

Пятая международная научная конференция PlantGen2019

Обнаружение CRISPR-кассет и генов *cas* в геноме *Arabidopsis thaliana*

Ю.М. Константинов^{1, 2} ✉, И.С. Петрушин²

¹ Сибирский институт физиологии и биохимии растений Сибирского отделения Российской академии наук, Иркутск, Россия

² Иркутский государственный университет, Иркутск, Россия

✉ e-mail: yukon@sifibr.irk.ru

Современный уровень знаний в области эволюции растительных вирусов позволяет отнести проблему генетических основ противовирусного иммунитета высших растений (в том числе важнейших сельскохозяйственных культур) к разряду наиболее актуальных проблем генетики и селекции. В соответствии с эндосимбиотической теорией принято считать, что митохондрии произошли от альфа-протеобактерий, которые были поглощены, но не подвергнуты деструкции клеткой-хозяином. В связи с открытием у прокариот CRISPR-Cas (clustered regularly interspaced short palindromic repeats – CRISPR-associated proteins) систем, выполняющих функцию адаптивного иммунитета, возникает вопрос, мог ли подобный механизм противовирусной защиты быть подхвачен эволюцией и использован представителями эукариот, например растениями. Задачей настоящей работы был анализ полных последовательностей ядерного, митохондриального и хлоропластного геномов *Arabidopsis thaliana* с целью поиска генетических элементов, сходных с таковыми в CRISPR-Cas системах у бактерий и архей. В результате методами *in silico* в митохондриальном геноме экотипов *A. thaliana* обнаружен локус регулярно перемежающихся коротких прямых повторов, соответствующий по своей организации CRISPR-локусу адаптивного CRISPR-Cas иммунитета прокариот. На вероятную связь обнаруженного в митохондриальном геноме высшего растения локуса с функцией адаптивного иммунитета указывает наличие у спейсерных последовательностей в составе найденной CRISPR-кассеты гомологии с геномом вируса мозаики цветной капусты, поражающего растения арабидопсиса. У линий арабидопсиса C24 и Ler последовательности повторов и спейсеров CRISPR-кассеты полностью идентичны. В то же время локализация самого CRISPR-локуса в митохондриальном геноме этих линий существенно различается. Установлено, что у линии Col-0 в результате четырех делеций и одной инсерции CRISPR-кассета полностью нарушена. Хотя гены *cas* в митохондриальном геноме исследуемых экотипов арабидопсиса не были найдены, установлено их наличие в ядерном геноме. В ядерном геноме экотипа Col-0 на всех пяти хромосомах обнаружены гены *cas* и многочисленные CRISPR-касеты. Полученные результаты позволяют предположить существование у растений системы адаптивного иммунитета, аналогичного CRISPR-иммунитету бактерий и архей.

Ключевые слова: *Arabidopsis thaliana*; экотипы; митохондриальный геном; ядерный геном; CRISPR-кассета; гены *cas*; гомология CRISPR-спейсеров; геном растительного вируса; адаптивный иммунитет; РНК-интерференция.

Для цитирования: Константинов Ю.М., Петрушин И.С. Обнаружение CRISPR-кассет и генов *cas* в геноме *Arabidopsis thaliana*. Вавиловский журнал генетики и селекции. 2019;23(7):809-816. DOI 10.18699/VJ19.554

Detection of CRISPR cassettes and *cas* genes in the *Arabidopsis thaliana* genome

Yu.M. Konstantinov^{1, 2} ✉, I.S. Petrushin²

¹ Siberian Institute of Plant Physiology and Biochemistry, SB RAS, Irkutsk, Russia

² Irkutsk State University, Irkutsk, Russia

✉ e-mail: yukon@sifibr.irk.ru

The state of the art in the evolution of plant viruses allows the genetic foundations of antiviral immunity in higher (including the most important crops) plants to be categorized as one of the most pressing issues of genetics and selection. According to the endosymbiotic theory, mitochondria descended from alphaproteobacteria that had been absorbed but not degraded by the host cell. The discovery of CRISPR-Cas systems (clustered regularly interspaced short palindromic repeats (CRISPR)-associated proteins), which implement the adaptive immunity function in prokaryotes, raises the question whether such a mechanism of antiviral protection could be caught up by evolution and used by representatives of eukaryotes (in particular, plants). The purpose of this work was to analyze the complete sequences of nuclear, mitochondrial, and chloroplast genomes of *Arabidopsis thaliana* in order to search for genetic elements similar to those in CRISPR-Cas systems of bacteria and archaea. As a result, *in silico* methods helped us to detect a locus of regularly intermittent short direct repeats in the mitochondrial genome of *A. thaliana* ecotypes. The structure of this locus corresponds to the CRISPR locus of the prokaryotic adaptive

antiviral immune system. The probable connection between the locus found in the mitochondrial genome of the higher plant and the function of adaptive immunity is indicated by a similarity between the spacer sequences in the CRISPR cassette found and the genome of *Cauliflower mosaic virus* affecting *Arabidopsis* plants. Sequences of repeats and spacers of CRISPR cassettes in *Arabidopsis* C24 and Ler lines are perfectly identical. However, the locations of the CRISPR locus in the mitochondrial genomes of these lines differ significantly. The CRISPR cassette in the Col-0 line was found to be completely broken as a result of four deletions and one insertion. Although *cas* genes were not detected in the mitochondrial genome of the studied *Arabidopsis* ecotypes, their presence was detected in the nuclear genome. Both *cas* genes and numerous CRISPR cassettes were found on all the five chromosomes in the nuclear genome of the Col-0 ecotype. The results suggest the existence of a system of adaptive immunity in plants, which is similar to the CRISPR immunity of bacteria and archaea.

Key words: *Arabidopsis thaliana*; ecotypes; mitochondrial genome; nuclear genome; CRISPR cassette; *cas* genes; homology of CRISPR spacers; plant virus genome; adaptive immunity; RNA interference.

For citation: Konstantinov Yu.M., Petrushin I.S. Detection of CRISPR cassettes and *cas* genes in the *Arabidopsis thaliana* genome. *Vavilovskii Zhurnal Genetiki i Seleksii* = Vavilov Journal of Genetics and Breeding. 2019;23(7): 809-816. DOI 10.18699/VJ19.554 (in Russian)

Введение

В настоящее время признается бесспорным, что одним из важнейших событий в возникновении эукариотической клетки стало приобретение архейного происхождения хозяином в качестве эндосимбионта альфа-протеобактерии, давшей впоследствии начало митохондриям (Archibald, 2015). В последние годы методами филогеномики удалось получить принципиально новые данные, демонстрирующие возможность нескольких эволюционных сценариев возникновения эукариотической клетки, включающих «позднее» или «раннее» приобретение митохондрий клеткой-хозяином (Poole, Gribaldo, 2014; Pittis, Gabaldon, 2016). В связи с обнаружением у огромного числа видов бактерий и архей CRISPR-Cas системы адаптивного иммунитета, основанного на явлении РНК-интерференции (Jansen et al., 2002; Mojica et al., 2005; Makarova et al., 2006; Barrangou et al., 2007; Lander, 2016), естественным образом возникает вопрос, возможно ли существование подобной защитной системы у эукариотических митохондрий – органелл, имеющих очевидное эволюционное родство со своим бактериальным происхождением предками. Особый интерес в этом отношении представляют митохондрии высших растений, имеющие геном чрезвычайно большого размера по сравнению с геномами животных и дрожжей.

Митохондриальный геном растений отличается, помимо размеров, необычной динамичностью, проявляющейся в виде высокой рекомбинационной активности, которая обусловлена повторяющимися последовательностями (Gualberto, Newton, 2017). Результатом этой рекомбинационной активности являются образование набора субгеномных форм и высокая геномная вариабельность даже в пределах одного вида. Подобные изменения в геномной структуре приводят к быстрой эволюции митохондриального генома растений. Кроме того, митохондриальный геном высших растений активно участвует в процессах горизонтального переноса генов, где может выступать в качестве как донора, так и акцептора генов (Kleine et al., 2009; Zhao et al., 2018).

Еще одной важной особенностью митохондриального генома высших растений является наличие в органеллах у многих исследованных в этом отношении видов растений видоспецифических наборов линейных и кольцевых плазмид, состав которых в пределах вида может существенно

различаться, например, у фертильных и стерильных форм (Esser et al., 1986; Thomas, 1986). Происхождение митохондриальных плазмид до сих пор остается неизвестным. Предполагают, что двуцепочечные плазмиды могли быть внесены в клетки высших растений симбиотическим или патогенным путем (Douce, Neuburger, 1989). В поддержку этой гипотезы свидетельствует тот факт, что митохондриальные линейные плазмиды своим 5'-концом ассоциированы с белком, что напоминает структуру некоторых вирусных нуклеиновых кислот (Douce, Neuburger, 1989). Кроме того, в пользу вероятного вирусного происхождения линейных плазмид S1 и S2 митохондрий кукурузы свидетельствует обнаружение в них генов, кодирующих белки нуклеинового обмена вирусного типа (Kuzmin, Levchenko, 1987; Kuzmin et al., 1988). В последние годы наблюдается значительный прогресс в изучении митовирусов – вирусов с простейшим РНК-геномом, которые специфически инфицируют митохондрии грибов (Shahi et al., 2019). Однако при этом появляются также данные в пользу существования растительных митовирусов, которые, как предполагают, могли возникнуть в результате горизонтального переноса соответствующих генов от инфицирующих растения грибов (Marienfeld et al., 1997; Bruenn et al., 2015; Nibert et al., 2018). Таким образом, если проводить аналогии между бактериями и растительными митохондриями, у последних в ходе эволюции, подобно прокариотам, также имела острая потребность в защите от инфицирующих нуклеиновых кислот вирусного и/или плазмидного происхождения.

Тем не менее вплоть до недавнего времени данных о существовании подобного механизма защиты от патогенной ДНК у представителей эукариот получено не было, за исключением единичного факта обнаружения в работе (Mojica et al., 2000) типичного CRISPR-локуса на митохондриальной плазмиде высшего растения *Vicia faba*. Дальнейшего развития в направлении поиска генов *cas* в митохондриальном, хлоропластном и ядерном геномах этого растительного вида данная работа, однако, не получила. Сведений об обнаружении генетических элементов CRISPR-Cas иммунитета в ядерном геноме растений ранее также не было получено (Jansen et al., 2002).

С учетом эволюционного происхождения митохондрий и особенностей организации митохондриального генома растений в настоящей работе методами *in silico* предпри-

нята попытка поиска генетических элементов, сходных с элементами CRISPR-Cas систем бактерий и архей, в митохондриальном, хлоропластном и ядерном геномах модельного растения *Arabidopsis thaliana*. Принимая во внимание высокую динамичность митохондриального генома растений, полногеномный анализ с целью поиска элементов, предположительно связанных с признаком адаптивного CRISPR-Cas иммунитета, проводили в митохондриальном геноме трех экотипов *A. thaliana* (C24, Ler и Col-0).

Материалы и методы

В качестве объекта исследования использованы полные последовательности ядерного генома (экотип Col-0), генома митохондрий (экотипы C24, Ler, Col-0) и генома хлоропластов (экотипы Ler, Col-0) модельного растения *Arabidopsis thaliana* (L.) Heynh. Последовательности ДНК взяты из базы GenBank (номера доступа ядерного генома NC_003070, NC_003071, NC_003074, NC_003075, NC_003076, митохондриального генома JF729200, JF729202, NC_037304 и генома хлоропластов KX551970, NC_000932). Изначально последовательности анализировали с помощью программы UGENE (Okonechnikov et al., 2012), она же применялась для построения иллюстраций (совместно с пакетом векторной графики Inkscape).

Поиск элементов CRISPR-Cas систем в геномах осуществлялся в онлайн-сервисе CRISPROne (Zhang, Ye, 2017). Для определения происхождения обнаруженных CRISPR-спейсеров выполнен поиск по базе NCBI BLAST (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) по вирусным таксонам с параметрами по умолчанию. В дальнейшем были отобраны случаи совпадения с числом отличий (mismatches) менее трех нуклеотидов.

Выравнивание последовательностей CRISPR-локуса в составе митохондриального генома экотипов *A. thaliana* проводилось программами Matcher (парное) (Rice et al., 2000) и MUSCLE (множественное) (Edgar, 2004). Анализ гомологии CRISPR-спейсеров с геномами видоспецифичных вирусов осуществлялся как в работе (Mihara et al., 2016), согласно Virus-Host DB (<https://www.genome.jp/virushostdb/3702>).

Результаты и обсуждение

К настоящему времени убедительно продемонстрировано, что критическими элементами CRISPR-Cas систем у бактерий и архей являются в общем случае CRISPR-локус, лидерная последовательность (an upstream leader sequence) и гены *cas* (Jansen et al., 2002; Richter et al., 2012). Основываясь на известном эволюционном родстве митохондрий и бактерий, мы осуществили поиск элементов CRISPR-Cas системы в митохондриальном геноме трех экотипов *A. thaliana* с использованием подходов и методов биоинформатики, широко применяемых в последние годы при изучении CRISPR-Cas систем прокариот (Jansen et al., 2002; Makarova et al., 2006, 2015; Grissa et al., 2007; Zhang, Ye, 2017; Couvin et al., 2018).

Контекстный анализ полной последовательности митохондриального генома *A. thaliana* таких экотипов, как C24 и Ler, позволил обнаружить сайт, структура которого полностью соответствует организации CRISPR-кассет

прокариотического происхождения. Особенности нуклеотидной организации CRISPR-подобного локуса в митохондриальном геноме данных экотипов представлены на рис. 1, а. Как видно из представленных данных, обнаруженная в митохондриальном геноме растительного происхождения CRISPR-касета образована тремя совершенными прямыми повторами длиной 20 п. н., разделенными двумя спейсерными последовательностями длиной 42 и 33 п. н. соответственно. В то же время полногеномный анализ митохондриальной ДНК экотипа Col-0 показал, что структура CRISPR-касеты в этом случае полностью нарушена в результате четырех делеций и одной инсерции в области повторов (см. рис. 1, б).

Примечательным, на наш взгляд, результатом проведенного анализа экотип-специфических особенностей характеристик митохондриальной CRISPR-касеты является тот факт, что при полном совпадении последовательности повторов и спейсеров локализация CRISPR-касеты (и ее нарушенного варианта) в митохондриальном геноме линий арабидопсиса C24, Ler и Col-0 существенно различается (рис. 2). Такие изменения локализации CRISPR-касеты в митохондриальной ДНК изучаемых экотипов арабидопсиса являются, по всей вероятности, результатом интенсивных перестроек митохондриального генома, обусловленных высокой рекомбинационной активностью, что характерно для митохондриального генома высших растений (Gualberto, Newton, 2017).

Специальный поиск позволил установить наличие многочисленных CRISPR-кассет в ядерном геноме *A. thaliana* (рис. 3). Их размеры и расположение на хромосомах представлены в электронном Приложении¹. Общее число спейсеров, содержащихся в 110 ядерных CRISPR-касетах, составляет 330. Подробный анализ гомологии спейсеров ядерных CRISPR-кассет с геномами растительных вирусов нами не проводился.

Результаты анализа спейсерных последовательностей в CRISPR-касете, локализованной в митохондриальной ДНК арабидопсиса с использованием базы данных растительных вирусов, суммированы в табл. 1. Установлено, что обнаруженные спейсеры имеют участки неслучайной гомологии с геномом трех штаммов вируса мозаики цветной капусты, который способен поражать растения *A. thaliana*. Дополнительно можно отметить, что для отдельных спейсеров были выявлены участки гомологии с несовпадающими участками генома разных штаммов этого вируса (данные не приводятся).

Поиск в митохондриальном геноме экотипов C24, Ler и Col-0 *A. thaliana* генов *cas* ни на последовательностях, непосредственно прилегающих к CRISPR-локусу, ни в остальных областях генома не дал положительных результатов. В то же время последовательности отдельных генов *cas* были обнаружены нами в ядерном геноме (табл. 2).

Проведенный *in silico* поиск позволил локализовать на трех хромосомах арабидопсиса (хромосомы 1, 2 и 3) ген *cas5*, входящий, в соответствии с существующей классификацией CRISPR-Cas систем (Makarova et al., 2015; Koonin et al., 2017), в состав эффекторного модуля CRISPR-Cas систем I типа. На хромосоме 3 помимо гена

¹ Приложение см. по адресу:

<http://www.bionet.nsc.ru/vogis/download/pict-2019-23/appx18.pdf>

а

| | | | |
|-----|--------|--|--------|
| C24 | 86224 | AACTCGACTGAAAGGAGAGG TTGTGAACACAACTCGACTGAAAGGAGAG | 86273 |
| Ler | 163198 | AACTCGACTGAAAGGAGAGG TTGTGAACACAACTCGACTGAAAGGAGAG | 163247 |
| C24 | 86274 | GTTGTGAACACAACT CGACTGAAAGGAGAGG TCCAAGGTAATTTATTAC | 86323 |
| Ler | 163248 | GTTGTGAACACAACT CGACTGAAAGGAGAGG TCCAAGGTAATTTATTAC | 163297 |
| C24 | 86324 | TCTTATAAAAGAGGG AACTCGACTGAAAGGAGAGG | 86358 |
| Ler | 163298 | TCTTATAAAAGAGGG AACTCGACTGAAAGGAGAGG | 163332 |

б

| | |
|-------|---|
| C24 | AACTCGACTGAAAGGAGAGGTTGTGAACACAACTCGACTGAAAGGAGAGGTTGTGAACA |
| Ler | AACTCGACTGAAAGGAGAGGTTGTGAACACAACTCGACTGAAAGGAGAGGTTGTGAACA |
| Col-0 | -ACTC-----AGAGATCAGGA----AACTTGATAGAAAG-----CCCTAGCT |
| | **** |
| C24 | CAAAC TCGACTGAA -----AGGAGAGGTCCAAGGTAATTTATTACTC |
| Ler | CAAAC TCGACTGAA -----AGGAGAGGTCCAAGGTAATTTATTACTC |
| Col-0 | CTAACTTGATAGAAATGCCCCAGCAATGGCCATCGGAGAGGTCCAACGTAATTTATTACTC |
| | * |
| C24 | TTATAAAAGAGGGAACTCGACTGAAAGGAGAGG |
| Ler | TTATAAAAGAGGGAACTCGACTGAAAGGAGAGG |
| Col-0 | TTATAAAAGAGGGAACTCGACTGAAAGGAGAGG |
| | ***** |

Рис. 1. Парное выравнивание (а) CRISPR-локуса экотипов C24 и Ler (повторы выделены жирным шрифтом) и множественное выравнивание (б) последовательностей CRISPR-локуса в митохондриальном геноме экотипов *A. thaliana*.

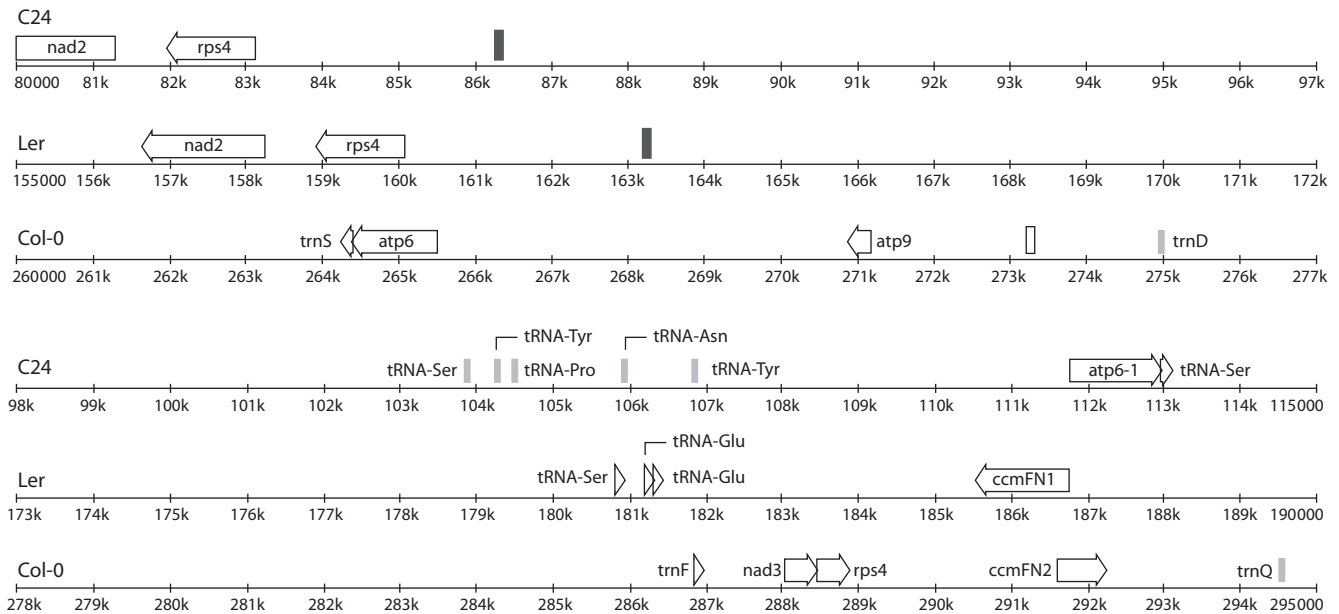


Рис. 2. Локализация CRISPR-кассеты в митохондриальном геноме экотипов *A. thaliana*.

Для линий C24 и Ler положение CRISPR-кассеты показано в виде черного прямоугольника. Для линии Col-0 положение нарушенной CRISPR-кассеты в митохондриальном геноме обозначено в виде незаполненного прямоугольника.

cas5 локализован ген *csm6*, кодирующий ассоциированную с CRISPR-Cas системой III-A типа РНКазу, которая у прокариот участвует в реализации противовирусного иммунитета путем деградации фаговых транскриптов (Jiang et al., 2016). В составе хромосомы 4 был обнаружен ген *csa5*, белковый продукт которого является универсальным компонентом CRISPR-Cas систем I-A типа (Daume et al., 2014). Данному белку приписывают роль в стабили-

зации R-петли в ходе протекания стадии интерференции (Daume et al., 2014). Наконец, в составе хромосомы 5 локализовано три участка разной длины, соответствующих гену *DEDDh*, который ранее был аннотирован в ядерном геноме арабидопсиса и является представителем генного семейства 3'→5' экзонуклеаз, участвующих в метаболизме малых некодирующих РНК (Chen et al., 2018). Отнесение *DEDDh* к генам *cas* может означать, что его белковый про-

дукт *in vivo* выполняет несколько функций, в том числе функцию, связанную с защитой растения от нуклеиновых кислот вирусов и плазмид.

Удалось установить, что ассоциированные с CRISPR-Cas системами типа I и III гены обратной транскриптазы (RT) представлены на всех пяти хромосомах арабидопсиса значительным числом копий (всего 43 гена) (см. табл. 2). Этим ферментам приписывается особо важная роль в функционировании CRISPR-Cas систем типа III, которая состоит во включении в уже существующую CRISPR-кассету новых спейсеров как с непосредственным участием RT, так и с участием продукта слияния RT с белком Cas-1 (RT-Cas1 fusion protein) (Silas et al., 2016; Togo et al., 2017).

Таким образом, проведенный нами поиск элементов CRISPR-Cas систем в митохондриальном и ядерном геномах *A. thaliana* позволил впервые выявить в геноме этого растения основные генетические элементы адаптивного иммунитета прокариотического типа, включающие CRISPR-локусы и гены *cas*. За исключением одной CRISPR-касеты в митохондриальном геноме, остальные структурные элементы системы локализованы в ядерном геноме. В целом пока неполный список найденных в геноме *A. thaliana* ассоциированных с CRISPR-Cas иммунитетом прокариотического типа генов (*cas5*, *cas6*, *cas5*, *cd06127*, RT) позволяет в соответствии с классификацией (Makarova et al., 2015; Koonin et al., 2017) предварительно отнести обнаруженную в данном модельном растении систему к системам класса 1, имеющим многосубъединичный эффекторный модуль.

В ходе исследования нам не удалось, однако, выявить ни в митохондриальном, ни в ядерном геноме *A. thaliana* структуры, соответствующей характеристикам лидерных последовательностей прокариотического типа (Alkhnabshi et al., 2016). Попытка поиска элементов CRISPR-Cas систем в хлоропластном геноме экотипов Leg и Col-0 *A. thaliana* также не дала положительных результатов.

В нашей работе впервые для представителей высших эукариот установлено наличие таких канонических элементов CRISPR-Cas систем прокариотического происхождения, как CRISPR-локусы и гены *cas* в геноме модель-

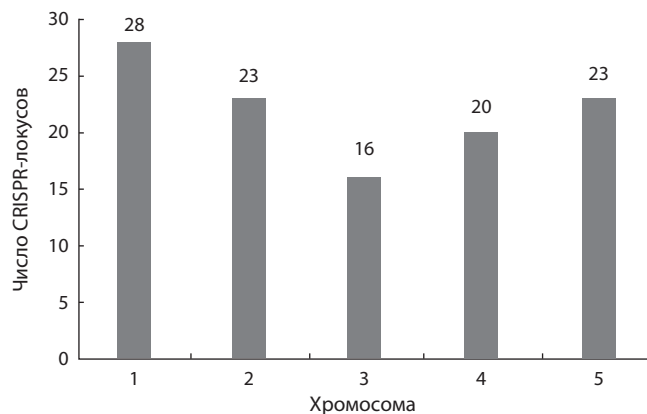


Рис. 3. Количественное распределение CRISPR-кассет по отдельным хромосомам ядерного генома *A. thaliana*.

ного растения *A. thaliana*. При этом в митохондриальном геноме арабидопсиса удалось обнаружить лишь единичную CRISPR-кассету, спейсерные последовательности которой проявляют неслучайную гомологию с геномом вируса мозаики цветной капусты (см. табл. 1). Важно подчеркнуть, что данный вирус способен инфицировать растения исследуемого вида. Для изученных линий арабидопсиса характерны существенные различия в геномной локализации CRISPR-локуса экотипов C24 и Leg. В случае экотипа Col-0 структура CRISPR-касеты полностью нарушена в результате четырех делеций и одной инсерции в области прямых повторов (см. рис. 1, б). Такой результат свидетельствует об активно идущих у высших растений процессах реорганизации митохондриального генома, проявляющихся в высокой скорости возникновения межлинейных различий на уровне митохондриальной ДНК.

Возможное существование CRISPR-Cas иммунитета у растений с эволюционной точки зрения представляется вполне оправданным, поскольку такие ДНК-содержащие органеллы растений, как митохондрии и хлоропласты, с очевидностью представляют собой привлекательные мишени для атак со стороны вирусов и плазмид чужеродного происхождения. Особенно уязвимы в этом отношении

Таблица 1. Выравнивание обнаруженных в митохондриальном геноме *A. thaliana* CRISPR-спейсеров с геномом вируса мозаики цветной капусты

| Выравнивание* | Изолят | Идентификатор доступа в NCBI GenBank |
|--|-----------|--------------------------------------|
| Спейсер 1 TCCAAGGTAATTTATTTACTCTTATAAAAAGAGGG | Cabb B-JI | KJ716236 |
| KJ716236 7512 AAGGGAAATTAGGGTTCTTATA 7533 | | |
| Спейсер 1 TCCAAGGTAATTTATTTACTCTTATAAAAAGAGGG | D/H | M10376 |
| M10376 3301 TACAAGAAAAAATATAAGGCTTATAAA 3327 | | |
| Спейсер 2 TTGTGAACACAAACTCGACTGAAAGGAGAGGTTGTGAACACA . . . | TUR239 | AB863182 |
| AB863182 1782 ATAAACTCGA-TCAAAGAAG 1800 | | |

* Цифрами обозначена локализация участка гомологии в геноме вируса.

Таблица 2. Характеристика обнаруженных в ядерном геноме *A. thaliana* генов *cas*

| Хромосома | Ген | Тип CRISPR-Cas системы | Начало–конец OPC* | Идентификация гена <i>cas</i> в соответствии с классификацией (Makarova et al., 2015; Koonin et al., 2017) |
|-------------------|--------------|------------------------|-------------------|--|
| 1 | <i>cas5</i> | I | 11918718–11919893 | COG1688 |
| 2 | <i>cas5</i> | I | 2540882–2541148 | cd09693 |
| 3 | <i>cas5</i> | I-B | 12023929–12024258 | mkCas0191 |
| | <i>csm6</i> | III-A, III-D | 10339967–10340668 | cd09742 |
| 4 | <i>csa5</i> | I-A | 5072131–5072607 | mkCas0163 |
| 5 | <i>DEDDh</i> | I | 9481429–9481962 | cd06127 |
| | | | 9482093–9482377 | |
| | | | 9482399–9482647 | |
| 1 | RT | I, III | 11916164–11916466 | pfam00078 |
| | | | 11916554–11917162 | |
| 2 | | | 10826–12517 | |
| | | | 17624–20293 | |
| | | | 23971–25653 | |
| | | | 30408–31532 | |
| | | | 3223463–3224872 | |
| | | | 3224999–3225685 | |
| | | | 3228729–3230522 | |
| | | | 5235099–5235800 | |
| | | | 5241327–5241869 | |
| | | | 5241873–5242298 | |
| | | | 5600424–5602496 | |
| | | | 5602506–5603024 | |
| | | | 5619463–5619846 | |
| | | | 5619956–5620156 | |
| 3 | | | 9435745–9436644 | |
| | | | 9436766–9437260 | |
| | | | 9444515–9444907 | |
| | | | 9445029–9445523 | |
| | | | 10341391–10343106 | |
| | | | 10343111–10343806 | |
| | | | 12021817–12022311 | |
| | | | 13167127–13168218 | |
| | | | 13174969–13175118 | |
| | | | 13175226–13175606 | |
| | | | 13467703–13469727 | |
| | | | 13469756–13470868 | |
| | | | 13476637–13477317 | |
| | | | 15489669–15490154 | |
| 15490132–15490506 | | | | |
| 15493988–15496318 | | | | |
| 4 | | | 3296005–3296457 | |
| | | | 3303483–3304247 | |
| | | | 3304260–3305717 | |
| | | | 3744233–3746887 | |
| | | | 3749914–3752901 | |
| | | | 3754635–3756176 | |
| | | | 5071003–5071551 | |
| | | | 5081577–5082176 | |
| 5 | | | 14576286–14579228 | |
| | | | 14582070–14582471 | |
| | | | 14582509–14584656 | |

* OPC – открытая рамка считывания.

митохондрии растений, если принять во внимание существование митовирусов, поражающих этот тип органелл (Marienfeld et al., 1997; Bruenn et al., 2015; Nibert et al., 2018), и природную компетентность растительных митохондрий к поглощению ДНК (Koulintchenko et al.,

2003). В целом, исходя из имеющейся на сегодняшний день информации, представляется преждевременным строить какие-либо гипотезы о возможном эволюционном происхождении элементов CRISPR-Cas системы в геноме арабидопсиса. Можно лишь заметить следующее. Сейчас

накоплено уже достаточно много экспериментальных фактов для реконструкции сценариев происхождения и эволюции CRISPR-Cas систем прокариот (Koonin, Makarova, 2019). Эти же данные, в свою очередь, могут оказаться полезными, если проблему происхождения CRISPR-Cas систем у растений попытаться рассмотреть с позиций эукариогенеза (Koonin, 2015; Lopez-Garia et al., 2017). В этом случае необходимо принять во внимание, что подобную систему защиты от патогенных нуклеиновых кислот мог иметь как альфа-протеобактериальный симбионт, давший начало митохондриям, так и, возможно, архейный хозяин протомитохондриального эндосимбионта. В принципе, наличия подобной защитной системы в эволюционном плане нельзя совершенно исключить и для цианобактериального предшественника современных хлоропластов.

На настоящем этапе изучения CRISPR-Cas систем у арабидопсиса нам не удалось идентифицировать известный для прокариот набор генов *cas*, продукты которых формируют модуль адаптации и эффекторный модуль CRISPR-Cas системы класса I и обеспечивают, таким образом, протекание стадий адаптации, экспрессии и интерференции (Koonin et al., 2017). Учитывая сведения о большом разнообразии найденных у бактерий и архей CRISPR-Cas систем (Westra et al., 2016; Koonin et al., 2017; Koonin, Makarova, 2019), естественно ожидать, что механизм адаптивного иммунитета у растений может существенно отличаться от такового у прокариот. В связи с этим, очевидно, что только использование комплексного подхода, включающего наряду с методами геномики методы транскриптомики и протеомики, позволит в дальнейшем составить более полное представление о генах и их белковых продуктах, образующих систему адаптивного иммунитета растений.

Заключение

В геноме высшего растения *A. thaliana* методами *in silico* впервые обнаружены такие элементы CRISPR-Cas системы прокариот, как CRISPR-касеты и гены *cas*, что может служить отправной точкой для проведения в дальнейшем детальных геномных, транскриптомных и протеомных исследований с использованием наряду с арабидопсисом более широкого набора видов растений (в том числе важнейших сельскохозяйственных) с целью определения групп генов, экспрессия которых может быть ассоциирована с работой природного механизма адаптивного иммунитета растений. Прикладное значение ожидаемых научных результатов по проблеме молекулярной природы адаптивного иммунитета растений трудно переоценить.

Список литературы / References

- Alkhnbashi O.S., Shah S.A., Garret R.A., Saundlers S.J., Costa F., Backhoven R. Characterizing leader sequences of CRISPR loci. *Bioinformatics*. 2016;32:i576-i585. DOI 10.1093/bioinformatics/btw454.
- Archibald J.M. Endosymbiosis and eukaryotic cell evolution. *Curr. Biology*. 2015;25:R911-R921. DOI 10.1016/j.cub.2015.07.055.
- Barrangou R., Freemaux C., Deveau H., Richards M., Boyaval P., Moineau S., Romero D.A., Horvath P. CRISPR provides acquired resistance against viruses in prokaryotes. *Science*. 2007;315:1709-1712. DOI 10.1126/science.1138140.
- Bruenn J.A., Warner B.E., Yerramsetty P. Widespread mitovirus sequences in plant genomes. *PeerJ*. 2015;3:e876. DOI 10.7717/peerj.876.
- Chen J., Liu L., You C., Gu J., Ruan W., Zhang L., Gan J., Cao C., Huang Y., Chen X., Ma J. Structural and biochemical insights into small RNA 3' end trimming by *Arabidopsis* SDN1. *Nat. Commun*. 2018;9:3585. DOI 10.1038/s41467-018-05942-7.
- Couvin D., Bernheim A., Toffano-Nioche C., Touchon M., Michalik J., Néron B., Rocha E., Vergnaud G., Gautheret D., Pourcel C. CRISPRCasFinder, an update of CRISPRFinder, includes a portable version, enhanced performance and integrates search for Cas proteins. *Nucleic Acids Res*. 2018;46:W246-W251. DOI 10.1093/nar/gky425.
- Daume M., Plagens A., Randau L. DNA binding properties of the small cascade subunit Csa5. *PLoS One*. 2014;9:e105716. DOI 10.1371/journal.pone.0105716.
- Douce R., Neuburger M. The uniqueness of plant mitochondria. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 1989;40:371-414. DOI 10.1146/annurev.pp.40.060189.002103.
- Edgar R.C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*. 2004;32(5):1792-1797. DOI 10.1093/nar/gkh340.
- Esser K., Kück V., Lang-Hinrichs C., Lemke P.A., Osiewacz H.D., Stahl P., Tudzynski U. *Plasmids of Eukaryotes. Fundamentals and Applications*. Berlin; Heidelberg: Springer-Verlag, 1986.
- Grissa I., Vergnaud G., Pourcel C. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res*. 2007;35:W52-W57. DOI 10.1093/nar/gkm360.
- Gualberto J.M., Newton K.J. Plant mitochondrial genomes: dynamics and mechanism of mutation. *Annu. Rev. Plant Biol.* 2017;68:225-252. DOI 10.1146/annurev-arplant-043015-112232.
- Jansen R., van Embden D.A., Gaastera W., Schouls L.M. Identification of genes that are associated with DNA repeats in prokaryotes. *Mol. Microbiol.* 2002;43:1565-1575. DOI 10.1046/j.1365-2958.2002.02839.x.
- Jiang V.V., Samai P., Marraffini L.A. Degratation of phage transcripts by CRISPR-associated RNases enables type III CRISPR-Cas immunity. *Cell*. 2016;164:710-721. DOI 10.1016/j.cell.2015.12.053.
- Kleine T., Maier U.W., Leister D. DNA transfer from organelles to the nucleus: the idiosyncratic genetics of endosymbiosis. *Annu. Rev. Plant Biol.* 2009;60:115-138. DOI 10.1101/cshperspect.a015990.
- Koonin E.V. Origin of eukaryotes from within archaea, archaeal eukaryote and bursts of gene gain: eukaryogenesis just made easier? *Phil. Trans. R. Soc. B*. 2015;370:201403333. DOI 10.1098/rstb.2014.0333.
- Koonin E.V., Makarova K.S. Origins and evolution of CRISPR-Cas systems. *Phil. Trans. R. Soc. B*. 2019;374(1772):20180087. DOI 10.1098/rstb.2018.0087.
- Koonin E.V., Makarova K.S., Zhang F. Diversity, classification and evolution of CRISPR-Cas systems. *Curr. Opin. Microbiol.* 2017;37:67-78. DOI 10.1016/j.mib.2017.05.008.
- Koulintchenko M., Konstantinov Y., Dietrich A. Plant mitochondria actively import DNA via the permeability transition pore complex. *EMBO J*. 2003;22:1245-1254. DOI 10.1093/emboj/cdg128.
- Kuzmin E.V., Levchenko I.V. S1 plasmid from cms-S-maize mitochondria encodes a viral type DNA-polymerase. *Nucleic Acids Res*. 1987;15:6758. DOI 10.1093/nar/15.16.6758.
- Kuzmin E.V., Levchenko I.V., Zaitseva G.N. S2 plasmid from cms-S-maize mitochondria potentially encodes a specific RNA polymerase. *Nucleic Acids Res*. 1988;16:4177. DOI 10.1093/nar/16.9.4177.
- Lander E.S. The heroes of CRISPR. *Cell*. 2016;164:18-28. DOI 10.1016/j.cell.2015.12.041.
- Lopez-Garcia P., Eme L., Moreira D. Symbiosis in eukaryotic evolution. *J. Theor. Biol.* 2017;434:20-33. DOI 10.1016/j.jtbi.2017.02.031.
- Makarova K.S., Grishin N.V., Shabalina S.A., Wolf Y.I., Koonin E.V. A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biol. Direct*. 2006;1:7. DOI 10.1186/1745-6150-1-7.

- Makarova K.S., Wolf Y.I., Alkhnbashi O.S., Costa F., Shah S.A., Saunders S.J., Barrangou R., Brouns S.J.J., Charpentier E., Haft D.H., Horvath R., Moineau S., Mojica F.J.M., Terns R.M., Terns M.P., White M.F., Yakunin A.F., Garrett R.A., van der Oost J., Backofen R., Koonin E.V. An updated evolutionary classification of CRISPR-Cas systems. *Nat. Rev. Microbiol.* 2015;13:1-15. DOI 10.1038/nrmicro3569.
- Marienfeld J.R., Unsel M., Brandt P., Brennicke A. Viral nucleic acid sequence transfer between fungi and plants. *Trends Genet.* 1997;13: 260-261. DOI 10.1016/S0168-9525(97)01168-2.
- Mihara T., Nishimura Y., Shimizu Y., Nishiyama H., Yoshikawa G., Uehara H. Linking virus genomes with host taxonomy. *Viruses.* 2016;8:66. DOI 10.3390/v8030066.
- Mojica F.J.M., Diez-Villasenor C., Soria E., Juez G. Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. *Mol. Microbiol.* 2000;36:244-246. DOI 10.1046/j.1365-2958.2000.01838.x.
- Mojica F.J., Garcia-Martinez J., Soria E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.* 2005;60:174-182. DOI 10.1007/s00239-004-0046-3.
- Nibert M., Vong M., Fugate K.K., Debat H.J. Evidence for contemporary plant mitoviruses. *Virology.* 2018;518:14-24. DOI 10.1016/j.virol.2018.02.005.
- Okonechnikov K., Golosova O., Fursov M. Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics.* 2012;28;1166-1167. DOI 10.1093/bioinformatics/bts091.
- Pittis A.A., Gabaldon T. Late acquisition of mitochondria by a host with chimeric prokaryotic ancestry. *Nature.* 2016;531(7592):101-104. DOI 10.1016/S0168-9525(00)02024-2.
- Poole A.M., Gribaldo S. Eukaryotic origins: how and when was the mitochondrion acquired? *Cold Spring Harb. Perspect. Biol.* 2014;6: a015990. DOI 10.1101/cshperspect.a015990.
- Rice P., Longden I., Bleasby A. EMBOSS: the European molecular biology open software suite. *Trends Genet.* 2000;16(6);276-277. DOI 10.1016/S0168-9525(00)02024-2.
- Richter C., Chang J.T., Fineran P. Function and regulation of clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR associated (Cas) systems. *Viruses.* 2012;4:2291-2311. DOI 10.3390/v4102291.
- Shahi S., Eusebio-Cope A., Kondo H., Hillman B.I., Suzuki N. Investigation of host range of and host defense against a mitochondrially replicating mitovirus. *J. Virol.* 2019;93(6):e01503-18. DOI 10.1128/JVI.01503-18.
- Silas S., Mohr G., Sidote D.J., Markham L.M., Sanchez-Amat A., Bhaya D., Lambowitz A.M., Fire A.Z. Direct CRISPR spacer acquisition from RNA by a natural reverse transcriptase-Cas1 fusion protein. *Science.* 2016;351(6276):aad4234. DOI 10.1126/science.aad4234.
- Thomas C.M. The nucleotide sequence and transcription of minicircular mitochondrial DNA's associated with fertile and cytoplasmic male-sterile lines of sugar beet. *Nucleic Acids Res.* 1986;14:9353-9370. DOI 10.1093/nar/14.23.9353.
- Toro N., Martinez-Abarca F., Gonzalez-Delgado A. The reverse transcriptases associated with CRISPR-Cas systems. *Sci. Rep.* 2017;7: 7089. DOI 10.1038/s41598-017-07828-y.
- Westra E.R., Dowling A.J., Broniewski J.M., van Houte S. Evolution and ecology of CRISPR. *Annu. Rev. Ecol. Evol. Syst.* 2016;47:307-334. DOI 10.1146/annurev-ecolsys-121415-032428.
- Zhang Q., Ye Y. Not all predicted CRISPR-Cas systems are equal: isolated *cas* genes and classes of CRISPR like elements. *BMC Bioinformatics.* 2017;18:92. DOI 10.1186/s12859-017-1512-4.
- Zhao N., Wang Y., Hua J. The roles of mitochondrion in intergenomic gene transfer in plants: a source and a pool. *Int. J. Mol. Sci.* 2018;19:547. DOI 10.3390/ijms19020547.

ORCID ID

Yu.M. Konstantinov orcid.org/0000-0002-0601-2788
I.S. Petrushin orcid.org/0000-0002-8788-5352

Благодарности. Авторы выражают благодарность Ian Max Møller и R. Shyama Prasad Rao за проявленный к работе интерес, критические замечания и полезное обсуждение. Отдельная благодарность А. Фадееву за ценные советы и полезные замечания по структуре статьи. Авторы чрезвычайно признательны анонимному рецензенту, чьи замечания и конструктивный анализ первоначального варианта статьи способствовали ее существенной переработке, приведшей к качественному улучшению изложения полученных результатов.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию 05.06.2019. После доработки 20.08.2019. Принята к публикации 21.08.2019.