

На правах рукописи



**ХАФИЗОВА**

**Галина Васильевна**

**«Особенности организации клеточной Т-ДНК  
у представителей рода *Nicotiana* L.»**

Специальность: 03.02.07 – Генетика

**АВТОРЕФЕРАТ**

**диссертации на соискание ученой степени  
кандидата биологических наук**

Санкт-Петербург

2022

Диссертационная работа выполнена на кафедре генетики и биотехнологии биологического факультета Федерального государственного бюджетного образовательного учреждения высшего образования «Санкт-Петербургский государственный университет»

**Научный руководитель:** **Матвеева Татьяна Валерьевна**  
доктор биологических наук,  
профессор кафедры генетики и биотехнологии  
Санкт-Петербургского государственного  
университета (СПбГУ), г. Санкт-Петербург

**Официальные  
оппоненты:** **Проворов Николай Александрович**  
доктор биологических наук,  
профессор, заместитель директора по научной  
работе Всероссийского научно-  
исследовательского института  
сельскохозяйственной микробиологии РАСХН,  
г. Санкт-Петербург

**Баранова Ольга Александровна**  
кандидат биологических наук,  
ведущий научный сотрудник лаборатории  
иммунитета растений к болезням  
Всероссийского научно-исследовательского  
института защиты растений, г. Санкт-Петербург

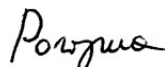
**Ведущая организация:** Федеральное государственное бюджетное  
учреждение науки «Институт общей генетики  
имени Н. И. Вавилова» Российской академии  
наук, г. Москва

Защита диссертации состоится «    »    2022 года в    часов на заседании диссертационного совета Д 006.041.02. при Всероссийском институте генетических ресурсов растений им. Н. И. Вавилова по адресу: 190031, г. Санкт-Петербург, ул. Большая Морская, д. 44.

С диссертацией можно ознакомиться в библиотеке Всероссийского института генетических ресурсов растений им. Н.И. Вавилова и на сайте института: <https://www.vir.nw.ru/dissertatsionnyj- sovet/>

Автореферат разослан «    »    2022 г.

Учёный секретарь  
диссертационного  
совета,  
доктор биологических наук



Елена Вячеславовна Рогозина

## ОБЩАЯ ХАРАКТЕРИСТИКА ДИССЕРТАЦИИ

**Актуальность темы.** Агробактериальная трансформация на сегодняшний день является самым распространенным методом для получения трансгенных растений в условиях лаборатории. При этом всё больший интерес приобретает изучение растений, геномы которых содержат последовательности агробактериального происхождения, некогда приобретенные в результате горизонтального переноса без участия человека. Такие растения называют природно-трансгенными, а последовательности, полученные от агробактерий - клеточной Т-ДНК (клТ-ДНК). История изучения природно-трансгенных растений началась с вида *Nicotiana glauca* Graham (White et al., 1983). Сегодня род *Nicotiana* насчитывает 16 видов, геномы которых содержат клТ-ДНК (Intrieri M.C., Buiatti M. 2001; Chen et al., 2014; Chen et al., 2018; Long et al., 2016). Для некоторых видов известен состав и количество клТ-ДНК в геноме (Chen et al., 2014; Chen et al., 2018). Изучение и сравнительный анализ клТ-ДНК в различных видах *Nicotiana* позволяет восстановить сценарий появления и распространения Т-ДНК в представителях *Nicotiana*, что, в свою очередь, помогает прояснить эволюционные процессы, происходившие в пределах рода *Nicotiana*. Так показано присутствие в геноме *N. tomentosiformis* L. четыре разных по составу клТ-ДНК, три из которых также найдены в геноме *N. tabacum* L., что одновременно служит доводом в пользу гипотезы о происхождении *N. tabacum* в результате гибридизации *N. tomentosiformis* и *N. sylvestris* Sp. & Comes, и свидетельствует о возможности передачи клТ-ДНК в ходе видообразования. Также известно, что некоторые гены в составе клТ-ДНК сохранили свою активность, например *rolC*, *orf13*, гены синтеза опинов (Furner et al., 1986; Mohajjel-Shoja et al., 2011; Chen et al., 2014; Chen et al., 2018), исходя из чего можно предположить, что, закрепляясь в геноме, Т-ДНК могла придавать растениям селективное преимущество. Таким образом, изучение клТ-ДНК у представителей рода *Nicotiana* позволяет установить эволюционные события, происходившие в истории рода, а также прояснить возможную роль Т-ДНК в растениях. Кроме того, анализ природно-трансгенных растений, существующих тысячи лет, будет полезен в построении моделей «поведения» трансгенов в геномах растений на протяженных временных интервалах, что поможет оценить отсроченные риски возделывания генно-модифицированных (ГМ) культур. В настоящий момент основные опасения, связанные с выращиванием ГМ растений, касаются возможной утечки трансгенов в природные популяции за счет межвидовой гибридизации с близкородственными видами. Межвидовая гибридизация является одним из основных способов видообразования в роду *Nicotiana* (Leitch et al., 2008), таким образом, природно-трансгенные представители рода *Nicotiana* могут послужить в качестве модели для изучения рисков передачи трансгенов и их закрепления в геномах родственных форм.

**Современное состояние исследований.** В настоящее время клТ-ДНК найдена у 16 представителей рода *Nicotiana* (Intrieri, Buiatti, 2001; Chen et al., 2014; Long et al., 2016; Chen et al., 2018). Для трех видов *Nicotiana*, для которых есть полногеномные данные, проведен детальный анализ клТ-ДНК, включающий информацию по количеству копий в геноме, составу

кЛТ-ДНК и активности входящих в нее генов. Род *Nicotiana* включает три подрода: *Tabacum*, *Rustica* и *Petunioides*. Все изученные подобным образом виды относятся к подроду *Tabacum*. Для всех трех – *N. tomentosiformis* Goodsp., *N. tabacum* L., *N. otophora* Griseb. – показано наличие нескольких различных по составу кЛТ-ДНК, что указывает на множественные акты трансформации в истории данных видов (Chen et al., 2014; Chen et al., 2018). Подробно изучены кЛТ-ДНК (ТА, ТВ, ТD) у трех сортов *N. tabacum* – TN90, K326 и Vasma/Xanthi, выявлены внутривидовые различия в структуре кЛТ-ДНК (Chen et al., 2014). По результатам анализа данных о составе и копияности кЛТ-ДНК сформулированы теории о том, как Т-ДНК распространялась по эволюционной ветви *Tabacum*. Для вида *N. glauca*, относящегося к ветви *Petunioides*, известен состав кЛТ-ДНК и проведена оценка активности генов. В литературе описана структура кЛТ-ДНК (Suzuki et al., 2002) и предполагается, что описанная вставка единственная в геноме *N. glauca*, однако отсутствие полногеномных данных не позволяет подтвердить, либо опровергнуть это предположение. Для остальных представителей рода *Nicotiana* был проведен только анализ отдельных генов методом ПЦР (Furner et al., 1986; Intrieri, Vuiatti, 2001; Chen et al., 2014). Таким образом, на сегодняшний день подробно изучены только природно-трансгенные виды *Nicotiana* подрода *Tabacum*. Подобный анализ, проведенный для представителей подрода *Petunioides*, позволит прояснить сценарии появления и распространения кЛТ-ДНК в данном подроде и сравнить их с известными сценариями для представителей подрода *Tabacum*, что поможет приблизиться к пониманию эволюционных событий рода *Nicotiana*.

**Цель** исследования – изучение структурно-функциональной организации клеточной Т-ДНК у представителей рода *Nicotiana*. В работе были поставлены следующие **задачи**:

1. Анализ структуры кЛТ-ДНК представителей подрода *Petunioides*
  - 1.1. Сборка генома *N. glauca* и уточнение структуры кЛТ-ДНК в геноме
  - 1.2. Сборка генома *N. noctiflora* и анализ структуры кЛТ-ДНК
  - 1.3. Анализ экспрессии генов в кЛТ-ДНК *N. noctiflora*
2. Изучение внутривидового полиморфизма кЛТ-ДНК *Nicotiana* на примере сортов *N. tabacum*
  - 2.1. Сравнительный анализ последовательностей различных по «возрасту» ТА и ТВ
  - 2.2. Анализ крупных перестроек в кЛТ-ДНК на примере ТА.

**Материал, методология и методы исследования.** Материалом исследования послужили асептические растения видов *Nicotiana glauca* (сорт 359 из коллекции ВНИИ табака и махорки) и *Nicotiana noctiflora* Hook (линия TW89 из коллекции Генбанка Института Северной Каролины, US *Nicotiana* Germplasm Collection, North Carolina State University), а также растения *Nicotiana tabacum* сортов отечественной селекции: Турецкий, Ориенталь, Virginia × Berley, Брянский 91, Suifu, Black Indian, Vuelta abajo, Havana 307, выращенные в открытом грунте. В работе использовали методы молекулярно-генетического анализа (выделение нуклеиновых кислот, ПЦР, секвенирование геномных фрагментов), и

биоинформатические методы (различные типы выравниваний, секвенирование и сборка генома).

**Научная новизна исследования.** Впервые получены данные полногеномного анализа для видов *N. glauca* и *N. noctiflora*, относящихся к подроду *Petunioides*. Геном *N. glauca* собран до уровня контигов, проведен анализ последовательностей, гомологичных агробактериальным, в результате чего уточнена структура клТ-ДНК gT - выявлен гомолог *orf13a*, ранее не отмеченный в левом плече повтора. Подтверждено отсутствие иных клТ-ДНК в геноме *N. glauca*, что свидетельствует об однократной агробактериальной трансформации в истории данного вида. Впервые получена полная последовательность gT.

Геном *N. noctiflora* собран до уровня контигов, в результате поиска последовательностей агробактериального происхождения в геноме найдены две различные по составу клТ-ДНК, NnT-ДНК1, и NnT-ДНК2, свидетельствующие о множественных актах трансформации данного вида. Впервые определен состав клТ-ДНК *N. noctiflora*. Определены возможные источники трансгенов - штаммы, максимально близкие к ныне охарактеризованным *Agrobacterium rhizogenes* 1724 и *Agrobacterium tumefaciens* CFBP1935.

В составе клТ-ДНК *N. noctiflora* выявлены интактные гены - *iaaM*, *iaaH*, *acs* в NnT-ДНК1 и *iaaM*, *acs*, *C* в NnT-ДНК2, измерены их относительные уровни экспрессии в различных органах растения. Показан высокий уровень экспрессии в корнях, что согласуется с данными по другим природно-трансгенным видам, а также с паттерном экспрессии генов Т-ДНК в бородачатых корнях трансформированных растений. Возможно, сохранение Т-ДНК в растительном геноме сопровождалось сохранением паттерна экспрессии генов.

Впервые проанализированы последовательности клТ-ДНК растений *N. tabacum* сортов Турецкий, Ориенталь, Virginia × Berley, Брянский 91, Suifu, Black Indian, Vuelta abajo, Havana 307. Как для генов, так и для межгенных пространств в клТ-ДНК был выявлен низкий уровень полиморфизма, сопоставимый с уровнем полиморфизма растительных генов, примером которых служит *PMT2*, отвечающий за синтез никотина. При этом у сортов Брянский 91 и Virginia × Berley показаны крупные структурные перестройки в клТ-ДНК ТА, которые могут быть использованы в качестве филогенетического маркера в эволюционных исследованиях культурного табака.

**Теоретическая и практическая значимость работы.** Собраны геномы видов *N. glauca* и *N. noctiflora*. Полученные в работе полногеномные данные могут быть использованы в различных исследованиях, в частности для изучения генетического контроля вторичных метаболитов и для исследований в сферах фарминдустрии и биотехнологии, поскольку содержат информацию о путях биосинтеза различных соединений, синтезируемых *N. glauca* и *N. noctiflora*.

Результаты анализа состава клТ-ДНК *N. noctiflora* и экспрессии ее генов вносят вклад в понимание возможной функции клТ-ДНК для растений, что расширяет наши знания в области горизонтального переноса генов между растениями и бактериями.

Анализ возможных источников Т-ДНК в природно-трансгенных видах растений позволяет

выявить штаммы агробактерий, которые наиболее часто трансформируют растения в природе. На основе этих данных могут быть созданы базы наиболее эффективных штаммов для дальнейшего их использования в генно-инженерных экспериментах по трансформации растений, в частности, для создания новых сортов сельскохозяйственных культур.

Сравнительный анализ состава и структуры клТ-ДНК в различных представителях рода *Nicotiana*, а также анализ сайтов их локализации в геномах позволяет прояснить филогенетические связи между видами и расширяет наши познания о том, как трансгены способны распространяться между представителями рода в ходе видообразования. Полученные в данной работе результаты анализа нескольких представителей *Nicotiana* говорят о том, что клТ-ДНК в них является, скорее, результатом множественных актов трансформации, чем следствием распространения трансгенов между близкими видами в результате, например, переопыления. Данная информация может быть использована при разработке подходов по оценке возможных экологических рисков возделывания ГМ культур, а также при подготовке лекций для студентов профильных специальностей и при написании научно-популярных материалов.

Разработан новый маркер для исследований внутривидового разнообразия *N. tabacum*, основанный на структурных различиях в клТ-ДНК ТА. Данный маркер может быть использован для изучения филогенетических отношений различных сортов культурного табака.

#### **Положения, выносимые на защиту:**

Геном *N. noctiflora*, относящегося к подроду *Petunioides*, содержит две разные по составу клТ-ДНК, отличающиеся от клТ-ДНК *N. glauca*, что свидетельствует о множественных актах агротрансформации данного вида. То же было описано для видов *N. tomentosiformis* и *N. otophora*, относящихся к подроду *Tabacum*. Таким образом, для представителей разных подродов показано сходство сценариев приобретения клТ-ДНК.

Крупные перестройки в клТ-ДНК ТА в геноме *N. tabacum* могут быть использованы в качестве филогенетического маркера в эволюционных исследованиях культурного табака, а также, в числе прочих маркеров, для паспортизации сортов, поскольку позволяют разделить сорта на группы. Наличие подобного маркера особенно актуально в связи с низким уровнем однонуклеотидного полиморфизма внутри вида *N. tabacum*, показанном на сортах Турецкий, Ориенталь, Virginia × Berley, Брянский 91, Suifu, Black Indian, Vuelta abajo, Havana 307, K326, TN90 и Basma/Xanthi.

**Апробация результатов.** Результаты работы были опубликованы в международных и российских научных журналах и сборниках тезисов, а также были представлены на российских и международных конференциях: III(XI) Международная Ботаническая конференция молодых ученых в Санкт-Петербурге (Санкт-Петербург, Россия, 2015); 2nd World Congress on Beneficial Microbes: Food, Pharma, Aqua & Beverages Industry (Финикс, США, 2016); Международная научная конференция “Растения и микроорганизмы: биотехнология будущего” (Уфа, Россия, 2018); Международная конференция “125 лет

прикладной ботаники в России” (Санкт-Петербург, Россия, 2019); VII съезд Вавиловского общества генетиков и селекционеров, посвященный 100-летию кафедры генетики СПбГУ (Санкт-Петербург, Россия, 2019); XXVIII Всероссийская молодежная научная конференция с элементами научной школы “Актуальные проблемы биологии и экологии” (Сыктывкар, Россия, 2020).

Работа выполнялась при поддержке следующих грантов:

Грант РФФИ 14-04-01480 - “Филогеографическое исследование видов рода *Linaria*, содержащих Т-ДНК *Agrobacterium rhizogenes* в геноме”, грант РФФИ 18-016-00118 - “Изучение роли гена *rolC* клеточной Т-ДНК в регуляции синтеза алкалоидов у табака”, грант РНФ 16-16-110 - “Организация генома природно-трансгенных растений *Linaria* и *Nicotiana*”, грант Министерства науки и высшего образования Российской Федерации в рамках соглашения No 075-15-2020-922 от 16.11.2020 на создание и развитие Научного центра мирового уровня “Агротехнологии будущего”.

**Публикации.** По материалам диссертации опубликованы 3 научные статьи в изданиях, рекомендованных ВАК и 13 научных работ в других изданиях.

**Личный вклад автора.** Основные результаты, изложенные в диссертации, получены автором самостоятельно на кафедре генетики и биотехнологии СПбГУ. Автор лично осуществлял анализ литературных данных по теме работы, проведение лабораторных исследований молекулярными методами, обработку экспериментальных данных, подготовку статей и докладов на конференциях. Планирование экспериментов и обсуждение результатов осуществлялось совместно с научным руководителем. Подготовка библиотек длинных прочтений для генома *N. noctiflora* и сборка генома была проведена коллегами из PMI (Philip Morris International), подготовку парно-концевых библиотек и прыжковых библиотек для *N. noctiflora* и *N. glauca* автор проводил совместно с Полевым Д. Е., что отражено в разделе «Благодарности».

**Структура работы.** Диссертационная работа изложена на 112 страницах, содержит 4 таблицы, иллюстрирована 25 рисунками и состоит из введения, обзора литературы, описания материала и методов исследований, результатов и обсуждения, заключения, списка сокращений и списка использованной литературы, включающего 140 источников, в том числе 137 ссылок на иностранном языке.

## СОДЕРЖАНИЕ РАБОТЫ

### Глава 1. ОБЗОР ЛИТЕРАТУРЫ

В обзоре литературы приведена характеристика рода *Nicotiana*, включающая сведения о появлении и распространении видов по современным ареалам произрастания, представлены сведения о составе рода и особенностях видообразования в его пределах, обобщены данные о природно-трансгенных видах в роде *Nicotiana* и о клТ-ДНК в данных видах. Также представлена характеристика культурного табака, *N. tabacum*, и описано его внутривидовое

разнообразии на уровне сортов и на генетическом уровне.

## Глава 2. МАТЕРИАЛ И МЕТОДЫ ИССЛЕДОВАНИЙ

**Растительный материал.** Материалом исследования послужили асептические растения видов *Nicotiana glauca* сорта '359' и *Nicotiana noctiflora* 'TW89', а также растения *Nicotiana tabacum* сортов: Турецкий, Ориенталь, Virginia × Berley, Брянский 91, Suifu, Black Indian, Vuelta abajo, Havana 307, выращенные в открытом грунте

### В работе был использован комплекс методов:

1) культивирование растений *in vitro* и в условиях теплицы в грунте;  
 2) молекулярно-биологические методы: различные варианты выделения тотальной ДНК растений, с 2ХСТАВ и с 10% СТАВ (Murray, Thompson, 1980), классическая ПЦР и ПЦР в реальном времени с последующей статистической обработкой результатов с помощью метода 2-ΔΔСТ (Livak and Schmittgen, 2001), секвенирование по Сенгеру, выделение РНК, обратная транскрипция, изготовление библиотек двух типов для полногеномного секвенирования;

3) биоинформатические методы включали анализ сырых данных в программе FastQC (Simons, 2010), сборку генома программами MaSuRCA-3.2.2 (Zimin et al., 2017), minimap2 2.17 (Li 2018), miniasm 0.3 (Li 2016), анализ клТ-ДНК в собранных последовательностях программой Last (<http://last.cbrc.jp/>), а также построение попарных (программа blastn из пакета ncbi-BLAST+) и множественных (программа MAFFT v.7) выравниваний.

## Глава 3. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

### Анализ структуры клТ-ДНК представителей подрода *Petunioides*

#### 1. Сборка генома *N. glauca* и уточнение структуры клТ-ДНК gT.

*N. glauca* является первым растением, в геноме которого была обнаружена клТ-ДНК, (White et al., 1983), названная gT (рис.1). Последовательность gT была собрана из фрагментов методом локального выравнивания и не была подтверждена экспериментально (Suzuki et al., 2002). Работа с повторами, в том числе сборка повторов, является одной из сложнейших задач в геномике. Метод локального выравнивания не позволяет получить достоверные результаты при определении точной структуры протяженного повтора с таким высоким уровнем сходства плеч как у gT (96,9%). Использование данного метода также не позволяет определить количество копий повтора в геноме растения. В свою очередь анализ данных полногеномного секвенирования позволяет ответить на эти вопросы, а также определить сайты локализации клТ-ДНК в растительном геноме, что уже было показано на примере *N. tomentosiformis* (Chen et al., 2014). Мы поставили перед собой задачу отсеквенировать и собрать геном природно-трансгенного вида *N. glauca* с целью уточнить структуру клТ-ДНК gT а также проверить наличие либо отсутствие иных клТ-ДНК в данном геноме.



Нами были подготовлены и секвенированы библиотеки *N. glauca* двух типов: pair-end или парно-концевые (размер вставки 350 п.н.) и mate-pair или прыжковые (размер вставки 2000–4000 п.н.). С использованием парно-концевых библиотек осуществляли основную сборку генома, второй тип позволил уточнить сборку, разрешить повторы, что особенно важно в случае клТ-ДНК, которая представляет собой инвертированный повтор. Была проведена сборка *de novo*, размер полученного генома *N. glauca* составил 3,2 млрд.п.н. В результате сборки было получено 514289 скаффолдов, N50 составил 3,1 т.п.н. Собранный геном был загружен в базу NCBI Genome WGS с присвоением Accession number PGPE000000001. *N. glauca* является диплоидным видом ( $2n = 24$ ). Размер полученного генома *N. glauca* приблизительно в 1,2 раза превосходит размеры ранее собранных геномов *N. sylvestris* и *N. tomentosiformis*, составляющие 2,636 млрд.п.н. и 2,682 млрд.п.н. соответственно (Sierra et al., 2013), что согласуется с данными о количестве ДНК в этих видах, полученными методом фотометрии с окрашиванием по Фельгену (Narayan 1987).

Поиск последовательностей гомологичных агробактериальным выполняли с помощью программы Last, осуществляющей полногеномное выравнивание, против предварительно составленной базы данных, использующей суффиксные деревья. База содержала все последовательности, обнаруженные на сегодняшний момент в составе описанных клТ-ДНК. В результате поиска в собранном геноме был найден контиг длиной 8180 п.н., содержащий правую растительно-бактериальную границу и последовательности, гомологичные агробактериальным генам *mis*, *orf14*, *orf13*, *orf13a*, *rolC*, *rolB* и вторую копию *rolC*, то есть одно плечо гТ и середину клТ-ДНК. *De novo* сборка повторов, содержащих последовательности с высоким уровнем сходства, не обладает достаточной точностью, в связи с чем было решено провести проверку результата молекулярным методом. Нами были подобраны праймеры к середине клТ-ДНК и к растительной ДНК, фланкирующей гТ слева. С помощью метода long PCR был получен недостающий фрагмент второго плеча гТ от *rolB*, то есть от середины клТ-ДНК до растительно-бактериальной границы слева от гТ с захватом растительной последовательности ДНК. Далее фрагмент был пошагово секвенирован по Сенгеру. Полученные сиквенсы плеч повтора согласуются с литературными данными (Suzuki, 2002), за исключением гомолога гена *orf13a*, который у Сузуки описан только в одном плече, в то время как по нашим данным оба плеча несут по одной копии *orf13a*. Новых клТ-ДНК в геноме *N. glauca* не обнаружено, что позволяет сделать вывод об однократной агробактериальной трансформации данного вида.

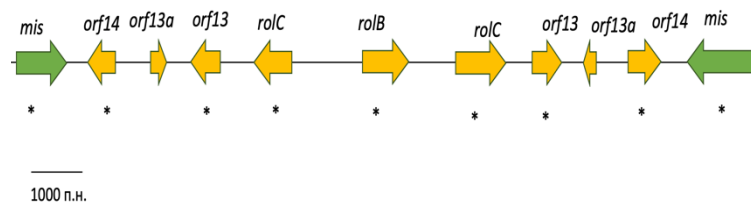


Рис.1 Структура клТ-ДНК в геноме *N. glauca* (Khafizova et. al., 2018), цветными стрелками показано направление открытых рамок считывания, тонкими стрелками показаны плечи повтора, звездочками отмечены сохранные открытые рамки.

Как было сказано ранее, анализ данных полногеномного секвенирования позволяет определить структуру и количество клТ-ДНК в растительном геноме, а также выявить сайты локализации. Подобный анализ ранее был проведен для представителей подрода *Tabacum* (Chen et al., 2014). *N. glauca* относится к подроду *Petunioides*, для представителей которого ранее не было получено полногеномных данных. Применение их для анализа клТ-ДНК представителя подрода *Petunioides* позволит прояснить сценарии появления и распространения клТ-ДНК в данном подрode и сравнить их с известными сценариями для представителей подрода *Tabacum*, что поможет приблизиться к пониманию эволюционных событий рода *Nicotiana*. Кроме того, *N. glauca* является источником широкого ряда вторичных метаболитов, в связи с чем собранный геном данного вида будет полезен для исследований в фармакологической индустрии, поскольку содержит информацию о путях биосинтеза различных соединений.

## 2. Сборка генома *N. noctiflora* и анализ структуры клТ-ДНК

Вид *N. noctiflora* является филогенетически близким к *N. glauca* (Knapp et al., 2004). При проведении скрининга представителей *Nicotiana* методом ПЦР на наличие генов, гомологичных агробактериальным, *N. noctiflora* была отнесена к видам, не содержащим клТ-ДНК в геноме (Intrieri M.C., Vuiatti M. 2001), однако позже данные транскриптома *N. noctiflora* показали обратное (Long et al., 2016). Нами был проведен предварительный анализ методом ПЦР, который подтвердил наличие в геноме *N. noctiflora* последовательностей, гомологичных агробактериальным *orf13*, *orf14*, *rolC* и *mis*. Схожий состав имеет клТ-ДНК близкого вида - *N. glauca* - в связи с чем возникло предположение, что клТ-ДНК в данных видах могла появиться в результате трансформации общего предкового вида. Для того, чтобы проверить данное предположение, необходима информация о точном составе клТ-ДНК в геноме *N. noctiflora*, в связи с чем была поставлена задача - провести полногеномное секвенирование и сборку генома *N. noctiflora*. Подготовку, секвенирование библиотек и оценку качества сырых данных для *N. noctiflora* проводили по аналогии с *N. glauca*. Размер вставки парно-концевой библиотеки составил 850 п.н. с длиной прочтения 250 п.н. Размер вставки прыжковой библиотеки составил около 4 т.п.н. с длиной прочтения 250 п.н. Для улучшения качества сборки нашими коллегами из РМИ, Николаем Ивановым и Николасом Сиерро, были подготовлены длинные прочтения с помощью технологии секвенирования третьего поколения – Oxford Nanopore. Полученные сырые данные также были отфильтрованы (FQ score >7, длина >5000), после чего была проведена гибридная сборка, в результате которой было получено 4468 скаффолдов, N50 составил 261 т.п.н. Таким образом, качество сборки генома *N. noctiflora* выше, чем для *N. glauca*. *N. noctiflora*, как и *N. glauca*, является диплоидом ( $2n = 24$ ). В геноме *N. noctiflora* были найдены 2 различающихся по составу клТ-ДНК. Обе клТ-ДНК организованы в виде несовершенных инвертированных повторов, что характерно для клТ-ДНК в геномах представителей *Nicotiana*. Длина первой, NnT-DNA1, составляет 21787 п.н., длина второй, NnT-DNA2, составляет 14611 п.н. (см. рис. 2; табл. 1). Оценка возраста клТ-ДНК путем сравнения плеч повтора показала, что NnT-DNA2 попала в геном *N. noctiflora* раньше, чем

NnT-DNA1. NnT-DNA1 и gT в геноме *N. glauca* несут одинаковый по структуре фрагмент, включающий гомологи генов *rolC*, *orf13*, *orf13a*, *orf14* и *mis*. Сходство данных фрагментов у двух видов на нуклеотидном уровне составляет 83%. Можно предположить, что данные виды или их предковые формы были трансформированы схожими штаммами агробактерий. Наличие в геноме *N. noctiflora* двух различающихся по составу клТ-ДНК свидетельствует о нескольких актах трансформации в ходе эволюции данного вида. Был проведен сравнительный анализ контигов, содержащих последовательности растительно-бактериальных границ в *N. glauca* и *N. noctiflora*. Было отмечено, что клТ-ДНК данных видов локализованы в разных сайтах в геномах. Различающиеся сайты локализации, а также различия в составе NnT-DNA1 и NnT-DNA2 от gT, свидетельствуют о том, что виды *N. noctiflora* и *N. glauca* были трансформированы независимо, а не получили клТ-ДНК от некой предковой формы до их возникновения.

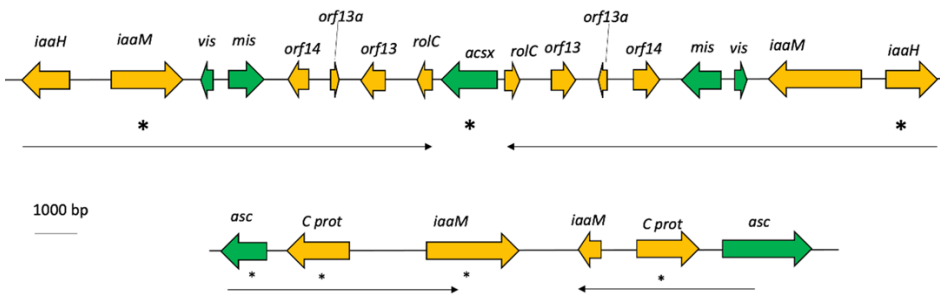


Рис. 2. Структура клТ-ДНК в геноме *N. noctiflora*. Сверху вниз: NnT-DNA1, NnT-DNA2. Цветными стрелками показано направление ОРС, тонкими стрелками показаны плечи повтора, звездочками отмечены сохраненные открытые рамки.

Таблица 1. Состав и анализ генов в клТ-ДНК *Nicotiana noctiflora*

Номер контига	Ген клТ-ДНК (гомолог)	Количество копий	ОРС	координаты	Сходство с последовательностью белка в NCBI		Уровень сходства между копиями (нукл), %
					% сходства	Организм, ID белка	
utg0013691 (NnT-DNA1)	<i>iaaH</i>	2	-	1847-3092	89	<i>Agrobacterium rhizogenes</i> , WP_174054129.1	98
			+	22372-23634	89		
	<i>iaaM</i>	2	+	4122-5789	95	<i>Agrobacterium rhizogenes</i> , WP_174054254.1	99
			-	19676-21892	93		
	<i>vis</i>	2	-	6234-6608	94	<i>Agrobacterium rhizogenes</i> , WP_174054130.1	94
			-	18871-19231	91		
	<i>mis</i>	2	-	6898-7880	77	<i>Agrobacterium rhizogenes</i> , WP_010900210.1	95
			-	17623-18581	79		
	<i>orf14</i>	2	+	8361-8912	74	<i>Agrobacterium rhizogenes</i> , WP_174075805.1	99
			-	16591-17142	75		
<i>orf13a</i>	2	-	9491-9727	50	<i>Agrobacterium rhizogenes</i> , ABS11827.1	100	
		-	15775-16011	50			

	<i>orf13</i>	2	-	10304-10861	75	<i>Agrobacterium rhizogenes</i> , WP_174075804.1	99
			-	14641-15198	73		
	<i>RolC</i>	2	-	11600-11878	85	<i>Agrobacterium rhizogenes</i> , WP_174075803.1	99
			-	13570-13905	70		
	<i>acs</i>	1	+	12088-13392	82	<i>Agrobacterium rhizogenes</i> , GAJ95539.1	-
utg0029481 (NnT-DNA2)	<i>acs</i>	2	+	1952-3104	87	<i>Agrobacterium rhizogenes</i> , WP_174054263.1	96
			-	14263-16562	86		
	<i>C</i>	2	+	3622-5193	76	<i>Agrobacterium rhizogenes</i> , WP_174054195.1	99
			+	12176-13731	74		
	<i>iaaM</i>	2	+	7002-9270	75	<i>Agrobacterium rhizogenes</i> , WP_174054196.1	97

Как для NnT-DNA1, так и для NnT-DNA2 на разных участках выявлено сходство с двумя различными штаммами агробактерий - с *Agrobacterium tumefaciens* CFBP1935 для гомологов генов *iaaH* и *iaaM*, и с *Agrobacterium rhizogenes* 1724 для участка *mis - acs* в NnT-DNA1 и для *acs - C* в NnT-DNA2. Таким образом, можно предположить, что штаммы, близкие к ныне охарактеризованным штаммам *Agrobacterium tumefaciens* CFBP1935 и *Agrobacterium rhizogenes* 1724 могли быть источником клТ-ДНК в геноме *Nicotiana noctiflora*. Мозаичное строение ранее уже было показано для клТ-ДНК ТВ, ТС и TD в геномах *N. tomentosiformis* (Chen et al., 2014), ТС и ТЕ в *N. otophora* (Chen et al., 2018). Предполагают, что подобная структура клТ-ДНК могла получиться в результате трансформации растения штаммом, несущим последовательности, полученные в результате горизонтального переноса генов от различных агробактерий, либо в результате коинфекции растения. Дальнейшее изучение различных клТ-ДНК, а также расширение знаний о разнообразии штаммов агробактерий помогут прояснить вопрос появления мозаичных клТ-ДНК в геномах представителей *Nicotiana*.

### 3. Анализ экспрессии генов в клТ-ДНК *N. noctiflora*

Некоторые из генов, входящих в состав клТ-ДНК, имеют интактную рамку считывания и экспрессируются в растении, в результате чего может меняться фенотип растения, его вторичный метаболизм и реакции на изменение условий внешней среды, в том числе на стресс (Chen and Otten 2017). Чаще всего в природно-трансгенных растениях ОРС сохраняются у генов синтеза опинов и генов, влияющих на морфогенез, в том числе *plast*-генов (Aoki and Syono 1999; Suzuki et al., 2002; Mohajjel-Shoja et al., 2011; Chen et al., 2014; Matveeva et al., 2012; Kyndt et al., 2015). Изучение активности генов клТ-ДНК могло бы помочь ответить на вопрос о функциях клТ-ДНК для растений.

В результате анализа клТ-ДНК *N. noctiflora* нами было обнаружено 7 сохранных ОРС (см. рис. 2), для которых мы измерили относительные уровни экспрессии. Для гомолога гена *acs* и обеих копий *C* в NnT-DNA2 были отмечены низкие уровни экспрессии, что затрудняет

детекцию сигнала, в связи с чем результаты по данным генам были исключены из работы. Значения относительных уровней экспрессии генов *iaaH*, *iaaM*, *acs* в NnT-DNA1, и *iaaM* - в NnT-DNA2 представлены на рисунке 3. Из графиков видно, что экспрессия всех изученных генов достигает максимального значения в тканях корня *N. noctiflora*, снижаясь в порядке лист - стебель. Подобная закономерность была отмечена и для других природно-трансгенных видов (Mohajjel-Shoja et al., 2011). Также известно, что гены Т-ДНК активно экспрессируются в бородачатых корнях трансформированных растений. Возможно, закрепление Т-ДНК в растительном геноме сопровождалось сохранением паттерна экспрессии генов. Интересно, что в стебле и листьях уровень экспрессии гена *acs* заметно выше, чем для генов *iaaH* и *iaaM*. На сегодняшний день неизвестно, какое значение имеют гены клТ-ДНК для растений. Существует ряд гипотез, согласно которым активность приобретенных генов могла придать растениям селективное преимущество. Существует ряд гипотез, согласно которым активность приобретенных генов могла придать растениям селективное преимущество. Например, гены *rolC* и *rolB* могли повышать регенерационные способности растения путем стимуляции клеточного деления и активация меристемы (Koltunow et al., 2001; Otten 2018), а гены *iaaH* и *iaaM*, кодирующие синтез индолилуксусной кислоты, могли влиять на ростовые процессы растений. Опины могли служить меткой, что конкретное растение уже трансформировано, и подавлять попытки повторной трансформации его другими штаммами агробактерий. Можно также предположить, что синтез растением опинов способствует формированию определенных микробных сообществ в ризосфере, а также на поверхности надземных органов. Для винограда было показано, что агробактерии способны перемещаться вверх по лозе от места поранения на расстояние до 30 см за сутки по сосудам ксилемы (Burr et al., 1998). Таким образом, синтез опинов в стебле и листьях трансформированных растений можно объяснить возможностью бактерий перемещаться по растению и ее потребностью в доступе к питательным веществам. В свою очередь природно-трансгенные растения могли сохранить подобную особенность. Функция генов клТ-ДНК для растений пока остается не установленной. Проведение функционального анализа трансгенов для большего количества природно-трансгенных растений помогло бы прояснить данный вопрос.

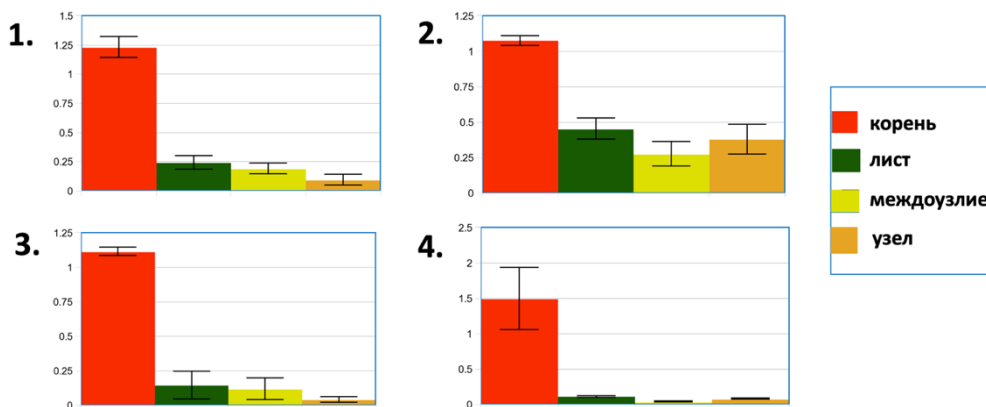


Рис. 3.  
Относительные уровни экспрессии генов клТ-ДНК *Nicotiana noctiflora*:  
1. - *iaaM*,  
2. - *acs*,  
3. - *iaaH* - в NnT-DNA1;  
4. - *iaaM* - в NnT-DNA2.

## Изучение внутривидового полиморфизма клТ-ДНК *Nicotiana* на примере сортов *N. tabacum*

**1. Сравнительный анализ последовательностей различных по «возрасту» ТА и ТВ** *Nicotiana glauca* и *Nicotiana noctiflora* относятся к подроду *Petunioides*, который является слабо охарактеризованной эволюционной ветвью рода *Nicotiana*. Малое видовое разнообразие данного подрода не позволяет корректно экстраполировать полученную информацию о клТ-ДНК в конкретном растении на всех представителей данного вида. В то же время внутри вида *Nicotiana tabacum* выделяют множество сортов, для большинства из которых не установлено происхождение. При широком многообразии фенотипов сорта *N. tabacum* характеризуются низким уровнем генетического разнообразия, что было показано методами RFLP, RAPD, AFLP (Lewis and Nicholson 2007). *Nicotiana tabacum* – это аллотетраплоид ( $2n = 48$ ), образованный путем межвидовой гибридизации видов *N. sylvestris* и *N. tomentosiformis*. *Nicotiana tabacum* является природно-трансгенным видом, поскольку содержит три различающиеся по составу и возрасту клТ-ДНК (ТА, ТВ, ТD), полученные от родительского вида *N. tomentosiformis*. Самой древней из них является ТВ, а самой молодой – ТА. С целью оценки внутривидового полиморфизма клТ-ДНК нами был проведен сравнительный анализ фрагментов самой древней и самой молодой клТ-ДНК в геномах 8 сортов *N. tabacum*: Брянский 91, Virginia × Berley 38, Vuelta abajo, Suifu, Black Indian, Havana 307, Турецкий, Ориенталь. Для анализа были выбраны участки генов, сохранивших рамку считывания, а также участки межгенных последовательностей. В ТА анализировали ген *rolC* (ТА.1) и последовательность между *rolC* и *orf13* (ТА.2), в ТВ – ген *orf14* (ТВ.1) и последовательность между *orf14* и *mis* (ТВ.1) (рис. 4).

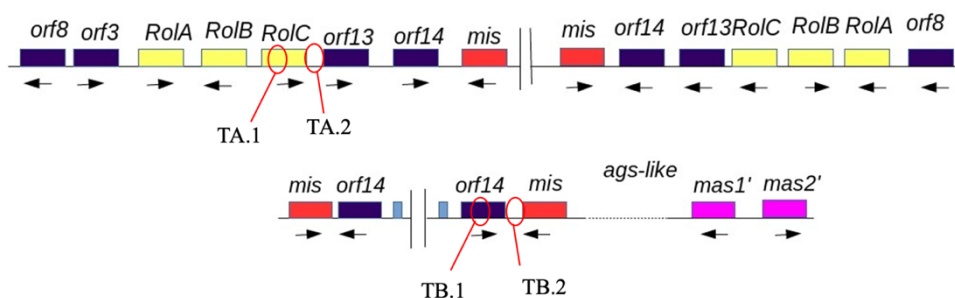


Рис. 4. КлТ-ДНК ТА (сверху) и ТВ (снизу), красными овалами отмечены фрагменты (не в масштабе), взятые в анализ. Длина фрагментов ТА.1, ТА.2, ТВ.1 составила порядка 600 п.н., длина фрагмента ТВ.2 составила 480 п.н.

На матрицах ДНК методом ПЦР были получены фрагменты, визуализация которых на агарозном геле показала единообразие длин (рис. 5). Все полученные фрагменты были секвенированы, после чего были построены 4 множественных выравнивания, показавшие сходство последовательностей как генов, так и межгенных промежутков в обеих клТ-ДНК на уровне 99–100%.



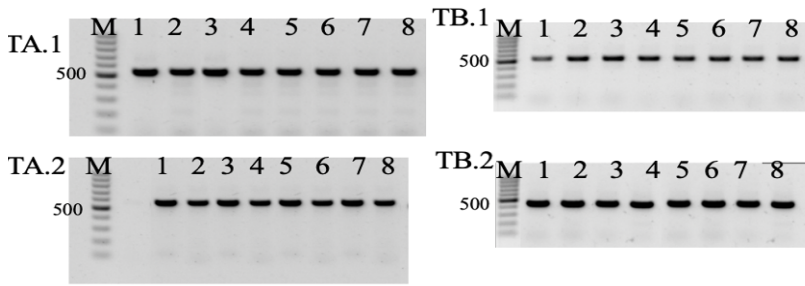


Рис. 5. Единообразие фрагментов гена *rolC* (ТА.1) и межгенного промежутка (ТА.2), гена *orf14* (ТВ.1) и межгенного промежутка (ТВ.2), полученных на матрицах ДНК 8 сортов *N. tabacum*

Возникает вопрос, является ли подобный уровень сходства характерной чертой для трансгенных последовательностей, которые могут сохраняться под действием каких-либо факторов, или данная тенденция распространяется на весь растительный геном. Для того, чтобы ответить на данный вопрос, нами был проведен анализ последовательностей «растительного происхождения» культурного табака на примере гена *PMT2*. *PMT2* - ген путресцин-N-метил-трансферазы - является одним из ключевых участников биосинтеза никотина. Представители различных сортов *N. tabacum* отличаются между собой не только морфологически, но и по ряду физиологических и биохимических показателей, таких как биосинтез различных алкалоидов. Никотин является основным алкалоидом культурного табака (Baranska et al. 2012), в его синтезе задействованы 4 *PMT*-гена. Для анализа был выбран именно *PMT2*, поскольку он содержит делецию 99 п.н. в первом экзоне, благодаря которой его можно достоверно отличить от *PMT1*, *PMT3* и *PMT4*. Материалом для анализа послужили отсеквенированные геномы сортов *N. tabacum* K326 (NCAA00000000.1, AWOJ00000000.1), TN90 (AYMY00000000.1) и Vasma Xanthi (AWOK00000000.1). Было построено множественное выравнивание последовательностей гена *PMT2* для трех сортов *N. tabacum*, которое показало, что последовательности сортов K326 и TN90 идентичны как в экзонах, так и в интронах, и отличаются от сорта Vasma Xanthi по двум нуклеотидным заменам в первом и в шестом экзонах соответственно, а также выпадениями нуклеотидов в двух интронах (Рис. 6). Таким образом, показан низкий уровень генетического разнообразия внутри вида *N. tabacum* как для трансгенных последовательностей, приобретенных в результате горизонтального переноса, так и для последовательностей «растительного происхождения». Полученные результаты согласуются с показанными ранее другими методами (Lewis and Nicholson 2007). Подобный уровень полиморфизма может быть следствием недавнего происхождения вида *N. tabacum*, а также результатом направленной селекции Moon et al., 2009). Опираясь на имеющиеся данные, невозможно сделать вывод, что послужило причиной появления существующего уровня полиморфизма в сортах *N. tabacum*, возможно, играло роль сочетание нескольких факторов. Дальнейшее изучение геномов сортов культурного табака, в том числе расширение выборки анализируемых сортов, может способствовать разьяснению данного вопроса.

```

PMT2_BasmaXanth atgaatggccaccaatggcacttccaacacccaaaacggccacaagaatgggacttcc
PMT2_K326 atgaatggccaccataatggcacttccaacacccaaaacggccacaagaatgggacttcc
PMT2_TN90 atgaatggccaccataatggcacttccaacacccaaaacggccacaagaatgggacttcc
*****

PMT2_BasmaXanth cgtaagtcctttaagggctctgtcaactatgcttggactactgttccaacatatccBacg
PMT2_K326 cgtaagtcctttaagggctctgtcaactatgcttggactactgttccaacatatccaacg
PMT2_TN90 cgtaagtcctttaagggctctgtcaactatgcttggactactgttccaacatatccaacg
*****

PMT2_BasmaXanth tattttctctctctctctctctctataaaaatggaagtttgattctataattgtca
PMT2_K326 tatttt--tctctctctctctctctataaaaatggaagtttgattctataattgtca
PMT2_TN90 tatttt--tctctctctctctctctataaaaatggaagtttgattctataattgtca
*****

PMT2_BasmaXanth tataattaactctttcaaatgtcttttttttttcagattcacaagcagcattca
PMT2_K326 tataattaactctttcaaatgtc--tttttttttcagattcacaagcagcattca
PMT2_TN90 tataattaactctttcaaatgtc--tttttttttcagattcacaagcagcattca
*****
    
```

Рис. 6. Множественное выравнивание последовательностей гена *PMT2* в геномах *N. tabacum* сортов K326 (NCAA00000000.1, AWOJ00000000.1), TN90 (AYMY00000000.1) и Basma Xanthi (AWOK00000000.1), показаны фрагменты с различиями.

## 2. Анализ крупных перестроек в клТ-ДНК на примере ТА

При изучении клТ-ДНК в различных сортах *Nicotiana tabacum* Chen с коллегами отметила, что сорта Basma Drama, Samsoun, и Xanthi несут в центральной части ТА делецию длиной 9684 п.н., расположенную между двумя неполными последовательностями *orf13*, в то время как сорта Wisconsin 38 и Havana 425 содержат полноразмерную ТА (см. рис. 7). В связи с тем, что для сортов культурного табака описан низкий уровень внутривидового разнообразия, для проведения исследований по выявлению родственных отношений между сортами данного вида необходимо разработать дополнительные маркеры. В качестве такого маркера можно использовать различия в структуре клТ-ДНК ТА. Кроме того, данный маркер можно применять для оценки полиморфизма клТ-ДНК на уровне одного вида на примере культурного табака.

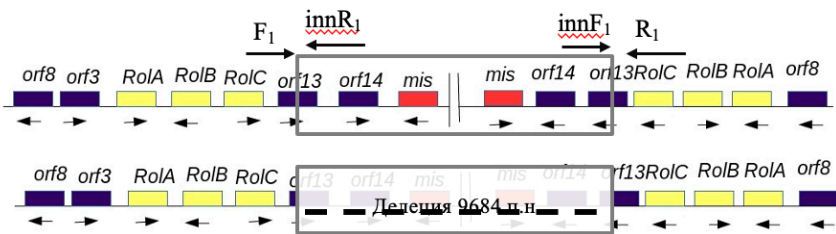


Рис. 7. схема ТА в *N. tabacum*, (сверху) клТ-ДНК ТА без делеции, описана в сортах Wisconsin 38, Havana 425; (снизу) клТ-ДНК ТА с делецией, описана у сортов Basma Drama, Samsoun, и Xanthi (Chen et al., 2014)

Нами был проведен анализ центральных участков ТА для 8 сортов культурного табака с целью выявить наличие либо отсутствие описанной делеции. Схема эксперимента заключалась в постановке на матрицах ДНК сортов *N. tabacum* ПЦР с сочетаниями праймеров к границам предполагаемой делеции (F + R), либо к границе делеции и к центральной части ТА (F + innR/innF+R) (см. рис. 7). В ходе работы для каждого сорта были получены фрагменты только с одним из сочетаний праймеров (рис. 8), таким образом, все исследованные образцы были поделены на две группы: несущие делецию (Брянский 91, Virginia × Berley 38) и сорта без делеции в центральном участке ТА (Vuelta abajo, Suifu, Black Indian, Havana 307, Турецкий, Ориенталь).



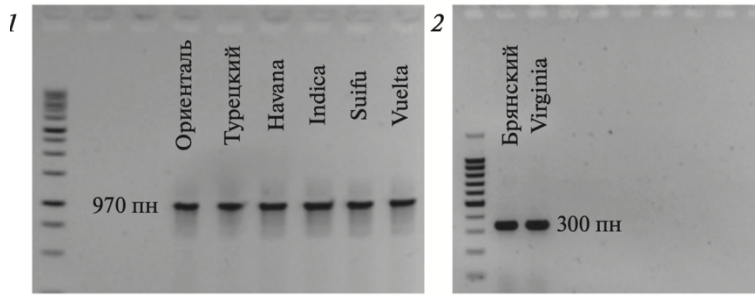


Рис. 8. Фрагменты, полученные на матрицах сортов *N. tabacum*: (1) – фрагменты, полученные с использованием праймеров F и innR, (2) – фрагменты, полученные с использованием праймеров F и R

Анализ сиквенсов полученных фрагментов показал, что положение делеции у сортов Брянский 91 и Virginia × Berley 38, а также у сорта Vasma Xanthi (AWOK01769803.1) совпадает с точностью до нуклеотида. Можно предположить, что данная делеция возникла у некой предковой формы до того, как были созданы сорта Virginia × Berley 38, Vasma Xanthi, Брянский 91, а также Vasma Drama, Samsoun, и Xanthi. Анализ сортов K326 (NCAA00000000.1, AWOJ00000000.1) и TN90 (AYMY00000000.1) показал отсутствие делеции в ТА. Согласно делению сортов по рыночным классам, а также согласно классификации Псаревой по подвидам, Vasma Drama, Samsoun, и Xanthi, в которых ранее была описана делеция, относятся к классу восточных табаков, так же, как и Vasma Xanthi, для которой наличие делеции было показано в данной работе. Если предположить, что данная делеция является маркером сортов, относящихся к восточным табакам, то сорта Virginia × Berley 38 и Брянский 91 также следует отнести к восточным табакам. Однако сорт Брянский 91 по коммерческой классификации относят к классу наполнителя для сигар, а по Псаревой - к Бразильской разновидности Южного подвида. Сорт Virginia × Berley 38 по коммерческой классификации так же относят к классу наполнителя для сигар, а по Псаревой он является гибридом Виргинской и Мерилендской разновидностей Американского подвида. Таким образом, пока невозможно однозначно определить отношение сортов с делецией к какой-либо группе внутри вида *N. tabacum* согласно имеющимся классификациям. Для разработки адекватной внутривидовой классификации культурного табака необходимо провести масштабную работу, охватывающую количество сортов, достаточное для отражения имеющегося их разнообразия. Учитывая сложности с подбором молекулярных маркеров, связанные с низким уровнем полиморфизма в сортах культурного табака, для построения классификации скорее всего потребуется разработка ряда маркеров, например микросателлиты, вариации числа копий генов и структурные варианты. В качестве одного из структурных вариантов может быть использована делеция в центральном участке ТА. Интересно, что как анализ структуры клТ-ДНК, так и анализ полиморфизма гена *PMT2* привели к одному и тому же результату - разделению сортов K326, TN90 и сорта Vasma Xanthi на разные группы по наличию/отсутствию делеции в центральном участке клТ-ДНК ТА и по уровню сходства последовательности гена *PMT2*. Результатов анализа трех сортов недостаточно, чтобы делать выводы, необходимо провести дополнительный эксперимент на большой выборке. Однако уже сейчас можно сказать, что структурный полиморфизм клТ-ДНК в сортах *N. tabacum* может быть использован для

изучения филогенетических отношений в пределах данного вида.

## ЗАКЛЮЧЕНИЕ

Собранные в данной работе геномы видов *N. glauca* и *N. noctiflora*, относящихся к эволюционной ветви *Petunioides*, позволили провести детальный анализ клТ-ДНК в этих видах. Результаты показали, что клТ-ДНК в *N. glauca* и *N. noctiflora* являются результатом независимых актов трансформации, на что указывают различия в составе клТ-ДНК и в сайтах локализации трансгенов в растительных геномах. В то время как для *N. glauca* присутствие в геноме единственной клТ-ДНК гТ свидетельствует об однократном акте агробактериальной трансформации данного вида, две различающиеся по составу клТ-ДНК в геноме *N. noctiflora* являются следствием множественных актов трансформации. Таким образом, для эволюционных ветвей *Petunioides* и *Tabacum*, к которой относятся виды *N. tomentosiformis* и *N. otophora*, показаны схожие сценарии появления клТ-ДНК. Интересно, что обе клТ-ДНК *N. noctiflora* имеют мозаичное строение, так же как клТ-ДНК *N. otophora* и некоторые из клТ-ДНК *N. tomentosiformis* (Chen et al., 2014; Chen et al., 2018).

Виды *N. glauca* и *N. noctiflora* являются первыми представителями подрода *Petunioides*, чьи геномы были отсеквенированы, что значительно расширяет возможности для проведения филогенетических исследований рода *Nicotiana*.

Мы показали, что гены клТ-ДНК *N. noctiflora* активнее всего экспрессируются в корнях, что согласуется с данными по другим видам природно-трансгенных растений. Возможно, при закреплении Т-ДНК в растительном геноме сохраняется паттерн экспрессии генов Т-ДНК в трансформированном растении.

Низкий уровень полиморфизма был показан для последовательностей ДНК сортов культурного табака, *N. tabacum*, как растительного, так и агробактериального происхождения. Это может быть следствием недавнего образования вида *Nicotiana tabacum* (Leitch et al., 2008), и/или результатом активного процесса селекции (Moon et al., 2009). Также есть мнение, что в формировании вида *N. tabacum* могла участвовать только небольшая часть генетического разнообразия, существовавшего в генофондах исходных видов-предшественников, *N. tomentosiformis* и *N. sylvestris*. (Lewis and Nicholson 2007).

Молекулярный маркер на основе делеции в центре ТА, разработанный для *Nicotiana tabacum*, позволяет разделить сорта на 2 группы, и может быть использован как для изучения филогенетических отношений различных сортов культурного табака, так и в качестве одного из маркеров для паспортизации сортов.

При изучении природно-трансгенных видов *Nicotiana* не обнаружено примеров распространения клТ-ДНК горизонтально между близкими видами в результате, например, переопыления. Об этом могло бы свидетельствовать присутствие у представителей разных подродов клТ-ДНК со схожей структурой при совпадении сайтов локализации клТ-ДНК в геномах. Наоборот, показано, что последовательности клТ-ДНК в филогенетически более

близких видах имеют большую степень сходства, нежели последовательности клТ-ДНК видов, относящихся к различным под родам (Intrieri, Buiatti, 2001). Так NnT-DNA1 в геноме *N. noctiflora* и gT в геноме *N. glauca* несут одинаковый по структуре фрагмент, включающий гомологи генов *rolC*, *orf13*, *orf13a*, *orf14* и *mis*. Сходство данных фрагментов у двух видов на нуклеотидном уровне составляет 83%. На основании полученных нами данных и результатов, ранее полученных другими научными группами, можно сделать вывод о низком уровне риска передачи трансгенов и их закрепления в геномах родственных форм, что является аргументом в защиту биобезопасности возделывания ГМО.

## ВЫВОДЫ

1. Анализ структуры генома *N. glauca* свидетельствует о том, что клТ-ДНК gT является единственной в геноме данного вида.
2. Геном *N. noctiflora* содержит 2 разные по составу клТ-ДНК (NnT-ДНК1, и NnT-ДНК2), вероятными источниками которых могут быть штаммы сходные с *Agrobacterium rhizogenes* 1724 и *Agrobacterium tumefaciens* CFBP1935.
3. В составе NnT-ДНК1 интактными являются гены *iaaM*, *iaaH*, *acs*, в составе NnT-ДНК2 - *iaaM*, *acs*, *C*, указанные гены активнее всего экспрессируются в корнях.
4. Последовательности ДНК сортов культурного табака, *N. tabacum*, агробактериального происхождения, а также гена *PMT2* «растительного происхождения» характеризуются одинаково низким уровнем полиморфизма.
5. Делеция в центральном участке ТА позволяет разделить сорта культурного табака на 2 группы по происхождению.

**Благодарности.** Хочу выразить благодарность своему научному руководителю Матвеевой Татьяне Валерьевне за помощь в планировании экспериментов, за обсуждение результатов и советы в подготовке докладов и статей по теме исследования, а также текста рукописи диссертации. Профессора Лутову Людмилу Алексеевну и профессора Леона Оттена (Leon Otten, Institut de Biologie Moléculaire du C.N.R.S.) благодарю за обсуждение результатов и ценные советы на протяжении всех лет работы над диссертацией. Особую благодарность выражаю сотруднику международной лаборатории “компьютерные технологии” ИТМО Добрынину Павлу Владимировичу за помощь в обработке данных полногеномного секвенирования, а также всестороннюю поддержку и консультации на протяжении всех этапов подготовки диссертации. Также благодарю сотрудника генетической лаборатории “SERBALAB” Полева Дмитрия Евгеньевича за помощь в подготовке библиотек и проведение секвенирования, коллег из РМИ – Николая Владимировича Иванова и Николаса Сьерро (Nicolas Sierro) за подготовку длинных прочтений и сборку генома *N. noctiflora*, сотрудницу Генбанка Института Северной Каролины Джессику Нифонг (Jessica Nifong, US Nicotiana Germplasm Collection, North Carolina State University) за предоставленные семена *N. noctiflora* и сотрудницу СПбГУ Наталью Цветкову за помощь. Также благодарю весь коллектив

лаборатории Генной и клеточной инженерии растений кафедры генетики СПбГУ, особенно Творогову Варвару Евгеньевну и Ганчеву Марию Семеновну, технические советы и моральная поддержка которых внесла неоценимый вклад в подготовку данной работы. И, наконец, благодарю коллег из Отдела генетических ресурсов масличных и прядильных культур ВИР, за терпение и поддержку на этапе написания рукописи и подготовки к защите.

В работе использовано оборудование ресурсных центров научного парка СПбГУ «Развитие молекулярных и клеточных технологий» и «Биобанк».

### Список публикаций по материалам диссертации:

#### В журналах, рекомендованных ВАК:

1. Galina Khafizova *Nicotiana glauca* whole-genome investigation for cT-DNA study/ **Galina Khafizova**, Pavel Dobrynin, Dmitrii Polev and Tatiana Matveeva// BMC Research Notes. 2018. V. 11. No. 18. P. 1. doi.org/10.1186/s13104-018-3127-x
2. Хафизова Г. В. Полиморфизм последовательностей ДНК агробактериального происхождения в сортах *Nicotiana tabacum*/ **Хафизова Г. В.**, Матвеева Т. В.// Генетика. 2020. Т. 56. No. 10. С. 1269. doi.org/10.31857/S0016675820100057
3. Khafizova G. Study of sequences in *Nicotiana* genomes acquired from *Agrobacterium* by horizontal gene transfer/ **Khafizova G.**// Biology and Medicine. 2016. V. 8. No. 6. P. 40. doi.org/10.4172/0974-8369.C1.003

#### В прочих изданиях:

1. Хафизова Г. В. Ген *rolC* агробактерий: на пути к пониманию функции/ **Хафизова Г. В.**, Матвеева Т. В.// Биотехнология и селекция растений. 2021. Т. 4. No. 1. С. 36–46. doi.org/10.30901/2658-6266-2021-1-04
2. Хафизова Г. В. Поиск клТ-ДНК в растительном геноме с помощью метода полногеномного выравнивания/ **Хафизова Г. В.**, Добрынин П. В., Матвеева Т. В.// Вестник защиты растений. 2016. Т. 89. №. 3. С. 177–178.
3. Хафизова Г. В. КлТ-ДНК как молекулярный маркер в филогении *Nicotiana tabacum*/ **Хафизова Г. В.**, Матвеева Т. В.// 125 years of Applied Botany in Russia. 2019. С. 62–62.
4. Хафизова Г. В. Природно-трансгенные растения и методы их изучения/ **Хафизова Г. В.**, Матвеева Т. В. //Актуальные проблемы биологии и экологии. 2020. С. 117–119.
5. Хафизова Г. В. Анализ полиморфизма клТ-ДНК у *N. tabacum*/ **Хафизова Г. В.**, Матвеева Т. В.// VII съезд Вавиловского общества генетиков и селекционеров, посвященный 100-летию кафедры генетики СПбГУ, и ассоциированные симпозиумы: Сборник тезисов конгресса. СПб.: Издательство «ВВМ», 2019. стр. 873