



УДК 582.751.92+575.113+575.89

Изучение эволюции цитрусовых растений на основании повторяющихся элементов генома, хлоропластного генома и 35S рДНК

М. Г. Монахос^{1,2*}, Д. С. Ульянов^{1,3}, Д. В. Романов^{1,4}, Ю. Н. Смотрова^{1,5}, В. А. Коробкова^{1,6}

¹ Всероссийский научно-исследовательский институт сельскохозяйственной биотехнологии, ул. Тимирязевская, д. 42, г. Москва, 127550, г. Россия.

² E-mail: mihaluch1864@gmail.com; ORCID iD: <https://orcid.org/0009-0003-6604-0822>

³ E-mail: uldas1508@gmail.com; ORCID iD: <https://orcid.org/0000-0002-5880-5931>

⁴ E-mail: akabos1987@gmail.com; ORCID iD: <https://orcid.org/0000-0002-7495-7572>

⁵ E-mail: juliksmotrova@gmail.com; ORCID iD: <https://orcid.org/0009-0005-9179-2203>

⁶ E-mail: bowlingistka@gmail.com; ORCID iD: <https://orcid.org/0009-0009-2330-3681>

*Автор для переписки

Ключевые слова: повторы, репитом, филогения, *Citrus*, LTR.

Аннотация. В нашем исследовании мы проанализировали репитом 12 растений рода *Citrus*, включая 3 предковых вида и 8 наиболее распространённых таксонов, и обнаружили у них высокое содержание LTR элементов Athila и Sire, нетипичное для цветковых растений. Для этих растений, на основании прочтений из открытых баз данных и впервые полученных прочтений из нашей коллекции, были построены филогенетические деревья на основе ядерного репитома, пластома и 35S рДНК цистрона. При этом было показано, что пластом несёт выраженный филогенетический сигнал, наиболее согласующийся с обновлённой классификацией рода *Citrus*, показывая картину, отличную от двух других филогенетических деревьев.

The study of citrus plant evolution based on repetitive genomic elements, chloroplast genome, and 35S rDNA

M. G. Monakhos, D. S. Ulyanov, D. V. Romanov, Y. N. Smotrova, V. A. Korobkova

All-Russian Research Institute of Agricultural Biotechnology, Timiryazevskaya St., 42, Moscow, 127550, Russian Federation

Keywords: *Citrus*, LTR, phylogeny, repeatome, repeats.

Summary. In our study, we analyzed the repeatomes of 12 *Citrus* plants, including 4 ancestral species and 8 the most widespread cultivated species. We detected an atypically high content of “Athila” and “Sire” LTR retrotransposons for angiosperms. For these plants, phylogenetic trees were constructed based on the nuclear repeatome, plastome, and 35S rDNA cistron, using sequencing data from open databases and, for the first time, sequencing data from our collection. The analysis demonstrated that the plastome carries a strong phylogenetic signal, which is most consistent with the updated classification of the *Citrus* genus, showing a distinct pattern compared to the other two phylogenetic trees.

Введение

Цитрусовые растения (род *Citrus* L.) крайне популярны и распространены по всему миру. За 2023 г. мировой урожай всех культивируемых цитрусовых культур (в основном лимон, лайм, апельсин, помело, грейпфрут и мандарин) составил более 169 млн тонн (FAOSTAT. URL: <https://www.fao.org>). Цитрусовые растения ценятся за счёт высоких вкусовых качеств, декоративности, содержания эфирных масел, микро- и макроэлементов, кроме того, они широко используются в народной медицине в регионах выращивания (Lu et al., 2023).

Несмотря на свою распространённость в мире, растения рода *Citrus* не имеют единой общепринятой классификации, а их систематика чрезвычайно запутана и разрознена. В ранних исследованиях Такака в 1954 г. выделил 157 видов *Citrus*, в 1977 г. – 162 вида рода (Tanaka, 1954, 1977). В период с 1943 по 1967 гг. Swingle основывал свою классификацию на морфологических и биохимических признаках, факторах вегетативного размножения, в частности, факультативном апомиксисе (Swingle, 1943, 1967). Таким образом, он и другие исследователи признали 17 видов (Bhattacharya, Dutta, 1956), 36 видов (Hodgson, 1961) и 31 вид (Singh, Nath, 1969). Ollitrault заявил, что такое количество видов является следствием того, что большинство из них представляют собой сортовые группы клонального происхождения внутри одного вида и появились вследствие факультативного апомиксиса. Ollitrault предложил триномиальную концепцию классификации, учитывающую ретикулярную эволюцию, частичный апомиксис и бесполое разнообразие от предковых гибридов, но при этом отметил, что в настоящее время геномные данные разрознены, и для окончательной классификации р. *Citrus* необходимы дальнейшие исследования (Ollitrault et al., 2020).

До сих пор не существует единой системы классификации р. *Citrus*. В течение длительного времени несколько исследователей классифицировали род по-разному (Swingle, 1943; Mabberley, 1997; Wu et al., 2018). Показателен пример *Citrus hindsii* Govaerts, 1999 (*Fortunella japonica* (Thunb.) Swingle, 1915), который разные авторы относят то к р. *Citrus* (Mabberley, 1998; Wang et al., 2022), то к р. *Fortunella* Swingle, 1915 (Swingle, 1915; Kashkush et al., 2003; Ma, Clemants, 2006). Сегодня продолжается масштабное изменение классификации цитрусовых, в рамках которого Ву и Оллитраутом (Wu et al., 2018; Ollitrault et

al., 2020) было предложено включить новокаледонский р. *Oxanthera* Montrouz., а также ранее исключённые р. *Microcitrus* Swingle, *Eremocitrus* Swingle, *Clymenia* Swingle, *Poncirus* Raf. и *Fortunella* в р. *Citrus*.

В настоящее время существует 5 видов, которые считаются родоначальниками р. *Citrus*: *C. cavaleriei* H. Lév., *C. maxima* (Burm.) Merr., *C. medica* Linnaeus, *C. micrantha* Wester, *C. reticulata* Blanco (Ollitrault et al., 2020). В нашем исследовании мы проанализировали 4 из них: *C. cavaleriei*, *C. maxima*, *C. medica* и *C. reticulata*, а также 8 наиболее распространённых таксонов *C. sinensis* Sauvage et Dabry de Thiersant, *C. unshiu* Marow., *C. hindsii* (Champ. ex Benth.) Govaerts, *C. clementina* Yu. Tanaka, *C. limon* (L.) Osbeck, *C. mangshanensis* S. W. He et G. F. Liu, *C. trifoliata* (L.) Raf. и цитрумело (*Citrus* × *paradisi* × *Citrus trifoliata*).

Хорошо известно, что большая часть генома растений состоит из повторяющейся ДНК (Puterova et al., 2017). Существует два типа повторяющейся ДНК: транспозируемые элементы и тандемные повторы (Ahmed, Liang, 2012), причём первые подразделяют на ретротранспозоны и ДНК-транспозоны. В свою очередь, ретротранспозоны, на основании наличия длинных терминальных повторов (от английского long terminal repeats, LTR), делятся на не-LTR ретротранспозоны и LTR ретротранспозоны. Транспозоны (от английского transposable elements, TE) – это последовательности ДНК, способные перемещаться в пределах генома. Так, у некоторых растений ретроэлементы с LTR составляют значительную часть ядерного генома, и их доля может превышать 50 % (Domingues et al., 2012; Vitte et al., 2014; Du et al., 2018; Neumann et al., 2019). TE обладают способностью распространяться в геноме и увеличивать свою численность. Однако у высших организмов развились регуляторные системы, такие как метилирование ДНК, которые ограничивают распространение TE: при этом транскрипция и транспозиция TE эпигенетически подавляются из-за их потенциального негативного влияния на организм.

Было установлено, что достоверные данные о репитеме могут быть получены при полногеномном секвенировании с низким охватом в диапазоне от 0,5 до 0,01 (Novák et al., 2010). Для работы с репитомом растений ввиду удобного интерфейса и широкого функционала можно использовать программу RepeatExplorer2, которая с помощью кластеризации аннотирует повторы и проводит их количественную оценку (Novák et

al., 2010, 2013, 2020). Dodsworth и др. использовали такую сравнительную кластеризацию для отслеживания филогенетических сигналов из повторов путём подсчёта количества прочтений по видам в каждом кластере (Dodsworth et al., 2015). Vitales и др. рассчитывали парные генетические расстояния между всеми кластерами, после чего строили консенсусное дерево на основании расстояний (Vitales et al., 2020). Herklotz и др. корректировали количество подаваемых прочтений для RepeatExplorer2 относительно размера генома, чтобы нивелировать разницу в обилии повторов у разных видов. Также они суммировали все различия для каждого кластера в основной матрице расстояний (Herklotz et al., 2021).

Несмотря на то, что большая часть видов цитрусовых является диплоидами с хромосомным числом $2n = 18$, по размеру генома данная группа достаточно вариативна. Единственные приведённые в базе королевских ботанических садов Кью (KEW. URL: <https://cvalues.science.kew.org/>) данные по размеру генома видов *Citrus* относятся к 5 видам: *C. medica* – 0,40 пг/1С, *C. limon* – 0,40 пг/1С, *C. reticulata* – 0,47 пг/1С (Marie, Brown, 1993), *C. sinensis* – 0,60 пг/1С (Guerra, 1984). Также, по нашим оценкам на основе данных, взятых из других источников, можно судить о размерах геномов следующих видов: *C. trifoliata* – 0,27 пг/1С (Peng et al., 2020), *C. unshiu* – 0,36 пг/1С (Isobe et al., 2023), *C. cavaleriei* – 0,37 пг/1С (Wang et al., 2017), *C. hindsii* – 0,38 пг/1С (Zhu et al., 2019), *C. maxima* – 0,39 пг/1С, *C. clementina* – 0,31 пг/1С, *C. mangshanensis* – 0,37 пг/1С (Wu et al., 2024), цитрумело – 0,27 пг/1С (Zhang et al., 2016). Средний размер генома составляет 0,35 пг; самый маленький геном у цитрумело, самый большой – у *C. sinensis*. Подобные различия у родственных видов как правило обусловлены именно накоплением повторяющихся последовательностей ДНК, причём в разных группах могут накапливаться разные классы повторов и с разной скоростью (Lozano et al., 2021; Kress et al., 2022; Nie et al., 2024).

В нашем исследовании мы сконструировали филогенетические деревья на основе репитома (Herklotz et al., 2021), 35S рДНК цистрона (Garcia et al., 2020) и пластома (Moreno-Aguilar et al., 2020) цитрусовых для проведения сравнительного анализа методик; а также проанализировали содержание LTR элементов в геноме каждого изученного растения.

Материалы и методы

Образцы

В исследовании использовались парные прочтения Illumina 12 видов *Citrus*. Прочтения образцов *C. unshiu* (SRR15099909), *C. cavaleriei* (SRR17023266), *C. maxima* (SRR25409037), *C. hindsii* (SRR22345194), *C. mangshanensis* (SRR17013816), цитрумело (*Citrus* × *paradisi* × *Citrus trifoliata*) (SRR1210645) были получены из библиотеки NCBI. Образцы *C. clementina*, *C. limon*, *C. medica*, *C. reticulata*, *C. sinensis*, *C. trifoliata* были отобраны из коллекции Всероссийского научно-исследовательского института сельскохозяйственной биотехнологии (ВНИИСБ) и секвенированы авторами. Секвенированные последовательности были загружены в NCBI (PRJNA1231762).

Растения р. *Citrus* из коллекции ВНИИСБ и их секвенирование

Во ВНИИСБ была собрана коллекция цитрусовых растений. Растения получили из коллекции Российского государственного аграрного университета – МСХА им. К. А. Тимирязева, а также коммерческих питомников: «Павловский питомник», «Павловский лимон», «Садовод Крыма».

Для выделения ДНК по протоколу СТАВ (Rogers, Bendich, 1985) использовали молодые, свежие листья растений. ДНК, растворённую в свободной от нуклеаз воде, использовали для полного геномного секвенирования. Концентрацию, количество и качество выделенной ДНК проверяли с помощью спектрофотометра NanoDrop OneC (Thermo Fisher Scientific Waltham, MA, United States). Образцы ДНК с показателями OD260/280 в диапазоне от 1,8 до 2,0 и OD260/230 в диапазоне от 1,8 до 2,0 считались качественными. Библиотеки кДНК синтезировали по технологии “Shotgun” с помощью набора Swift 2S® Turbo DNA Library Kit (Swift Bioscience, Ann Arbor, MI, USA) в соответствии с протоколом производителя. Секвенирование проводили на приборе Illumina NextSeq с набором NextSeq 500/550 Mid Output Kit v2.5 (300 циклов), как описано в протоколах Illumina для “paired-end” прочтений. Секвенирование проводилось в ООО «Геномед» (Москва, Россия).

Контроль качества прочтений

Отобранные прочтения загружались на платформу Galaxy (GALAXY. URL: <https://galaxy-elixir.>

cerit-sc.cz/), после чего были проверены на качество, очищены от адаптеров и прочтений короче 148 п. н. Оставшиеся прочтения были объединены и конвертированы в формат FASTA с помощью инструментов “FastQC”, “Trimmomatic” и “Preprocessing of FASTQ paired-end reads” с параметрами по умолчанию.

Кластеризация и аннотация повторов и рДНК в Repeat Explorer

Анализ состава и долей в геноме повторов был проведён с помощью инструмента RepeatExplorer2 в соответствии с методом, описанным Novák и др. (Novák et al., 2020). Для индивидуального анализа нами отбирался случайным образом один миллион пар прочтений для достижения рекомендованного покрытия генома (0,1×–0,5×). Для сравнительного анализа отбирались количества прочтений, обратно пропорциональные ожидаемым размерам генома видов (Herklotz et al., 2021). Кластеризация проводилась с использованием базовых настроек RepeatExplorer2 (90 % сходство, минимальное покрытие = 55; порог размера кластера = 0,01 %) с параметром времени обработки “extra long”.

Сборка пластома и цистрона

Мы провели *de novo* сборку пластома и 35S рДНК цистрона для всех 12 образцов.

Для сборки пластома использовали прочтения, которые прошли контроль качества и были объединены в один fasta-файл. Сборка проводилась с помощью программы Novoplasty версии 4.3.5 (Dierckxsens et al., 2016), в качестве референсного генома использовали пластом *C. limon* (MT880608.1), в качестве начального участка сборки – ген RuBP (V00171.1). Сборки хлоропластного генома приложены к статье (см. прил. 1 на сайте журнала).

Сборку 35S рДНК цистрона проводили в программе QIAGEN CLC Genomics Workbench 24.0.2 (<https://digitalinsights.qiagen.com/>). В сборке также использовали объединённые прочтения, прошедшие контроль качества, с картированием на референсную последовательность 35S рДНК *C. reticulata* (ON479678.1) (Moreno-Aguilar et al., 2020, 2022).

Построение филогенетического дерева на основе репитома

Филогенетический анализ проводился с использованием данных о копийности кластеров, полученных в результате сравнительного кла-

стерного анализа RepeatExplorer2. Были удалены данные о кластерах митохондриальной и хлоропластной ДНК, а также кластеры с низкой копийностью (менее 50 прочтений в кластере). Данные были нормализованы – количество прочтений каждого кластера было переведено в процент от общего количества прочтений всех кластеров, без учёта кластеров с низкой копийностью. Филогенетическое дерево по методу ближайших соседей было построено в R 4.4.0 (R core team, 2024): сначала из нормализованных данных была получена матрица евклидовых расстояний с использованием функции distance пакета phylentropy (Drost, 2018), далее с помощью функции NJ (“neighbor joining”, метод ближайших соседей) пакета ape (Paradis et al., 2004) было получено дерево, которое отредактировали в iTol версии 6 (Letunic, Bork, 2024).

Построение филогенетических деревьев на основе последовательностей пластома и цистрона

Собранные пластома выровняли с помощью программы MAFFT версии 7 (Katoh, Standley, 2013), после построили в ней филогенетическое дерево методом ближайших соседей. Также на основе пластома было построено дерево ML (“maximum likelihood”, метод максимального правдоподобия) в R 4.4.0 с помощью функционала пакета phangorn (Schliep, 2011). Сборки рДНК цистрона выравняли с помощью функционала CLC, в этой же программе было построено NJ дерево на основе выравниваний. Деревья отредактировали для визуализации в iTol.

Результаты и их обсуждение

Состав репитома

В наших образцах мы идентифицировали 8 групп повторяющихся последовательностей из 16, принадлежащих к суперсемейству Copia: Ale, Alesia, Angela, Ikeros, Ivana, SIRE, Tar, Tork; и 6 групп из 14, относящихся к Gypsy: CRM, Galadriel, Reina, Tekay, Athila, Retand.

Группа SIRE имеет наибольшее содержание в надсемействе Copia (1,91–2,74 %). Группа Alesia имеет наименьшее содержание (0,02–0,05 %) и присутствует только в таксонах цитру-мело, *C. limon*, *C. reticulata*, *C. trifoliata*, *C. unshiu*. В надсемействе Gypsy наиболее представленной группой является Athila (5,24–9,59 %), а наименее представленной – Reina (0,02–0,09 %), которая не представлена в таксонах *C. cavaleriei*,

C. clementina и цитрумело (табл. 1). При этом, по литературным данным, у цветковых растений, как правило, преобладают группы Retand и Angela (Burke et al., 2002). У цитрусов из нашего

анализа эти линии имели высокое, но не максимальное содержание в геноме: Retand (0,86–2,17 %) и Angela (0,66–0,97 %).

Таблица 1. Содержание LTR элементов в геноме растений р. *Citrus*

Lineage	<i>Citrus cavaleriei</i>	Цитрумело	<i>Citrus clementina</i>	<i>Citrus hindsii</i>	<i>Citrus limon</i>	<i>Citrus mangshanensis</i>	<i>Citrus maxima</i>	<i>Citrus medica</i>	<i>Citrus reticulata</i>	<i>Citrus sinensis</i>	<i>Citrus trifoliata</i>	<i>Citrus undshiu</i>
Ale	0,09	0,63	0,05	0,06	1,02	0,09	0,07	0,23	0,18	0,21	0,17	0,26
Alesia	0,00	0,04	0,00	0,00	0,05	0,00	0,00	0,00	0,05	0,00	0,05	0,02
Angela	0,93	0,69	0,67	0,97	0,82	0,95	0,76	0,86	0,64	0,66	0,67	0,71
Ikeros	0,00	0,23	0,10	0,04	0,24	0,08	0,05	0,12	0,10	0,14	0,06	0,02
Ivana	0,07	0,14	0,16	0,11	0,21	0,14	0,12	0,15	0,16	0,17	0,20	0,14
SIRE	2,10	1,75	1,91	2,68	3,79	2,45	1,96	2,27	2,73	2,05	2,74	2,29
Tar	0,09	0,22	0,26	0,23	0,33	0,21	0,23	0,25	0,21	0,29	0,22	0,28
Tork	0,60	1,22	1,06	0,57	1,83	0,82	0,74	1,20	1,16	1,26	1,10	1,24
Copia total	3,88	4,92	4,21	4,66	8,29	4,74	3,93	5,08	5,23	4,78	5,21	4,96
CRM	0,54	0,59	0,63	0,22	0,42	0,19	0,51	0,87	0,17	0,72	0,12	0,62
Galadriel	0,17	0,18	0,25	0,15	0,25	0,16	0,13	0,20	0,38	0,19	0,32	0,17
Reina	0,00	0,00	0,00	0,02	0,09	0,02	0,00	0,06	0,02	0,03	0,02	0,03
Tekay	0,19	0,33	0,49	0,04	0,55	0,24	0,01	0,26	0,54	0,49	0,56	0,52
Athila	5,76	5,24	7,35	5,66	9,59	6,17	6,08	7,82	5,99	7,55	6,07	6,54
Retand	0,91	0,86	1,60	0,95	2,17	1,15	0,95	1,56	1,22	1,59	1,38	1,14
Gypsy Total	7,57	7,20	10,32	7,04	13,07	7,93	7,68	10,77	8,32	10,57	8,47	9,02
Unclass	1,06	0,72	0,80	0,63	0,21	0,71	2,75	1,07	1,20	0,93	1,18	2,06
TOTAL	12,5	12,84	15,33	12,33	21,57	13,38	14,36	16,92	14,75	16,28	14,86	16,04

Филогенетические деревья

На основе данных репитома и последовательностей 35S рДНК цистрона были построены филогенетические деревья методом NJ. Для пластома дополнительно использовали метод ML. Все полученные деревья укоренились от *C. trifoliata*. Деревья, построенные на основе пластома разными методами (NJ и ML), продемонстрировали высокую степень сходства (рис. 1). Расхождения между ними были незначительными: в 4 узлах с поддержкой bootstrap < 100 % наблюдались различия в значениях в пределах 20,5 %; виды одинаково группировались в клады, однако расположение видов внутри своих клад менялось. Для последующего анализа мы выбрали NJ-дерево на базе пластома (рис. 1а), так как в нём использовался тот же алгоритм (NJ), что и при создании деревьев на основе репитома и 35S рДНК цистрона, а различия между NJ и ML-топологиями пластома были статистически несущественны.

Филогенетические деревья, построенные методом NJ на базе репитома, цистрона и пластома были сопоставлены на рис. 2. Деревья на основании разных данных имеют существенные различия по распределению таксонов, это связано с разной интенсивностью филогенетического сигнала, в зависимости от исходных генетических данных и методики определения филогенетических связей. Наибольшее сходство с существующими положениями о таксономии цитрусов, по нашим данным, имеет филогенетическое дерево на основе хлоропластного генома (рис. 2В) (Ollitrault et al., 2020), выраженное, например, близким расположением *C. maxima* и *C. sinensis*. Нахождение *C. reticulata* и *C. clementina* в одном узле дерева (рис. 2В) соответствует результатам предыдущих исследований пластома р. *Citrus* (Sun, Lin, 2021). Однако есть и несоответствие существующей классификации: так, близкие виды *C. sinensis* и *C. reticulata* находятся, по на-

шим данным, на отдалении друг от друга. Такое отдаление может быть следствием неполной достоверности филогенетического сигнала, полученного из пластома, недостаточной выборки растений, или же возможных недостатков существующей систематики, которая интенсивно дополняется и корректируется.

Филогенетическое дерево на основе 35S рДНК цистрона (рис. 2С) может быть недо-

верным ввиду относительно коротких прочтений (до 429 пар нуклеотидов), полиморфизмы в которых оказывают чрезмерное или недостаточное влияние при филогенетическом анализе. Расхождение образцов у дерева на основе репитома (рис. 2С) не показало связи с размером генома и также, как и дерево на основе 35S рДНК имеет низкое соответствие с существующими положениями о систематике р. *Citrus*.

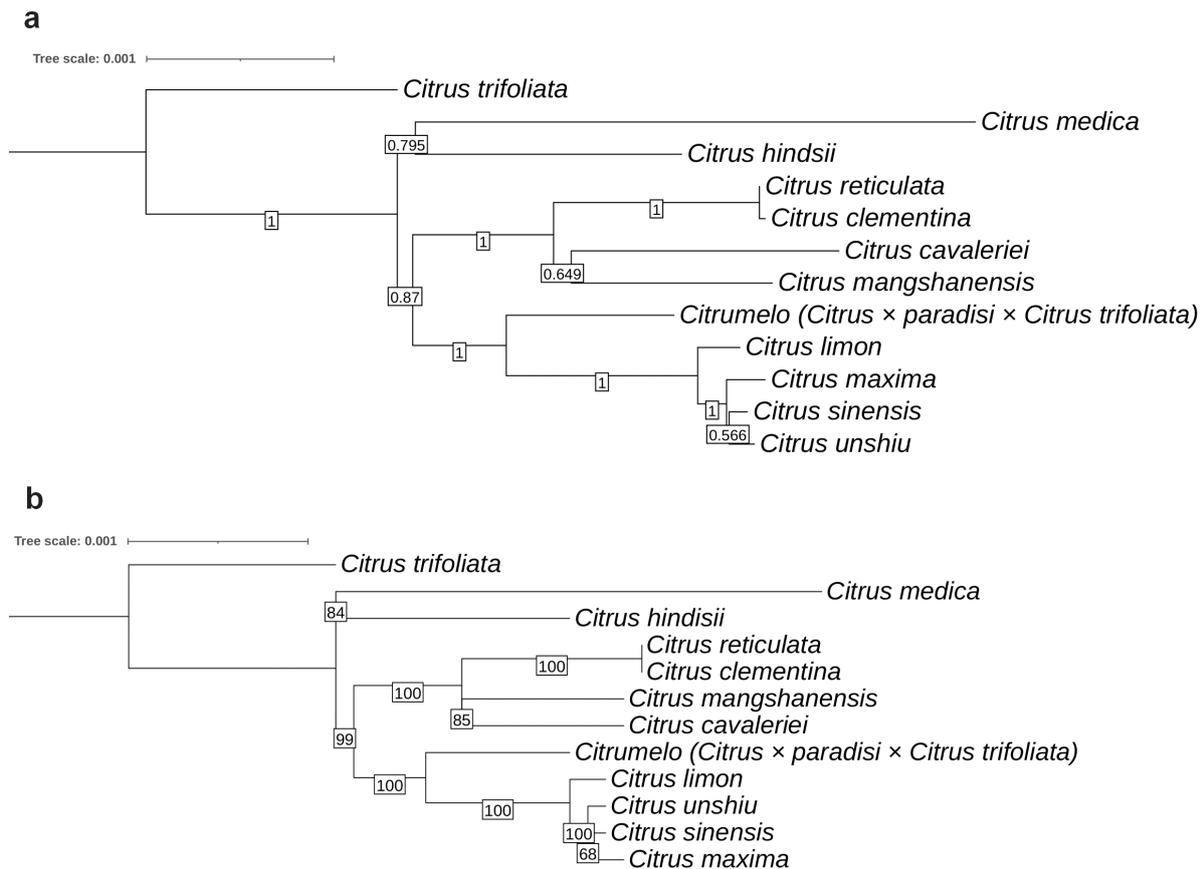


Рис. 1. Филогенетические деревья двенадцати представителей р. *Citrus* на основе пластома, полученные ML методом (a) и NJ методом (b).

Несмотря на ограниченный филогенетический сигнал данных о 35S рДНК цистроне и репитоме генома для реконструкции эволюционных деревьев, эти элементы сохраняют ценность при изучении внутривидовой и межвидовой изменчивости в цитогенетике. Так, последовательности 35S рДНК цистрона широко применяются в качестве FISH-проб для хромосомного картирования (Chen et al., 2008; Garcia et al., 2010; He et al., 2021). Однако их использование сопряжено с рядом проблем: отсутствием унифицированной терминологии, техническими ограничениями методов и субъективностью интерпретации ре-

зультатов (Rosselló et al., 2022). Отдельный интерес представляют данные о репитоме – видоспецифичных повторах, которые также служат основой для разработки FISH-маркеров. Например, тандемные повторы, характерные для конкретных таксонов, успешно используются для идентификации хромосомных перестроек и анализа геномной архитектуры (Lee et al., 2018; Novák et al., 2020). Таким образом, интеграция данных о цистронах и репитоме с другими цитогенетическими подходами позволяет преодолеть ограничения отдельных методов, обеспечивая комплексный анализ эволюционных процессов.

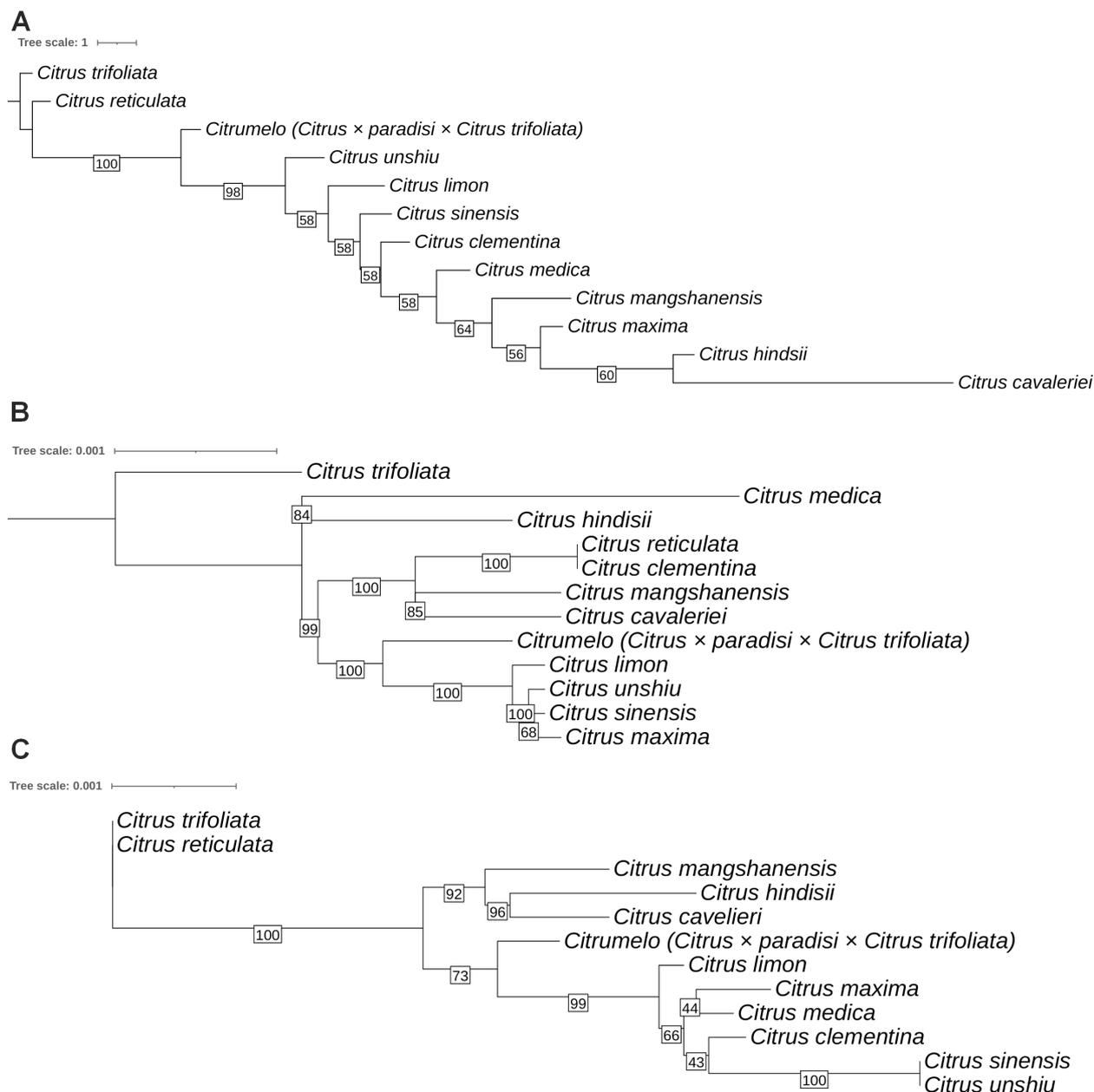


Рис. 2. Филогенетические деревья двенадцати представителей р. *Citrus* на основе репитома (А), пластома (В) и 35S рДНК цитрона (С).

Заключение

Полученные нами филогенетические деревья на основе копииности повторов и 35S рДНК цитрона имели низкое соответствие с существующими положениями таксономии р. *Citrus*. Данные на основе хлоропластного генома показали большее соответствие с существующей систематикой. Пропорции разных групп LTR элементов в геноме растений р. *Citrus* имеют отличия, от-

носительно других цветковых растений. Если у большинства покрытосеменных растений преобладают LTR элементы групп Retand и Angela, то у описанных растений р. *Citrus* преобладают группы Athila и SIRE.

Благодарности

Работа выполнена при финансовой поддержке гранта РНФ № 23-16-00234.

REFERENCES / ЛИТЕРАТУРА

- Ahmed M., Liang P.** 2012. Transposable elements are a significant contributor to tandem repeats in the human genome. *International Journal of Genomics* 1: 1–7. <https://doi.org/10.1155/2012/947089>
- Bhattacharya S., Dutta S.** 1956. *Classification of citrus fruits of Assam*. New Dehli: ICAR. 110 pp.
- Burke W. D., Malik H. S., Rich S. M., Eickbush T. H.** 2002. Ancient lineages of non-LTR retrotransposons in the primitive eukaryote, *Giardia lamblia*. *MBE* 19: 619–630. <https://doi.org/10.1093/oxfordjournals.molbev.a004121>
- Chen G. F., Wang G. C., Zhang C. Y., Zhang B. Y., Wang X. K., Zhou B. C.** 2008. Development of rRNA and rDNA-targeted probes for fluorescence *in situ* hybridization to detect *Heterosigma akashiwo* (Raphidophyceae). *J. Exp. Mar. Biol. Ecol.* 355: 66–75.
- Dierckxsens N., Mardulyn P., Smits G.** 2016. NOVOPlasty: *de novo* assembly of organelle genomes from whole genome data. *Nucleic Acids Res.* 45, 4: e18–e18. <https://doi.org/10.1093/nar/gkw955>
- Dodsworth S., Chase M. W., Kelly L. J., Leitch I. J., Macas J., Novak, P., Piednoel M., Weiss-Schneeweiss H., Leitch A. R.** 2015. Genomic repeat abundances contain phylogenetic signal. *System. Biolog.* 64: 112–126. <https://doi.org/10.1093/sysbio/syu080>
- Domingues D. S., Cruz G. M., Metcalfe C. J., Nogueira F. T., Vicentini R., de S Alves C., Van Sluys M.-A.** 2012. Analysis of plant LTR-retrotransposons at the fine-scale family level reveals individual molecular patterns. *BMC Genom.* 13: 1–14.
- Drost H. G.** 2018. Philentropy: information theory and distance quantification with R. *JOSS* 3: 765. <https://doi.org/10.21105/joss.00765>
- Du D., Du X., Mattia M. R., Wang Y., Yu Q., Huang M., Yu Y., Grosser J. W., Gmitter F. G.** 2018. LTR retrotransposons from the *Citrus × clementina* genome: characterization and application. *Tree Genet. Genomes* 14: 1–14.
- FAO [2025]. FAOSTAT: Agricultural Production Data. URL: <https://www.fao.org/>.
- Garcia S., Panero J. L., Siroky J., Kovarik A.** 2010. Repeated reunions and splits feature the highly dynamic evolution of 5S and 35S ribosomal RNA genes (rDNA) in the Asteraceae family. *BMC Plant Biol.* 10: 1–18.
- Garcia S., Wendel J. F., Borowska-Zuchowska N., Ainouche M., Kuderova, A., Kovarik A.** 2020. The utility of graph clustering of 5S ribosomal DNA homoeologs in plant allopolyploids, homoploid hybrids, and cryptic introgressants. *Front. Plant Sci.* 11: 41. <https://doi.org/10.3389/fpls.2020.00041>
- Guerra M. dos S.** 1984. Cytogenetics of rutaceae. II. Nuclear DNA content. *Caryologia* 37: 219–226.
- He J., Lin S., Yu, Z., Song A., Guan Z., Fang W., Chen S., Zhang F., Jiang J., Chen F.** 2021. Identification of 5S and 45S rDNA sites in *Chrysanthemum* species by using oligonucleotide fluorescence *in situ* hybridization (Oligo-FISH). *Mol. Biol. Rep.* 48: 21–31.
- Herklotz V., Kovařík A., Wissemann V., Lunerová J., Vozárová R., Buschmann S., Olbricht K., Groth M., Ritz C. M.** 2021. Power and weakness of repetition-evaluating the phylogenetic signal from repeatomes in the family Rosaceae with two case studies from genera prone to polyploidy and hybridization (*Rosa* and *Fragaria*). *Front. Plant Sci.* 12: 738119. <https://doi.org/10.3389/fpls.2021.738119>
- Hodgson R. W.** 1961. Taxonomy and nomenclature in citrus. In: *IOCV Conference Proceedings (Gainesville, 1961)*. Vol 2(2). Riverside: University of California. Pp. 1–7.
- Isobe S., Fujii, H., Shirasawa K., Kawahara Y., Endo T., Shimada T.** 2023. Haploid-resolved and chromosome-scale genome assembly in *Citrus unshiu* and its parental species, *C. nobilis* and *C. kinokuni*. *bioRxiv* 2023.06.02.543356. <https://doi.org/10.1101/2023.06.02.543356>
- Kashkush K., Feldman M., Levy A. A.** 2003. Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat. *Nat. Genet.* 33: 102–106.
- Katoh K., Standley D. M.** 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *MBE* 30: 772–780.
- Kress W. J., Soltis D. E., Kersey P. J., Wegrzyn J. L., Leebens-Mack J. H., Gostel M. R., Liu X., Soltis P. S.** 2022. Green plant genomes: What we know in an era of rapidly expanding opportunities. *Proc. Natl. Acad. Sci. U.S.A.* 119, 4: e2115640118. <https://doi.org/10.1073/pnas.2115640118>
- Lee Y.-I., Yap J. W., Izan S., Leitch I. J., Fay M. F., Lee Y.-C., Hidalgo O., Dodsworth S., Smulders M. J. M., Gravendeel B., Leitch A. R.** 2018. Satellite DNA in *Paphiopedilum* subgenus *Parvisepalum* as revealed by high-throughput sequencing and fluorescent *in situ* hybridization. *BMC Genom.* 19: 578. <https://doi.org/10.1186/s12864-018-4956-7>
- Letunic I., Bork P.** 2024. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Res.* 52, 1: W78–W82. <https://doi.org/10.1093/nar/gkae268>
- Lozano R., Gazave E., Dos Santos J. P. R., Stetter M. G., Valluru R., Bando N., et al.** 2021. Comparative evolutionary genetics of deleterious load in sorghum and maize. *Nat. Plants* 7: 17–24. <https://doi.org/10.1038/s41477-020-00834-5>

- Lu X., Zhao C., Shi H., Liao Y., Xu F., Du H., Xiao H., Zheng J.** 2023. Nutrients and bioactives in citrus fruits: Different citrus varieties, fruit parts, and growth stages. *Crit. Rev. Food Sci. Nutr.* 63: 2018–2041. <https://doi.org/10.1080/10408398.2021.1969891>
- Ma J., Clemants S.** 2006. A history and overview of the Flora Reipublicae Popularis Sinicae (FRPS, Flora of China, Chinese edition, 1959–2004). *Taxon* 55: 451–460. <https://doi.org/10.2307/25065592>
- Mabberley D. J.** 1997. *The plant-book: a portable dictionary of the vascular plants*. Cambridge: Cambridge university press. 161 pp.
- Mabberley D. J.** 1998. Australian *Citreae* with notes on other *Aurantioideae* (Rutaceae). *Telopea* 7: 333–344.
- Marie D., Brown S. C.** 1993. A cytometric exercise in plant DNA histograms, with 2C values for 70 species. *Biol. Cell* 78: 41–51.
- Moreno-Aguilar M. F., Arnelas I., Sánchez-Rodríguez A., Viruel J., Catalán P.** 2020. Museomics unveil the phylogeny and biogeography of the neglected Juan Fernandez Archipelago *Megalachne* and *Podophorus* endemic grasses and their connection with relict Pampean-Ventanian fescues. *Front. Plant Sci.* 11: 819. <https://doi.org/10.3389/fpls.2020.00819>
- Moreno-Aguilar M. F., Inda L. A., Sánchez-Rodríguez A., Arnelas I., Catalán P.** 2022. Evolutionary dynamics of the repeatome explains contrasting differences in genome sizes and hybrid and polyploid origins of grass loliinae lineages. *Front. Plant Sci.* 13: 901733. <https://doi.org/10.3389/fpls.2022.901733>
- Neumann P., Novák P., Hošťáková N., Macas J.** 2019. Systematic survey of plant LTR-retrotransposons elucidates phylogenetic relationships of their polyprotein domains and provides a reference for element classification. *Mob. DNA* 10: 1–17.
- Nie Y., Liu X., Zhao L., Huang Y.** 2024. Repetitive element expansions contribute to genome size gigantism in *Pamphagidae*: A comparative study (Orthoptera, *Acridoidea*). *Genomics* 116: 110896. <https://doi.org/10.1016/j.ygeno.2024.110896>
- Novák P., Neumann P., Macas J.** 2010. Graph-based clustering and characterization of repetitive sequences in next-generation sequencing data. *BMC Bioinform.* 11: 378. <https://doi.org/10.1186/1471-2105-11-378>
- Novák P., Neumann P., Macas J.** 2020. Global analysis of repetitive DNA from unassembled sequence reads using RepeatExplorer2. *Nat. Protoc.* 15: 3745–3776. <https://doi.org/10.1038/s41596-020-0400-y>
- Novák P., Neumann P., Pech J., Steinhaisl J., Macas J.** 2013. RepeatExplorer: a Galaxy-based web server for genome-wide characterization of eukaryotic repetitive elements from next-generation sequence reads. *Bioinformatics* 29: 792–793. <https://doi.org/10.1093/bioinformatics/btt054>
- Ollitrault P., Curk F., Krueger R.** 2020. Chapter 4 – *Citrus* taxonomy. In: J. M. Talon, M. Caruso, F. G. Gmitter (eds.). *The Genus Citrus*. Woodhead Publishing. Pp. 57–81. <https://doi.org/10.1016/B978-0-12-812163-4.00004-8>
- Paradis E., Claude J., Strimmer K.** 2004. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20: 289–290. <https://doi.org/10.1093/bioinformatics/btg412>
- Peng Z., Bredeson J. V., Wu G. A., Shu S., Rawat N., Du D., Parajuli S., Yu Q., You Q., Rokhsar D. S., Gmitter F. G., Deng Z.** 2020. A chromosome-scale reference genome of trifoliolate orange (*Poncirus trifoliata*) provides insights into disease resistance, cold tolerance and genome evolution in *Citrus*. *Plant J.* 104: 1215–1232. <https://doi.org/10.1111/tbj.14993>
- POWO [2025]. *Plants of the World Online*. Facilitated by the Royal Botanic Gardens, Kew. URL: <https://powo.science.kew.org/> (Accessed 10 March 2025).
- Puterova J., Razumova O., Martinek T., Alexandrov O., Divashuk M., Kubat Z., Hobza R., Karlov G., Kejnovsky E.** 2017. Satellite DNA and transposable elements in seabuckthorn (*Hippophae rhamnoides*), a dioecious plant with small Y and large X chromosomes. *Genom. Biol. Evol.* 9: 197–212.
- Rogers S. O., Bendich A. J.** 1985. Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant tissues. *Plant Mol. Biol.* 5: 69–76. <https://doi.org/10.1007/BF00020088>
- Rosselló J. A., Maravilla A. J., Rosato M.** 2022. The nuclear 35S rDNA world in plant systematics and evolution: A primer of cautions and common misconceptions in cytogenetic studies. *Front. Plant Sci.* 13: 788911. <https://doi.org/10.3389/fpls.2022.788911>
- Schliep K. P.** 2011. phangorn: phylogenetic analysis in R. *Bioinformatics* 27, 4: 592–593. <https://doi.org/10.1093/bioinformatics/btq706>
- Singh R., Nath N.** 1969. Practical approach to the classification of *Citrus*. In: *Proceedings of International Citrus Symposium (Riverside, March 16–26, 1968)*. Vol 1. Riverside: University of California. Pp. 435–440.
- Sun C., Lin H.** 2021. The complete chloroplast genome and phylogenetic analysis of *Citrus clementina* (Rutaceae). *Mitochondrial DNA, Part B. Resour.* 6: 2926–2927. <https://doi.org/10.1080/23802359.2021.1972860>
- Swingle W. T.** 1915. A new genus, *Fortunella*, comprising four species of kumquat oranges. *J. Wash. Acad. Sci.* 5: 165–176.

-
- Swingle W. T.** 1943. The botany of *Citrus* and its wild relatives in the orange subfamily. *The Citrus Industry* 1: 128–474.
- Swingle W. T.** 1967. The botany of *Citrus* and its wild relatives. *The Citrus Industry* 1: 190–430.
- Tanaka T.** 1954. Species problem in citrus: A critical study of wild and cultivated units of citrus based upon field studies in their native homes (Revisio Aurantiacearum IX). *JSPS* 3: 141.
- Tanaka T.** 1977. Fundamental discussion of *Citrus* classification. *Stud. Citrol.* 14: 1.
- Vitales D., Garcia S., Dodsworth S.** 2020. Reconstructing phylogenetic relationships based on repeat sequence similarities. *Mol. Phylogenetics Evol.* 147: 106766. <https://doi.org/10.1016/j.ympev.2020.106766>
- Vitte C., Fustier M.-A., Alix K., Tenailon M. I.** 2014. The bright side of transposons in crop evolution. *Brief. Funct. Genom.* 13: 276–295.
- Wang T., Chen L.-L., Shu H.-J., You F., Liang X.-L., Li J., Ren J., Wanga V. O., Mutie F. M., Cai X. Z.** 2022. *Fortunella venosa* (Champ. ex Benth.) C. C. Huang and *F. hindsii* (Champ. ex Benth.) swingle as independent species: evidence from morphology and molecular systematics and taxonomic revision of *Fortunella* (Rutaceae). *Front. Plant Sci.* 13: 867659.
- Wang X., Xu Y., Zhang S., Cao L., Huang Y., Cheng J., et al.** 2017. Genomic analyses of primitive, wild and cultivated *Citrus* provide insights into asexual reproduction. *Nat. Genet.* 49: 765–772. <https://doi.org/10.1038/ng.3839>
- Wu G. A., Terol J., Ibanez V., López-García A., Pérez-Román E., Borredá C., et al.** 2018. Genomics of the origin and evolution of *Citrus*. *Nature* 554: 311–316. <https://doi.org/10.1038/nature25447>
- Wu Y., Wang F., Lyu K., Liu R.** 2024. Comparative analysis of transposable elements in the genomes of *Citrus* and *Citrus*-related genera. *Plants* 13: 2462. <https://doi.org/10.3390/plants13172462>
- Zhang Y., Barthe G., Grosser J. W., Wang N.** 2016. Transcriptome analysis of root response to *Citrus* blight based on the newly assembled Swingle citrumelo draft genome. *BMC Genom.* 17: 485. <https://doi.org/10.1186/s12864-016-2779-y>
- Zhu C., Zheng X., Huang Yu., Ye J., Chen P., Zhang C., et al.** 2019. Genome sequencing and CRISPR /Cas9 gene editing of an early flowering Mini-Citrus (*Fortunella hindsii*). *Plant Biotechnol. J.* 17(11): 2199–2210. <https://doi.org/10.1111/pbi.13132>