов, включающие фамилию, имя, отчество, год рождения, сведения об образовании, служебные адреса, адрес электронной почты, ученую степень, ученое звание, должность, область научных интересов.

Рукописи авторства сторонних организаций предоставляются вместе с сопроводительным письмом, в структуре которого должны содержаться сведения о том, что:

- рукопись не находится на рассмотрении в другом издании;
- не была ранее опубликована;
- автор (авторы) несут ответственность за достоверность представленных в рукописи материалов.

Также необходимо представить акт экспертизы о возможности опубликования в открытой печати. Научная статья должна иметь 2 рецензии от специалистов, имеющих ученую степень доктора или кандидата наук.

1. Сдаваемый документ должен быть представлен в электронном виде в формате MS Word. Название файлов — фамилия первого автора латинскими буквами.
2. Формат бумаги A4 (297 × 210 мм), ориентация — книжная.
3. Поля: верхнее — 2,5 см, нижнее — 2,5 см, левое — 2,5 см, правое — 2,5 см.
4. Основной текст статьи набирается шрифтом Times New Roman, размером 12 пт, в одну колонку с одинарным межстрочным интервалом. Не допускается использование табуляции или пробелов для обозначения первой строки абзаца.
5. Автоматическая расстановка переносов обязательна. Запрещается ручная расстановка принудительных переносов.
6. Название статьи набирать полужирным начертанием шрифта по центру. Переносы в заголовках не допускаются.
7. Все таблицы, содержащиеся в документе, должны быть реализованы средствами работы с таблицами редактора MS Word. Не допускается вложение таблиц, созданных в других программах. Таблицы и графики должны быть пронумерованы и иметь названия. Не допускается размещение таблиц и рисунков в конце статьи (непосредственно перед списком литературы).

8. Вставка в текст символов (например, β , ϵ) производится только через опцию «Вставка\Символ». Выключку вверх и вниз (C^2 , C_4) выполнять через меню «Формат\Шрифт\Верхний индекс\Нижний индекс». Греческие символы должны быть прямыми, латинские буквы набираются курсивом. Математические формулы ($\lim$, $\sum$, $\sin$, и т. д.) и цифры набираются прямым начертанием.
9. Сложные слова набираются через дефис (минерал-индикатор, К-пространство). Тире отбиваются с обеих сторон неразрывным пробелом как знак препинания между словами: система «человек – машина», «май – июнь». Тире между цифрами, напр., 20–30 чел. — не отбивается.
10. Кавычки по всему тексту должны быть одного «рисунка». Кавычки не отбиваются от заключенных в них слов.
11. При подготовке к печати графиков, блок-схем, диаграмм, файлы должны быть названы таким образом, чтобы было понятно, к какой статье они принадлежат и какими по порядку рисунками статьи являются. Графики должны иметь толщину всех линий не менее 0,2 пункта для четкого воспроизведения. Все надписи на рисунках должны быть набраны на компьютере и сгруппированы с рисунком, не допускается использование сканированного текста.
12. Фотоматериалы предоставляются в виде отдельных электронных файлов, также необходимы распечатки на лазерном принтере всех иллюстраций на листе формата А4. Отсканированные фотоиллюстрации серой, черно-белой цветовой модели должны иметь разрешение 600 dpi и формат TIFF.
13. Список цитированных источников располагается в конце текста, ссылки нумеруются согласно порядку цитирования в тексте. Порядковые номера ссылок должны быть написаны внутри квадратных скобок (напр.: [1]).

ЦЕНТР КОЛЛЕКТИВНОГО ПОЛЬЗОВАНИЯ «ГЕНОМ»

Направления деятельности:

- Оказание услуг по молекулярно-генетическому анализу образцов ДНК.
- Оказание научно-методической помощи и консультаций по различным областям применения геномных технологий.

Оборудование:

Полногеномный секвенатор серии NextSeq (Illumina, Inc, США)

Позволяет решать как небольшие задачи секвенирования, так и секвенирование целых геномов. Дает возможность секвенировать геном человека за один запуск (30-кратное покрытие за 29 часов работы), осуществлять таргетное секвенирование (экзомное секвенирование, секвенирование митохондриальных генов, секвенирование ампликонов), а также секвенирование транскриптомов. Позволяет проводить исследования по анализу метилирования. Объем получаемых данных — 120 Гб. Доступны «быстрая химия» AmpliSeq, пользовательские custom-панели.

Полногеномный секвенатор MiSeq (Illumina, Inc, США)

Позволяет решать специализированные задачи, такие как исследование целевых участков генома, метагеномика, секвенирование малых геномов, оценка экспрессии групп генов, секвенирование ампликонов и HLA типирование. Новое поколение реагентов MiSeq позволяет достичь производительности в 15 млрд п. н. и 25 млн отдельных прочтений, при этом длина прочтения составляет 2×300 п. н.

Генетические анализаторы 3500 (Applied Biosystems, Japan)

Предназначены для определения нуклеотидной последовательности фрагментов ДНК, для проведения SSR, LOH, SNP, MLPA, FLP и t-RFLP анализов. Возможно проведение секвенирования *de novo* и ресеквенирования небольших геномов до 1000 п. н. Количество капилляров — 8. Количество красителей — 6.

Возможность считывания меток: TAMRA, ROX, R6G, FAM, LIZ. Скорость секвенирования 700 п. н с точностью 98,5% менее 40 мин.

Система для капельной цифровой ПЦР QX200™ Droplet Digital PCR

Предназначена для цифровой капельной ПЦР нуклеиновых кислот, для анализа экспрессии единичных генов, обнаружения и оценки возбудителей заболеваний; определения количества копий гена (CNV), определения вирусной нагрузки, выявления ГМО, работы с микроРНК а также валидации и количественной оценки библиотек NGS. Помимо работы с Taq-man ПЦР реализована возможность использования интеркалирующего красителя EvaGreen. Анализируется ~ 20 000 капель, сформированных из 20 мкл реакционной смеси.

Центр также оснащен следующим оборудованием: амплификаторы, амплификаторы с возможностью регистрации продуктов ПЦР в режиме реального времени CFX96 (Bio-Rad, США), роботизированная станция для масштабного выделения ДНК Freedom EVO 75 (TECAN, Швейцария), молекулярный сканер Pharos FX Plus Molecular Imager (BioRad, США), спектрофотометр NanoDrop 8000 (США), система очистки воды PURELAB FLEX 3, PF3XXXXM1 (Elga, Великобритания), генератор льда IMS-30 (Тесеус Лаб, Чехия), ПЦР-бокс, вортексы, центрифуги и др. оборудование, необходимое для проведения пробоподготовок (ПЦР, NGS и др.).

Центр коллективного пользования «ГЕНОМ»
тел. +375 17 284 19 41; +375 17 332 16 13

Научное издание

МОЛЕКУЛЯРНАЯ И ПРИКЛАДНАЯ ГЕНЕТИКА

СБОРНИК НАУЧНЫХ ТРУДОВ

Том 25

Ответственный за выпуск *А.Л. Богданова*
Переводчик *М.Г. Пыжова*

Подписано в печать 17.09.2018. Формат 60×84^{1/8}. Бумага офсетная. Гарнитура Times.
Печать цифровая. Усл. печ. л. 15,11. Уч.-изд. л. 9,50. Тираж 100 экз. Заказ № 15439.
Отпечатано в ООО «Колорград».

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