

## ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ



УДК 575.174+575.113

**Изучение структурных особенностей нетранскрибируемых спейсеров 5S-рДНК *Citrus sinensis* и *C. reticulata***О.С. Александров\* , Д.В. Романов 

Всероссийский научно-исследовательский институт сельскохозяйственной биотехнологии,  
Россия, 127550, г. Москва, ул. Тимирязевская, д. 42  
\*e-mail: olegsandrov@gmail.com

Апельсин (*Citrus sinensis*) и мандарин (*C. reticulata*) являются экономически важными сельскохозяйственными растениями, которые выращиваются во многих странах мира, в том числе и на юге России. В связи с этим широко проводятся их генетические исследования, включающие как секвенирование отдельных локусов, так и полногеномное секвенирование. Одними из важнейших локусов генома являются массивы генов 5S рибосомной РНК, которые представляют собой тандемные повторы. Их мономеры состоят из консервативной кодирующей 120-нуклеотидной части и нетранскрибируемого спейсера (NTS), который часто имеет различную длину и последовательность у разных видов. В данной работе были секвенированы и изучены 8 NTS апельсина (сорт Гамлин) и 16 NTS мандарина (мандарин иволлистный – 3, сорт Карликовый Уншиу – 8, Тиахара – 4 и Уншиу – 1). Полученные NTS имели различную длину и были разделены на два класса – NTS\_218 (длиной 217–221 п.о.) и NTS\_381 (длиной 381 п.о.). Была изучена внутренняя структура NTS обоих классов и распределение характерных для NTS ДНК-мотивов. Полученные результаты расширяют теоретическое понимание природы NTS и знание о строении NTS у близкородственных видов. Они могут быть использованы при анализе гибридов и сортов цитрусовых, выведенных с участием апельсина и мандарина.

**Ключевые слова:** апельсин, мандарин, NTS, поли-T-мотивы, поли-G-мотивы, ДНК повторы

**DOI:** 10.55959/MSU0137-0952-16-79-4-3

**Введение**

Подтриба Citrinae Engl. относится к семейству Рутовые (Rutaceae Juss.) и входит в трибу Aurantieae подсемейства Поморанцевые (Aurantioideae Eaton.) [1]. Одним из самых известных родов этой подтрибы является род *Citrus* L., виды которого (прежде всего, апельсин (*C. sinensis* (L.) Osbeck), грейпфрут (*C. paradisi* Macfad.), лимон (*C. limon* (L.) Osbeck), мандарин (*C. reticulata* Blanco) и некоторые другие) – это широко распространенные садовые плодовые деревья [2]. Согласно современным генетическим исследованиям, цитрусовые появились около 8 млн лет назад в регионе, включающем северо-восточную Индию, Мьянму и юго-восточный Китай. Выделяют четыре предковых вида – помело (*C. maxima* (Burm.) Merr.), цитрон (*C. medica* L.), лайм (*C. aurantiifolia* Christm.) и мандарин. В процессе одомашнивания, спонтанных и направленных скрещиваний получались известные на сегодняшний день формы. Так, скрещивание мандарина и помело привело к появлению апельсина и некоторых подвидов мандарина, а гибридизация апельсина с помело привела к появлению грейпфрута [1].

Апельсин и мандарин возделываются как сельскохозяйственные растения во многих странах мира и имеют для некоторых из них важное экономическое значение, поскольку торговля плодами этих видов составляет значительную долю в структуре их экспорта [3–6]. В связи с этим ведутся селекционная работа и широкое генетическое изучение данных и родственных им видов цитрусовых. На данный момент получен огромный массив генетических и геномных данных представителей подтрибы Citrinae, состоящий из результатов секвенирования отдельных локусов и целых геномов [7–11].

Среди часто изучаемых локусов генома важное теоретическое и прикладное значение имеют гены, кодирующие 5S рибосомную РНК (5S-рРНК), которая у эукариот входит в состав большой 60S субъединицы рибосом [12]. У большинства эукариот, в том числе и у высших семенных растений, эти гены не входят в кластер с генами 18S, 5,8S и 26S рРНК, но присутствуют в геноме в виде отдельных от них массивов, которые представляет собой, по сути, тандемные ДНК-повторы [13–15]. Мономер 5S рибосомной ДНК (5S-рДНК) состоит из двух

частей: консервативного 120-нуклеотидного участка, транскрибируемого РНК-полимеразой III, и нетранскрибируемого спейсера (NTS) [16, 17]. Длина и нуклеотидная последовательность NTS часто являются геном- или субгеномспецифичными и могут иметь различия даже у близкородственных видов. В связи с этим полиморфизм 5S-рДНК широко используется в исследованиях, связанных с установлением филогенетических отношений между видами, анализе гибридов, а также при разработке видо- и субгеномспецифичных ДНК-маркеров [18–27]. Несмотря на вариабельность, NTS многих организмов имеют некоторые общие структурные элементы. Прежде всего, это содержащиеся в начале NTS Т-богатые мотивы и ТАТА-подобные мотивы в положении около -25–-30 п.о. с конца NTS. Также NTS ряда видов может содержать внутри поли-Т- или микросателлитные блоки [17, 28–35].

Благодаря тому, что 120-нуклеотидная транскрибируемая часть мономера 5S-рДНК является консервативной, подобранная на ее основе пара праймеров позволяет проводить амплификацию NTS у широкого спектра видов. Использование таких универсальных праймеров неоднократно встречается в ранее опубликованных работах [36–39], и можно сказать, что они показали себя в качестве эффективных инструментов для изучения 5S-рДНК даже в малоизученных, не полностью секвенированных, геномах [28].

В данной работе впервые были секвенированы и изучены NTS апельсина и мандарина (24 последовательности). Полученные результаты и выводы имеют теоретическое значение (расширяют понимание о структуре NTS, знание о строении NTS у близкородственных видов и т.д.) и могут быть использованы в практической селекционной работе при анализе гибридов и сортов цитрусовых, выведенных с участием апельсина и мандарина.

### Материалы и методы

#### *Растительный материал и выделение ДНК.*

В работе использовались пять растений апельсина (сорта Вашингтон Навел (питомник «Павловский лимон»), Гамлин (Лаборатория плодводства РГАУ-МСХА им. Тимирязева), Голден Бакай (питомник «Садовод Крым»), Навелина (питомник «Павловский лимон»), Томсон Навелин (питомник «Садовод Крым»)) и семь растений мандарина (мандарин иволистный (питомник «Павловский лимон»), Карликовый Уншиу (питомник «Садовод Крым»), Кавано-Васе (питомник «Павловский лимон»), Михо-Васе (питомник «Садовод Крым»), Сацума (питомник «Павловский лимон»), Тиэхара (питомник «Павловский лимон») и Уншиу (Лаборатория плодводства РГАУ-МСХА им. Тимирязева)). С данных растений, выращиваемых в оранжерее ФГБНУ ВНИИСБ, были собраны молодые листья и выделена ДНК по методике Дойл и Дойл (1990) с модификациями [40, 41].

#### *Амплификация, клонирование и секвенирование.*

Для амплификации NTS изучаемых растений проводилась полимеразная цепная реакция (ПЦР) с праймерами 5S1/5S2 [39] на амплификаторе C1000 Touch™ Thermal Cycler (Bio-Rad, США) при следующих условиях: 1) 94°C – 5 мин; 2) 30 циклов (94°C – 20 с, 60°C – 20 с, 72°C – 20 с); 3) 72°C – 10 мин. Ампликоны разделяли в 2,5%-ном агарозном геле при 5 В/см и детектировали с помощью системы гель-документации Gel Doc XR+ (Bio-Rad, США).

Ампликоны клонировали в вектор pAL2-T (ЗАО «Евроген», Россия) согласно инструкции производителя. Для химической трансформации использовали компетентные клетки XL1-Blue (ЗАО «Евроген», Россия) согласно инструкции их применения. Отбор целевых колоний проводили с помощью бело-голубой селекции. Анализ вставок проводили методом ПЦР с праймерами M-13 (forward 5'-GTAAAACGACGGCCAGT-3'; reverse 5'-GGAAACAGCTATGACCATG-3') при следующих условиях: 1) 94°C – 5 мин; 2) 30 циклов (94°C – 30 с, 55°C – 30 с, 72°C – 1 мин 30 с); 3) 72°C – 10 мин. Секвенирование вставок проводилось на генетическом анализаторе 3500xL Applied Biosystems (США).

#### *Обработка последовательностей и их анализ.*

Конвертация секвенированных последовательностей в FASTA формат производилась с помощью программы «ab1 viewer» ([https://www.chiplot.online/static/ChiBioTools/src/ab1\\_file.html](https://www.chiplot.online/static/ChiBioTools/src/ab1_file.html)). Последовательности NTS выделялись из вставок при выравнивании с последовательностью гена 5S-рРНК льна (X01531). Обработка последовательностей, выравнивания, поиск ДНК-мотивов и расчет уровня идентичности осуществлялись с помощью программы «GenDoc» [42].

Поиск микросателлитных мотивов осуществлялся с помощью программы «DNA Microsatellites Finder» (<https://www.novoprolabs.com/tools/microsatellites-finder>).

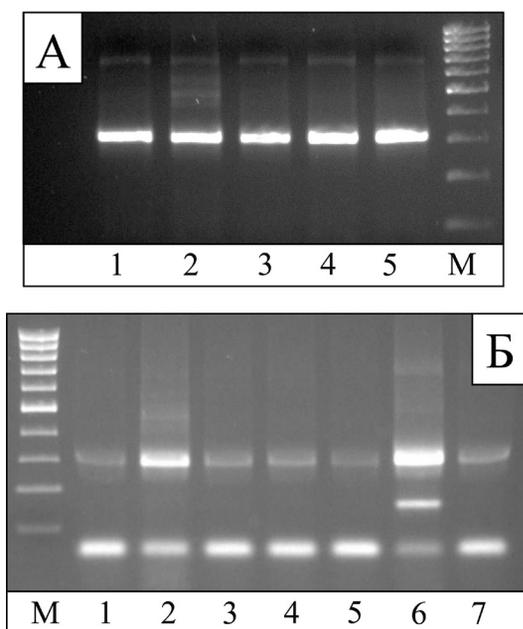
### Результаты и обсуждение

#### *Результаты (особенности) амплификации NTS апельсина и мандарина.*

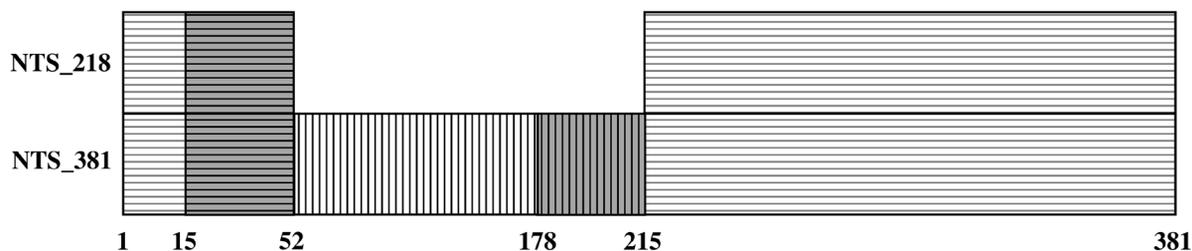
ДНК пяти растений апельсина (сорта Вашингтон Навел, Гамлин, Голден Бакай, Навелина и Томсон Навелин) и семи растений мандарина (Мандарин иволистный, сорта мандарина Карликовый Уншиу, Кавано-Васе, Михо-Васе, Сацума, Тиэхара и Уншиу) была использована в качестве матрицы в ПЦР-эксперименте с праймерами 5S1/5S2. Электрофоретическое разделение ампликонов показало во всех паттернах наличие мажорных фрагментов длиной 316–320 п.о. (рис. 1), включающих в себя NTS длиной 218–220 п.о. (система праймеров подобрана так, что вместе с NTS амплифицируются участки гена 5S-рРНК длиной 98 п.о.). В нескольких профилях также присутствовал минорный компонент длиной около 700–800 п.о., оче-

видно, включающий в себя несколько мономеров «ген 5S-рРНК + NTS», что часто наблюдается при амплификации tandemных повторов [26]. Профили апельсина сорта Гамлин и мандарина сортов Карликовый Уншиу и Тиахара включали дополнительные фрагменты длиной порядка 480 п.о., что соответствует длине содержащихся в них NTS около 380 п.о. Кроме того, профиль мандарина сорта Тиахара содержал фрагменты около 150 п.о. с NTS длиной около 50 п.о.

В целом, приблизительные длины NTS апельсина и мандарина, полученные при проведении ПЦР-экспериментов с помощью праймеров 5S1/5S2, находятся в диапазоне 100–700 п.о., как и NTS многих других организмов [17]. Короткие же NTS, сходные по длине с дополнительными NTS-фрагментами мандарина сорта Тиахара, встречаются нечасто и были описаны для тополя белого и некоторых видов диких злаков [43–45].



**Рис. 1.** Результаты амплификации NTS апельсина и мандарина с помощью праймеров 5S1/5S2 [39]: **А** – ампликоны апельсина сортов Вашингтон Навел (1), Гамлин (2), Голден Бакай (3), Навелина (4), Томсон Навелин (5), **Б** – ампликоны мандарина иволлистного (1) и сортов Карликовый Уншиу (2), Кавано-Васе (3), Михо-Васе (4), Сацума (5), Тиахара (6) и Уншиу (7). М – маркер молекулярного веса с шагом 100 п.о.



**Рис. 2.** Сравнительная схема последовательностей классов NTS\_218 и NTS\_381. Горизонтальной штриховкой обозначены участки с высоким уровнем идентичности. Вертикальной штриховкой обозначена инсерция в NTS\_381. Серым цветом обозначены идентичные 38-нуклеотидные участки перед инсерцией и на ее конце. Границы участков обозначены порядковыми номерами соответствующих нуклеотидов под схемой.

**Клонирование, секвенирование и особенности NTS апельсина и мандарина.** Ампликоны апельсина сорта Гамлин, мандарина иволлистного и мандарина сортов Карликовый Уншиу, Тиахара и Уншиу были очищены и клонированы в АТ-векторе. Было получено 24 клона, содержащих целевые вставки (апельсин сорта Гамлин – 8, мандарин иволлистный – 3, мандарин сортов Карликовый Уншиу – 8, Тиахара – 4 и Уншиу – 1). ПЦР-продукты вставок, полученные при амплификации с праймерами M13, были секвенированы. Все вставки содержали последовательности NTS, которые были проанализированы и размещены в базе данных GenBank (PQ141631–PQ141654).

Среди последовательностей NTS апельсина сорта Гамлин шесть имели длину 217–218 п.о. (L13, L16, L17, L18, L21, L23), одна – 250 п.о. (L19) и одна – 381 п.о. (L14). Анализ последовательности L14 (381 п.о.) показал высокий уровень идентичности с последовательностями 217–218 п.о. (90–98%) на участках с 1 по 52 п.о. и с 216 по 381 п.о. BLAST-анализ инсерции L14 53\_215 показал наличие на ее конце 38-нуклеотидного участка, который был практически полностью идентичен участку L14 15\_52 (имелась G/A замена в положении 30 п.о.), который граничит с данной инсерцией вверх по цепи (рис. 2). Ввиду существенных отличий по длине и структуре NTS длиной 217–218 п.о. и 381 п.о. были отнесены к двум классам: NTS\_218 и NTS\_381. Наличие двух классов NTS в одном геноме (или субгеноме) ранее было описано и у других растений – например, у элимусов, псевдорогнерий и ячменей [45–47]. Данный факт может быть косвенным свидетельством гибридного происхождения таких видов.

Анализ последовательности L19 (250 п.о.) показал наличие 31-нуклеотидной дупликации в начале этого NTS. Выравнивание участка L19 32\_250 данного NTS показало высокий уровень идентичности (88–95%) с остальными NTS\_218 апельсина сорта Гамлин. Возникновение дупликаций в регионе 5S-рДНК ранее отмечалось и у других видов. В частности, Сколз и др. [48], анализируя последовательности 5S-рДНК некоторых злаков, описывают наличие дупликаций как в кодирующей части мономера, так и в нетранскрибируемом спейсере.

Пытаясь объяснить, каким образом появляются наблюдаемые ими дупликации и делеции, авторы выдвигают гипотезу, которая основывается на ошибках репликации в данной области из-за наличия близко расположенных друг к другу идентичных коротких мотивов (TTGGG и ATGGG). В дублированном начале последовательности L19 таких мотивов найдено не было, но были обнаружены шесть мотивов СТТТТ в положениях 4\_8, 11\_15, 36\_40, 43\_47, 63\_67 и 85\_89 п.о., которые вероятно могли бы стать причиной запуска механизма, идентичного описанному Сколзом и др. [48].

Среди секвенированных NTS мандарина большинство относилось к классу NTS\_218 с вариацией длины в пределах 217–222 п.о. (мандарин иволлистный – три последовательности 217 п.о. (L57, L59, L60); сорт Карликовый Уншиу – пять последовательностей 218 п.о., одна 220 п.о. (L61\_1, L61\_2, L62, L64, L67, L69); сорт Тиахара – две последовательности 221 п.о. (L77, L83) и две последовательности 222 п.о. (L78, L84); сорт Уншиу – одна последовательность 217 п.о. (L90)). Клон L61 содержал вставку длиной 808 п.о., включающую в себя, кроме частей гена 5S-rРНК по краям, два целых гена и два NTS. Такой результат стал возможным благодаря тому, что массивы 5S-рДНК представляют собой тандемные повторы и с помощью универсальных праймеров на их матрице могут амплифицироваться фрагменты, содержащие два мономера и более (схема такой амплификации представлена, например, в работе Александрова и др. (2018) [26]). Особенностью NTS сорта Тиахара являлся мотив GAAT в положении 28\_31 п.о. (приложение, Б). Выравнивание всех полученных NTS мандарина показало, что у остальных NTS вместо данного мотива присутствует мотив СС. Выявленный полиморфизм может послужить основанием для разработки специфичных молекулярных маркеров, позволяющих идентифицировать сорт мандарина Тиахара.

Последовательности L68 и L70 сорта Карликовый Уншиу относились к классу NTS\_381. Выравнивание данных последовательностей с L14 апельсина сорта Гамлин показало, что 38-нуклеотидные участки 15\_52 и 178\_215 L68 полностью идентичны таковым у L14, а участок 178\_215 L70 не имеет описанной выше замены G/A и полностью идентичен участку 15\_52 L14. Однако участок 15\_52 L70 имел три замены (две G/A на 8 и 23 п.о. данного участка и одна C/T на 33 п.о. данного участка).

**Поиск характерных мотивов в NTS апельсина и мандарина.** В ранее опубликованных работах было отмечено, что начальные участки NTS 5S-рДНК часто содержат поли-Т-мотивы [28, 44–47]. Такие мотивы были найдены в 16-17-нуклеотидных начальных участках обоих классов изучаемых NTS апельсина и мандарина. Среди последовательностей класса NTS\_218 наиболее распространенным

был вариант C<sub>1</sub>T<sub>1</sub>C<sub>2</sub>T<sub>4</sub>A<sub>1</sub>T<sub>1</sub>C<sub>1</sub>T<sub>4</sub> (обнаружен у 17-ти NTS\_218, частота встречаемости 80,96%). В единичных NTS этого класса также встречались варианты T<sub>3</sub>C<sub>1</sub>T<sub>4</sub>A<sub>1</sub>T<sub>1</sub>C<sub>1</sub>T<sub>4</sub>, C<sub>1</sub>T<sub>1</sub>G<sub>1</sub>C<sub>1</sub>T<sub>4</sub>A<sub>1</sub>T<sub>1</sub>C<sub>1</sub>T<sub>4</sub>, C<sub>1</sub>T<sub>2</sub>C<sub>1</sub>T<sub>5</sub>A<sub>1</sub>T<sub>1</sub>C<sub>1</sub>T<sub>4</sub> и C<sub>1</sub>T<sub>1</sub>C<sub>1</sub>T<sub>5</sub>A<sub>1</sub>T<sub>1</sub>C<sub>1</sub>T<sub>4</sub> (частота встречаемости каждого из них составила 4,76%). Последний вариант также встречался во всех трех секвенированных последовательностях класса NTS\_381.

Анализ секвенированных NTS апельсина и мандарина также показал, что поли-Т-мотивы встречаются не только в начальных 16-17-нуклеотидных участках (приложение). Все NTS\_381 содержали T<sub>6</sub>-мотив в положении 102\_107 п.о.; T<sub>5</sub>-мотивы в положениях 54\_58, 124\_128 и 145\_149 п.о.; T<sub>4</sub>-мотивы в положениях 32\_35, 135\_138, 195\_198, 217\_220 п.о. Кроме того, дополнительные поли-Т-мотивы встречались в отдельных последовательностях NTS\_381: T<sub>7</sub> в положении 160\_166 п.о. у L14 и L68; T<sub>5</sub> в положении 54\_58 п.о. у L14 и L70; T<sub>4</sub> в положениях 117\_120, 130\_133 и 160\_163 п.о. у L70. Последовательности класса NTS\_218 не содержали T<sub>7</sub>-, T<sub>6</sub>- и T<sub>5</sub>-мотивы, но имели T<sub>4</sub>-мотивы в положении 32\_35 п.о. (33\_36 п.о. у L61\_2; 34\_37 п.о. у L77, L78, L83 и L84) и 54\_57 п.о. (у всех, кроме L13; в положении 55\_58 п.о. у L61\_2; 56\_59 п.о. у L77 и L83; 57\_60 п.о. у L78 и L84). Дополнительные T<sub>4</sub>-мотивы встречались в положении 51\_54 п.о. у L78 и L84, а также 73\_76 п.о. у L59 (74\_77 п.о. у L61\_2). Наличие поли-Т-мотивов не только в начале NTS, но и в иных участках ранее было описано также и для NTS разных видов семейства Elaeagnaceae L. [28]. Функции данных мотивов не вполне ясны, но можно сделать предположение, что в случае мутационных изменений внутри участка «конец гена 5S рРНК/начало NTS» наличие поли-Т-мотивов ниже по цепи может обеспечить синтез иных РНК-продуктов.

Возможно также влияние поли-Т-мотивов (точнее, соответствующих им полиадениновых блоков в антипараллельной цепи) на физический изгиб молекулы ДНК NTS, что может быть важно для формирования конститутивного гетерохроматина. Ранее у различных видов эукариот было выявлено наличие изгибов ДНК сателлитных повторов конститутивного гетерохроматина [49–52]. Определяющая роль полиадениновых блоков в появлении изгибов ДНК, в свою очередь, была подтверждена в экспериментах по сайт-направленному мутагенезу у табака [53] и сравнительными исследованиями на таких видах, как *C. limon*, *Poncirus trifoliata* (L.) Raf., *Ricinus communis* L. и др. [54, 55].

Помимо поли-Т-мотивов в NTS апельсина и мандарина были обнаружены поли-Г-мотивы (приложение). У всех секвенированных NTS\_381 было выявлено несколько таких мотивов: G<sub>6</sub>-мотив в положении 341\_346 п.о.; G<sub>5</sub>-мотив в положении 95\_99 п.о. и четыре G<sub>4</sub>-мотива в положениях

252\_255 п.о., 261\_264 п.о. (отсутствует у L70), 280\_283 п.о. и 287\_290 п.о. Большинство же последовательностей класса NTS\_218 содержало G<sub>6</sub>-мотив в положении 178\_183 п.о. и несколько G<sub>4</sub>-мотивов, которые встречались не во всех последовательностях сразу. У некоторых NTS\_218 G<sub>6</sub>-мотиву соответствовали мотивы с большим или меньшим количеством G (G<sub>7</sub>-мотив в положении 176\_182 п.о. у L16; G<sub>5</sub>-мотив в положении 178\_182 п.о. у L59; G<sub>4</sub>-мотив в положении 178\_181 п.о. у L61\_1). У NTS мандарина сорта Тиахара G<sub>6</sub>-мотив был смещен ниже по цепи и находился в положении 181\_186 п.о. у L77 и L83 или в положении 182\_187 п.о. у L78 и L84. О роли поли-G-мотивов в NTS трудно утверждать однозначно, но можно предположить их вклад в обеспечение сохранности такой жизненно важной структуры, как 5S-рДНК. В нескольких работах встречается мнение, что длинные последовательности из гуаниновых нуклеотидов являются сенсорами повреждения ДНК, так как ряд агентов (прежде всего окислителей) при повреждении воздействуют именно на поли-G-последовательности [56, 57]. Кроме того, окисление или иная модификация G-мотивов в промоторной области могут влиять на регуляцию экспрессии, являясь частью эпигенетических механизмов [58–62]. Конечно, данные выводы актуальны в большей степени для случаев, когда речь идет о работе РНК-полимеразы II, регуляторные элементы которой находятся выше по цепи относительно транскрибируемого гена. Транскрипция же 5S-рРНК осуществляется РНК-полимеразой III, регуляторные последовательности которой находятся внутри самого гена 5S-рРНК.

Помимо поли-G-мотива в NTS 5S-рДНК часто имеются и другие регуляторные последовательности. Например, в положении -25–-30 п.о. с конца NTS встречаются ТАТА-подобные мотивы (ТАТА-боксы необходим для регуляции работы РНК-полимеразы II) [17, 28, 43]. В изучаемых в данной работе 24 NTS канонической последовательности ТАТА-боксов (ТАТА<sup>A</sup>/<sub>T</sub>А<sup>A</sup>/<sub>T</sub>) найдено не было. Однако в зоне -25–-30 п.о. с конца NTS ТАТА-подобные мотивы были обнаружены: АТТАА – во всех четырех NTS мандарина сорта Тиахара; GTТАА – в остальных последовательностях класса NTS\_218 и последовательностях L14 и L68 класса NTS\_381; СТТАА – последовательности L70 класса NTS\_381. Факт наличия таких мотивов позволяет выдвинуть гипотезу о том, что ранее, до превращения ТАТА-боксов в ТАТА-подобные мотивы, последовательности 5S-рРНК могли синтезироваться и РНК-полимеразой II. Факт же наличия серьезных мутационных изменений, приведших к образованию ТАТА-подобных мотивов, подтверждает независимость РНК-полимеразы III от наличия или сохранности ТАТА-боксов, в то время как даже одонуклеотидные замены в ТАТА-боксе могут при-

вести к существенному снижению эффективности работы РНК-полимеразы II [63].

У некоторых видов (тополь дельтовидный, облепиха, чечевица, несколько видов рыб) имеется такая особенность NTS, как наличие внутри их последовательностей микросателлитных мотивов [29–35, 64]. Поиск микросателлитов в NTS апельсина и мандарина показал отсутствие в них зрелых микросателлитных субповторов. Почти во всех случаях отмечались лишь двукратные повторения микросателлитных мотивов. Только у последовательности L61\_1 в положении 57\_62 п.о. было отмечено троекратное повторение микросателлитного мотива TG. Данный мотив также был единственным динуклеотидным из 83 найденных (частота встречаемости 1,21%). Остальные были три- (CCG, GCG, ATC, GGT, TCG, TCC, CTT, TTG, GGA, GCC, ATT, ACT), тетра- (TGCG, GGAT, CGGC, CGAC, GGGA) или гексануклеотидными (CGCGGG), с частотами встречаемости 67,47%, 27,71% и 3,61% соответственно (пентануклеотидных мотивов найдено не было). Вполне возможно, что найденные «зарождающиеся» микросателлиты, находящиеся в самом начале своего эволюционного пути, являются неким резервом изменчивости NTS и могут впоследствии расширяться за счет увеличения числа копий мотивов вследствие различных мутационных процессов (проскальзывание фермента при репликации, ошибки при кроссинговере и др.) [65, 66].

### Заключение

В данной работе были секвенированы и изучены 24 последовательности NTS 5S-рДНК одного сорта апельсина и четырех сортов мандарина. Последовательности были разделены на два класса (NTS\_318 и NTS\_218), так как между ними имелись существенные отличия как по длине, так и по внутренней структуре. Поиск различных ДНК-мотивов (поли-T-, поли-G-, ТАТА-подобные и микросателлитные мотивы) выявил некоторые особенности изучаемых NTS. Полученные результаты являются очередным шагом на пути расширения теоретического понимания природы NTS в целом, а также знания о строении NTS у близкородственных видов и могут найти практическое применение при анализе гибридов и сортов цитрусовых, включающих генетический материал апельсина и мандарина.

Работа выполнена при финансовой поддержке Российского научного фонда (проект № 23-16-00234 «Молекулярно-цитогенетическое изучение рода Цитрус (*Citrus*) для использования в селекции»). Работы проводили без использования животных и без привлечения людей в качестве испытуемых. Авторы заявляют об отсутствии конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. Wu G.A., Terol J., Ibanez V., et al. Genomics of the origin and evolution of *Citrus*. *Nature*. 2018;554(7692):311–316.
2. Жбели А., Сорокопудов В.Н. Перспективы выращивания цитрусовых культур в Сирии. *Вест. Ланд. Арх.* 2021;28:18–21.
3. Budiarto R., Pratita D. *Citrus* export performances of Southeast Asian Countries: A comparative analysis. *Teknotan*. 2022;16:7–12.
4. Каруя Т., Chinembiri E.K., Kalaba M.W. Identifying strategic markets for South Africa's *Citrus* exports. *Agrekon*. 2014;53(1):124–158.
5. Márquez O., Borda R.C., Márquez F., Bravo L.R. Production and export of peruvian mandarin to the Russian Federation market, as a business opportunity, period 2013–2017. *J. Glob. Manag. Sci.* 2018;1(1):22–29.
6. Lachman J., Tacsir E., Pereyra M. Public–private cooperation and the provision of public goods for lemons and citrus exports from Argentina and Uruguay. *J. Agribus. Dev. Emerg. Econ.* 2022;12(4):604–619.
7. *The Citrus Genome*. Eds. A. Gentile, S. La Malfa, Z. Deng. Cham.: Springer Nature Switzerland AG; 2020. 294 pp.
8. Gmitter F.G., Chen C., Machado M.A. de Souza A.A., Ollitrault P., Froehlicher Y., Shimizu T. Citrus genomics. *Tree Genet. Genomes*. 2021;8(3):611–626.
9. Su H.J., Hogenhout S.A., Al-Sadi A.M., Kuo C.H. Complete chloroplast genome sequence of Omani Lime (*Citrus aurantiifolia*) and comparative analysis within the Rosids. *PLoS One*. 2014;9(11):e113049.
10. Xu Q., Chen L.L., Ruan X. et al. The draft genome of sweet orange (*Citrus sinensis*). *Nat. Genet.* 2013;45:59–66.
11. Sun L., Nasrullah, Ke F., Nie Z., Wang P., Xu J. Citrus genetic engineering for disease resistance: past, present and future. *Int. J. Mol. Sci.* 2019;20(21):5256.
12. Moraleva A.A., Deryabin A.S., Rubtsov Yu.P., Rubtsova M.P., Dontsova O.A. Eukaryotic Ribosome Biogenesis: The 60S Subunit. *Acta Naturae*. 2022;14(2):39–49.
13. Барков А.Н., Трубникова Е.В., Стабровская Н.В. Молекулярные особенности организации и транскрипции рибосомных генов. *Уч. зап. Эл. Науч. Журн. КГУ*. 2007;1:54–73.
14. Little R.D., Braaten D.C. Genomic organization of human 5 S rDNA and sequence of one tandem repeat. *Genomics*. 1989;4(3):376–383.
15. Petrová I., Fajkus J. The rDNA loci – intersections of replication, transcription, and repair pathways. *Int. J. Mol. Sci.* 2021;22(3):1302.
16. Wicke S., Costa A., Muñoz J., Quandt D. Restless 5S: the re-arrangement(s) and evolution of the nuclear ribosomal DNA in land plants. *Mol. Phylogenet. Evol.* 2011;61(2):321–332.
17. Sastri D.C., Hilu K., Appels R., Lagudah E.S., Playford J., Baum B.R. An overview of evolution in plant 5S DNA. *Pl. Syst. Evol.* 1992;183(3–4):169–181.
18. Yang Y.W., Tseng P.F., Tai P.Y., Chang Ch.J. Phylogenetic position of *Raphanus* in relation to *Brassica* species based on 5S rRNA spacer sequence data. *Bot. Bull. Acad. Sin.* 1998;39(3):153–160.
19. Brown G.R., Carlson J.E. Molecular cytogenetics of the genes encoding 18s-5.8s-26s rRNA and 5s rRNA in two species of spruce (*Picea*). *Theoret. Appl. Genet.* 1997;95(1–2):1–9.
20. Liu Z.L., Zhang D., Wang X.Q., Ma X.F., Wang X.R. Intragenomic and interspecific 5S rDNA sequence variation in five Asian pines. *Am. J. Bot.* 2003;90(1):17–24.
21. Matyáček R., Fulneček J., Lim K.Y., Leitch A.R., Kovarik A. Evolution of 5S rDNA unit arrays in the plant genus *Nicotiana* (Solanaceae). *Genome*. 2002;45(3):556–562.
22. Baum B.R., Edwards T., Johnson D.A. Codependence of repetitive sequence classes in genomes: phylogenetic analysis of 5S rDNA families in *Hordeum* (Triticeae: Poaceae). *Genome*. 2010;53(3):180–202.
23. Alexandrov O.S., Karlov G.I. Development of 5S rDNA-based molecular markers for the identification of *Populus deltoides* Bartr. ex-Marshall, *Populus nigra* L., and their hybrids. *Forests*. 2018;9(10):604.
24. Alexandrov O.S., Karlov G.I. The development of *Populus alba* L. and *Populus tremula* L. species specific molecular markers based on 5S rDNA nontranscribed spacer polymorphism. *Forests*. 2019;10(12):1092.
25. Alexandrov O.S., Karlov G.I. The development of new species-specific molecular markers based on 5S rDNA in *Elaeagnus* L. species. *Plants*. 2021;10(12):2713.
26. Alexandrov O.S., Divashuk M.G., Karlov G.I. Development of the St/J/Vgenome specific molecular marker on basis of 5S-rDNA polymorphism. *Moscow Univ. Biol. Sci. Bull.* 2018;73(1):18–23.
27. Alexandrov O.S., Kroupin P.Yu., Karlov G.I., Divashuk M.G. The improvement of the CAPS-marker for St, J and V subgenome identification in Triticeae tribe plants using the 5S non-transcribed spacer polymorphism. *Res. Crops*. 2024;25(1):12–19.
28. Alexandrov O.S., Razumova O.V., Karlov G.I. A comparative study of 5S rDNA non-transcribed spacers in Elaeagnaceae species. *Plants*. 2021;10(1):4.
29. Negi M.S., Rajagopal J., Chauhan N., Cronn R., Lakshmikumaran M. Length and sequence heterogeneity in 5S rDNA of *Populus deltoides*. *Genome*. 2002;45(6):1181–1188.
30. Fernández M., Ruiz M.L., Linares C., Fominaya A., Pérez de la Vega M. 5S rDNA genome regions of *Lens* species. *Genome*. 2005;48(5):937–942.
31. Cruz V.P., Oliveira C., Foresti F. An intriguing model for 5S rDNA sequences dispersion in the genome of fresh water stingray *Potamotrygon motoro* (Chondrichthyes: Potamotrygonidae). *Mol. Biol.* 2015;49(3):466–469.
32. Campo D., Machado-Schiaffino G., Horreo J.L., Garcia-Vazquez E. Molecular organization and evolution of 5S rDNA in the genus *Merluccius* and their phylogenetic implications. *J. Mol. Evol.* 2009;68(3):208–216.
33. Deiana A.M., Cau A., Salvadori S., Coluccia E., Cannas R., Milia A. Tagliavini J. Major and 5S ribosomal sequences of the largemouth bass *Micropterus salmoides* (Perciformes, Centrarchidae) are localized in GC-rich regions of the genome. *Chromosome Res.* 2000;8(3):213–218.
34. Moran P., Garcia-Vazquez E. Identification of highly prized commercial fish using a PCR-based methodology. *Biochem. Mol. Biol. Educ.* 2006;34(2):121–124.

35. Gornung E., De Innocentilis S., Annesi F., Sola L. Zebrafish 5S rRNA genes map to the long arms of chromosome 3. *Chromosome Res.* 2000;8(4):362.
36. Williamson R., Brownlee G.G. The sequence of 5S ribosomal RNA from two mouse cell lines. *FEBS Lett.* 1969;3(5):306–308.
37. Bhatia S., Singh K., Jagannathan V., Lakshmikumar, M. Organization and sequence analysis of the 5S rRNA genes in *Brassica campestris*. *Plant Sci.* 1993;92(1):47–55.
38. Fernández-Tajes, J., Méndez, J. Identification of the razor clam species *Ensis arcuatus*, *E. siliqua*, *E. directus*, *E. macha*, and *Solen marginatus* using PCR-RFLP analysis of the 5S rDNA region. *J. Agr. Food Chem.* 2007;55(18):7278–7282.
39. Falistocco E., Passeri V., Marconi G. Investigations of 5S rDNA of *Vitis vinifera* L.: sequence analysis and physical mapping. *Genome.* 2007;50(10):927–938.
40. Doyle J.J., Doyle J.L. Isolation of plant DNA from fresh tissue. *Focus.* 1990;12(1):13–15.
41. Razumova O.V., Alexandrov O.S., Divashuk M.G., Sukhorada T.I., Karlov G.I. Molecular cytogenetic analysis of monoecious hemp (*Cannabis sativa* L.) cultivars reveals its karyotype variations and sex chromosomes constitution. *Protoplasma.* 2016;253(3):895–901.
42. GeneDoc: Analysis and visualization of genetic variation. [Электронный ресурс] URL: <http://www.nrbsc.org/gfx/genedoc/ebinet.htm> (дата обращения: 09.08.2024).
43. Wilson N. *Genome analysis of Populus species: assessment of genetic diversity of P. deltoides, characterization of wide hybrids and phylogenetic analysis using molecular markers*. New Delhi: Teri University, 2013. 177 pp.
44. Baum B.R., Bailey L.G. The 5S rDNA units in *Kengyilia* (Poaceae: Triticeae): Diversity of the nontranscribed spacer and genomic relationships. *Can. J. Bot.* 2000;78(12):1571–1579.
45. Baum B.R., Edwards T., Johnson D.A. Diversity within the genus *Elymus* (Poaceae: Triticeae) as investigated by the analysis of the nr5S rDNA variation in species with St and H haplotypes. *Mol. Genet. Genomics.* 2015;290(1):329–342.
46. Baum B.R., Johnson D.A. The molecular diversity of the 5S rRNA gene in barley (*Hordeum vulgare*). *Genome.* 1994;37(6):992–998.
47. Baum B.R., Edwards T., Johnson D.A. Loss of 5S rDNA units in the evolution of *Agropyron*, *Pseudoroegneria*, and *Douglasdeweya*. *Genome.* 2008;51(8):589–598.
48. Scoles G.J., Gill B.S., Xin Z.-Y., Clarke B.C., McIntyre C.L., Chapman C., Appels R. Frequent duplication and deletion events in the 5S RNA genes and the associated spacer regions of the Triticeae. *Pl. Syst. Evol.* 1988;160(1–2):105–122.
49. Beridze T.G., Tsirekidze N.I., Roitberg M.A. On the tertiary structure of satellite DNA. *Biochimie.* 1992;74(2):187–194.
50. Murphy M.R., Fowlkes D.M., Fitzgerald-Hayes M. Analysis of centromere function in *Saccharomyces cerevisiae* using synthetic centromere mutants. *Chromosoma.* 1991;101(3):189–197.
51. Martinez-Balbas A., Rodrigues-Campo A., Garcia-Ramirez, M., Sainz J., Carrera P., Aymami J., Azorin F. Satellite DNAs contain sequences that induced curvature. *Biochemistry.* 1990;29(9):2342–2348.
52. Matyasek R., Gazdova B., Fajkus J., Bezděk M. NTRS, a new family of highly repetitive DNAs specific for the T1 chromosome of tobacco. *Chromosoma.* 1997;106(6):369–379.
53. Kralovics R., Fajkus J., Kovařík A., Bezděk, M. DNA curvature of the tobacco GRS repetitive sequence family and its relation to nucleosome positioning. *J. Biomol. Struct. Dyn.* 1995;12(5):1103–1119.
54. Fann J.-Y., Kovarik A., Hemleben V., Tsirekidze N.I., Beridze T.G. Molecular and structural evolution of *Citrus* satellite DNA. *Theor. Appl. Genetics.* 2001;103():1068–1073.
55. Alexandrov O.S., Karlov G.I. Molecular cytogenetic analysis and genomic organization of major DNA repeats in castor bean (*Ricinus communis* L.). *Mol. Genet. Genomics.* 2016;291(2):775–787.
56. Murray V. The frequency of poly(G) tracts in the human genome and their use as a sensor of DNA damage. *Comput. Biol. Chem.* 2015;54:13–17.
57. Spiegel J., Adhikari S., Balasubramanian S. The structure and function of DNA G-quadruplexes. *Trends Chem.* 2020;2(2):123–136.
58. Fleming, A.M., Zhou J., Wallace S.S., Burrows C.J. A role for the fifth G-track in G-quadruplex forming oncogene promoter sequences during oxidative stress: do these “spare tires” have an evolved function? *ACS Cent. Sci.* 2015;1(5):226–233.
59. Fleming, A.M., Ding Y., Burrows C.J. Oxidative DNA damage is epigenetic by regulating gene transcription via base excision repair. *Proc. Natl. Acad. Sci. U.S.A.* 2017;114(10):2604–2609.
60. Cogo S., Ferino A., Miglietta G., Pedersen E.B., Xodo L.E. The regulatory G4 motif of the Kirsten ras (KRAS) gene is sensitive to guanine oxidation: implications on transcription. *Nucleic Acids Res.* 2018;46(2):661–676.
61. Bielskutė S., Plavec J., Podbevšek P. Impact of oxidative lesions on the human telomeric G-quadruplex. *J. Am. Chem. Soc.* 2019;141(6):2594–2603.
62. Lee H.T., Bose A., Lee C.Y., Opresko P.L., Myong S. Molecular mechanisms by which oxidative DNA damage promotes telomerase activity. *Nucleic Acids Res.* 2017;45(20):11752–11765.
63. Bjornsdottir G., Myers L.C. Minimal components of the RNA polymerase II transcription apparatus determine the consensus TATA box. *Nucleic Acids Res.* 2008;36(9):2906–2916.
64. Alexandrov O.S., Evtukhov A.V., Kiselev I.I., Karlov G.I. Molecular genetic features of 5S rDNA nontranscribed spacer in *Hippophae rhamnoides* L. *Moscow Univ. Biol. Sci. Bull.* 2016;71(4):57–60.
65. Животовский Л. А. Микросателлитная изменчивость в популяциях человека и методы ее изучения. *Инф. Вест. ВОГУС.* 2006;10(1):74–96.
66. Anmarkrud J.A., Kleven, O., Bachmann, L., Lifjeld J.T. Microsatellite evolution: Mutations, sequence variation, and homoplasy in the hypervariable avian microsatellite locus HrU10. *BMC Evol. Biol.* 2008;8:138.

Поступила в редакцию 12.08.2024

После доработки 14.10.2024

Принята в печать 27.11.2024

## RESEARCH ARTICLE

## Study of structural features of 5S rDNA non-transcribed spacers of *Citrus sinensis* and *C. reticulata*

O.S. Alexandrov\* , D.V. Romanov 

All-Russia Research Institute of Agricultural Biotechnology, Timiryazevskaya 42, 127550 Moscow, Russia  
\*e-mail: olegsandrov@gmail.com

Orange (*Citrus sinensis*) and mandarin (*C. reticulata*) are economically important agricultural plants grown in many countries of the world, including southern Russia. In this regard, their genetic studies are widely carried out, including both sequencing of individual loci and whole genome sequencing. Some of the most important loci in a genome are the 5S ribosomal RNA gene arrays, which are tandem repeats. Their monomers consist of a conservative 120-nucleotide coding part and a non-transcribed spacer (NTS), which often has different length and sequence in different species. In this work, 8 NTSs of orange (Hamlin variety) and 16 NTSs of mandarin (3 willow-leaved mandarin, as well as 8 Dwarf Unshiu, 4 Tiahara, and 1 Unshiu varieties) were sequenced and studied. The obtained NTSs had different lengths and were divided into two classes – NTS\_218 (217–221 bp long) and NTS\_381 (381 bp long). The internal structure of NTSs of both classes was studied, including such characteristics as the sequence of the 16–17 bp start region involved in transcription termination and polyadenylation, the presence of poly-T and poly-G motifs, TATA-like motifs at position –25––30 bp from the end of NTS, the presence of microsatellite motifs, etc. The obtained results expand the theoretical understanding of the NTS nature, knowledge of the NTS structure in closely related species and can be used in the analysis of *Citrus* hybrids and varieties bred with the participation of orange and mandarin.

**Keywords:** orange, mandarin, NTS, poly-T motifs, poly-G motifs, DNA repeats

**Funding:** The research was funded by Russian Science Foundation, project number 23-16-00234.

### Сведения об авторах

Александров Олег Сергеевич – канд. биол. наук, ст. науч. сотр. Лаборатории клеточной инженерии ФГБНУ ВНИИСБ. Тел.: 8-499-976-65-44; e-mail: olegsandrov@gmail.com; ORCID: <https://orcid.org/0000-0002-7146-4094>

Романов Дмитрий Викторович – канд. биол. наук, ст. науч. сотр. Лаборатории прикладной геномики и частной селекции сельскохозяйственных растений ФГБНУ ВНИИСБ. Тел.: 8-499-976-65-44; e-mail: akabos1987@gmail.com; ORCID: <https://orcid.org/0000-0002-7495-7572>