

Достижения и перспективы использования методов высокопроизводительного секвенирования в генетике и селекции картофеля

И.В. Быкова¹, Н.А. Шмаков^{1, 2}, Д.А. Афонников^{1, 2}, А.В. Кочетов^{1, 2}, Е. К. Хлесткина^{1, 2}

¹ Федеральное государственное бюджетное научное учреждение «Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук», Новосибирск, Россия

² Федеральное государственное автономное образовательное учреждение высшего образования «Новосибирский национальный исследовательский государственный университет», Новосибирск, Россия

Для повышения эффективности процесса селекции картофеля в последние годы активно используются методы отбора с помощью ДНК-маркеров. Применение маркер-ориентированной селекции, а также геномной селекции и новейших методов создания форм картофеля с заданными свойствами (геномное редактирование) требует широкомасштабного изучения геномов и генов картофеля на основе высокопроизводительных методов секвенирования нового поколения. Цель настоящего обзора – рассмотреть направления исследований в картофелеводстве, основанные на применении данных методов, и систематизировать связанные с этой темой сведения об интернет-ресурсах. Уделяется также внимание особенностям используемых моделей и подходов в генетических исследованиях картофеля ввиду сложной организации его генома и высокого уровня гетерозиготности. В генетических исследованиях нередко прибегают к использованию диплоидных форм, к которым относятся диплоидные виды картофеля, искусственно полученные гетерозиготные дигаплоиды и гомозиготные удвоенные моноплоиды. Наличие искусственно созданных диплоидных форм сыграло важную роль в успешном секвенировании генома картофеля, завершеном в 2011 г. Работа Международного консорциума по секвенированию генома картофеля включала не только конструирование геномных библиотек, секвенирование, сборку и аннотирование генома, но и фундаментальные исследования на основе полученной последовательности, направленные на выявление особенностей эволюции генома картофеля, идентификацию SNP (single nucleotide polymorphism – однонуклеотидный полиморфизм), анализ генов и генных сетей, контролирующих устойчивость к фитопатогенам и технологические свойства. Важным следствием секвенирования генома стала идентификация более 8 тыс. SNP и апробация метода GBS (genotyping-by-sequencing – генотипирование путем секвенирования) на картофеле, что является основой для геномной селекции этой культуры и выявления новых хозяйственно ценных генов картофеля методом полногеномного анализа ассоциаций (genome-wide association studies – GWAS). Ведется работа по оптимизации существующих биоинформатических методов анализа для осуществления работ в этом направлении, учитывающая особенности организации генома картофеля. В обзоре приведен анализ доступных электронных баз данных, содержащих результаты работ по секвенированию генома и транскриптома картофеля, а также сопутствующих ресурсов. Эта информация будет полезна для планирования работ по сравнительному анализу транскриптома картофеля и исследований с использованием геномных ДНК-маркеров. Результаты секвенирования генома, а также сравнительных исследований, базирующихся на анализе транскриптома и микроРНК

Achievements and prospects of applying high-throughput sequencing techniques to potato genetics and breeding

I.V. Bykova¹, N.A. Shmakov^{1, 2}, D.A. Afonnikov^{1, 2}, A.V. Kochetov^{1, 2}, E.K. Khlestkina^{1, 2}

¹ Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia
² Novosibirsk State University, Novosibirsk, Russia

In recent years, marker-assisted selection (MAS) has been intensively used to increase potato breeding efficiency. Large-scale studies of the potato genome and genes exploiting next-generation sequence (NGS) approaches are required for broad application of MAS, genomic selection as well as genomic editing (the newest approach for creating potato with desired properties). In this review, trends in potato NGS-based research are overviewed, and related Internet resources are systematized. Special attention is given to peculiarities of the models and the approaches used in potato genetic studies, taking into account the complex organization of its genome and a high level of heterozygosity. In genetic studies diploids are used often, including diploid potato species, artificially obtained heterozygous dihaploids and homozygous double monoploids. The availability of artificially created diploid forms played an essential role in potato genome sequencing, which was completed in 2011. The Potato Genome Sequencing Consortium activities included not only constructing genome libraries, sequencing, assembling and annotation of the genome, but also genome sequence-based investigations uncovering features of potato genome evolution, SNP identification, analysis of genes and gene networks regulating resistance to phytopathogens and technological characteristics. An important outcome of the genome sequencing was further identification of more than 8 thousand SNPs and approbation of the Genotyping-by-sequencing (GBS) method on potato, which is the basis for potato genomic selection and for discovery of economically important genes using genome wide association studies (GWAS). Optimization of existing bioinformatic tools to support these studies, taking into account potato genome organiza-

образцов культивируемого картофеля и его дикорастущих сородичей, с одной стороны, представляют фундаментальный интерес, способствуя выявлению особенностей эволюции генома, онтогенетического развития растений картофеля и механизмов ответа на различные стимулы окружающей среды, а с другой – позволяют осуществлять широкий спектр прикладных работ, направленных на разработку эффективных геномных и ген-специфичных маркеров и получения с их помощью новых сортов картофеля с заданными свойствами.

Ключевые слова: RNAseq; *Solanum tuberosum*; картофель; геном, транскриптом; секвенирование; гены; маркеры; базы данных.

peculiarities, are carried out. This review gives analysis of databases containing potato genome and transcriptome sequencing results, as well as accompanying resources. This information should prove useful while planning comparative assays of potato transcriptome or application of DNA-markers. Sequencing of the genome as well as transcriptomes and microRNomes of cultivated potato and its wild relatives, on one hand, is of fundamental interest, assisting in detecting features of genome evolution, ontogenetic development and mechanisms of various environmental stresses responses. On the other hand, it is the basis for a wide range of practical applications for developing effective genomic and gene-specific markers and marker-assisted breeding of new potato cultivars with desired properties.

Key words: databases; genes; genome; markers; potato; RNAseq; *Solanum tuberosum*; sequencing; transcriptome.

КАК ЦИТИРОВАТЬ ЭТУ СТАТЬЮ:

Быкова И.В., Шмаков Н.А., Афонников Д.А., Кочетов А.В., Хлесткина Е.К. Достижения и перспективы использования методов высокопроизводительного секвенирования в генетике и селекции картофеля. Вавиловский журнал генетики и селекции. 2017;21(1):96-103. DOI 10.18699/VJ17.227

HOW TO CITE THIS ARTICLE:

Bykova I.V., Shmakov N.A., Afonnikov D.A., Kochetov A.V., Khlestkina E.K. Achievements and prospects of applying high-throughput sequencing techniques to potato genetics and breeding. Vavilovskii Zhurnal Genetiki i Selekcii = Vavilov Journal of Genetics and Breeding. 2017;21(1):96-103. DOI 10.18699/VJ17.227

Картофель по объемам производства и потребления входит в четверку наиболее важных продовольственных культур наряду с пшеницей, рисом и кукурузой (www.fao.org). Селекция картофеля – динамичный процесс, отвечающий постоянно меняющимся требованиям в области переработки картофеля и возникновению новых технологических операций, а в последнее время – также потребности в разработке продуктов для функционального и диетического питания. Кроме того, традиционные задачи селекции картофеля – повышение урожайности и устойчивости к заболеваниям и неблагоприятным климатическим условиям (Gebhardt, 2013).

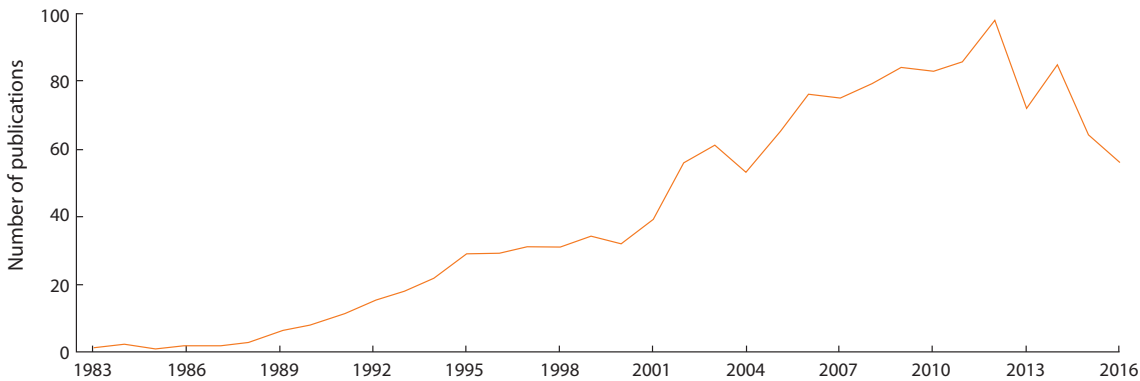
С целью повышения эффективности процесса селекции картофеля в последние годы активно используются схемы маркер-ориентированной селекции, подходящие для отбора по моно- и олигогенным признакам (Gebhardt, 2013; Ramakrishnan et al., 2015; Хлесткина и др., 2016; Колчанов и др., 2017); экономическая целесообразность применения данного метода обоснована (Slater et al., 2013). Согласно оценке публикационной активности (см. рисунок), вплоть до 2012 г. отмечалась выраженная положительная динамика использования ДНК-маркеров в картофелеводстве. Некоторое ее снижение в последние три-четыре года требует дополнительного анализа и может быть связано с переключением интереса на новейший метод получения генотипов с заданными свойствами (геномное редактирование с помощью системы CRISPR/Cas9).

Для улучшения сельскохозяйственных культур, в том числе картофеля, по моно- и олигогенным признакам также перспективны современные методы геномного редактирования, а для отбора по полигенным признакам – геномная селекция (Bortesi, Fischer, 2015; Butler et al., 2015; Wang et al., 2015; Хлесткина, Шумный, 2016; Andersson et al., 2016; Butler, Douches, 2016; Slater et al., 2016).

Применение маркер-ориентированной селекции и геномного редактирования требует предварительного проведения интенсивных исследований, направленных на выявление хозяйственно ценных генов, установление их аллельных вариантов на уровне ДНК. Для геномной селекции, помимо этого, необходимы экономичные высокопроизводительные методы анализа полиморфизма ДНК. И то и другое требует вовлечения методов секвенирования нового поколения. Цель настоящего обзора – рассмотреть направления исследований в картофелеводстве, основанные на применении методов высокопроизводительного секвенирования, и систематизировать сведения об интернет-ресурсах, связанные с данными исследованиями.

Организация генома картофеля и особенности используемых моделей в генетических исследованиях

Основной вид картофеля, находящийся в производстве, – *Solanum tuberosum* – относится к семейству Solanaceae, которое включает в себя и другие важные сельскохозяйственные культуры (томат, перец, баклажан, табак). *S. tuberosum* – автотетраплоид ($2n = 4x = 48$) с базовым числом хромосом 12 и размером (гаплоидного) генома $0,84 \cdot 10^9$ п. н. (Visser et al., 2009). Для сортов *S. tuberosum* характерна высокая степень гетерозиготности, что обеспечивает повышенную жизнеспособность растений картофеля при полиплоидном уровне организации. Из-за этих особенностей (полиплоидный геном и гетерозиготность) анализ отдельных генов картофеля, так и генома в целом затруднен. В генетических исследованиях нередко прибегают к использованию диплоидных форм, к которым относятся: диплоидные сородичи *S. tuberosum*; искусственно полученные гетерозиготные диплоидные формы (дигаплоиды); гомозиготные диплоидные формы



Trend in the number of publications dealing with the use of molecular markers in potato research and breeding. Website: <http://scopus.com>; search results for “marker+potato+Solanum” in article titles, abstracts and key words are presented (“Solanum” was added to exclude sweet potato research).

(получаемые за счет удвоения моноплоидов). Вовлечение диплоидных сородичей в генетические исследования не только упрощает анализ, но и позволяет изучить потенциал этих видов в качестве доноров селекционно значимых генов, особенно генов устойчивости к фитопатогенам и вредителям (Nowicki et al., 2012; Рогозина и др., 2012, 2013; Limantseva et al., 2014).

Помимо естественных диплоидных форм в генетических исследованиях используют специально созданные диплоидные линии картофеля. Например, Н. van Os et al. (2006) сконструировали генетическую карту картофеля, плотно насыщенную ДНК-маркерами (более 10 тыс. AFLP-локусов; amplified fragment length polymorphism – полиморфизм длины амплифицированных фрагментов) при использовании картирующей популяции, полученной от скрещивания двух диплоидных гетерозиготных клоновых линий. Одна из родительских форм данной популяции в дальнейшем была включена в проект по полногеномному секвенированию картофеля (The Potato... Consortium, 2011). Дигаметоидные формы создают за счет понижения уровня пloidности с тетра- до диплоидного (Hougas et al., 1958), при этом создаваемые диплоидные формы являются гетерозиготными. Между тем для решения ряда генетических задач, а также для практического применения в гибридной селекции необходимы гомозиготные диплоидные формы. Такие формы получают за счет удвоения моноплоидов, и именно удвоенный моноплоид стал основной моделью для проведения полногеномного секвенирования картофеля (The Potato... Consortium, 2011). Моноплоиды получают из диплоидных форм через культуру пыльников (Veilleux et al., 1985). Возможность создания таких форм ограничена генетически детерминированной способностью генотипов к культивированию *in vitro*.

Секвенирование генома картофеля и результаты работы консорциума

Для проведения сложной и масштабной работы по секвенированию генома картофеля был организован Международный консорциум (The Potato Genome Sequencing Consortium – PGSC), включающий 16 исследовательских групп из Аргентины, Бразилии, Великобритании,

Китая, Индии, Ирландии, Италии, Нидерландов, Новой Зеландии, Перу, Польши, России, США и Чили. С целью получения качественной геномной последовательности основным объектом полногеномного секвенирования был выбран гомозиготный генотип – искусственно созданный удвоенный моноплоид (DM1-3 516R44), произведенный из диплоидной формы *S. tuberosum* группы Phureja (*S. phureja*). Последовательность генома удвоенного моноплоида DM1-3 516R44, в свою очередь, стала основой для интеграции данных, полученных при секвенировании гетерозиготной диплоидной культурной формы *S. tuberosum* ssp. *tuberosum* (Visser et al., 2009; The Potato... Consortium, 2011). Именно этот гетерозиготный диплоидный образец (RH-89-039-16) изначально был взят в качестве модели для полногеномного секвенирования – отчасти ввиду того, что являлся родительской формой наиболее информативной на тот момент картирующей популяции, с использованием которой была создана молекулярно-генетическая карта картофеля высокой плотности (van Os et al., 2006). Однако с самого начала анализ данных для гетерозиготной формы оказался затруднен, поэтому в процессе работы консорциума в качестве основной модели была выбрана упомянутая выше гомозиготная диплоидная форма DM1-3 516R44 (далее – DM).

Деятельность консорциума включала следующие этапы: 1) собственно секвенирование и построение хромосомных последовательностей (конструирование геномных библиотек на основе ВАС-клонов (ВАС – bacterial artificial chromosome, бактериальная искусственная хромосома) и фосмид, секвенирование, сборка и аннотирование генома), в результате чего была определена нуклеотидная последовательность, составляющая 86 % от всего генома и установлено 90.3 % от ранее предсказанного числа генов; 2) фундаментальные исследования на основе полученной последовательности генома (изучение эволюции генома картофеля, идентификация SNP, анализ генов и генных сетей, контролирующих устойчивость к фитопатогенам и технологические свойства картофеля).

Полученная сборка генома образца DM имеет длину $7.26 \cdot 10^8$ п. н., 86 % из которых привязано к генетической карте, и содержит 39031 аннотированный ген.

Table 1. Internet resources for potato genome sequences (The Potato...consortium 2011)

Database name*	Assembling type	Assembly of pseudomolecules	BAC clones	Nonnuclear genomes	Other sequences
Sol Genome	Scaffolds	Yes	Yes	Chloroplast, mitochondrial	CDS, proteins
Spud		Yes	Yes		CDS, cDNA, proteins, SSR, DaT, OPA, SNP
NCBI		No	Yes	Chloroplast	CDS
Ensembl	Chromosomes	No	No		CDS, cDNA, microRNA
Plaza		No	No		CDS

*Websites and references are given in Supplementary material 1.

Table 2. SNP and indel frequencies according to the results of sequencing the whole potato genome or its fragments

Samples compared	SNP	indel	Reference
11 diploids and 17 tetraploids; genomic fragments of total length 71 kb	1/21	1/243	Rickert et al., 2003
47 accessions; genomic fragments of total length > 25 kb	1/23	1/441	Simko et al., 2006
2 genotypes (heterozygous diploid and doubled monoplod), whole genome sequences	1/40	1/394	The Potato... Consortium, 2011
Homologous sequences of a single heterozygous diploid	1/29	1/253	»
83 cultivars of tetraploid potato (807 resequenced genes)	1/18	1/163	Uitdewilligen et al., 2013
83 cultivars of tetraploid potato (807 resequenced genes – coding regions only)	1/24	1/1448	»

Сведения о сборке генома картофеля и сопутствующие данные представлены в пяти основных интернет-ресурсах (табл. 1). Их подробный анализ дан в Приложении (Доп. материалы 1)¹.

Сравнение двух секвенированных форм картофеля (DM и RH) показало уровень идентичности 97.5 % (повторяющаяся фракция генома при этом не учитывалась). Встречаемость SNP в среднем составила один на 40 п. н., а инсерций/делаций (indels) – одну на 394 п. н. При сравнении последовательностей гетерозиготного образца RH установлен уровень идентичности в гомологичных участках 96.5 %, встречаемость SNP – один на 29 п. н., indel – одна на 253 п. н. Сравнение частот SNP и indel по данным консорциума, а также предшествующих и последующих исследований приведено в табл. 2.

Данные геномного секвенирования послужили основой при разработке микрочипа картофеля для анализа более 8 тыс. SNP-локусов (Hamilton et al., 2011; Felcher et al., 2012; SolCAP, 2016) и применении на картофеле относительно нового метода высокопроизводительного генотипирования – GBS (genotyping-by-sequencing – генотипирование путем секвенирования) (Uitdewilligen et al., 2013). Эти разработки являются основой для будущего освоения метода геномной селекции картофеля и выявления новых

хозяйственно ценных генов картофеля методом GWAS (genome-wide association studies – полногеномный анализ ассоциаций). Однако на пути к широкому применению этих подходов в анализе культивируемых форм картофеля стоит вопрос разработки биоинформатических методов оценки получаемых данных с учетом высокой степени гетерозиготности полиплоидных образцов. Проводятся работы, направленные на решение этой сложной задачи (Rosyara et al., 2016).

Сравнение полученной последовательности генома картофеля с геномом томата – представителя того же семейства, но не образующего клубни, в отличие от картофеля, а также проведенный консорциумом анализ транскриптома клубней и столонов позволили расширить представление о механизмах клубнеобразования. При клубнеобразовании усиливалась транскрипция 333 генов, в наибольшей степени – генов, кодирующих запасные белки, ингибиторы протеиназ и пататин. При этом выявлено, что большое семейство ингибиторов протеаз КТІ у картофеля вдвое больше по численности, чем у томата. Предполагается, что КТІ нацелены на ингибирование экзогенных протеиназ, защищая тем самым клубни от воздействия патогенов и вредителей (The Potato... Consortium, 2011).

Другая группа генов, активность которой существенно различается в клубнях и столонах, – гены биосинтеза крахмала. Сравнение уровня экспрессии генов у двух

¹ Дополнительные материалы 1–2 см. в Приложении 2 по адресу: <http://www.bionet.nsc.ru/vogis/download/pict-2017-21/appx2.pdf>

секвенированных образцов показало, что у гетерозиготной формы, полученной от современного коммерческого сорта, гены биосинтеза крахмала в среднем в 3–8 раз активнее, чем у удвоенного моноплоида (DM), производного образца из группы Phureja. У DM, наоборот, выше уровень мРНК альфа- (в 10–25 раз) и бета-амилазы (в 5–10 раз). В целом различия между двумя генотипами хорошо согласуются с гипотезой о том, что отбор на повышенную продуктивность клубней – это отчасти результат снижения активности генов гидролиза крахмала. Выявление дифференциально экспрессирующихся генов, влияющих на метаболизм крахмала, имеет важное значение для получения в перспективе сортов картофеля, продуцирующих крахмал с заданными свойствами, который будет представлять собой оптимальное сырье для тех или иных отраслей промышленности (Хлесткин и др., 2017).

Поскольку картофель поражается разнообразными вредителями и возбудителями болезней, то особое внимание по завершении секвенирования было уделено идентификации генов устойчивости (*R*-генов). Большинство изученных до этого генов *R* Solanaceae, как было показано, несут NBS- и LRR-домены (NBS – nucleotide-binding site, LRR – leucine-rich-repeat). У секвенированного образца DM обнаружено 438 NBS-LRR-кодирующих генов (Jupe et al., 2012), среди них известные ранее *R1*, *RB*, *R2*, *R3a*, *Rpi-blb2* и *Rpi-vnt1.1*. Однако 39.4 % выявленных генов несли функциональные мутации (сдвиг рамки считывания, преждевременные стоп-кодоны и т. д.), в том числе эти нарушения выявлены в генах *R1*, *R3a* и *Rpi-vnt1.1*. Такой высокий уровень изменчивости наблюдается параллельно быстрой эволюции эффективных генов некоторых фитопатогенов картофеля. Изучение выделенных новых *R*-генов позволит выявить в дальнейшем гены-кандидаты для использования в селекции картофеля.

Таким образом, данные, полученные в результате секвенирования генома картофеля, служат основой для широкого спектра как фундаментальных исследований, так и прикладных работ, направленных на получение новых сортов с заданными свойствами, разработку новых генетических и ген-специфических маркеров картофеля для селекции по хозяйственно значимым признакам.

Транскриптомные исследования как ключ к пониманию механизмов формирования хозяйственно значимых признаков картофеля

Сравнительный анализ транскриптомов генотипов картофеля, контрастных по проявлению того или иного признака, позволяет понять, какие молекулярно-генетические механизмы задействованы в формировании фенотипа, найти и маркировать гены-мишени для селекции новых улучшенных форм картофеля.

Данные об исследованиях транскриптома картофеля методами RNAseq и анализа ДНК-микрочипов содержатся в репозиториях Gene Expression Omnibus NCBI – GEO (Clough, Barrett, 2016) и ArrayExpress EBI (Kolesnikov et al., 2015).

GEO (Gene Expression Omnibus Database; <https://www.ncbi.nlm.nih.gov/geo/>) – это база данных, в которой собраны результаты высокопроизводительного секвенирования

транскриптомов (RNAseq) и анализа экспрессии генов с использованием микрочипов, в том числе для *S. tuberosum* (Rensink et al., 2005; Carvallo et al., 2011; Evers et al., 2012; Zhang et al., 2013; Stare et al., 2015; Gálvez et al., 2016). База данных GEO включает несколько разделов, в том числе DataSets, который содержит описание экспериментов по секвенированию транскриптомов. В этом разделе суммированы данные 1583 экспериментов, посвященных анализу транскриптома картофеля. Результаты транскриптомных экспериментов в БД GEO представлены в 30 публикациях, в число которых входят исследования по изменению транскриптома картофеля в ответ на заражение вирусами (Baebler et al., 2014; Petek et al., 2014; Stare et al., 2015), на воздействие факторами абиотического стресса (Gálvez et al., 2016), различия в уровне экспрессии генов miRNA в зависимости от стадии развития (Lakhotia et al., 2014). Первичные данные секвенирования транскриптомов в экспериментах RNA-seq представлены в базе данных NCBI SRA (Sequence Read Archive), основном мировом архиве результатов высокопроизводительного секвенирования нуклеотидных последовательностей. SRA содержит информацию по 2004 экспериментам, связанным с высокопроизводительным секвенированием *S. tuberosum*, из которых 1438 относятся к секвенированию генома и 564 – к секвенированию транскриптома (Massa et al., 2011; Zhang et al., 2013; Gong et al., 2015; Liu et al., 2015; Zuluaga et al., 2015).

Еще одна полезная база данных – ArrayExpress (<https://www.ebi.ac.uk/arrayexpress/>). Этот ресурс Европейского института биоинформатики содержит результаты 77 экспериментов по анализу экспрессии генов картофеля (Ali et al., 2014; Vurga et al., 2014). Основные типы экспериментов, данные по которым опубликованы в ArrayExpress для *S. tuberosum*, это высокопроизводительное секвенирование транскриптома ('RNA-seq of coding RNA', 9 экспериментов) и анализ экспрессии с помощью микрочипов ('transcription profiling by array', 62 эксперимента).

Наиболее часто встречающейся постановкой транскриптомного эксперимента у картофеля является сравнение уровней экспрессии генов растений для двух контрастных условий. Такие эксперименты позволяют выявить ранее неизвестные детали механизма взаимодействия растений картофеля с патогеном. В отдельных случаях может быть идентифицирован новый ген устойчивости, что позволяет разработать диагностические ДНК-маркеры для ускоренного отбора селекционного материала по целевому гену.

Так, на основе сравнительного анализа транскриптома чувствительных и устойчивых к фитофторе образцов дикого и культивируемого картофеля среди дифференциально экспрессирующихся генов были выявлены не только ожидаемые гены, имеющие отношение к известным механизмам устойчивости (гиперчувствительный ответ, программируемая клеточная гибель, активность эндопептидаз), но и гены с функциями, ранее не ассоциированными с защитой от данного патогена (ацилирование белков, трансмембранный транспорт и др.) (Frades et al., 2015). Кроме того, авторы описали несколько новых потенциальных генов устойчивости (*R*) – кандидатов для дальнейшего использования в селекционном процессе.

Другая группа исследований – сравнение паттернов транскрипции генов в разных частях растений или на разных стадиях развития. Их результаты важны для понимания молекулярных механизмов, лежащих в основе эволюционных и онтогенетических изменений, что представляет прежде всего фундаментальный интерес. Кроме того, в свете интенсивно осваиваемой технологии геномного редактирования такие исследования открывают новые возможности для выявления генов-мишеней для направленного мутагенеза. Например, обнаруженные при сравнении транскриптомов столонов и клубней дифференциально экспрессирующиеся гены, связанные с клубнеобразованием (The Potato ... Consortium, 2011), представляют собой потенциальные мишени для применения стратегии доместикации *de novo* (Хлесткина, 2016).

Идентификация микроРНК, участвующих в регуляции физиологических процессов

МикроРНК (miRNA) представляет собой класс коротких некодирующих эндогенных РНК, которые играют важную роль в посттранскрипционной регуляции экспрессии генов и участвуют в контроле онтогенетических, физиологических и биохимических процессов. Это короткие (21–24 нуклеотида) одноцепочечные молекулы, в образовании которых участвуют РНК полимеразы II и эндорибонуклеазы DCL1. Идентификация микроРНК и предсказание их мишеней выполняются с помощью существующих и активно развивающихся методов биоинформатического анализа секвенированных фракций микроРНК, выделяемых в ходе сравнительных экспериментов из разных частей растений, на тех или иных этапах развития, а также при воздействии различных внешних стимулов. Для секвенирования используют как экспериментальные подходы, связанные с высокопроизводительным секвенированием (Zhang et al., 2013), так и биоинформатические подходы для предсказания структуры микроРНК по уже известным последовательностям (Zhang et al., 2009). Исследования направлены на изучение микроРНК, которые регулируют экспрессию конкретных генов (Chi et al., 2015), участвующих в формировании стрессового ответа на засуху (Zhang et al., 2014) и играющих роль в процессе клубнеобразования (Martin et al., 2009; Bhogale et al., 2014; Lakhota et al., 2014).

Данные по микроРНК картофеля аннотированы в нескольких базах данных: miRBase (Kozomara, Griffiths-Jones, 2013), miRNEST 2.0 (Szcześniak, Makałowska, 2014), RNAcentral (The RNAcentral Consortium, 2017) и Spud DB (Hirsch et al., 2014). Подробный анализ этих электронных ресурсов дан в Доп. материалах 2.

Если до применения методов NGS на картофеле работы по изучению его микроРНК сводились к проверке функций микроРНК, известных у представителей растительного мира, то с применением методов высокопроизводительного секвенирования стало возможным выявление новых, специфических для картофеля семейств. Так, в работе (Zhang et al., 2013) в листьях и столонах картофеля было идентифицировано 28 консервативных семейств микроРНК растений и выявлено 120 специфических для картофеля семейств. Для большинства из них были предсказаны гены-мишени.

В дальнейших исследованиях сравнительный анализ микроРНК (совокупности всех микроРНК) выполнялся в разрезе онтогенетического развития (Lakhota et al., 2014) и ответа на засуху (Zhang et al., 2014). В результате в работе (Lakhota et al., 2014) были выявлены микроРНК картофеля, мишенями для которых являются гены, кодирующие известные факторы, важные для развития растений (ARF16 – auxin response factor 16, NAM – no apical meristem, RAP1 – relative to APETALA2 1, HAM – hairy meristem), а в работе (Zhang et al., 2014) выявлены четыре микроРНК, участвующие в регуляции ответа на данный вид стресса, и предсказаны мишени для этих микроРНК.

Заключение

Результаты секвенирования генома, а также сравнительных исследований, базирующихся на анализе транскриптома и микроРНК образцов культивируемого картофеля и его дикорастущих сородичей, с одной стороны, представляют фундаментальный интерес, способствуя выявлению особенностей эволюции генома, онтогенетического развития растений картофеля и механизмов ответа на различные стимулы окружающей среды, а с другой стороны, являются основой для широкого спектра прикладных работ, направленных на разработку эффективных геномных и ген-специфичных маркеров и получения с их помощью новых сортов картофеля с заданными свойствами.

Acknowledgments

This work was supported by the Russian Science Foundation, project 16-16-04073.

Conflict of interest

The authors declare no conflict of interest.

References

- Ali A., Alexandersson E., Sandin M., Resjö S., Lenman M., Hedley P., Levander F., Andreasson E. Quantitative proteomics and transcriptomics of potato in response to *Phytophthora infestans* in compatible and incompatible interactions. BMC Genomics. 2014;15(1):497. DOI 10.1186/1471-2164-15-497.
- Andersson M., Turesson H., Nicolai A., Falt A., Samuelsson M., Hofvander P. Efficient targeted multiallelic mutagenesis in tetraploid potato (*Solanum tuberosum*) by transient CRISPR-Cas9 expression in protoplasts. Plant Cell Rep. 2016;36(1):117-128. DOI 10.1007/s00299-016-2062-3.
- Baebler Š., Witek K., Petek M., Stare K., Tušek-Žnidarič M., Pompe-Novak M., Renaut J., Szajko K., Strzelczyk-Zyta D., Marczewski W., Morgiewicz K., Gruden K., Hennig J. Salicylic acid is an indispensable component of the *Ny-1* resistance-gene-mediated response against *Potato virus Y* infection in potato. J. Exp. Bot. 2014; 65(4):1095-1109.
- Bhogale S., Mahajan A.S., Natarajan B., Rajabhoj M., Thulasiram H.V., Banerjee A.K. MicroRNA156: A potential graft-transmissible microRNA that modulates plant architecture and tuberization in *Solanum tuberosum* ssp. *andigena*. Plant Physiol. 2014;164(2):1011-1027. DOI 10.1104/pp.113.230714.
- Bortesi L., Fischer R. The CRISPR/Cas9 system for plant genome editing and beyond. Biotechnol. Advance. 2015;33(1):41-52. Available at <http://dx.doi.org/10.1016/j.biotechadv.2014.12.006>.
- Burra D.D., Berkowitz O., Hedley P.E., Morris J., Resjö S., Levander F., Liljeroth E., Andreasson E., Alexandersson E. Phosphite-induced changes of the transcriptome and secretome in *Solanum tuberosum* leading to resistance against *Phytophthora infestans*. BMC Plant Biol. 2014;14:254. DOI 10.1186/s12870-014-0254-y.

- Butler N.M., Atkins P.A., Voytas D.F., Douches D.S. Generation and inheritance of targeted mutations in potato (*Solanum tuberosum* L.) using the CRISPR/Cas System. *PLoS ONE*. 2015;10(12). DOI 10.1371/journal.pone.0144591.
- Butler N.M., Douches D.S. Sequence-specific nucleases for genetic improvement of potato. *Am. J. Potato Res.* 2016;93(4):303-320. DOI 10.1007/s12230-016-9513-9.
- Carvalho M.A., Pino M.T., Jeknić Z., Zou C., Doherty C.J., Shiu S.H., Chen T.H.H., Thomashow M.F. A comparison of the low temperature transcriptomes and CBF regulons of three plant species that differ in freezing tolerance: *Solanum commersonii*, *Solanum tuberosum*, and *Arabidopsis thaliana*. *J. Exp. Bot.* 2011;62(11):3807-3819.
- Chi M., Liu C., Su Y., Tong Y., Liu H. Bioinformatic prediction of upstream microRNAs of PPO and novel microRNAs in potato. *Can. J. Plant Sci.* 2015;95(5):871-877. DOI <http://dx.doi.org/10.1139/CJPS-2014-308>.
- Clough E., Barrett T. The Gene Expression Omnibus database. *Methods Mol. Biology.* 2016;1418:93-110. DOI 10.1007/978-1-4939-3578-9_5.
- Evers D., Legay S., Lamoureux D., Hausman J.F., Hoffmann L., Renault J. Towards a synthetic view of potato cold and salt stress response by transcriptomic and proteomic analyses. *Plant. Mol. Biol.* 2012;78(4-5):503-14. DOI 10.1007/s11103-012-9879-0.
- Felcher K.J., Coombs J.J., Massa A.N., Hansey C.N., Hamilton J.P., Veilleux R.E., Buell C.R., Douches D.S. Integration of two diploid potato linkage maps with the Potato Genome Sequence. *PLoS ONE*. 2012;7(4). DOI 10.1371/journal.pone.0036347.
- Frades I., Abreha K.B., Proux-Wéra E., Lankinen Å., Andreasson E., Alexandersson E. A novel workflow correlating RNA-seq data to *Phytophthora infestans* resistance levels in wild *Solanum* species and potato clones. *Front. Plant Sci.* 2015;6:718. DOI 10.3389/fpls.2015.00718.
- Gálvez J.H., Tai H.H., Lagüe M., Zebarth B.J., Strömviik M.V. The nitrogen responsive transcriptome in potato (*Solanum tuberosum* L.) reveals significant gene regulatory motifs. *Sci. Reports.* 2016;6:26090. DOI 10.1038/srep26090.
- Gebhardt C. Bridging the gap between genome analysis and precision breeding in potato. *Cell.* 2013;29(4):248-256. DOI 10.1016/j.tig.2012.11.006.
- Gong L., Zhang H., Gan X., Zhang L., Chen Y., Nie F., Shi L., Li M., Guo Z., Zhang G., Song Y. Transcriptome profiling of the potato (*Solanum tuberosum* L.) plant under drought stress and water-stimulus conditions. *PLoS ONE*. 2015;10(5):e0128041. DOI 10.1371/journal.pone.0128041.
- Hamilton J.P., Hansey C.N., Whitty B.R., Stoffel K., Massa A.N., Deynze A., De Jong W., David S., Douches D.S., Buell C.R. Single nucleotide polymorphism discovery in elite north american potato germplasm. *BMC Genomics.* 2011;12:302. DOI 10.1186/1471-2164-12-302.
- Hirsch C.D., Hamilton J.P., Childs K.L., Cepela J., Crisovan E., Vaillancourt B., Hirsch C.N., Habermann M., Neal B., Buell C.R. Spud DB: A resource for mining sequences, genotypes, and phenotypes to accelerate potato breeding. *Plant Genome.* 2014;7(1). DOI 10.3835/plantgenome2013.12.0042.
- Hougas R.W., Peloquin S.J., Ross R.W. Haploids of the common potato. *J. Heredity.* 1958;49:103-106.
- Jupe F., Pritchard L., Etherington G.J., MacKenzie K., Cock P.J.A., Wright F., Sharma S.K., Bolser D., Bryan G.J., Jones J.D.G., Hein I. Identification and localization of the NB-LRR gene family within the potato genome. *BMC Genomics.* 2012;13:75. DOI 10.1186/1471-2164-13-75.
- Khlestkin V.K., Peltek S.E., Kolchanov N.A. Target genes for development of potato (*Solanum tuberosum* L.) cultivars with desired starch properties. *Sel'skokhozyaistvennaya biologiya=Agricultural Biology.* 2017;52(1):25-36. DOI 10.15389/agrobiol.2017.1.25. (in Russian).
- Khlestkina E.K. Genomnoe redaktirovanie kak mashina vremeni, ili domestikatsiya za paru let. *Nauka iz pervykh ruk [Genome editing as 'time machine' or domestication within a couple of years. Science First Hand].* 2016;71-72(5-6):72-75. (in Russian).
- Khlestkina E.K., Shumny V.K. Prospects for application of breakthrough technologies in breeding: The CRISPR/Cas9 system for plant genome editing. *Genetika = Genetics.* 2016;52(7):774-787. (in Russian).
- Khlestkina E.K., Shumny V.K., Kolchanov N.A. Marker-assisted selection and examples of its application in world potato growing. *Dostizheniya nauki i tehniki = Achievements of Science and Technology of AIC.* 2016;30(10):5-8. (in Russian).
- Kolchanov N.A., Kochetov A.V., Salina E.A., Pershina L.A., Khlestkina E.K., Shumny V.K. The state and trends of marker-assisted and genomic selection application in plants. *Vestnik Rossiiskoi Akademii Nauk = Herald of the Russian Academy of Sciences.* 2017;87:40-46. (in Russian).
- Kolesnikov N., Hastings E., Keays M., Melnichuk O., Tang Y.A., Williams E., Dylag M., Kurbatova N., Brandizi M., Burdett T., Megy K., Pilicheva E., Rustici G., Tikhonov A., Parkinson H., Petryszak R., Sarkans U., Brazma A. ArrayExpress update—simplifying data submissions. *Nucl. Acids Res.* 2015;43(D1):D1113-D1116. DOI 10.1093/nar/gku1057.
- Kozomara A., Griffiths-Jones S. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucl. Acids Res.* 2013;42(D1):D68-D73. DOI <https://doi.org/10.1093/nar/gkt1181>.
- Lakhota N., Joshi G., Bhardwaj A.R., Bhardwaj A.R., Katiyar-Agarwal S., Agarwal M., Jagannath A., Goel S., Kumar A. Identification and characterization of miRNAome in root, stem, leaf and tuber developmental stages of potato (*Solanum tuberosum* L.) by high-throughput sequencing. *BMC Plant Biology.* 2014;14:6. DOI 10.1186/1471-2229-14-6.
- Limantseva L., Mironenko N., Shuvalov O., Antonova O., Khiutti A., Novikova L., Afanassenko O., Spooner D., Gavrilenko T. Characterization of resistance to *Globodera rostochiensis* pathotype Ro1 in cultivated and wild potato species accessions. *Plant Breeding.* 2014; 133:660-665. DOI 10.1111/pbr.12195.
- Liu Y., Lin-Wang K., Deng C., Warran B., Wang L., Yu B., Yang H., Wang J., Espley R.V., Zhang J., Wang D., Allan A.C. Comparative transcriptome analysis of white and purple potato to identify genes involved in anthocyanin biosynthesis. *PLoS ONE.* 2015;10(6):e0129148. DOI 10.1371/journal.pone.0129148.
- Martin A., Adam H., Diaz-Mendoza M., Zurczak M., González-Schain N.D., Suárez-López P. Graft-transmissible induction of potato tuberization by the microRNA miR172. *Development.* 2009; 136(17):2873-2881. DOI 10.1242/dev.031658.
- Massa A.N., Childs K.L., Lin H., Bryan G.J., Giuliano G., Buell C.R. The transcriptome of the reference potato genome *Solanum tuberosum* group phureja clone DM1-3 516R44. *PLoS ONE.* 2011;6(10):e26801. DOI 10.1371/journal.pone.0026801.
- Nowicki M., Foolad M.R., Nowakowska M., Kozik E.U. Potato and tomato late blight caused by *Phytophthora infestans*: An overview of pathology and resistance breeding. *Plant Disease.* 2012;96(1):4-17. Available at <http://dx.doi.org/10.1094/PDIS-05-11-0458>.
- Petek M., Rotter A., Kogovšek P., Baebler Š., Mithöfer A., Gruden K. *Potato virus Y* infection hinders potato defence response and renders plants more vulnerable to Colorado potato beetle attack. *Mol. Ecol.* 2014;23(21):5378-5391. DOI 10.1111/mec.12932.
- Ramakrishnan A.P., Ritland C.E., Blas Sevillano R.H., Riseman A. Review of potato molecular markers to enhance trait selection. *Am. J. Potato Res.* 2015;92(4):455-472. DOI 10.1007/s12230-015-9455-7.
- Rensink W.A., Lee Y., Liu J., Iobst S., Ouyang S., Buell C.R. Comparative analyses of six solanaceous transcriptomes reveal a high degree of sequence conservation and species-specific transcripts. *BMC Genomics.* 2005;6:124. DOI 10.1186/1471-2164-6-124.
- Rickert A.M., Kim J.H., Meyer S., Nagel A., Ballvora A., Oefner P.J., Gebhardt C. First generation SNP/InDel markers tagging loci for pathogen resistance in the potato genome. *Plant Biotechnol.* 2003;1: 399-410. DOI 10.1046/j.1467-7652.2003.00036.x.
- Rogozina E.V., Khavkin E.E., Sokolova E.A., Kuznetsova M.A., Gavrilenko T.A., Limantseva L.A., Biryukova V.A., Chalaja N.A.,

- Jones R.W., Deahl K.L. Clone collection of wild species and interspecific hybrids of potato studied phytopathologically and by means of DNA markers. *Trudy po prikladnoi botanike, genetike i selektsii*=Proceedings on Applied Botany, Genetics and Breeding. 2013;174:23-32. (in Russian).
- Rogozina E.V., Shuvalov O.Yu., Antonova O.Yu., Gavrilenko T.A. Interspecific and intraspecific diversity on resistance to Y virus in potato. *Sel'skokhozyaistvennaya biologiya = Agricultural Biology*, 2012;5:64-69. DOI 10.15389/agrobiology.2012.5.64. (in Russian).
- Rosyara U.R., De Jong W.S., Douches D.S., Endelman J.B. Software for genome-wide association studies in autopolyploids and its application to potato. *Plant Genome*. 2016;9(2). DOI 10.3835/plantgenome2015.08.0073.
- Simko I., Haynes K.G., Jones R.W. Assessment of linkage disequilibrium in potato genome with single nucleotide polymorphism markers. *Genetics*. 2006;173(4):2237-2245. DOI 10.1534/genetics.106.060905.
- Slater A.T., Cogan N.O.I., Forster J.W. Cost analysis of the application of marker-assisted selection in potato breeding. *Mol. Breeding*. 2013;32(2):299-310. DOI 10.1007/s11032-013-9871-7.
- Slater A.T., Cogan N.O.I., Forster J.W., Hayes B.J., Daetwyler H.D. Improving genetic gain with genomic selection in autotetraploid potato. *Plant Genome*. 2016;9(3). DOI 10.3835/plantgenome2016.02.0021.
- SolCAP 2016. Available at http://solcap.msu.edu/potato_infinium.shtml. Доступ 08.10.2016.
- Stare T., Ramšak Ž., Blejec A., Stare K., Turnšek N., Weckwerth W., Wienkoop S., Vodnik D., Gruden K. Bimodal dynamics of primary metabolism-related responses in tolerant potato–*Potato virus Y* interaction. *BMC Genomics*. 2015;16(1):716. DOI 10.1186/s12864-015-1925-2.
- Szcześniak M.W., Makałowska I. miRNEST 2.0: a database of plant and animal microRNAs. *Nucl. Acids Res.* 2014;42:D74-D77. DOI <https://doi.org/10.1093/nar/gkt1156>.
- The Potato Genome Sequencing Consortium. Genome sequence and analysis of the tuber crop potato. *Nature*. 2011;475:189-195. DOI 10.1038/nature10158.
- The RNAcentral Consortium. RNAcentral: a comprehensive database of non-coding RNA sequences. *Nucl. Acids Res.* 2017;45(D1):D128-D134. DOI <https://doi.org/10.1093/nar/gkw1008>.
- Uitdewilligen J.G.A.M.L., Wolters A.-M.A., D'hoop B.B., Borm T.J.A., Visser R.G.F., van Eck H.J. A next-generation sequencing method for genotyping-by-sequencing of highly heterozygous autotetraploid potato. *PLoS ONE*. 2013;8(5):e62355. DOI 10.1371/journal.pone.0062355.
- Van Os H., Andrzejewski S., Bakker E., Barrena I., Bryan G.J., Carmel B., Ghareeb B., Isidore E., de Jong W., van Koert P., Lefebvre V., Milbourne D., Ritter E., Ruppe van der Voort J.N.A.M., Rousselle-Bourgeois F., van Vliet J., Waugh R., Visser R.G.F., Bakker J., van Eck H.J. Construction of a 10,000-marker ultradense genetic recombination map of potato: providing a framework for accelerated gene isolation and a genomewide physical map. *Genetics*. 2006;173:1075-1087. DOI 10.1534/genetics.106.055871.
- Veilleux R.E., Boozé-Daniels J., Pehu E. Anther culture of a 2n pollen producing clone of *Solanum phureja* Juz. & Buk. *Can. J. Genet. Cytology*. 1985;27:559-564.
- Visser R.G.F., Bachem C.W.B., de Boer J.M., Bryan G.J., Chakrabarti S.K., Feingold S., Gromadka R., van Ham R.C.H.J., Huang S., Jacobs J.M.E., Kuznetsov B., de Melo P.E., Milbourne D., Orjeda G., Sagredo B., Tang X. Sequencing the potato genome: outline and first results to come from the elucidation of the sequence of the world's third most important food crop. *Am. J. Potato Res.* 2009;86:417-429. DOI 10.1007/s12230-009-9097-8.
- Wang S., Zhang S., Wang W., Xiong X., Meng F., Cui X. Efficient targeted mutagenesis in potato by the CRISPR/Cas9 system. *Plant Cell Rep.* 2015;34(9):1473-1476. DOI 10.1007/s00299-015-1816-7.
- Zhang N., Yang J., Wang Z., Wen Y., Wang J., He W., Liu B., Si H., Wang D. Identification of novel and conserved microRNAs related to drought stress in potato by deep sequencing. *PLoS ONE*. 2014;9(4):e95489. DOI 10.1371/journal.pone.0095489.
- Zhang R., Marshall D., Bryan G.J., Hornyik C. Identification and characterization of miRNA transcriptome in potato by high-throughput sequencing. *PLoS ONE*. 2013;8(2):e57233. DOI 10.1371/journal.pone.0057233.
- Zhang W., Luo Y., Gong X., Zeng W., Li S. Computational identification of 48 potato microRNAs and their targets. *Comput. Biol. Chem.* 2009;33(1):84-93. Available at <http://dx.doi.org/10.1016/j.compbiolchem.2008.07.006>.
- Zuluaga P.A., Solé M., Lu H., Góngora-Castillo E., Vaillancourt B., Coll N., Buell C.R., Valls M. Transcriptome responses to *Ralstonia solanacearum* infection in the roots of the wild potato *Solanum comersonii*. *BMC Genomics*. 2015;16(1):246. DOI 10.1186/s12864-015-1460-1.