

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ «САНКТ-ПЕТЕРБУРГСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
КАФЕДРА ГЕНЕТИКИ И БИОТЕХНОЛОГИИ

На правах рукописи



Хафизова Галина Васильевна
**ОСОБЕННОСТИ ОРГАНИЗАЦИИ КЛЕТОЧНОЙ Т-ДНК
У ПРЕДСТАВИТЕЛЕЙ РОДА *NICOTIANA L.***

Специальность: 03.02.07 Генетика

Диссертация
на соискание ученой степени
кандидата биологических наук

Научный руководитель:
доктор биологических наук, профессор
Матвеева Татьяна Валерьевна

Санкт-Петербург
2022

Оглавление

СПИСОК СОКРАЩЕНИЙ.....	4
ВВЕДЕНИЕ	5
ГЛАВА 1. ОБЗОР ЛИТЕРАТУРЫ.....	16
1.1. Общая характеристика, происхождение и география рода <i>Nicotiana</i> L..	16
1.2. Особенности видообразования в роде <i>Nicotiana</i>	21
1.3. Природно-трансгенные виды в роду <i>Nicotiana</i>	26
1.4. Ботаническое описание и характеристика генома вида <i>N. tabacum</i>	38
1.5. Внутривидовое разнообразие <i>N. tabacum</i>	43
ГЛАВА 2. МАТЕРИАЛЫ И МЕТОДЫ.....	51
2.1. Анализ структуры клТ-ДНК представителей подрода <i>Petunioides</i>	51
2.1.1. Растительный материал.....	51
2.1.2. Методы культивирования растений, использованные в работе..	51
2.1.3. Молекулярно-биологические методы.....	51
2.1.4. Биоинформатические методы.....	54
2.2. Изучение внутривидового полиморфизма клТ-ДНК <i>Nicotiana</i> на примере сортов <i>N. tabacum</i>	55
2.2.1. Растительный материал.....	55
2.2.2. Методы культивирования растений, использованные в работе..	55
2.2.3. Молекулярно-биологические методы.....	56
2.2.4. Биоинформатические методы.....	59
ГЛАВА 3. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ.....	60
3.1. Анализ структуры клТ-ДНК представителей подрода <i>Petunioides</i>	60
3.1.1. Сборка генома <i>N. glauca</i> и уточнение структуры клТ-ДНК gТ..	60
3.1.2. Сборка генома <i>N. noctiflora</i> и анализ структуры клТ-ДНК.....	64
3.1.3. Анализ экспрессии генов в клТ-ДНК <i>N. noctiflora</i>	69

3.2. Изучение внутривидового полиморфизма клТ-ДНК <i>Nicotiana</i> на примере сортов <i>N. tabacum</i>	73
3.2.1. Сравнительный анализ последовательностей ГА и ТВ, различных по «возрасту».....	73
3.2.2. Анализ крупных перестроек в клТ-ДНК на примере ГА.....	79
ЗАКЛЮЧЕНИЕ.....	83
ВЫВОДЫ.....	85
СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ.....	86
БЛАГОДАРНОСТИ.....	104
ПРИЛОЖЕНИЯ.....	106

СПИСОК СОКРАЩЕНИЙ

КлТ-ДНК — клеточная Т-ДНК

млн — миллион

млрд.п.н. — миллиард пар нуклеотидов

ОРС — открытая рамка считывания

п.н. — пар нуклеотидов

пГМО — природные генно-модифицированные организмы

ПЦР — полимеразная цепная реакция

Т-ДНК — transferred DNA, перенесенная ДНК

т.п.н. — тысяча пар нуклеотидов

NGS — next generation sequencing, секвенирование следующего поколения

ТА, ТВ, ТС, ТD, ТЕ, гТ — названия клеточных Т-ДНК в геномах видов *Nicotiana L.*

ВВЕДЕНИЕ

T-ДНК представляет собой последовательности, локализованные на больших плаزمиде почвенных бактерий «*Agrobacterium*». «*Agrobacterium*» — это сборная группа бактерий, объединяющая *A. rhizogenes*, *A. tumefaciens* и *A. vitis*, которые относятся к различным родам. Ранее они относились к одному роду, *Agrobacterium*, внутри которого штаммы были поделены на группы в зависимости от того, какой тип плазмиды с T-ДНК они содержали. Бактерий, несущих Ti плазмиду, относили к *A. tumefaciens* и *A. vitis*, а штаммы с Ri плазмидой были отнесены к *A. rhizogenes*. Однако в настоящее время выстраивается новая классификация, основанная на результатах филогенетических исследований. Согласно новой таксономической системе, некоторые представители *A. tumefaciens* классифицированы как *Rhizobium radiobacter* (Young et al., 2001), *A. rhizogenes* переименован в *Rhizobium rhizogenes* (Young et al., 2001), а *A. vitis* переименован в *Allorhizobium vitis* (Mousavi et al., 2015). Поскольку классификация *Agrobacterium/Rhizobium* все еще находится в стадии разработки и в большинстве литературных источников используются традиционные названия — *A. tumefaciens*, *A. rhizogenes* и *A. vitis*, в данной работе также будут использоваться эти названия.

Агробактерии способны переносить T-ДНК в клетки растений и встраивать их в растительный геном, в результате чего на растениях развиваются косматые корни либо корончатый галл (Cardarelli et al., 1987; Akiyoshi et al., 1984). В норме в природе трансформированное растение погибает, однако в 1982 году было обнаружено растение *Nicotiana glauca* Graham., геном которого содержал гомологичные T-ДНК последовательности (клеточную T-ДНК), при этом растение имело нормальный фенотип без характерных признаков агробактериальной инфекции (White et al., 1983). Подобные растения, в геномах которых содержатся последовательности иных организмов, являются трансгенными. В случае, когда трансформация произошла в

природе, без участия человека, говорят о природно-трансгенных растениях или природных генно-модифицированных организмах (пГМО). Первое природно-трансгенное растение, геном которого содержал клеточную Т-ДНК, было обнаружено случайно (White et al., 1983). Далее, в результате уже целенаправленного скрининга рода *Nicotiana* L., было описано еще 14 природно-трансгенных видов (Intrieri and Buiatti, 2001), а спустя еще десятилетие клТ-ДНК была найдена в геномах льнянки (Matveeva et al., 2012) и батата (Kyndt et al., 2015). Расширению списка пГМО способствовало в том числе появление новых методов, таких как NGS (next generation sequencing, секвенирование нового поколения) (Chen et al., 2014), а также накопление данных о геномных последовательностях растений, что позволило применять биоинформатические методы поиска трансгенов и обрабатывать большие массивы данных (Matveeva and Otten, 2019). На сегодняшний день в списке пГМО более 40 видов двудольных, два вида однодольных, и список продолжает увеличиваться (Matveeva, 2021). Известные природно-трансгенные растения можно разделить на 3 группы в зависимости от того, какие гены входят в состав их клеточной Т-ДНК (Matveeva and Otten, 2019). В состав клТ-ДНК могут входить гены синтеза опинов (*mis*, *acs*, *mas* и др.), *plast*-гены (*rolC*, *rolB*, *orf13*, *orf14* и др.), а также гены синтеза ауксина (*iaaH*, *iaaM*) и гены с пока неустановленной функцией (*orf511*, *orf13a*). Многие гены в составе клТ-ДНК сохраняют рамку считывания (Matveeva and Otten, 2019). Так называемые *plast*-гены (от phenotypic plasticity) влияют на развитие растения, вызывая различные морфологические эффекты, в том числе разрастание корневой массы и образование корончатого галла (Otten, 2018). Помимо морфологических эффектов *plast*-гены могут влиять на вторичный метаболизм (Bulgakov, 2008) и на углеводный обмен растений (Mohajjel-Shoja et al., 2011). Предполагают, что активность данных генов может изменять гормональный баланс растения (Otten, 2018), однако молекулярные механизмы данных процессов на сегодняшний день не установлены. Опины — это продукты конденсации аминокислот

с кетокислотами, либо продукты конденсации сахаров, которые используются агробактериями в качестве источника азота и углерода. Таким образом, внедряя гены синтеза опинов в растительный геном, бактерия создает себе «фабрику» по производству питательных веществ (Petit et al., 1979). Для некоторых видов природно-трансгенных растений (культурный табак, повилика) показан синтез опинов (Chen et al., 2016; Zhang et al., 2020). Самая многочисленная группа пГМО включает виды, клТ-ДНК которых состоит только из генов синтеза опинов. В самой малочисленной, куда входит всего четыре вида - *Eugenia uniflora* и *Vaccinium macrocarpon*, - клТ-ДНК содержит только *plast*-гены. Наибольший интерес представляет группа растений с «композитными» клТ-ДНК со сложным составом, в которые входят как гены синтеза опинов, так и *plast*, и иные гены. К данной группе относятся, помимо прочих, представители рода *Nicotiana*.

Актуальность темы исследования

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. История изучения природно-трансгенных растений началась с вида *Nicotiana glauca* (White et al., 1983), а сегодня в роде *Nicotiana* насчитывают 16 видов, геномы которых содержат клТ-ДНК (Intrieri, Buiatti, 2001; Chen et al., 2014; Long et al., 2016; Chen et al., 2018). 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 среди представителей

данного рода, что, в свою очередь, помогает прояснить эволюционные процессы, происходившие в роду *Nicotiana*. Так показано присутствие в геноме *N. tomentosiformis* четыре разных по составу клТ-ДНК (ТА, ТВ, ТС, TD), три из которых также найдены в геноме *N. tabacum*, что одновременно служит доводом в пользу гипотезы о происхождении *N. tabacum* в результате гибридизации *N. tomentosiformis* и *N. sylvestris*, 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 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 у трех сортов *N. tabacum* – TN90, K326 и Basma/Xanthi, выявлены внутривидовые различия в структуре клТ-ДНК (Chen et al., 2014). По результатам анализа данных о составе и копийности 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; Intriери, Buiatti, 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 подрode и сравнить их с известными сценариями для представителей подрода *Tabacum*, что поможет приблизиться к пониманию эволюционных событий рода *Nicotiana*.

Целью данной работы является изучение структурно-функциональной организации клеточной Т-ДНК у представителей рода *Nicotiana*. В работе поставлен ряд задач:

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

Научная новизна диссертационной работы

Впервые получены полногеномные данные 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. Данные полногеномного секвенирования и сборки генома *N. glauca* были размещены в базе данных NCBI GenBank и в настоящее время находятся в открытом доступе.

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*: NnT-ДНК1 содержит гомологи агробактериальных генов *iaaH*, *iaaM*, *vis*, *mis*, *orf14*, *orf13a*, *orf13*, *RolC* и *acs*, в состав NnT-ДНК2 входят гомологи *acs*, *C* и *iaaM*. Определены возможные источники трансгенов - штаммы, максимально близкие к ныне охарактеризованному *Agrobacterium rhizogenes* 1724 и *Agrobacterium*

tumefaciens CFBR1935.

В составе клТ-ДНК *N. noctiflora* выявлены интактные гены - *iaaM*, *iaaH*, *acs* в NnТ-ДНК1 и *iaaM*, *acs*, *C* в NnТ-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. Эти факты могут свидетельствовать как о недавнем происхождении сортов, так и о прохождении видом “бутылочного горлышка” в ходе своей эволюции, что согласуется с современными представлениями об истории вида *N. tabacum*. При этом у сортов Брянский 91 и Virginia × Berley показаны крупные структурные перестройки в клТ-ДНК ТА, ранее описанные для сорта Vasma/Xanthi. Подобные перестройки могут быть использованы в качестве филогенетического маркера в эволюционных исследованиях культурного табака.

Основные положения, выносимые на защиту

Геном вида *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*, показанного на сортах Турецкий, Ориенталь, Virginia × Verley, Брянский 91, Suifu, Black Indian, Vuelta abajo, Havana 307, K326, TN90 и Basma/Xanthi.

Практическая и теоретическая 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.

Материал и методы исследования. Материалом исследования послужили асептические растения видов *Nicotiana glauca* (сорт 359 из коллекции ВНИИ табака и махорки) и *Nicotiana noctiflora* (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. Асептические растения поддерживали *in vitro* в стерильных условиях. В работе использовали методы молекулярно-генетического анализа (выделение нуклеиновых кислот, ПЦР, секвенирование геномных фрагментов), и биоинформатические методы (различные типы выравниваний, секвенирование и сборка генома *de novo*, гибридная сборка).

Апробация работы

Результаты работы были опубликованы в международных и российских научных журналах и сборниках тезисов, а также были представлены на следующих 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:

1. Грант РФФИ 14-04-01480 - “Филогеографическое исследование видов рода *Linaria*, содержащих T-ДНК *Agrobacterium rhizogenes* в геноме”.
2. Грант РФФИ 18-016-00118 - “Изучение роли гена *rolC* клеточной T-ДНК в регуляции синтеза алкалоидов у табака”.
3. Грант РФФИ 16-16-110 - “Организация генома природно-трансгенных растений *Linaria* и *Nicotiana*”.
5. 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.

ГЛАВА 1. ОБЗОР ЛИТЕРАТУРЫ

1.1. Общая характеристика, происхождение и география рода *Nicotiana*

Род *Nicotiana* относится к семейству пасленовых, Solanaceae, порядку *Solanales* (Olmstead et al., 2008). Среди представителей данного рода встречаются различные жизненные формы от невысоких однолетних трав (*N. linearis*) до кустарников, достигающих 7 метров в высоту (*N. glauca*). Также виды значительно различаются по форме и размеру листовой пластинки, по размеру цветка и окраске венчика, по типам опыления (Knapp 2020). Среди представителей *Nicotiana* описаны как самоопылители (*N. plumbaginifolia*), так и опыляемые ветром/насекомыми виды (*N. longiflora*, *N. alata*), а для некоторых (*N. glauca*) возможны оба варианта опыления (Lee et al., 2008; Schueller, 2004). Будучи различными морфологически и физиологически, представители рода *Nicotiana* освоили широкий спектр условий произрастания от болотистых регионов (*N. longiflora*) до пустынь (*Nicotiana gossei*). Род *Nicotiana* включает в себя 82 вида, из которых 40 являются диплоидами, 35 - аллополиплоидами. Для 7 видов число хромосом не определено, в основном к ним относятся представители, произрастающие в Австралии (Knapp, 2020). Так как почти половина видов *Nicotiana* являются результатом межвидовой гибридизации, вопрос построения филогении рода *Nicotiana* и восстановление его демографической истории представляют собой сложные задачи для исследователей. В 1954 году в своей монографии «*Nicotiana*» Goodspeed предположил, что предковые формы, давшие начало роду *Nicotiana*, появились путем межвидовой гибридизации представителей наиболее близких к нему родов - *Cestrum* и *Petunia*. Он выделял два предковых генетических пула: более «примитивный» *Cestroid* и более «прогрессивный» *Petunioid*, от которых взяли начало две отдельные ветви рода *Nicotiana*, обладающие четко различимыми морфологическими признаками, и ставшие началом всего

современного разнообразия. Goodspeed также предположил, что исходно хромосомный набор в роду равнялся 12, а эволюция рода и появление многообразия видов шли в основном за счет гибридизации и удвоения числа хромосом (Goodspeed, 1954). Позже анализ последовательностей хлоропластных генов *ndhF* и *rbcL* представителей семейства Solanaceae показал, что ни *Cestrum*, ни *Petunia* не являются филогенетически близкими родами к *Nicotiana*. По последним данным род *Nicotiana* входит в ранее не идентифицированную кладу «X = 12», объединяющую виды с хромосомным набором равным 24 (Olmstead and Palmer, 1992; Olmstead and Sweere, 1994), и является сестринской группой для эндемичных австралийских родов трибы Anthocercidae, с которыми он образует подсемейство Nicotianoideae (Olmstead et al. 2008; Clarkson et al., 2004; Särkinen et al., 2013). Внутривидовая классификация *Nicotiana* впервые была предложена в 1838 году. George Don выделил 4 секции: *Tabacum*, *Rustica*, *Petunioides*, и *Polydiclia*, опираясь на размеры и цвет венчика (Knapp et al., 2004). Позже, дополнив морфологические характеристики цветка данными о географическом распространении растений и их хромосомном наборе, Goodspeed выделил в роду *Nicotiana* три подрода (*Tabacum*, *Rustica*, *Petunioides*) и 14 секций (Goodspeed, 1954). На сегодняшний день количество подродов по-прежнему равно трем, а вот состав и количество секций были пересмотрены в соответствии с результатами молекулярно-филогенетического анализа (см. рис.1, Приложение 1) (Knapp et al., 2004; Clarkson et al., 2004).

Сегодня в роду выделяют 13 секций, самая крупная (по количеству видов) - *Suaveolentes*, в которую входят представители Австралии и Африки. Более половины всех видов рода произрастает на территории Южной Америки, которая, таким образом, является очагом наибольшего видового разнообразия. Два вида – *N. tabacum* и *N. rustica* – являются сельскохозяйственными техническими культурами и возделываются повсеместно (Knapp, 2020). Впервые виды *Nicotiana* были найдены в 1753 году Карлом Линнеем в тропиках Америки. Это были *N. glutinosa*, *N. paniculata*,

N. rustica, и *N. tabacum*. Именно Южную Америку позже Goodspeed называл предполагаемым местом происхождения рода, приводя в качестве доказательств морфологические и цитогенетические данные различных видов *Nicotiana* (Goodspeed, 1954). Современные очаги произрастания, по теории Goodspeed, были сформированы постепенным расселением видов.

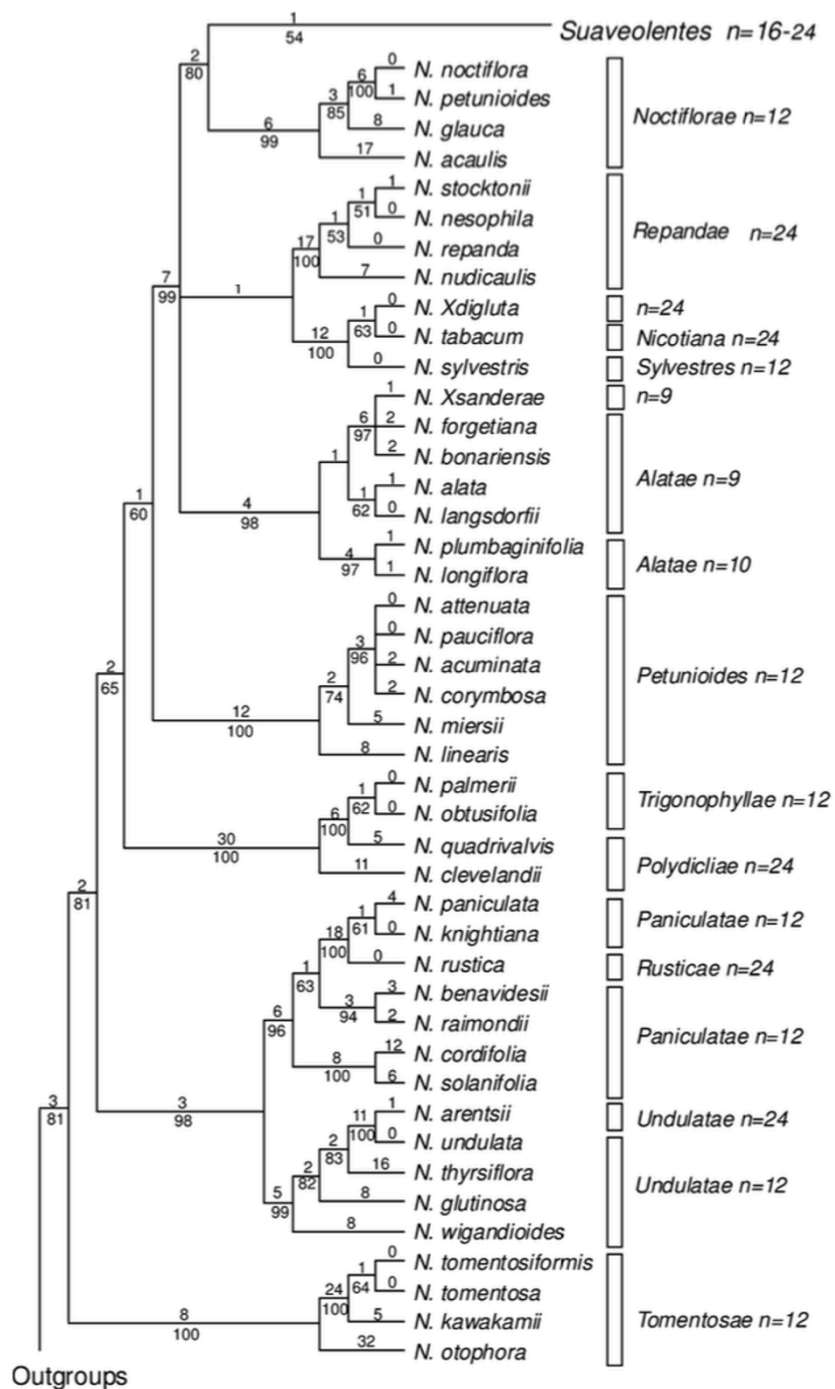


Рис. 1. Филогенетическое древо рода *Nicotiana*, построенное на основании анализа пластидных последовательностей *matK*, *trnL-F*, *trnS-G*, *ndhF* (Clarkson et al., 2004). Справа обозначены секции по Кларк и др., 2004 и хромосомные наборы видов.

Также была высказана идея, что австралийские виды появились вследствие одиночной колонизации, поскольку анализ сайтов рестрикции пластидной ДНК показал небольшие различия в паттерне данных видов (Olmstead and Palmer, 1992). Единственный найденный в Африке вид табака, *N. africana* (Merxmüller and Buttler, 1975), определили как сестринский к австралийской секции *Suaveolentes* (Chase et al., 2003; Olmstead and Palmer, 1992). Таким образом, в настоящее время представители рода *Nicotiana* обитают на территориях Южной и Северной Америки, Австралии и Африки (Chase et al., 2003). Зарождение предковых форм *Nicotiana* происходило на территории древнего суперконтинента Гондвана, предположительно в том регионе, который соответствует современной Южной Америке (Crisp et al., 2009). Раскол суперконтинента привел к отделению Африки, Аравии и Южной Америки (Западная Гондвана) от Австралии, Антарктиды, Мадагаскара и Индостана (Восточная Гондвана) (Jokat et al., 2003). Виды, попавшие на территорию Восточной Гондваны, положили начало австралийской секции *Suaveolentes* (Olmstead et al. 2008). Эволюцию и распространение видов данной секции напрямую связывают с колебаниями климата. Ранние виды сформировались до начала аридизации Австралии в среднем миоцене (10–7 млн лет назад) по периферии континента, во влажных местообитаниях. Далее эволюция секции *Suaveolentes* шла по пути приспособления растений к более засушливому климату, что способствовало расширению ареалов произрастания вглубь Австралии. По мере того, как виды эволюционировали и заполняли новые экологические ниши в засушливой центральной зоне, происходило сокращение числа хромосом (Byrne et al., 2008). Помимо австралийских видов в состав секции *Suaveolentes* входит *N. africana* - единственный произрастающий в Африке представитель рода *Nicotiana* (Chase et al., 2003). *N. africana* - эндемик, который

встречается только на скалистых обнажениях Центрального плато Намибии (Merxmüller, Butler, 1975). На сегодняшний день не установлено, как происходило расселение по континентам предковых форм, давших начало австралийской и африканской ветвям секции *Suaveolentes*. Рассматривают два варианта: через Атлантический океан в Африку, затем в Австралию, либо путем пересечения Тихого океана по островам с последующим вымиранием (Aoki and Ito, 2000). Восстановление точной демографической истории секции *Suaveolentes* осложнено сроком давности события переноса предковых форм *Nicotiana*, а также локальным вымиранием видов и расширением популяции в засушливой зоне Австралии (Chase et al., 2018). На расселение и дальнейшее видообразование предковых форм *Nicotiana* по территории Западной Гондваны повлияли процессы горообразования в Андах. Поднятие и расширение Андского горного хребта, привело к возникновению географических барьеров, что на разных временных промежутках нашло отражение как в образовании различных видов, так и в появлении внутривидовых различий у представителей, чей ареал пересекается с Андами. Эволюционный сценарий появления морфологических различий в разных популяциях внутри одного вида подробно описан для колокольчиков подсемейства *Lobelioideae* (Lagomarsino et al., 2016) и люпинов рода *Lupinus* (Hughes and Eastwood, 2006), произрастающих в Андах. Эти виды часто заселяют изолированные долины, что позволяет предположить независимое течение процессов эволюции в отдельных изолированных группах (Särkinen et al., 2012). Тем же можно объяснить сильную дифференциацию морфологии цветков *N. glutinosa*, ареал которого проходит через Анды от южного Эквадора до северной Боливии (Goodspeed, 1954). В пределах Нового Света секции *Nicotiana* занимают отчетливые непересекающиеся центры видового разнообразия. Так виды секции *Alatae* произрастают в основном на пастбищах Южного конуса, представители секций *Repandae* и *Polydichiae* встречаются в засушливых областях Северной Америки, многие виды сосредоточены преимущественно в Андах (секции *Tomentosae*,

Paniculatae, и *Undulatae*). Наиболее широкие ареалы отмечены для представителей секций *Noctiflorae* и *Petunioides*, первые встречаются в Аргентине и Чили по обе стороны Андских Кордильер, в том числе на больших высотах, а вторые имеют разрозненное распространение в Южной и Северной Америке (Кнарр, 2020). Предполагают, что ранее ареалы произрастания видов были шире и ареалы различных секций пересекались, поскольку сегодня почти половину рода *Nicotiana* составляют аллополиплоиды, возникшие путем межвидовой гибридизации, в том числе гибридизации видов, относящихся к разным секциям, и даже разным под родам. Так, например, *N. tabacum* — аллотетраплоид, произошедший в результате гибридизации *N. sylvestris* и *N. tomentosiformis*, относящихся к под родам *Petunioides* и *Tabacum* соответственно (Murad et al., 2002). Ранее предполагали, что *N. tabacum* произошел путем межвидовой гибридизации *N. tomentosiformis* и *N. otophora*, относящихся к одному под роду — *Petunioides*, и имеющих равные хромосомные наборы (Kenton et al., 1993; Riechers and Timko, 1999; Ren and Timko, 2001), однако результаты более поздних исследований говорят в пользу *N. sylvestris* (Murad et al., 2002; Moon et al., 2008). Более подробно вопрос установления предковых видов *N. tabacum* будет освещен в разделе 1.5 обзора литературы. Работа по восстановлению родственных связей в роду *Nicotiana* является нетривиальной задачей для исследователей в связи с особенностями видообразования, присущими данному роду. Рассмотрим подробнее вопросы филогении *Nicotiana* и подходы к ее изучению.

1.2. Особенности видообразования представителей в роде *Nicotiana*

Филогенетические отношения внутри *Nicotiana* отражают ретикулярный или сетчатый тип эволюции, когда отдельные ветви возникают посредством частичного

слияния двух предковых ветвей. Таким образом, филогенетическое древо *Nicotiana* может местами напоминать скорее сеть, чем разветвляющееся дерево (см. Рис. 2).

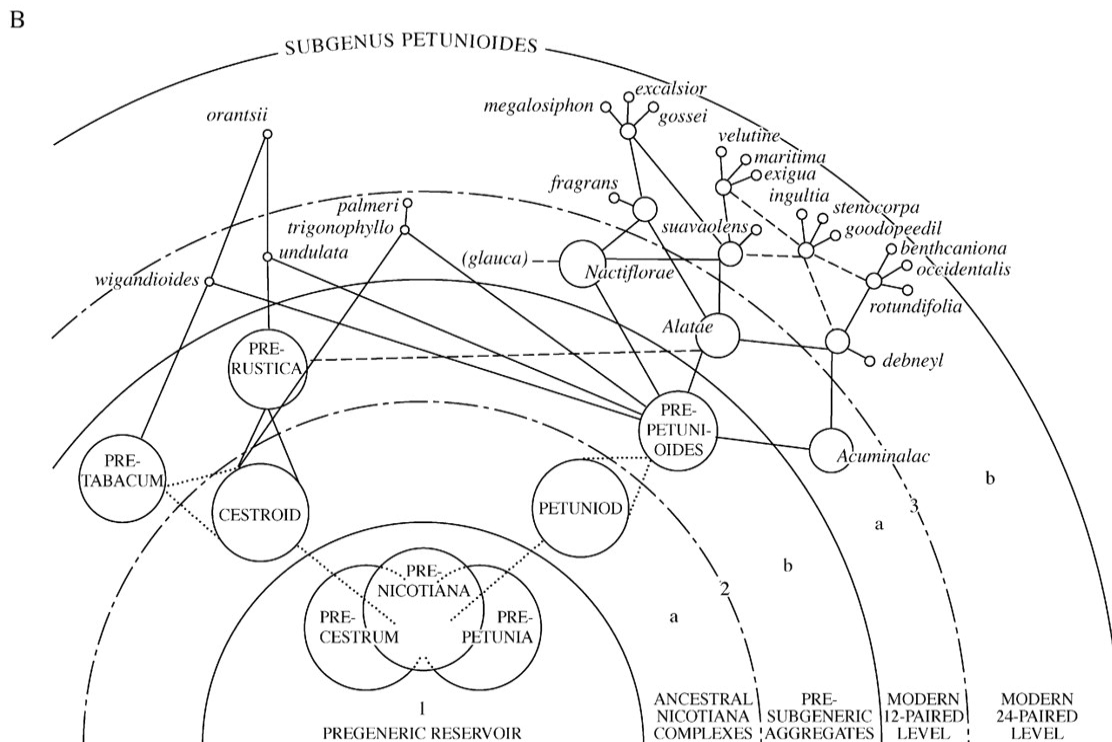
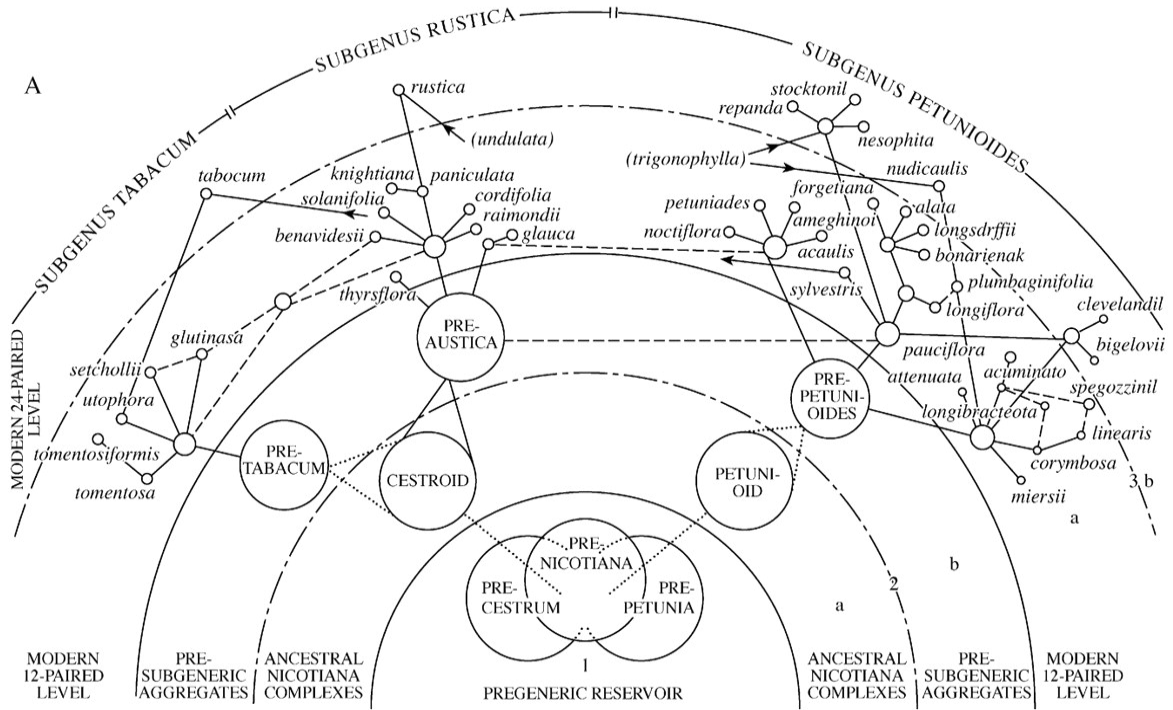


Рис. 2. Филетическая диаграмма рода *Nicotiana*, отражающая сетчатый тип эволюции рода (Goodspeed, 1954); А) подроды *Tabacum*, *Rustica* и *Petunioides*; В) расширенная схема подрода *Petunioides*.

Видовое разнообразие в секциях, содержащих только диплоидные или преимущественно диплоидные виды, находится на относительно низком уровне. Так секции *Alatae*, *Paniculatae* и *Petunioides* насчитывают по восемь видов каждая, в то время как монофилетическая аллополиплоидная секция *Suaveolentes* включает больше трех десятков видов (Knapp et al., 2004). Goodspeed предполагал, что большое количество полиплоидов в Австралии является результатом нескольких различных событий полиплоидии, однако результаты филогенетических исследований, проведенных разными научными коллективами, говорят об одном первичном событии полиплоидии, за которым последовало обширное видообразование (Chase et al., 2018; Clarkson et al. 2017). Основным способом видообразования для *Nicotiana* является межвидовая гибридизация, реализуемая как путем аллополиплоидии (Aoki and Ito, 2000; Chase et al., 2003), так и посредством гомоплоидного гибридного видообразования, при котором происходит частичное слияние и рекомбинация геномов двух родительских видов без увеличения уровня ploidy (Kelly et al., 2010). Современных методов классической цитогенетики бывает недостаточно для обнаружения диплоидных гибридных видов и изучения их геномов, а молекулярные подходы к их анализу только разрабатываются (Posada, 2002; Wuif et al., 2001). Возможно поэтому отношения между секциями с диплоидными видами и внутри них изучены не так хорошо. Для изучения аллополиплоидов с целью установления предковых видов используют различные методы цитогенетики, включая геномную *in situ* гибридизацию (GISH), и молекулярные данные (Kenton et al., 1993; Clarkson et al., 2005; Kelly et al., 2013). Акты аллополиплоидизации случались в роду *Nicotiana* неоднократно в разных секциях с последующей независимо проходящей эволюцией. По времени образования аллополиплоиды можно разделить на молодые (*N. tabacum*,

<600 тысяч лет назад), средние (*N. quadrivalvis*, около 1,2 млн лет назад) и древние (большая часть австралийских видов секции *Suaveolentes*, приблизительно 6 млн лет назад). Чем больше времени прошло с момента межвидовой гибридизации, тем сложнее разделить субгеномы и идентифицировать предковые виды, поскольку происходит накопление изменений в организации генетического материала (Leitch et al., 2008). Так для аллополиплоидов секции *Suaveolentes* ближайшие родственные виды пока не установлены (Clarkson et al. 2017; Chase et al. 2018). Возможность различить субгеномы молодых аллотетраплоидов с помощью GISH была показана для видов *N. tabacum*, *N. rustica* и *N. arentsii*, возраст самого молодого из которых составляет менее 200 тысяч лет (Kenton et al. 1993; Lim et al., 2004). Подобное исследование, проведенное для древних полиплоидов возрастом более 5 млн лет, не позволило оценить вклад каждого из родительских геномов (Clarkson et al., 2005). Lim высказал предположение, что у подобных древних видов успела произойти почти полная перестройка геномного материала с воссозданием консервативной структуры кариотипа, включающая замену некодирующих последовательностей некими новыми последовательностями (Lim et al., 2007). Обычно за полиплоидизацией у покрытосеменных следуют процессы уменьшения размера генома и диплоидизация, так рис *Oryza sativa* L. утратил примерно 80% дублированных генов, кукуруза *Zea mays* L. - более 50% (Leitch and Bennett, 2004; Leitch et al. 2008; Dodsworth et al., 2016). Увеличение количества ДНК после полиплоидизации наблюдается гораздо реже: на сегодняшний день такое явление показано для ячменя (Jakob et al., 2004). Среди представителей рода *Nicotiana* встречаются как «уменьшившие», как и «увеличившие свой геном» виды, причем некоторые из них могут иметь общий предковый геном, и при этом по-разному реагировать на произошедшую амфиплоидизацию. В то же время наблюдается прямая зависимости роста потери количества генетического материала от возраста гибрида (Leitch et al., 2008). Интересно, что редукция генетического материала у представителей рода *Nicotiana*

отмечена не только для полиплоидных видов, но и для диплоидов, например, для представителей секции *Alatae* (Lim et al., 2006). То, что уменьшение числа хромосом, изменение их морфологии и, очевидно, связанная с этим перестройка тандемных повторов в рДНК показаны как для полиплоидов (*N. tabacum*, *N. rustica*) (Kovarik et al., 2004; Skalicka et al., 2005), так и для диплоидных представителей секции *Alatae*, говорит о наличии закономерностей в процессах эволюции геномов *Nicotiana*, не зависящих от их пloidности. Перестройки в генетическом материале, происходившие при видообразовании, могли иметь важное значение для дальнейшей эволюции видов, предоставляя задел для образования новых генов и появления новых признаков. А это в свою очередь могло способствовать расселению на новые территории, повышая шансы на выживание и закрепление нового вида в занятой им экологической нише. В защиту этого предположения можно привести тот факт, что у многих аллополиплоидов венчики цветков в среднем более короткие и широкие, чем у их предковых форм, что может свидетельствовать об их направленной эволюции в сторону расширения круга опылителей (McCarthy et al., 2016).

Относительно эволюции и видообразования в роде *Nicotiana* остается достаточно много вопросов и предположений. Не всегда удается однозначно установить отношения между видами, не для всех аллополиплоидов на сегодня определены предковые формы. Это связано как с процессами, происходящими с геномом в ходе видообразования, так и с разрешающей способностью применяемых методов. Открытие новых молекулярных маркеров позволило бы продвинуться в изучении взаимосвязей внутри рода. В качестве подобного маркера, например, можно использовать клеточную Т-ДНК. Это последовательности, некогда полученные растениями в результате агробактериальной трансформации, и закрепившиеся в геноме растения. Клеточная Т-ДНК присутствует в геномах 16-ти видов *Nicotiana*, принадлежащих к разным секциям и подродам. Анализ последовательностей, входящих в состав клТ-ДНК представителей *Nicotiana*, показал более высокий

уровень сходства клТ-ДНК видов, относящихся к одному подроду (Intrieri, Buiatti, 2001), что наводит на мысль о возможной взаимосвязи эволюционных процессов в роду *Nicotiana* и событий приобретения клТ-ДНК представителями данного рода. Видам *Nicotiana*, в геномах которых была обнаружена клТ-ДНК, а также деталям ее строения посвящена следующая глава обзора.

1.3. Природно-трансгенные виды в роду *Nicotiana*

Впервые клеточная Т-ДНК была обнаружена в геноме табачного дерева, *N. glauca* (White et al., 1983). При проверке результатов трансформации растений *N. glauca* бактериями штамма *A. rhizogenes* А4 методом Саузерн-блот гибридизации позитивный сигнал был обнаружен как в ДНК трансформированных растений, так и в контрольных. Дальнейший анализ подтвердил, что в геноме нетрансформированных растений *N. glauca* содержатся последовательности, гомологичные Т-ДНК, которые White предложил назвать клеточной Т-ДНК (White et al., 1983). КлТ-ДНК *N. glauca* представляет собой несовершенный инвертированный повтор (так называемые правое и левое плечи) суммарной длиной порядка 10 kb, содержащий открытые рамки считывания *orf11*, *orf12*, *orf13*, *orf13a*, и *orf14* (Рис. 3). Различными научными группами (Aoki et al., 1994; Furner et al., 1986) были получены сиквенсы фрагментов клТ-ДНК *N. glauca*. В 2002 году Suzuki досеквенировал недостающие участки и методом локального выравнивания объединил все сиквенсы в единую последовательность. Это позволило более подробно изучить данную клТ-ДНК, позже названную gТ, оценить сходство между агробактериальными генами *rolB* и *rolC*, принадлежащими различным штаммам, и последовательностями *orf11* и *orf12* соответственно. Между *rolB* и *orf11* сходство на нуклеотидном уровне составило 82,7%, между *rolC* и *orf12* - 87,5%. Также между *orf14* и границей Т-ДНК были найдены еще две открытые рамки считывания, названные *NgmisL* и *NgmisR*,

гомологичные гену микимопинсинтазы *mis* (Suzuki et al., 2002). Этот ген был обнаружен годом ранее Suzuki и соавторами на плазмиде pRi1724 в агробактериальном штамме *A. rhizogenes* MAFF301724 (Suzuki et al., 2001).

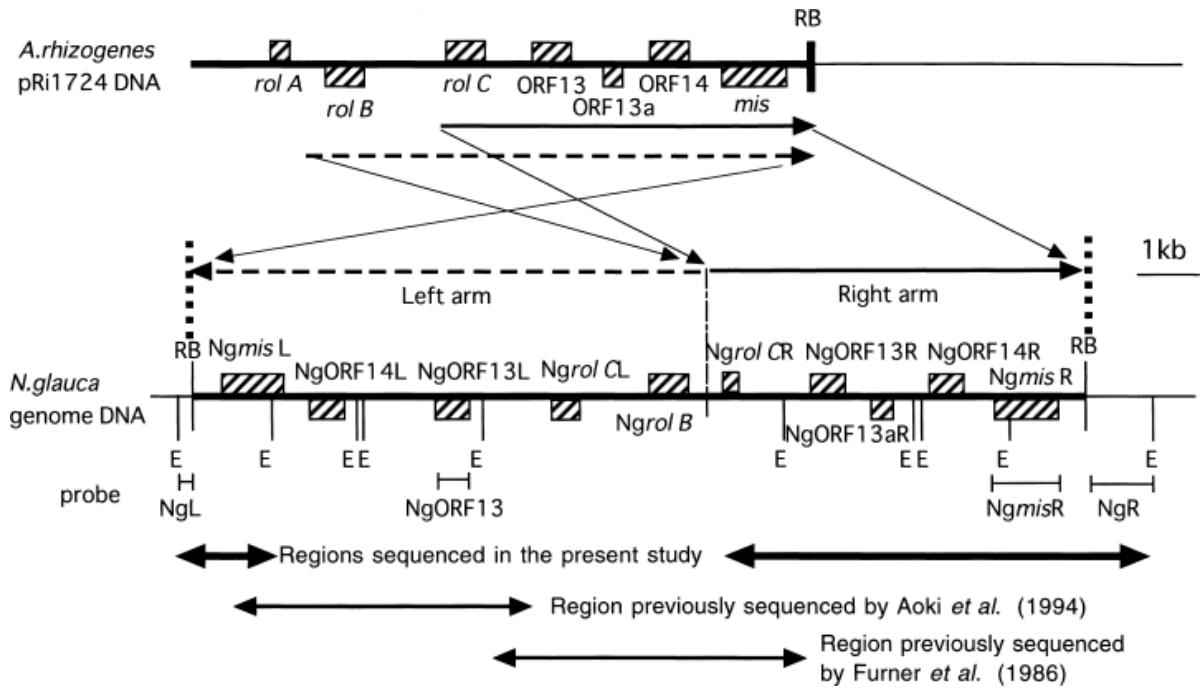


Рис. 3. Структура клТ-ДНК в геноме *N. glauca* (Suzuki et al., 2002); двунаправленными стрелками отмечены фрагменты, секвенированные различными научными группами, ссылки на работы размещены в разделе «список использованной литературы».

Длина левого плеча клТ-ДНК, содержащего гомологи *rolB* и *rolC*, составляет 7968 п.н., в то время как правое плечо длиной 5778 п.н. не содержит гомолог *rolB*. Уровень сходства нуклеотидной последовательности между левым и правым плечами превышает 97,9%, а порядок генов сохраняется, что указывает на происхождение данного повтора в результате дупликации (Suzuki et al., 2002). Путем сравнения полученных сиквенсов правого и левого плеча клТ-ДНК *N. glauca* с Т-ДНК четырех разных Ri плазмид: Ri1724 микимопинового типа, Ri2659 кукумопинового типа, RiA4b агропинового типа и Ri8196 маннопинового типа, было доказано, что гТ

является интегрированной в растительный геном Т-ДНК Ri плазмиды микимопинового типа, схожей с Ri1724 (Suzuki et al., 2002).

Пытаясь обнаружить возможную связь закрепления Т-ДНК в растениях с географическим фактором, Furner собрал образцы *N. glauca* в Перу, Боливии и Парагвае, и во всех исследованных популяциях идентифицировал клТ-ДНК в растительных геномах (Furner et al., 1986). Присутствие гТ во всех проанализированных растениях позволило предположить, что клТ-ДНК попала в геном *N. glauca* до образования этого вида, а значит, велика вероятность, что и другие родственные виды могут содержать кТ-ДНК. С целью обнаружения подобных видов был проведен скрининг представителей рода *Nicotiana* методом ПЦР с праймерами, специфичными для отдельных генов, входящих в состав гТ. В результате природно-трансгенные виды были обнаружены в различных секциях и подродах, при этом в ранних исследованиях больше всего их было отмечено в составе секций подрода *Tabacum*, в то время как *N. glauca* относится к секции *Noctiflorae* подрода *Petunioides* (Furner et al., 1986, Suzuki et al., 2002, Intrieri, Buiatti, 2001). Параллельно с поиском новых видов, содержащих клТ-ДНК, изучали и сравнивали состав уже обнаруженных вставок в различных видах. Ранние исследования различных представителей рода *Nicotiana* показали, что только у двух (*N. glauca*, *N. cordifolia*) сохранился набор онкогенов, аналогичный таковому в Ri плазмиде. У *N. tabacum*, *N. tomentosiformis*, и *N. otophora* отсутствует гомолог гена *rolB*, и ещё у ряда представителей обнаружен только гомолог *rolC* (*N. debneyi*, *N. suaveolentes*, *N. acuminata*, *N. arentsi*, *N. gossei*, *N. exigua*, *N. setchelli*), либо только *rolB*, изначально ошибочно принятый за *rolC* (*N. bigelovi*, *N. miersi*) (Intrieri, Buiatti, 2001). Однако результаты более поздних исследований, проведенные с применением полногеномного секвенирования, опровергают часть данных, полученных ранее методом ПЦР и секвенирования по Сенгеру. Например, показано, что гомолог гена *rolB* присутствует в двух из четырех клТ-ДНК *N. tomentosiformis*, а также в клТ-ДНК *N. otophora* и *N. tabacum* (сорта К326,

TN90, *Vasma/Xanthi*), который, видимо, получил его от *N. tomentosiformis* как от одного из предковых видов (Chen et al., 2014; Chen et al., 2018). Секвенирование транскриптомов ряда видов *Nicotiana* также привело к пересмотру некоторых результатов, полученных ранее методом ПЦР. Так Intrieri и Buiatti отнесли *N. noctiflora* к видам, не содержащим клТ-ДНК в геноме, однако спустя 15 лет в последовательностях транскриптома *N. noctiflora* были найдены фрагменты гомологов генов *orf13*, *orf14*, *RolC*, *iaaH*, *iaaM*, *mis* (Intrieri, Buiatti, 2001; Long et al., 2016). Наличие данных секвенирования нового поколения (NGS) позволяет не только уточнить состав вставок, но также определить их копийность и обнаружить сайт встраивания Т-ДНК в растительный геном. Так в результате анализа полногеномных данных *N. tomentosiformis* были выявлены четыре различающиеся между собой по составу клТ-ДНК (ТА, ТВ, ТС, TD), каждая из которых отличалась от ранее описанной гТ в геноме *N. glauca* (рис. 3–7). Последовательность ТА соответствует фрагменту Т-ДНК штамма А4 *A. rhizogenes*, и представляет собой инвертированный повтор, состоящий из плеч ТА-1 и ТА-2, различие между которыми составляет всего 1,2% (Рис. 3). Между ТА-1 и ТА-2 находится богатый повторами участок.

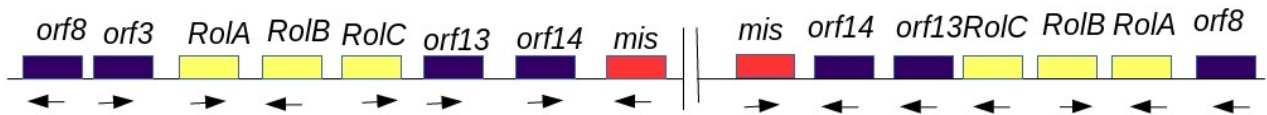


Рис. 4. Схема последовательности клТ-ДНК (ТА) в геноме *N. tomentosiformis* (Chen et al., 2014). Тонкими стрелками показано направление ОРС. Здесь и далее размеры гомологов генов даны не в масштабе.

ТВ также представляет собой несовершенный инвертированный повтор и состоит из двух различающихся плеч — ТВ-1 и ТВ-2 (Рис. 5). Плечо ТВ-1 содержит гомологи генов *orf14* и *mis*, а также участок длиной 122 п.н. на 86% совпадающий с геном *iaaM* штамма А4. Между последовательностями ТВ-1 и ТВ-2 наблюдается высокий

уровень сходства (97,4%), однако в дополнение к генам *orf14* и *mis* в правой части ТВ-2 (более длинной, чем у ТВ-1) находится *ags*-подобный ген синтеза агропинов, совпадающий с *ags*-подобным геном *Nectria haematococca*, и на 65% совпадающий с участком последовательности *A. tumefaciens*, а также гены *mas1'* и *mas2'*, отвечающие за синтез маннопинов, последовательности которых на 85% совпали с таковыми у штамма *A. rhizogenes* 8196. Есть предположение (Chen et al., 2014), что данный маннопино-агропиновый регион, как и фрагмент *iaaM*, был получен растением с другой Т-ДНК (относительно последовательностей генов *orf14*, *mis*).

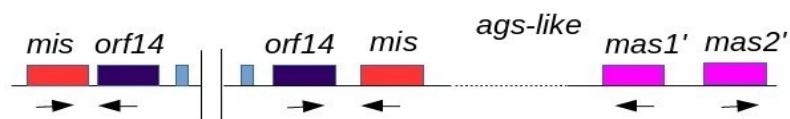


Рис. 5. Схема последовательности клТ-ДНК (ТВ) в геноме *N. tomentosiformis* (Chen et al., 2014)

КлТ-ДНК ТС включает как области, гомологичные участкам штамма *A. rhizogenes* А4, так и последовательности, для которых была показана разной степени гомология с генами грибов базидиомицетов, плесневых грибов, эндофитного гриба и олиготрофных бактерий (Рис. 6). Гомология с последовательностью штамма А4 прослеживается от *rolB* до *orf2* в обоих плечах ТС. Последовательность на левом конце ТС-1 транслируется в белок, подобный октопин/нопалиндегидрогеназе, который демонстрирует сходство на 40% с последовательностью белка *Aspergillus nidulans*, на 34% – с *Melampsora larici-populina* и на 28% – с *Bradyrhizobium oligotrophicum*. Поиск наиболее близкой последовательности в представителях рода *Agrobacterium* указал на октопинсинтазу штаммов *Agrobacterium tumefaciens* А6 и *Agrobacterium vitis* Тм4 (сходство на уровне аминокислот для обоих штаммов составляет 25%). На конце плеча ТС-2 расположена последовательность, транслирующая белок, на 42% схожий с последовательностью белка С штамма *A. tumefaciens* С58 и на 41% – с белком *A. vitis* Тм4. Плечи ТС-1 и ТС-2 различаются

между собой на 5,8%, при этом вся последовательность ТС на 68–75% схожа с соответствующей последовательностью ТА, и на 69–77% - с таковой последовательностью в штамме А4.

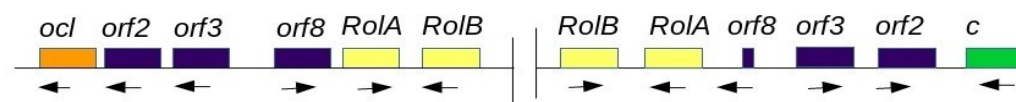


Рис. 6. Схема последовательности клТ-ДНК (ТС) в геноме *N. tomentosiformis* (Chen et al., 2014)

TD также является несовершенным инвертированным повтором и состоит из двух плеч - TD-1 и TD-2. Оба плеча содержат гомологи гена *orf14*, при этом последовательность TD-Orf14 сильно отличается от ТА-Orf14 и ТВ-Orf14. Аминокислотная последовательность TD-Orf14 на 36% совпадает с последовательностью белка Orf14 штамма *A. rhizogenes* 8196, в то время как для нуклеотидной последовательности *orf15* в плече TD-2 наибольшее сходство показано со штаммом *A. rhizogenes* А4. В плече TD-2 также расположен гомолог *orf511*, усеченная версия которого, *orf455*, находится в более коротком плече TD-1 (Рис. 7).

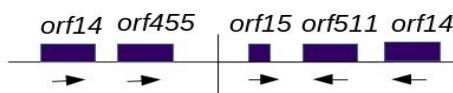


Рис. 7. Схема последовательности клТ-ДНК (TD) в геноме *N. tomentosiformis* (Chen et al., 2014)

Для вида *N. otophora*, который относится к той же секции, что и *N. tomentosiformis*, также были получены полногеномные данные, анализ которых позволил установить, что геном *N. otophora* содержит две копии ТС (рис. 6), ранее описанных в *N. tomentosiformis*, и два варианта клТ-ДНК ТЕ, уникальной для *N. otophora* (рис. 8) (Sierro et al., 2014; Chen et al., 2014). ТЕ1 и ТЕ2 различаются в основном делецией

длиной примерно 12 т.п.н. в центральной части. ТЕ представляет собой сложно организованный повтор, для разных частей которого показана гомология с различными штаммами агробактерий. Так ближайшим гомологом ТЕ-6b является Т-6b из *A. vitis* Тm4, в то время как последовательность повтора от *mas1'* до *rolC* на 78% схожа с правой частью Т-ДНК штамма *A. rhizogenes* 8196. По своей организации ТЕ также может быть разделена на части: *6b-6b* и *mas1'-mas1'* (отделены вертикальной пунктирной чертой на рис. 8). Повтор *6b-6b* похожа на Т-ДНК *A. tumefaciens* и *A. vitis* типа LB-*acs-5-iaaH-iaaM-ipt-6a-6b-ocs/vis*-RB, в которой произошла перестройка с утратой генов *5*, *ipt* и *6a*. Участок *mas1'-mas1'* состоит из другого инвертированного повтора, подобного правой части Т-ДНК штамма *A. rhizogenes* 8196, - *rolB-rolC-orf13-orf13a-orf14-mas2'-mas1'*, - утратившей *rolB* (Chen et al., 2018). Chen и соавторы предполагают, что подобная клТ-ДНК произошла вследствие слияния фрагментов Т-ДНК двух разных штаммов агробактерий. Подобная картина показана для ТВ в геноме *N. tomentosiformis*.

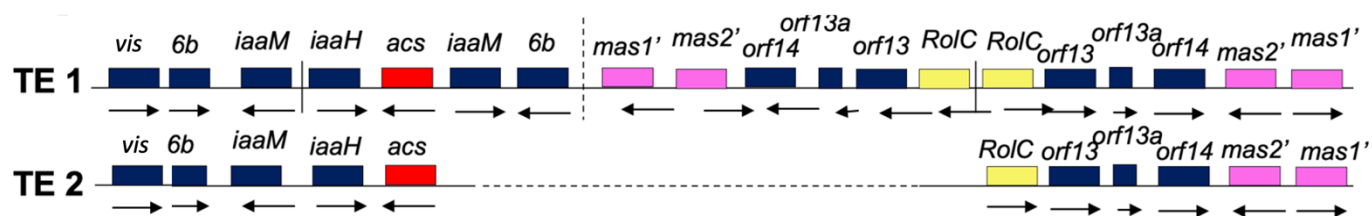


Рис. 8. Схема последовательности клТ-ДНК (ТЕ) в геноме *N. otophora* (Chen et al., 2018)

В геноме *N. tabacum*, произошедшего путем межвидовой гибридизации *N. tomentosiformis* и *N. sylvestris*, было найдено три из четырех клТ-ДНК предкового вида (ТА, ТВ, ТD), что говорит о том, что вид *N. tabacum* образовался уже после того, как геном *N. tomentosiformis* был трансформирован агробактериями. Также для *N. tabacum* было показано отсутствие ТС (Chen et al., 2014). При этом ТС присутствует в геноме вида *N. otophora*, который относится к той же секции, что и *N.*

tomentosiformis. Сайт интеграции ТС в растительный геном у *N. otophora* и *N. tomentosiformis* совпадает, исходя из чего предполагают, что ТС попала в геном предковой формы до расхождения *N. tomentosiformis* и *N. otophora* (рис. 9) (Chen et al., 2018).

В настоящее время вопрос попадания и распространения клТ-ДНК в геномах *Nicotiana* остается открытым. Существует несколько гипотез, описывающих возможные варианты развития событий.

1. Множественные акты трансформации представителей рода.
2. Попадание Т-ДНК в геном одного из видов с последующим случайным распространением в пределах рода за счет гибридизации между видами.
3. Однократное встраивание Т-ДНК в предковую форму на ранних этапах эволюции рода, а затем потеря вставки у различных представителей рода *Nicotiana*.

В пользу первой гипотезы говорит присутствие в геномах *N. tomentosiformis* и *N. otophora* нескольких различных по своему составу клТ-ДНК. Мозаичное строение некоторых клТ-ДНК, то есть различное происхождение участков также может говорить о том, что трансформация растения агробактериями происходила неоднократно в несколько этапов. Еще одним доводом в пользу гипотезы о множественных актах является то, что у видов *N. tabacum* и *N. glauca*, относящихся к разным эволюционным ветвям рода *Nicotiana*, различаются сайты локализации клТ-ДНК в геномах. Это было ранее показано с использованием тест-системы, основанной на ПЦР, разработанной нашей научной группой на прошлых этапах работы (Хафизова и др., 2014). В поддержку второй гипотезы выступает тот факт, что для рода *Nicotiana* одним из основных способов видообразования является межвидовая гибридизация (Leitch et al., 2008). Возможность передачи клТ-ДНК в

ходе межвидовой гибридизации доказана на примере вида *N. tabacum*. Согласно третьей гипотезе, клТ-ДНК была получена в результате однократного акта агробактериальной трансформации, после чего некоторые виды эту вставку утратили. Примером элиминации клТ-ДНК из генома является ТС. Chen и коллеги предположили, что после попадания в растительный геном предковой формы *Nicotiana*, из которой позже будут образованы виды *N. tomentosiformis* и *N. otophora*, произошла ее дупликация. Позже в *N. tomentosiformis* происходит элиминация одной копии ТС, а затем оставшуюся копию теряет *N. tabacum* (рис. 9) (Chen et al., 2018).

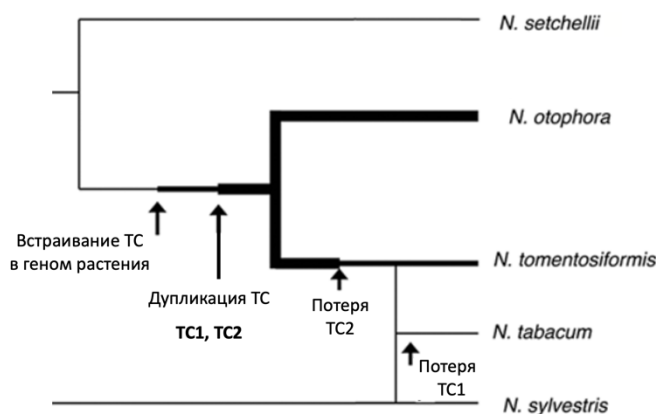


Рис. 9. Модель эволюции ТС в геномах *Nicotiana*. Ширина линий отражает присутствие в геноме вида 1, 2 копий ТС, или ее отсутствие (Chen et al., 2018).

Сведения о составе клТ-ДНК и о сайтах ее локализации в растительных геномах, на основании которых возможно выстраивать и подтверждать подобные гипотезы, получены в основном для видов, принадлежащих к подроду *Tabacum*, для которых имеются полногеномные данные. Таким образом, только для этой эволюционной ветви рода *Nicotiana* могут быть построены модели распространения клТ-ДНК. Для вида *N. glauca*, принадлежащего к подроду *Petunioides*, известна структура клТ-ДНК, которая, однако не была подтверждена молекулярными методами. Остальные 13 видов, в которых найдены последовательности, гомологичные агробактериальным,

изучали только с помощью ПЦР, в связи с чем для них имеются только данные по отдельным генам. Данные о наличии и составе Т-ДНК у различных видов *Nicotiana* обобщены в таблице 1.

Таблица 1. Т-ДНК в различных видах *Nicotiana*

Подрод	Секция согласно Knapp et al., 2004	Виды	Гены Т-ДНК	Acc. number	Метод исследования	Ссылка на источник
<i>Rustica</i>	<i>Paniculatae</i>	<i>N.paniculata</i>	-		ПЦР	2, 5
		<i>N.knightiana</i>	-		ПЦР, секвенирование транскриптома	2, 5, 7
		<i>N.solanifolia</i>	-		ПЦР	5
		<i>N.benavidesii</i>	+(<i>rolC</i>)	нет	ПЦР	2, 5
		<i>N.cordifolia</i>	+(<i>rolB-ORF14</i>)	AF281252.1; AF281244.1; AF281248.1	ПЦР, секвенирование транскриптома	2, 5, 7
	<i>N.raimondi</i>	-		ПЦР	5	
	<i>Rusticae</i>	<i>N.rustica</i>	-		ПЦР	2, 5
<i>Tabacum</i>	<i>Tomentosae</i>	<i>N.tomentosa</i>	+(<i>orf3-mis, ags-mas1'-mas2', orf2, ocl, c, orf15, orf511</i>)	нет	ПЦР	2, 4, 6
		<i>N.tomentosiformis</i>	+(<i>orf3-mis, ags-mas1'-mas2', orf2,</i>	AF281249.1; AF281241.1; AF281245.1; SRX1100400	Полногеномное секвенирование, секвенирование транскриптома	2, 4, 5, 6, 7

			<i>ocl, c, orf15, orf511)</i>			
		<i>N.otophora</i>	<i>+(orf2-rolB, rolC-ORF14, mas1, mas2, iaaM, iaaH, acs, vis)</i>	AF281250.1; AF281243.1; AF281247.1	Полногеномное секвенирование	2, 5, 6
		<i>N. kawakamii</i>	<i>(orf3-mis, ags-mas1'-mas2', orf2, ocl, c, orf15, orf511)</i>	нет	ПЦР	6
		<i>N.setchelli</i>	<i>+(rolC)</i>	Нет	ПЦР, секвенирование транскриптома	2, 7
	<i>Nicotiana</i>	<i>N.tabacum</i>	<i>+(orf3-mis, ags-mas1'-mas2', c, orf15, orf511)</i>	AF281242.1; AF281246.1	Полногеномное секвенирование	2, 4, 5, 6
<i>Petunioides</i>	<i>Undulatae</i>	<i>N.glutinosa</i>	-		ПЦР	2, 5
		<i>N.undulata</i>	-		ПЦР	5
		<i>N.arentsii</i>	<i>+(rolC)</i>	нет	ПЦР	5
	<i>Trigonophyllae</i>	<i>N.trigonophylla</i>	-		ПЦР	5
	<i>Sylvestris</i>	<i>N.sylvestris</i>	-		ПЦР	5
	<i>Alatae</i>	<i>N.langsdorffi</i>	-		ПЦР	2, 5
		<i>N.alata</i>	-		ПЦР	5
		<i>N.longiflora</i>	-		ПЦР	5
		<i>N.forgetiana</i>	-		ПЦР	5
		<i>N.sanderiae</i>	-		ПЦР	5

		<i>N.plumbaginifolia</i>	-		ПЦР	5
<i>Repandae</i>		<i>N.nesophila</i>	-		ПЦР	5
		<i>N.stocktonii</i>	-		ПЦР	5
		<i>N.repanda</i>	-		ПЦР	5
		<i>N.nudicaulis</i>	-		ПЦР	5
<i>Noctiflorae</i>		<i>N.noctiflora</i>	+ (<i>rolC</i> , <i>orf14</i> , <i>orf13</i> , <i>iaaM</i> , <i>iaaH</i> , <i>mis</i>)?	SRX1100384	ПЦР, секвенирование транскриптома	5, 7
		<i>N.glauca</i>	+(<i>rolB-mis</i>)	X03432.1; D16559.1; AB071334.1; AB071335.1; SRX110016	ПЦР, секвенирование транскриптома	1, 2, 3, 4, 7
<i>Petunioides</i>		<i>N.acuminata</i>	(<i>rolC</i>)	нет	ПЦР	5
		<i>N.pauciflora</i>	-		ПЦР	5
		<i>N.attenuata</i>	-		ПЦР	5
		<i>N.miersi</i>	+(<i>rolB</i>)	нет	ПЦР	5
<i>Bigelovianae</i>		<i>N.bigelovi</i>	+(<i>rolB</i>)	нет	ПЦР	5
<i>Polydiclae</i>		<i>N.clevelandi</i>	-		ПЦР	5
<i>Suaveolentes</i>		<i>N.umbratica</i>	-		ПЦР	5
		<i>N.gossei</i>	+(<i>rolC</i>)	Нет	ПЦР	5
		<i>N.rotundifolia</i>	-		ПЦР	5
		<i>N.suaveolens</i>	+(<i>rolC</i>)	Нет	ПЦР	5
		<i>N.exigua</i>	+(<i>rolC</i>)	Нет	ПЦР	5
		<i>N.goodspeedii</i>	-		ПЦР	5
1 - White et al., 1983; 2 - Furner et al., 1986; 3 - Aoki et al., 1994; 4 - Suzuki et al., 2002; 5 - Intriери and Buiatti, 2001; 6 — Chen et al., 2014; 7 - Long et al., 2016						

Для большинства представителей рода *Nicotiana* анализ клТ-ДНК был проведен на уровне отдельных генов методом ПЦР. Однако уже показано, что данный метод может давать ошибочные результаты. Вероятно, это происходит вследствие накопления мутаций, в том числе делеций участков генов, что препятствует отжигу праймеров на матрице ДНК и ведет к получению ложноотрицательных результатов, как было показано на примере *N. noctiflora*. *N. noctiflora* как и *N. glauca* относится к подроду *Petunioides*, и про клТ-ДНК в данных видах известно гораздо меньше, чем про клТ-ДНК подрода *Tabacum*. Более подробное их изучение 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*.

Изучение структуры клТ-ДНК отдельно взятого вида ограничено фактором низкого уровня внутривидового разнообразия, характерного для большинства представителей рода *Nicotiana*. Оценка степени внутривидового полиформизма клТ-ДНК, проведенная на нескольких растениях, может не достоверно отражать картину, характерную для всего вида. Вид *Nicotiana tabacum* культивируют повсеместно, для него получено множество сортов, при этом он содержит клТ-ДНК. Таким образом, на примере различных сортов *N. tabacum* может быть проведен анализ внутривидовой изменчивости клТ-ДНК.

1.4. Ботаническое описание и характеристика генома вида *N. tabacum*

Nicotiana tabacum L. — культурный табак — это однолетнее травянистое растение высотой от 0,2 до 3 метров в зависимости от сорта, с крупными зелеными листьями. Форма листовой пластинки также зависит от сорта и варьирует от широко-яйцевидной до продолговато-ланцетовидной, лист цельный, получерешковый, листорасположение очерёдное. Стебель прямостоячий. Цветки трубчатые или

воронковидные, актиноморфные, пятилопастные, собраны в метельчатые кисти. Венчик спайнолепестный длиной 5—6 см, лепестки бело-розоватые, розовые или красноватые. Плод - сухая коробочка 1,5–1,8 см длиной, немного длиннее чашечки. Все надземные вегетативные органы *N. tabacum* покрыты короткими липко-железистыми волосками, выделяющими желтый секрет, содержащий различные алкалоиды, в том числе никотин (Kishore, 2014).



Рис. 10. *N. tabacum* L., фото из коллекции Королевских ботанических садов Кью (Royal Botanic Gardens, Kew), Великобритания

Местом образования данного вида считают Южную Америку (Goodspid, 1954). В настоящее время *N. tabacum* выращивают повсеместно, при этом считается, что он встречается только в культуре, поскольку популяции дикого *N. tabacum* не были обнаружены (Lewis and Nicholson, 2007). Основным назначением возделывания культурного табака в коммерческих целях является изготовление сырья для

табакокурения, также некоторые сорта выращивают в качестве декоративных растений (Kishore, 2014). Кроме того, *N. tabacum* является классическим объектом генетики растений, геной инженерии и биотехнологии растений, а также источником одной из наиболее часто используемых клеточных линий BY-2 (Nagata et al., 2004; Edwards et al., 2017). Несмотря на научную значимость *N. tabacum* и ценность его как экономически важной культуры, до недавнего времени генетические ресурсы, доступные для анализа генома, генетического картирования и селекции с помощью молекулярных маркеров, были относительно ограничены, по крайней мере, по сравнению с другими экономически важными культурами (Davis and Nielsen, 1999). Большой размер генома (4,5 млрд.п.н.) (Leitch et al., 2008), а также его насыщенность повторами (>70%) (Zimmerman and Goldberg, 1977; Renny-Byfield et al., 2011) значительно осложнили процесс получения полногеномных данных и непосредственно сборку, в связи с чем первый референсный геном *N. tabacum* был опубликован только в 2014 году. Анализ геномов трех сортов *N. tabacum* позволил сопоставить составленные ранее физические и генетические карты и наложить их на последовательность генома, идентифицировать и охарактеризовать пути биосинтеза различных алкалоидов, а также прояснить вопрос происхождения вида, долгое время служивший предметом для споров (Sierra et al., 2014).

Nicotiana tabacum - классический амфидиплоид ($2n = 48$), который возник в результате межвидовой гибридизации предковых форм видов *Nicotiana sylvestris* Spegazzini and Comes ($2n = 24$) и *Nicotiana tomentosiformis* Goodspeed ($2n = 24$) предположительно на восточных склонах Анд в Боливии или северной Аргентине (Goodspeed 1954; Gerstel and Sisson, 1995). Геном *Nicotiana tabacum* содержит два цитологически различных родительских субгенома, S-геном и T-геном. Молекулярные и биохимические данные свидетельствуют о том, что источником S-генома, а также цитоплазмы, послужил вид *N. sylvestris*, который, соответственно, считается материнским родительским видом (Bland et al., 1985; Olmstead and Palmer,

1992; Yukawa et al., 2006). Возможным источником Т-генома долгое время называли как вид *N. tomentosiformis* (Bogani et al., 1997; Volkov et al., 1999; Lim et al., 2000), так и вид *N. otophora* Grisebach ($2n = 24$) (Kenton et al., 1993; Riechers and Timko, 1999), относящиеся к одной секции, *Tomentosae*, а также их возможную гибридную форму (Kenton et al., 1993; Parokonny and Kenton, 1995). Результаты поиска отцовского родительского вида, полученные при использовании различных методов, не согласовывались между собой. Так, анализ хромосом *N. tomentosiformis*, *N. otophora* и *N. tabacum* методом *in situ* гибридизации (Kenton et al., 1993) а также структурный анализ и анализ экспрессии генов синтеза вторичных метаболитов (*PMT*) (Riechers and Timko, 1999) указывали на сродство Т-генома с геномом *N. otophora*. В пользу *N. otophora* выступал также географический критерий - показано, что современные ареалы произрастания видов *N. otophora* и *N. sylvestris* пересекаются, в то время как *N. tomentosiformis* произрастает значительно северней, чем *N. sylvestris* (Gray et al., 1974). Однако стоит учитывать, что за несколько сотен лет с момента видообразования культурного табака (Okamoto and Goldberg, 1985) могли произойти изменения в климате и микрорельефе, что могло повлиять на современную картину. В поддержку гипотезы об участии *N. tomentosiformis* в происхождении *N. tabacum* говорят результаты анализа рибосомных повторов (Volkov et al., 1999) и RAPD (случайная амплификация полиморфной ДНК) (Yi, 1997). Анализ тандемных повторов в Т-геноме *N. tabacum* также предоставил доводы в пользу *N. tomentosiformis* как источника Т-генома и, кроме того, позволил установить наиболее близкую к *N. tabacum* линию *N. tomentosiformis*, - cv. NIC 479/84 (Murad et al., 2002). Полученные позже полногеномные данные подтвердили гипотезу о происхождении *N. tabacum* в результате гибридизации *N. tomentosiformis* и *N. sylvestris* (Sierro et al., 2014). Генетический материал *N. tabacum*, полученный от родительских видов, в ходе эволюции претерпел изменения. Современные формы культурного табака содержат 20 S хромосом, 18 Т хромосом и 10 гибридных (Kenton et al., 1993; Parokonny and

Kenton, 1995). Гибридные хромосомы состоят из фрагментов T и S хромосом, примечательно, что у *N. tabacum* их количество значительно выше, чем у *N. rustica* и *N. arentsii*, которые также являются межвидовыми гибридами (Lim et al., 2004). Возможно, это связано со значительным филогенетическим расстоянием между предковыми видами *N. tabacum*. В таком случае большое количество межгеномных перестроек можно объяснить «геномным стрессом» аллополиплоида, возникающим вследствие различий в геномах предковых форм, и последующей активацией транспозонов и ретротранспозонов, возникновению разрывов ДНК и появлению межгеномных транслокаций (McClintock, 1984; Melayah et al., 2004). Предковые виды *N. rustica* и *N. arentsii* относятся к одной секции (Chase et al., 2003), таким образом, объединение в гибридах более близких геномов могло сопровождаться меньшим количеством перестроек (Lim et al., 2004). Полученный в сборке размер генома *N. tabacum* на 4–8% меньше суммы размеров геномов *N. sylvestris* и *N. tomentosiformis*, что согласуется с данными, ранее полученными методом проточной цитометрии и окрашиванием по Фельгену (<3.7%) (Leitch et al., 2008) и говорит об уменьшении размера генома после полиплоидизации (Sierro et al., 2014). Подобная реакция на полиплоидизацию часто встречается у растений, что подробнее описано в разделе 1.2 обзора литературы. Интересно, что элиминация ДНК у гибрида происходила неравномерно, T-геном утратил больше повторов, чем S-геном. Таким образом, в настоящее время вклад родительских видов в геном *N. tabacum* оценивают как 53% (*N. sylvestris*) : 47% (*N. tomentosiformis*) (Renny-Byfield et al., 2011, Sierro et al., 2014). Насколько сократилась доля кодирующих последовательностей для каждого из родительских видов, еще предстоит выяснить.

Еще до того, как были получены полногеномные данные для культурного табака, его активно исследовали с применением различных методов, среди которых были методы цитогенетики (Parokony and Kenton, 1995), анализ полиморфизма ферментов (пероксидаз, эстераз, каталаз и др.) (Sheen, 1972), анализ полиморфизма

длин рестрикционных фрагментов (Bland et al., 1985), различные вариантами блоттинга (Okamoto and Goldberg, 1985; Koukalova et al., 1989). Затем были секвенированы и собраны геномы органелл *N. tabacum* (Yukawa et al., 2006). Для получения качественной сборки ядерного генома требовались алгоритмы, способные определить принадлежность гомологичных последовательностей к тому или иному субгеному, полученному от родительских видов. С развитием биоинформатических инструментов подобное разделение стало возможным. В настоящее время собраны геномы трех сортов *N. tabacum* - K326, TN90 и Vasma/Xanti (Sierro et al., 2014). Поскольку *N. tabacum* является важной экономической культурой - для него выведено множество сортов, общее число которых неизвестно. Сорта различаются по морфологическим и физиологическим признакам, таким как темпы роста в различные периоды вегетации, засухоустойчивость, а также по содержанию вторичных метаболитов (Псарева Е., 1969). Селекция культурного табака ведется сразу по нескольким направлениям, при этом происхождение большинства сортов не установлено. Вопросами классификации культурного табака и внутривидовой филогении активно занимаются научные коллективы США (Fricano et al., 2012), Китая (Lu et al., 2013) и Турции (Yaldiz et al., 2018). Следующий раздел обзора затрагивает исторический аспект возникновения современного разнообразия сортов *Nicotiana tabacum*, а также охватывает исследования, посвященные внутривидовому полиморфизму культурного табака.

1.5. Внутривидовое разнообразие *N. tabacum*

Nicotiana tabacum является одной из важнейших не пищевых сельскохозяйственных культур, возделываемой во многих странах (Moon et al., 2009). Крупнейшие коллекции культурного табака поддерживаются в США, Китае, Индии, Бразилии (Yaldiz et al., 2017) и России. Российская коллекция *N. tabacum* во

Всероссийском НИИ табака, махорки и табачных изделий насчитывает более 3 тысяч сортообразцов (по данным из Отчета о научно-исследовательской работе за 2013–2020 годы ФГБНУ ВНИИТТИ). История возделывания и селекции культурного табака начинается с коренных племен Южной Америки. До открытия Америки Колумбом ареал *N. tabacum* простирался на большую часть Южной Америки, в Мексику и на Карибские острова (Setchell, 1921; Gerstel and Sisson, 1995). Генетический дрейф в сочетании с естественным и управляемым человеком отбором привел к разделению существовавших форм *N. tabacum* на отдельные типы. Хотя записей о ранних типах немного, четыре различные формы были описаны в Вест-Индии еще в 1667 г. (Garner et al., 1936). Сохранившихся описаний ранних форм культурного табака, однако, недостаточно для того, чтобы сопоставить формы *N. tabacum*, выращиваемые коренными американцами, с теми, что были перевезены для возделывания в страны Старого Света. На сегодняшний день существуют десятки тысяч сортов культурного табака, однако родственные отношения между ними установлены далеко не во всех случаях (Moon et al., 2009; Fricano et al., 2012; Sierro et al., 2014). Описанные Comes (1899) и Anastasia (1914) ранние попытки разработать внутривидовую классификацию *N. tabacum* на основе анализа морфологических признаков не принесли результатов (Lewis and Nicholson, 2007). Распространение культурного табака вследствие Европейской колонизации и рост потребления табачной продукции в течение 17-19 веков привели к формированию рыночных классов *N. tabacum*, основанных на предпочтениях потребителей в отношении различных способов употребления табака (Garner et al., 1936; Tilley, 1948; Spinden, 1950; Brooks, 1952; Tso, 1990). В настоящее время существует восемь рыночных классов: Берли, наполнитель для сигар, табак для закручивания сигар, темный табак воздушной сушки, темный табак паровой сушки, табак дымовой сушки, Мэриленд и восточный (Moon et al., 2009). Представители данных классов сильно различаются по своим морфологическим и химическим характеристикам, при этом генетический контроль признаков, по которым

представители были поделены на классы, в настоящее время не очень хорошо изучены (Lewis and Nicholson, 2007). Другой вариант внутривидовой классификации *N. tabacum*, активно используемый в России, был предложен Еленой Псаревой (Псарева, 1969). В основе данной классификации лежат морфо-биологические характеристики растения, свойства получаемого сырьевого продукта и ареал, в котором возделывают данную форму культурного табака. По Псаревой внутри вида *N. tabacum* выделяют 5 подвидов: азиатский, южный, островной, восточный, и американский. Каждый из подвидов включает две разновидности и ряд сортоформ (рис.11). Подобный тип классификации построен на принципах агроэкологической классификации культурных растений, разработанных Н. И. Вавиловым (Вавилов, 1957). Морфо-биологическая характеристика включает фенотипические признаки растения, среди которых высота побега, форма и размер листовой пластины, цвет листа, а также особенности биологии, такие как период созревания и темпы роста, засухоустойчивость, содержание никотина в листьях, и другие (Псарева, 1969). Содержание никотина в листьях, а затем в табачном сырье, является одним из ключевых критериев, по которому велась селекция. *N. tabacum* является одним из видов рода *Nicotiana*, производящих наибольшее количество алкалоидов. Еще одним видом с очень многообещающим биосинтетическим потенциалом урожайности, у которого никотина может накапливаться в девять раз больше, чем у *N. tabacum* (Stanfill et al. 2015), и который также выращивают для производства, является *N. rustica*. Алкалоиды в культурном табаке составляют 2% от сухого веса, при этом основную часть составляет никотин (95%), а на норникотин, анабазин, котинин и анатабин приходится большая часть оставшихся 5%. Помимо основных алкалоидов, *N. tabacum* синтезирует их сложные смеси, например, миосмин, 2',3'-дегидроникотин, никотириин, N-метиланатабин или N-метиланабазин (Baranska et al., 2012). В настоящее время усиливается интерес производителей к изготовлению табачной продукции со сниженным содержанием вредных веществ, в связи с чем,

соответственно, повышается внимание исследователей как к анализу регуляторных путей биосинтеза, так и к генетическому разнообразию и изменчивости внутри вида *N. tabacum* для выведения сортов культурного табака с пониженным уровнем биосинтеза различных алкалоидов (Gavilano et al., 2006).

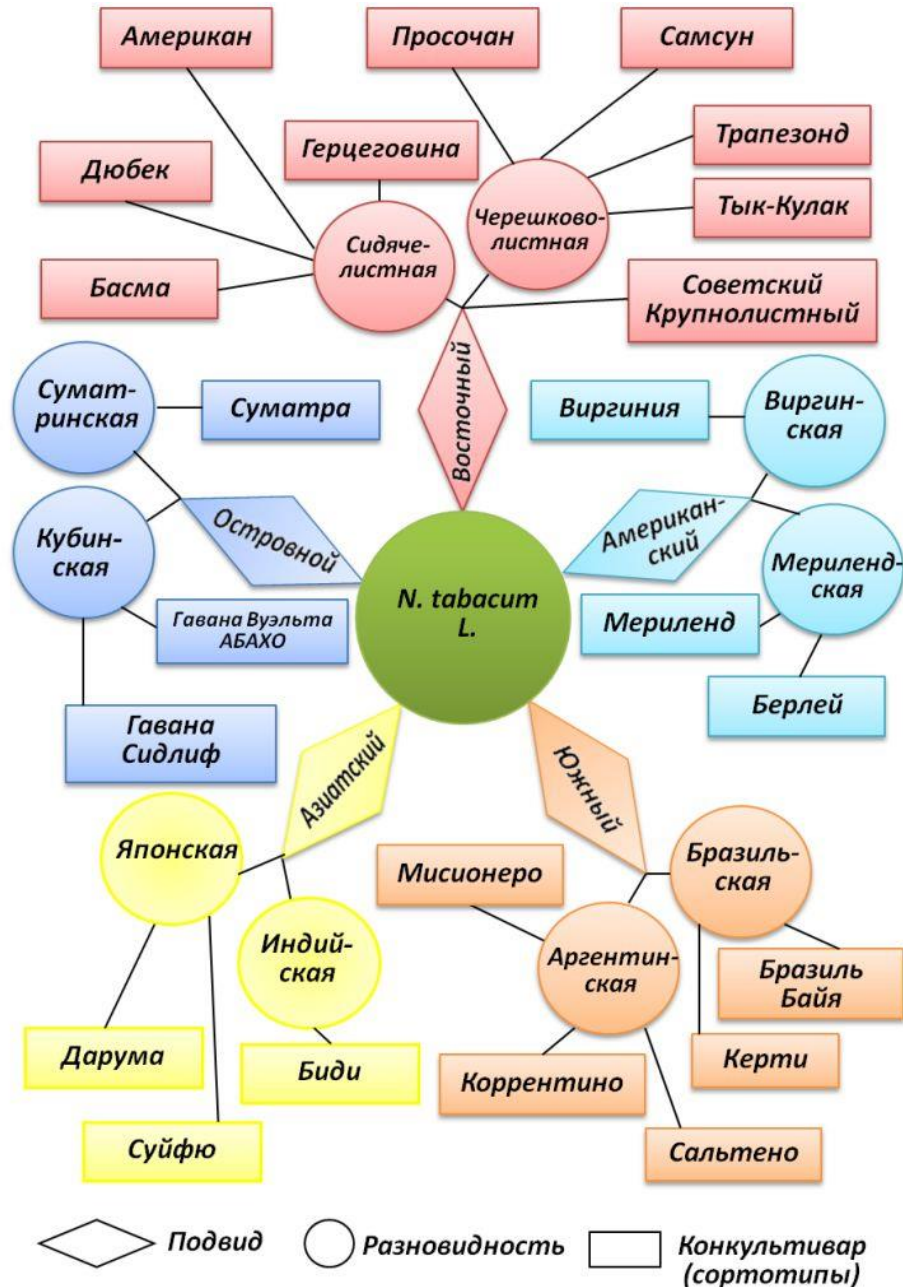


рис. 11. Схема внутривидовой классификации культурного табака (Е. Н. Псарева, 1969 г.)

Несмотря на то, что для поступающих в коллекционный фонд образцов указывают особенности их происхождения, для большинства сортов культурного табака их родственные отношения не установлены, в отличие от, например, подсолнечника, для многих сортов и образцов которого генеалогия относительно недавно была восстановлена (Gavrilova, Anisimova, 2017). Научными группами разных стран были предприняты попытки установить внутривидовые взаимоотношения сортов *N. tabacum* с помощью методов молекулярной биологии (Zhang et al., 2006; Darvishzadeh et al., 2013), однако исследователи столкнулись с проблемой низкого генетического разнообразия представителей культурного табака. При том, что среди сортов *N. tabacum* существует значительная фенотипическая изменчивость, нуклеотидная изменчивость, выявленная методами полиморфизма длины рестрикционного фрагмента (RFLP), случайно амплифицированной полиморфной ДНК (RAPD) и маркерами полиморфизма длины амплифицированного фрагмента (AFLP), оказалась сравнительно низкой (Brandle, Bai, 1999; Ren, Timko, 2001; Rossi et al., 2001). Предполагают несколько возможных причин, объясняющих наблюдаемую картину. Во-первых, указывают на недавнее время образования вида *N. tabacum*. В поддержку данного предположения выступают данные транскриптома культурного табака. Показано, что в большинстве случаев копии гомологичных генов, полученные от обоих родительских видов, не только сохраняются, но и экспрессируются (Vombarely et al., 2012). Подобная избыточность и полученные ранее доказательства ограниченной неофункционализации генов *N. tabacum* позволяют предположить, что этот вид может быть слишком молодым с точки зрения эволюции, в связи с чем такие механизмы, как потеря генов или неофункционализация, еще не успели в значительной степени повлиять на формирование его генома. В отличие от культурного табака вид *N. benthamiana*, с момента полиплоидизации которого прошел более длительный период времени, содержит большее количество однокопийных генов по сравнению с *N. tabacum* (Bally et al., 2015; Edwards et al., 2017). Второй

возможной причиной называют участие в формировании генетического пула *N. tabacum*, вероятно, только небольшой части генетического разнообразия, существовавшего в генофондах исходных видов-предшественников, сформировавших вид *N. tabacum* (Lewis, Nicholson, 2007). Это предположение пока не может быть доказано или опровергнуто в силу отсутствия данных об уровнях полиморфизма ДНК в генофондах диплоидных предшественников, *N. tomentosiformis* и *N. sylvestris*. Однако подобная картина описана для культурного арахиса, *Arachis hypogaea* L., где наблюдается низкий уровень полиморфизма молекулярных маркеров внутри культурного аллотетраплоида при гораздо более выраженном полиморфизме в генофондах его вероятных предков (Milla et al., 2005). В качестве третьей возможной причины, которая могла бы повлиять на формирование существующего сегодня уровня разнообразия, называют направленную селекцию. Согласно историческим документам, большинство ранних производителей табака в Латинской Америке, как правило, не вели направленную селекцию своих сортов, что, возможно, ограничивало снижение генетического разнообразия по сравнению с современными активно селекционными популяциями. Частота ауткроссинга, то есть, скрещивания представителей разных форм, для *N. tabacum* достигала 18% (Litton, Stokes, 1964). Выращивание табака в США началось с различных местных сортов, которые постепенно улучшались в течение 19 века путем народной селекции (Shamel, Cobey, 1907). Когда в начале 20-го века начали применять современные методы селекции, обнаружили, что скрещивание растений из разных рыночных классов, приводит к образованию линий неприемлемого качества, в связи с чем резко снизили процент ауткроссинга (Garner et al., 1936; Clayton, 1958). Активная селекция, проводимая на протяжении века, могла стать причиной снижения генетического разнообразия внутри *N. tabacum*. Анализ 117 сортов табака дымовой сушки из Коллекции зародышевой плазмы *Nicotiana* США продемонстрировал потерю более 50% аллельного разнообразия у сортов, выпущенных в период с 1930-х по 2005 год (Moon et al., 2009).

Высокая степень генетического родства между современными коммерческими сортами культурного табака вызывает озабоченность по поводу степени генетической изменчивости, остающейся в пулах зародышевой плазмы, на основе которой будут вести дальнейшую селекцию сортов для улучшения урожайности, устойчивости к болезням и качества табачного сырья. Значительные фенотипические различия по морфо-биологическим признакам, в том числе устойчивости к болезням, среди культивируемых *N. tabacum* демонстрируют потенциал генотипической селекции для улучшения сортов (Darvishzadeh et al., 2013; Elliott et al., 2008). Возникает вопрос, чем объяснить подобную фенотипическую изменчивость сортов культурного табака, существующую при низком уровне полиморфизма. Возможно, это результат влияния эпигенетических модификаций и/или структурной реорганизации генома при сохранении низкого уровня SNP. Для генома *N. tabacum* показаны примеры реорганизации генома посредством межгеномных перестроек (Kenton et al., 1993), и есть предположения, что среди различных сортов культурного табака могут присутствовать некоторые вариации в отношении структуры генома, такие как делеции, транслокации и иные (Mallah, 1943; Cameron, 1952; Kenton et al., 1993).

N. tabacum является природно-трансгенным растением, геном которого содержит клеточную Т-ДНК, полученную от предкового вида *N. tomentosiformis* (Chen et al., 2014). Изучение клТ-ДНК трех сортов *N. tabacum*, для которых есть полногеномные данные, позволило выявить различия в структуре клТ-ДНК ТА, которые заключаются в протяженной делеции, расположенной по центру ТА у сорта Basma/Xanthi (рис. 12). У сортов K326 и TN90 подобной делеции нет (Chen et al., 2014). Ранее методом ПЦР уже было показано, что клТ-ДНК *N. tabacum* сортов Basma Drama 2 и Samsoun and Xanthi содержат неполные последовательности *orf13* в отличие от *N. tomentosiformis* и сортов *N. tabacum* Wisconsin 38 и Havana 425 (Mohajjel-Shoja et al., 2011). Все сорта, у которых была обнаружена делеция в ТА, относятся к рыночному классу восточных табаков. Возможно, делеция возникла у формы, которая дала начало

данному классу. В таком случае наличие данной делеции может служить маркером, по которому сорт может быть отнесен к восточным табакам. Однако на сегодняшний день выборка сортов *N. tabacum*, исследованных по данному признаку, составляет всего 7 образцов, и недостаточна для формирования выводов.

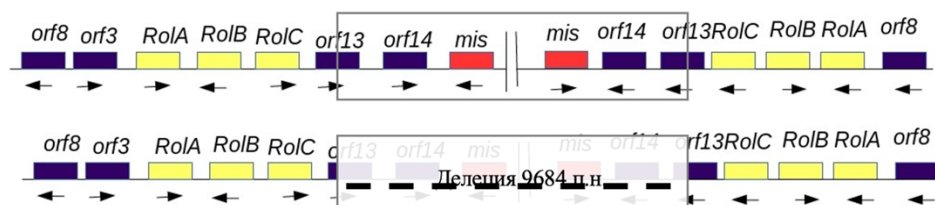


рис.12. Структурные варианты клТ-ДНК ТА в геноме *N. tabacum*. Вверху ТА без делеции, показана у сортов Wisconsin 38, Havana 425, K326, TN90; внизу ТА с делецией, показана у сортов Basma Drama 2, Samsoun и Xanthi, Basma/Xanti (Chen et al., 2014).

Разработка внутривидовой филогении *N. tabacum* в настоящее время находится в зачаточном состоянии и объединяет в себе различные подходы, такие как анализ морфо-биологических признаков, анализ характеристики сырья, географических ареалов выращивания, групп сцепления генов (Fricano et al., 2012; Moon et al., 2009; Псарева, 1969). Различия в клТ-ДНК могут послужить в качестве дополнительного филогенетического маркера для разработки внутривидовой классификации культурного табака.

ГЛАВА 2. МАТЕРИАЛЫ И МЕТОДЫ

2.1. АНАЛИЗ СТРУКТУРЫ КЛЕТОЧНОЙ Т-ДНК ПРЕДСТАВИТЕЛЕЙ ПОДРОДА *PETUNIOIDES*

2.1.1. Растительный материал

Материалом для работы послужили асептические растения видов *N. glauca* сорт '359' и *N. noctiflora* 'TW89'. Семена *N. glauca* были получены в Федеральном государственном бюджетном научном учреждении «Всероссийский научно-исследовательский институт табака, махорки и табачных изделий», семена *N. noctiflora* были любезно предоставлены Генбанком Университета Северной Каролины "US Nicotiana Germplasm Collection".

2.1.2. Методы культивирования растений, использованные в работе

Проращивание семян и поддержание растений проводили в асептических условиях на среде Мурасиге-Скуга с 20 г/л сахарозы (Murashige, Skoog, 1962) при температуре 23°C и фотопериоде 16 часов день/ 8 часов ночь.

Стерилизацию семян проводили, выдерживая семена в смеси H₂O₂ (30%) с EtOH (96%) в пропорции 1:1 в течение 10 минут, периодически встряхивая и переворачивая пробирку с раствором и семенами. После семена подсушивали в условиях ламинарного бокса на проавтоклавированной фильтровальной бумаге.

2.1.3. Молекулярно-биологические методы

Выделение ДНК для полногеномного секвенирования осуществляли из свежих листьев асептических растений *N. glauca* и *N. noctiflora* по модифицированному (с 10% СТАВ-буфером) протоколу макровыделения ДНК из растительной ткани (Murray, Thompson, 1980). В процессе выделения использовали стерильные наконечники для пипеток с отрезанными кончиками, избегая пипетирования, чтобы получить на выходе высокомолекулярную ДНК (фрагменты ~20 тысяч пар оснований).

Подготовку библиотеки типа pair-end производили по протоколу “TruSeq® Nano DNA Library Prep Reference Guide”. Подготовку библиотеки типа mate-pair производили по протоколу “Nextera® Mate Pair Library Prep Reference Guide”.

Секвенирование библиотек производили в Биобанке СПбГУ на секвенаторах Illumina HiSeq2500 и MiSeq Sequencing System. Дополнительно нашими коллегами из института РМІ (Филип Моррис Интернэшнл) были получены длинные прочтения с использованием технологии Oxford Nanopore для генома *N. noctiflora*.

Для уточнения структуры гТ использовали метод long PCR («дальнобойной ПЦР») с помощью набора “LONG PCR enzyme Mix” (Thermo Scientific™) в соответствии с протоколом производителя. Праймеры для ПЦР были подобраны с использованием программы Primer3 v.0.4.0 (Untergasser et al., 2012). Последовательности использованных праймеров приведены в таблице 1. Все праймеры синтезированы в компании “Евроген”.

Выделение тотальной РНК из асептических растений *N. noctiflora*. Для выделения были выбраны 3 растения в вегетирующей стадии, из каждого растения было выделено 4 образца РНК: из листьев, корня, узлов и междоузлий. Гомогенизацию растительных тканей проводили с жидким азотом на приборе TissueLyser (QIAGEN) в течение 30–60 секунд до порошкообразного состояния образца. Выделение РНК осуществляли набором RNeasy Mini Kit (QIAGEN) в соответствии с инструкцией к

набору. ДНКазную обработку РНК проводили с помощью набора RapidOut DNA Removal Kit (Thermo Scientific™) по протоколу производителя.

Измерение концентрации РНК проводили на приборе спектрофотометре NanoDrop 2000 (Thermo Scientific™).

Обратную транскрипцию осуществляли набором iScript™ cDNA Synthesis Kit (Bio-Rad) в соответствии с инструкцией.

Уровень экспрессии генов кЛТ-ДНК *N. noctiflora* измеряли с помощью ПЦР в реальном времени. В качестве референсного гена использовали ген глицеральдегидфосфатдегидрогеназы *gapdh*, который характеризуется стабильным уровнем экспрессии в разных тканях растений (Iskandar et al., 2004). Количество технических повторностей было не менее трех.

Состав смеси для ПЦР: 1 мкл кДНК *N. tabacum*, 10 мкл смеси iTaq Universal SYBR Green Supermix (Bio-Rad) (2X) и по 10 пмоль праймеров. Итоговый объем смеси составлял 20 мкл. ПЦР проводили на амплификаторе Bio-Rad CFX96 по следующей программе:

- 5 мин при 95⁰С — 1 цикл
- 15с при 95⁰С, 30с при 60⁰С, 30с при 72⁰С — 50 циклов
- 10 мин при 72⁰С – 1 цикл

Данные по расчету пороговых циклов (Ct) были получены с помощью программы CFX-Manager (BioRad), входящей в стандартное программное обеспечение прибора. Оценку экспрессии генов для каждой пробы проводили в трех аналитических повторностях, данные по которым затем усредняли. Обработку результатов проводили с помощью метода 2-ΔΔСТ (Livak, Schmittigen, 2001). Данные количественной оценки экспрессии генов представляли в относительных единицах, рассчитанных при сравнении с уровнями экспрессии референсного гена *gapdh*. Статистическую обработку полученных результатов проводили в Microsoft Excel.

Праймеры для ПЦР были подобраны с использованием программы Primer3 v.0.4.0 (Untergasser et al., 2012). Последовательности использованных праймеров приведены в таблице 2. Все праймеры синтезированы в компании “Евроген”.

Таблица 2. Последовательности праймеров для проведения ПЦР в реальном времени и long PCR

Праймер	T отжига	Последовательность	Назначение праймера
NglgapdhF1	60	5' GGTGCCAAGAAGGTTGTGAT	Анализ уровня экспрессии генов клТ-ДНК
NglgapdhR1	60	5' CAAGGCAGTTGGTAGTGCAA	
Nn1acsF2	60	5' AAGAGGAGGGCATCACGTTG	
Nn1acsR2	60	5' TCGAGACCAATCCAAACGCA	
Nn2acsF1	60	5' TGATCTGCCATGAATGCCCT	
Nn2acsR1	60	5' GCACAGCCGATTAGAGTGGT	
Nn2C1F1	60	5' TGTCGCCTTCCACGCAATAT	
Nn2C1R1	60	5' CGGTTAGCACGTTTCTTGGC	
Nn2iaaMF1	60	5' AACGCAAATGTAGCCGAGGA	
Nn2iaaMR1	60	5' GCTTGATCGTCGCCTGGATA	
NniaaMF2	60	5' CGGCTCTCGGACACTAAGTA	
NniaaMR1	60	5' CCATAAGCCACCATCTCAA	
NniaaHF1	60	5' TGCATGACTGAGCAATCCTG	
NniaaHR1	60	5' TGCTTTAAAGGGAATATTGC	
NgLngRightRolB	75	5' CTCTCCCATAAAGGTTGGTGACTTTTAG	Уточнение структуры гТ
NgLngLeftRolB	75	5'CTGGCTGAGCTTCTAAAGAACAAAATC	

2.1.4. Биоинформатические методы

Расшифровку данных с секвенатора MinION Oxford Nanopore производили с помощью программы guppy 4.2.2 (<https://community.nanoporetech.com>).

Проверку качества полученных данных с секвенатора MinION Oxford Nanopore производили с помощью программы seqkit 0.14.0 (Shen et al., 2016).

Проверку качества полученных данных с секвенаторов Illumina HiSeq2500 и MiSeq Sequencing System осуществляли в программе FastQC (Simons, 2010).

Сборку генома *N. glauca* осуществляли с помощью программы MaSuRCA-3.2.2 (Zimin et al., 2017). Сборку генома *N. noctiflora* осуществляли с помощью программ minimap2 2.17 (Li 2018), miniasm 0.3 (Li ,2016), а также ntEdit 1.3.2 (Vaser et al., 2017).

Оценку размера генома осуществляли при помощи анализа распределения частот к-меров размером 23 нуклеотида (Marcaris, 2011).

Проверку качества сборки генома *N. noctiflora* проводили с помощью программ уак (Cheng et al., 2021) и BUSCO 4.1.4. (Seppey et al., 2019).

Поиск сохранных генов в клТ-ДНК *N. noctiflora* проводили в программе UNIPRO Ugene ver.38 (Okonechnikov et al., 2012).

Поиск последовательностей, гомологичных агробактериальным, осуществляли с помощью программы Last (<http://last.cbrc.jp/>) против построенной базы данных.

2.2. ИЗУЧЕНИЕ ВНУТРИВИДОВОГО ПОЛИМОРФИЗМА клТ-ДНК *NICOTIANA* НА ПРИМЕРЕ СОРТОВ *N. TABACUM*

2.2.1. Растительный материал

В работе были использованы растения *Nicotiana tabacum* сортов отечественной селекции: Турецкий, Ориенталь, Virginia × Berley 38, Брянский 91, Suifu, Black Indian, Vuelta abajo, Havana 307, выращенные в открытом грунте из семян, приобретенных в магазине семян табака «Табакъ&К».

2.2.2. Методы культивирования растений, использованные в работе

Выращивание растений в грунте. Высадку семян растений вида *N. tabacum* осуществляли в смесь садового грунта и вермикулита (50/50), предварительно за сутки пропитанного фунгицидом. После посева семена выдерживали в сухом теплом месте (25⁰С) в течение 7–10 суток до появления всходов в емкостях, накрытых пленкой, и дополнительно контейнером из непрозрачного пластика для создания эффекта затенения. После этого емкости переносили на стеллаж с системой досветки. Полив проростков осуществляли в поддоны емкостей, пленку не снимали. По прошествии месяца с момента посадки с емкостей, содержащих растения, сняли пленку и выращивали в открытых горшках при температуре 23⁰С и фотопериоде 16 часов день/ 8 часов ночь.

2.2.3. Молекулярно-биологические методы

Выделение ДНК для дальнейшего использования в ПЦР осуществляли из свежих листьев выращенных в грунте сортовых растений *N. tabacum* по стандартному протоколу для выделения ДНК из растительной ткани с 2хСТАВ-буфером (Murray, Thompson, 1980). Для выделения ДНК брали материал трех растений каждого сорта.

Подбор праймеров осуществляли на основе сиквенсов клТ-ДНК *N. tomentosiformis*, полученных Chen в 2014 году (Accession N. KJ599826, KJ599827), с помощью программы Primer3 v.0.4.0 (Untergasser et al., 2012). Праймеры синтезированы в компании “Евроген”. Последовательности использованных праймеров приведены в таблице 3.

Для анализа последовательностей различных по «возрасту» ТА и ТВ использовали праймеры к генам (*orf14* в ТВ, *rolC* в ТА) и к межгенным спейсерам (см. рис.13).

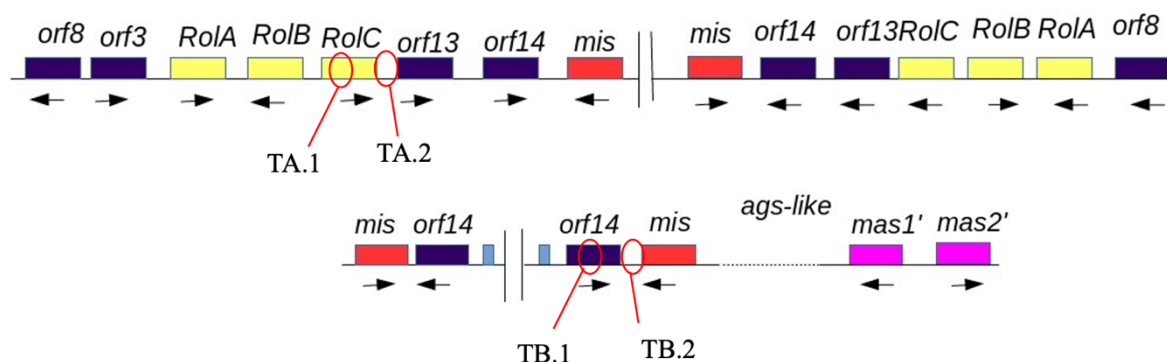


Рис. 13. Схема эксперимента для анализа последовательностей различных по «возрасту» ТА и ТВ, красными овалами отмечены фрагменты (не в масштабе), взятые в анализ. Длина фрагментов ТА.1, ТА.2, ТВ.1 составила порядка 600 п.н., длина фрагмента ТВ.2 составила 480 п.н.

Для анализа крупных перестроек в клТ-ДНК использовали 2 сочетания праймеров: либо прямой праймер к последовательности *orf13* в одном плече повтора (F на схеме эксперимента, рис.14) и обратный - к последовательности *orf13* во втором плече (R на схеме эксперимента), либо прямой праймер к последовательности *orf13* в одном плече повтора (F на схеме эксперимента) и обратный - к центральной части того же плеча (innR на схеме эксперимента). Работа сводилась к подбору комбинаций праймеров для выявления делеции либо ее отсутствия в центральной части клТ-ДНК ТА в геномах разных сортов *Nicotiana tabacum*. Расположения праймеров показаны на следующей схеме:

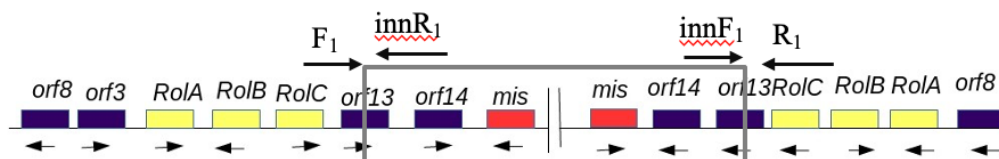


рис.14. Схема эксперимента для анализа крупных перестроек в клТ-ДНК, серой рамкой обозначена делеция, подписаны праймеры с указанием направления синтеза цепи в ПЦР

Таблица 3. Последовательности праймеров для анализа клТ-ДНК ТА и ТВ

Праймер	T отжига	Последовательность	Назначение праймера
Ntab_TARolC_F1	57	5' GAGCTCAAACCCTTACTGGA	Сравнительный анализ последовательностей ТА и ТВ
Ntab_TARolC_R1	57	5' TTGCTGATCCAACCTTCTGA	
Ntab_TAspac_F1	57	5' CCGCAATTGTTTAATTCCT	
Ntab_TAspac_R1	57	5' GCGCCGATGTATGATAATTT	
Ntab_TBorf14_F1	57	5' CGATAACTCACCGGTCTTCT	
Ntab_TBorf14_R1	57	5' GATCTCGACCTCTGGAAAAG	
Ntab_TBspac_F1	57	5' GTTTCCGGACTACAAGAGGA	
Ntab_TBspac_R1	57	5' TACGTTTTTATCCGCCTTTT	
NtabTA_innR1	55	5' ATACGTGCTCACTCGTCTCT	
NtabTA_F1	55	5' CAGGTAGGTCTCAACAGCTC	
NtabTA_R1	55	5' GTAGAAGTGGCCTGTCTCAG	
NtabTA_innF1	55	5' ATGGCTAAAGAATCGAGTGA	

ПЦР использовали для амплификации интересующих фрагментов. Состав смеси для ПЦР: 1 мкл ДНК *N. tabacum*, 10 мкл смеси DreamTaq Green PCR Master Mix (Thermo Scientific™) (2X) и по 10 пмоль праймеров. Итоговый объем смеси составлял 20 мкл. ПЦР проводили на амплификаторе Bio-Rad C1000 Thermal Cycler по следующей программе:

- 5 мин при 93⁰С — 1 цикл
- 17с при 93⁰С, 40с при X⁰С, 40с при 72⁰С — 40 циклов
- 5 мин при 72⁰С – 1 цикл

Где X – температура отжига, ее устанавливали в соответствии с необходимой температурой для конкретной пары праймеров. Праймеры синтезированы в компании “Евроген”.

Электрофорез в агарозном геле использовали для проверки качества выделенной ДНК, также для разделения и визуализации продуктов амплификации. Пробы разделяли в 1% агарозном геле на однократном SB буфере (0.4 г NaOH, 2.25 г борной кислоты на 1 л дистиллированной воды) с добавлением EtBr (1%). Объем каждой из

наносимых проб — 3 мкл. Длительность электрофореза составила 30–35 мин при напряжении 90 В. В качестве маркера молекулярного веса использовали маркер «GeneRuler 100 bp DNA Ladder» либо «GeneRuler 1 kb DNA Ladder» (1.5 мкл). Изображение геля было получено с помощью системы гель-документирования «Bio-Rad».

Секвенирование по Сенгеру было выполнено ресурсным центром РМИКТ.

2.2.4. Биоинформатические методы

Анализ сиквенсов проводили в программе UNIPRO Ugene ver.38 (Okonechnikov et al., 2012).

Множественное выравнивание полученных нуклеотидных последовательностей сортов *N. tabacum* осуществляли в программе MAFFT v.7 (<http://mafft.cbrc.jp/alignment/server/index.html>) (Kato et al., 2019).

Выравнивание (локальное) нуклеотидных последовательностей генов *PMT* сортов *N. tabacum* осуществляли с помощью программы blastn из пакета ncbi-BLAST+ (<https://blast.ncbi.nlm.nih.gov/>) против базы данных database Nucleotide collection (NCBI nr/nt) со стандартными параметрами.

ГЛАВА 3. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

3. 1. АНАЛИЗ СТРУКТУРЫ КЛТ-ДНК ПРЕДСТАВИТЕЛЕЙ ПОДРОДА *PETUNIOIDES*

3.1.1. Сборка генома *N. glauca* и уточнение структуры клТ-ДНК gТ

Nicotiana glauca или табачное дерево - первое растение, в геноме которого была обнаружена клеточная Т-ДНК (White et al., 1983), названная gТ. Разными исследователями были получены отдельные участки последовательности gТ (Aoki et al., 1994; Furner et al., 1986; Suzuki et al., 2002), которые затем с помощью метода локального выравнивания были объединены в протяженный фрагмент (Suzuki et al., 2002). Структура gТ, полученная Сузуки с соавторами, не была проверена экспериментально. 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 gТ (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 было показано на примере *Nicotiana 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 gТ а также проверить наличие либо отсутствие иных клТ-ДНК в данном геноме.



Рис. 15. Растение *Nicotiana glauca* в теплице, фото Матвеевой Т. В.

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. Исходные протоколы подготовки библиотек были оптимизированы в ходе

рабочего процесса для получения фрагментов необходимого размера: подбирали условия (время, количество циклов) для фрагментации закольцованной ДНК ультразвуком, соотношения магнитных частиц и рабочей смеси с ДНК. Обе библиотеки были секвенированы на Illumina HiSeq 2500. Всего было получено 210 млрд.п.н. сырых данных. В ходе проведения оценки качества полученных данных не было найдено примесей адаптеров, линкеров, и прочих искусственных олигонуклеотидов. Распределение GC% унимодальное составило 40%, что говорит об отсутствии контаминации человеческими и бактериальными последовательностями. Фрагменты низкого качества (FQ score <30) были отфильтрованы на подготовительном этапе и в сборку не вошли. 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 PGPE00000000.1. 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). На момент сборки нами генома *N. glauca* (2017) в открытом доступе были собранные геномы *N. sylvestris*, *N. tomentosiformis*, *N. otophora*, *N. benthamiana* и 3 сорта *N. tabacum*, из них по качеству сборки, отраженному в таких параметрах как N50, количество скаффолдов и др., наилучшее качество было показано для генома *N. tabacum* сорта TN90 (Sierra et al., 2014). Данный геном был выбран в качестве референсного для проведения полногеномного выравнивания генома *N. glauca*. В результате выравнивания на геном *N. tabacum* сорта TN90 скаффолды *N. glauca* совпадали в среднем на 88% (Khafizova et al., 2018), что говорит о довольно сильных различиях. *N. tabacum* относится к подроду *Tabacum* секции *Nicotiana*, в то время как *N. glauca* относится к подроду *Petunioides* секции *Noctiflorae*, что объясняет

полученный уровень различий при выравнивании. Выбрать для подобного анализа вид секции *Noctiflorae* не представлялось возможным в связи с отсутствием полногеномных данных для представителей данной секции, что еще раз подчеркивает значимость получения полногеномных данных *N. glauca*. 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 и в области биосинтеза, поскольку содержит информацию о путях биосинтеза различных соединений.

Поиск последовательностей, гомологичных агробактериальным, выполняли с помощью программы Last, осуществляющей полногеномное выравнивание. Предварительно нами была составлена база данных, использующая суффиксные деревья, содержащая все последовательности, гомологичные агробактериальным, обнаруженные на сегодняшний момент в составе описанных клТ-ДНК. Подобная структура данных обеспечивает быстрый доступ и значительно сокращает требования к вычислительным ресурсам, делая возможным проведение дальнейшего анализа на обычном компьютере. 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 центральному участку 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 центра 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* (рис. 16). Новых клТ-ДНК в геноме *N. glauca* не обнаружено, что позволяет сделать вывод об однократной агробактериальной трансформации данного вида.

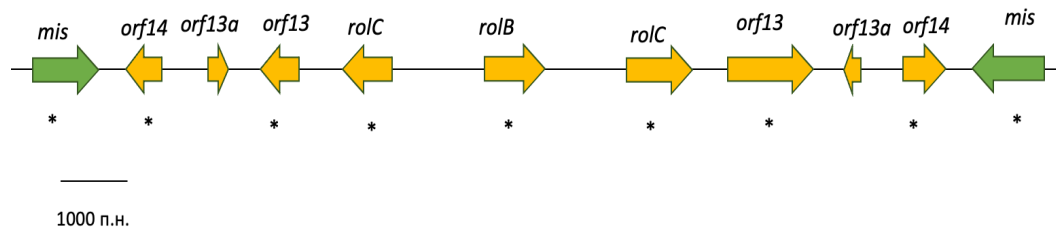


Рис.16. Структура клТ-ДНК в геноме *N. glauca* (Khafizova et. al., 2018), Цветными стрелками показано направление открытых рамок считывания, тонкими стрелками показаны плечи повтора, звездочками отмечены сохраненные открытые рамки.

Для большинства природно-трансгенных видов *Nicotiana* анализ клТ-ДНК был проведен на уровне отдельных генов методом ПЦР, однако данный способ анализа имеет ряд ограничений. Накопление мутаций в последовательностях клТ-ДНК, делеции участков генов негативно влияют на отжиг праймеров на матрице, что приводит к получению ложноотрицательных результатов. Так, например, *N. noctiflora* по результатам ПЦР-анализа числился в списке видов, не содержащих клТ-ДНК в геноме (Intrieri, Buiatti, 2001), в то время как анализ транскриптома показал обратное (Long et al., 2016). Вид *N. noctiflora* является филогенетически близким к *N. glauca* (Knapp et al., 2004), для него, так же как для *N. glauca*, отсутствуют полногеномные данные.

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

Nicotiana noctiflora (ночной табак), как и *N. glauca*, - диплоидный вид из подрода *Petunioides*. В результате массового скрининга представителей рода *Nicotiana* методом ПЦР с праймерами к генам *rolC*, *rolB*, *orf13* и *orf14* ночной табак был отнесен к видам, геномы которых не содержат клТ-ДНК (Intrieri, Vuiatti, 2001). Позже в 2016 году при изучении генов-участников молекулярных механизмов защиты растений в отсековенном транскриптом *N. noctiflora* были обнаружены транскрипты генов, гомологичных агробактериальным (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* для последующего поиска и анализа клТ-ДНК в его составе.

Нами были подготовлены и отсековены библиотеки *N. noctiflora* двух типов. 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. Исходные протоколы подготовки библиотек были оптимизированы по примеру с библиотеками для *N. glauca*. Обе библиотеки были отсековены на Illumina HiSeq 2500. Анализ качества данных, проведенный в программе FastQC показал отсутствие отклонений в качестве прочтений. Значительных примесей адаптеров не присутствовало. Исходные файлы были отфильтрованы по качеству и по наличию линкерных последовательностей, чтобы разделить прочтения из разных библиотек. Распределение GC% обеих библиотек унимодальное в диапазоне 42%, что подтверждает отсутствие контаминаций.

Прочтениями низкого качества считались фрагменты со значением FQ score ниже 30, или с большим количеством неидентифицированных нуклеотидов. Для дальнейшей работы прочтения низкого качества были отфильтрованы из сырых данных. Для улучшения качества сборки нашими коллегами из PMI (Филип Моррис Интернэшнл), 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 программами minimap2 и miniasm. В результате сборки было получено 4468 скаффолдов, N50 составил 261 т.п.н., таким образом, качество сборки генома *N. noctiflora* выше, чем для *N. glauca*.

Для поиска и анализа клТ-ДНК в геноме *N. noctiflora* предварительно была составлена база данных, содержащая белковые последовательности генов, гомологичных агробактериальным генам, входящих в состав Т-ДНК. С помощью алгоритма BLAST в геноме *N. noctiflora* были найдены два скаффолда, содержащие две различающихся по составу клТ-ДНК. Обе клТ-ДНК организованы в виде несовершенных инвертированных повторов, что характерно для клТ-ДНК в геномах представителей *Nicotiana*, а также других природно-трансгенных растений. Длина первой, NnT-DNA1, составляет 21787 п.н., в каждом плече расположены гомологи агробактериальных генов *iaaH*, *iaaM*, *vis*, *mis*, *orf14*, *orf13a*, *orf13*, *RolC*, по центру NnT-DNA1 находится гомолог *acs* (см. рис. 17; табл. 4). Длина второй клТ-ДНК, NnT-DNA2, составляет 14611 п.н., в состав плеч входят гомологи *acs*, *C* и *iaaM* (см. рис. 17; табл. 4). Поскольку клТ-ДНК представляет собой повтор, можно приблизительно оценить время его попадания в геном растения, сравнив плечи повтора. Сходство плеч NnT-DNA2 на нуклеотидном уровне составляет 98%, в то время как для NnT-DNA1 – 99%, из чего можно сделать вывод, что NnT-DNA2 попала в геном *Nicotiana noctiflora* раньше, чем NnT-DNA1. NnT-DNA1 и гТ в геноме *Nicotiana glauca* несут одинаковый по структуре фрагмент, включающий гомологи генов *rolC*, *orf13*, *orf13a*, *orf14* и *mis*.

Степень сходства 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, говорит о том, что виды *N. noctiflora* и *N. glauca* были трансформированы независимо, а не получили клТ-ДНК от некой предковой формы до их возникновения.

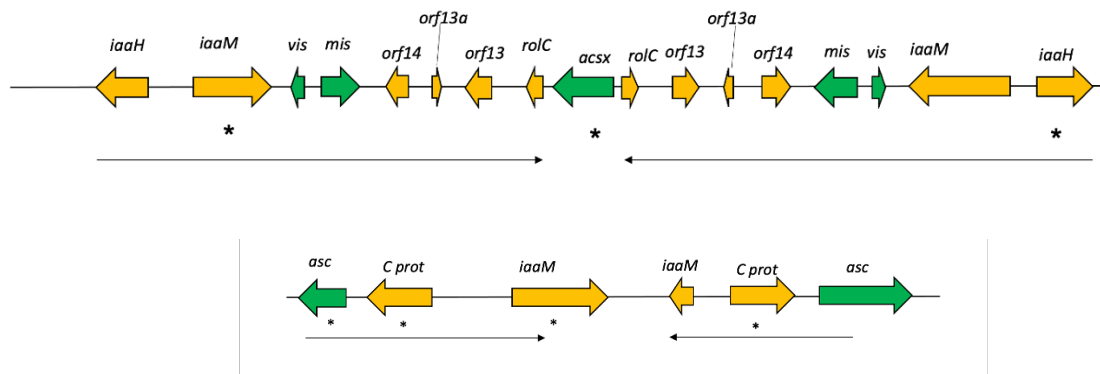


Рис. 17. Структура клТ-ДНК в геноме *Nicotiana noctiflora*. Сверху вниз: NnT-DNA1, NnT-DNA2. Цветными стрелками показано направление ОРС, тонкими стрелками показаны плечи повтора, звездочками отмечены сохраненные ОРС.

Попытка найти возможный предковый штамм для NnT-DNA1 показала, что NnT-DNA1, видимо, несет в себе сочетание фрагментов разных агробактериальных штаммов, так как на разных участках для данной клТ-ДНК выявлено сходство с различными агробактериями. Так, участок *iaaH* - *iaaM* в обоих плечах NnT-DNA1 на 74% схож с последовательностью штамма *Agrobacterium tumefaciens* CFBP1935, в то

время как участок *mis - acs* имеет 76–86% сходства с последовательностью штамма *Agrobacterium rhizogenes* 1724. Для NnT-DNA2 также показано мозаичное строение - так уровень сходства фрагмента *iaaM* в одном плече повтора с нуклеотидной последовательностью *A. tumefaciens* CFBR1935 составляет 68%, *iaaM* в другом плече - 71% с плазмидой *A. tumefaciens* pTi82143 и 68% с *A. tumefaciens* CFBR1935, в то время как для фрагмента *acs - C* в обоих плечах показано сходство со штаммом *A. rhizogenes* 1724. Таким образом, можно предположить, что источником клТ-ДНК в геноме *N. noctiflora* могли быть штаммы, близкие к ныне охарактеризованным штаммам *A. tumefaciens* CFBR1935 и *A. rhizogenes* 1724. 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 возникнуть в результате трансформации растения штаммом, несущим последовательности, полученные в результате горизонтального переноса генов от различных агробактерий. Известно, что у бактерий горизонтальный перенос генов встречается очень часто, являясь основным способом получения новых комбинаций генетического материала и, как следствие, разнообразия. Однако на сегодняшний день штаммы со структурой Т-ДНК, подобной клТ-ДНК ТВ, ТС, ТЕ, TD, а также NnT-DNA1 и NnT-DNA2, в литературе не описаны, хотя важно отметить, что современные штаммы не полностью соответствуют тем, которые участвовали в трансформации растений. Еще одной возможной причиной образования подобной мозаичной структуры является коинфекция растения агробактериями различных штаммов и слияние фрагментов Т-ДНК непосредственно перед встраиванием их в растительный геном. В условиях лаборатории показано множество примеров коинфекции растений несколькими различными штаммами (Otten et al., 1984; Owens, Smigocki, 1988; Charest et al., 1989), однако при этом для агробактерий показан механизм, угнетающий как иные штаммы агробактерий, так и других прокариот, находящихся поблизости от сайта инфицирования на растении, повышая таким образом конкурентоспособность

конкретного штамма (Ma et al., 2014). В настоящий момент имеющихся данных недостаточно, чтобы сказать, насколько вероятна коинфекция в природе. Дальнейшее изучение различных клТ-ДНК, а также расширение знаний о разнообразии штаммов агробактерий помогут прояснить вопрос появления мозаичных клТ-ДНК в геномах представителей *Nicotiana*.

Таблица 4. 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	
		-	10804-11304	75			

3.1.3. 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, 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, 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. 17), для которых мы измерили относительные уровни экспрессии. Анализ проводили в трех биологических повторностях, для каждой из которых было получено не менее трех технических повторностей. Далее данные по биологическим повторностям усредняли и рассчитывали количественные значения по методу 2-ΔΔСТ (Livak, Schmittgen, 2001). В качестве референсного гена был использован ген глицеральдегидфосфатдегидрогеназы *gapdh*, который характеризуется стабильным уровнем экспрессии в разных тканях растений (Iskandar et al., 2004). В результате были получены значения относительных уровней экспрессии генов *iaaH*, *iaaM*, *acs* в NnT-DNA1, и *iaaM* - в NnT-DNA2, которые представлены на рисунке 18. Для 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. Из графиков (рис.18) видно, что экспрессия всех изученных генов достигает максимального значения в тканях корня *N. noctiflora*, снижаясь в порядке лист - стебель.

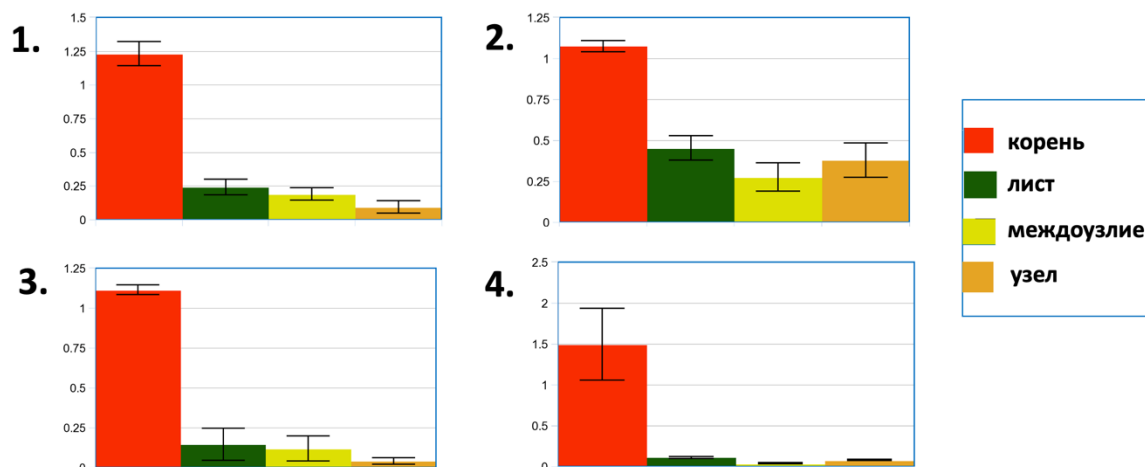


Рис. 18. Относительные уровни экспрессии генов клТ-ДНК *Nicotiana noctiflora*:

1. - *iaaM*, 2. - *acs*, 3. - *iaaH* - в NnT-DNA1; 4. *iaaM* - в NnT-DNA2

При этом интересно, что в стебле и листьях уровень экспрессии гена *acs* заметно выше, чем для генов *iaaH* и *iaaM* (рис. 18).

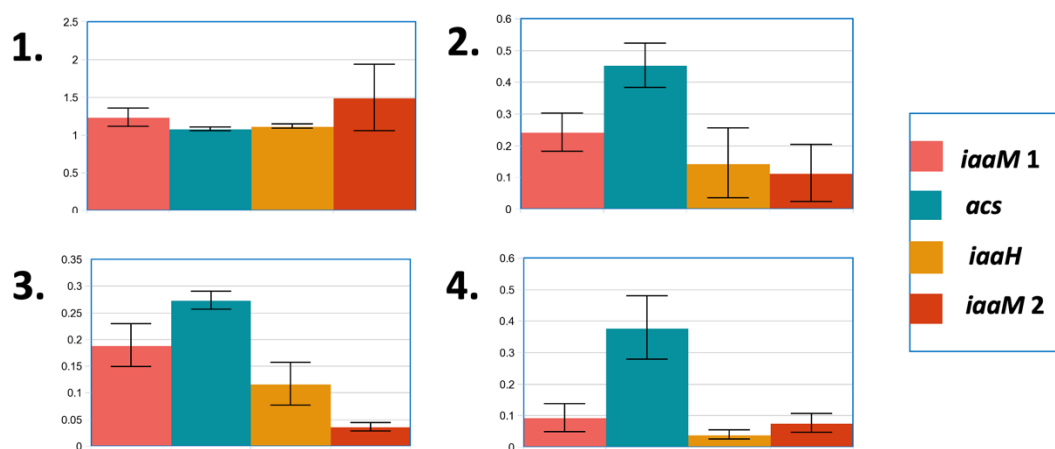


Рис. 18. Относительные уровни экспрессии генов клТ-ДНК *Nicotiana noctiflora*:

1. - корень, 2. - лист, 3. - междоузлие, 4. - узел

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. Например, разрастание

корневой системы, которое запускают гены Т-ДНК *A. rhizogenes*, могло способствовать выживанию в засушливых условиях за счет увеличения поверхности всасывания (Chen, Otten, 2017). Стимуляция клеточного деления и активация меристемы генами *rolC* и *rolB* могли повышать регенерационные способности растения (Koltunow et al., 2001; Otten, 2018), а гены *iaaH* и *iaaM*, кодирующие синтез индолилуксусной кислоты, влиять на ростовые процессы растений. Гены *rolA*, *rolC* и *orf13* вызывают укорочение побегов (Lemcke, Schmulling, 1998; Guivarch et al., 1996), что также могло служить адаптивным преимуществом. Обсуждают также влияние *plast*-генов на вторичный метаболизм и на углеводный обмен растений (Otten, 2018; Mohajjel-Shoja et al., 2011). Попытки объяснить роль опинов для растения нашли отражение в иммунной гипотезе, согласно которой секретируемые растением опины служат как метка, что растение уже трансформировано, и подавляют попытки повторной трансформации его другими штаммами агробактерий. Иммунная гипотеза также предлагает механизм сайленсинга генов вновь полученной Т-ДНК за счет экспрессии генов уже имеющейся клТ-ДНК (Escobar et al., 2001). Критикой данной гипотезы служит присутствие нескольких клТ-ДНК у представителей рода *Nicotiana* секции *Tomentosae*, которые были трансформированы несколько раз (Chen et al., 2014; Chen et al., 2016). Возможно, в первичной и в повторных трансформациях участвовали штаммы, несущие гены синтеза и катаболизма различных опинов, в связи с чем не происходило распознавание чужой метки. Подобное предположение согласуется с присутствием разных генов опинсинтаз в клТ-ДНК некоторых представителей *Nicotiana*. 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. Таким образом, функция генов клТ-ДНК для растений пока остается не установленной. Возможно, активность генов, полученных от агробактерий, была важна при трансформации растения и на ранних этапах его регенерации, после чего привнесенные гены могли утратить свою функцию. Гены Т-ДНК Ri-плазмид экспрессируются в бородатых корнях, для генов клТ-ДНК также показан более высокий уровень экспрессии в корнях по сравнению с другими органами растения, из чего можно предположить, что паттерн экспрессии генов Т-ДНК у растений-трансформантов мог сохраниться при закреплении клТ-ДНК в растительном геноме.

Анализ клТ-ДНК представителей подрода *Tabacum*, проведенный с использованием полногеномных данных, позволил установить количество и состав клТ-ДНК в видах *N. tomentosiformis*, *N. tabacum* и *N. otophora*. На основе этих данных были разработаны сценарии появления и распространения клТ-ДНК по подроду *Tabacum*, включающие множественные акты трансформации растений агробактериями, элиминация клТ-ДНК из растительного генома в ходе его эволюции и наследование клТ-ДНК при видообразовании путем межвидовой гибридизации. Анализ полногеномных данных *N. glauca* и *N. noctiflora*, относящихся к подроду *Petunioides*, позволяет предположить подобные сценарии для данного подрода и сравнить их с описанными для подрода *Tabacum*, что в результате поможет приблизиться к пониманию эволюционных событий рода *Nicotiana*.

3.2. ИЗУЧЕНИЕ ВНУТРИВИДОВОГО ПОЛИМОРФИЗМА КЛТ-ДНК *NICOTIANA* НА ПРИМЕРЕ СОРТОВ *N. TABACUM*

3.2.1. Сравнительный анализ последовательностей ТА и ТВ, различных по «возрасту»

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, Nicholson, 2007). *N. tabacum* – это природно-трансгенный вид, который содержит три различающиеся по составу и возрасту клТ-ДНК (ТА, ТВ, TD) (рис. 4, 5, 7), 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.2) (рис. 13). На матрицах ДНК методом ПЦР были получены фрагменты, визуализация которых на агарозном геле показала единообразие длин (см. рис. 20).

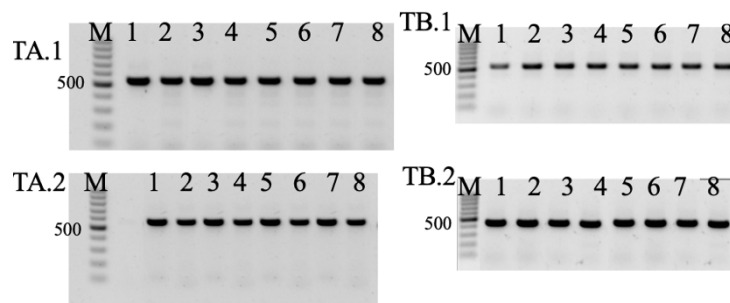


Рис. 20. Единообразие фрагментов гена *rolC* (ТА.1) и межгенного промежутка (ТА.2), гена *orf14* (ТВ.1) и межгенного промежутка (ТВ.2), полученных на матрицах ДНК 8 сортов *N. tabacum*

Все полученные фрагменты были секвенированы, после чего были построены 4 множественных выравнивания, показавшие высокий уровень сходства последовательностей как генов, так и межгенных промежутков в обеих клТ-ДНК (рис. 22, 22). Все изученные фрагменты имеют сходство 99–100% на нуклеотидном уровне. 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 растительных последовательностей 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*. В литературе описана взаимосвязь замены нуклеотида в позиции 75 последовательности *PMT2* с уровнем синтеза никотина в сортах культурного табака (Reddy, Saiprasad, 2015). Материалом для анализа послужили отсеквенированные геномы сортов *N. tabacum* K326 (NSAA00000000.1, 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 (см. рис. 23).

brya	ggtagtgtagcagcgaatgggctcaagaagcattatggaaaacgccaagatgagctcaaaacc	brya	ggtttgatgtaacagctcctgggttttgaatcctaattgtgtgtgtgttttaacggaagaat
vu	ggtagtgtagcagcgaatgggctcaagaagcattatggaaaacgccaagatgagctcaaaacc	vir	ggtttgatgtaacagctcctgggttttgaatcctaattgtgtgtgtgttttaacggaagaat
hav	ggtagtgtagcagcgaatgggctcaagaagcattatggaaaacgccaagatgagctcaaaacc	vu	ggtttgatgtaacagctcctgggttttgaatcctaattgtgtgtgtgttttaacggaagaat
ind	ggtagtgtagcagcgaatgggctcaagaagcattatggaaaacgccaagatgagctcaaaacc	hav	ggtttgatgtaacagctcctgggttttgaatcctaattgtgtgtgtgttttaacggaagaat
ori	ggtagtgtagcagcgaatgggctcaagaagcattatggaaaacgccaagatgagctcaaaacc	ind	ggtttgatgtaacagctcctgggttttgaatcctaattgtgtgtgtgttttaacggaagaat
ori	ggtagtgtagcagcgaatgggctcaagaagcattatggaaaacgccaagatgagctcaaaacc	ori	ggtttgatgtaacagctcctgggttttgaatcctaattgtgtgtgtgttttaacggaagaat
tur	ggtagtgtagcagcgaatgggctcaagaagcattatggaaaacgccaagatgagctcaaaacc	su	ggtttgatgtaacagctcctgggttttgaatcctaattgtgtgtgtgttttaacggaagaat
su	ggtagtgtagcagcgaatgggctcaagaagcattatggaaaacgccaagatgagctcaaaacc	tur	ggtttgatgtaacagctcctgggttttgaatcctaattgtgtgtgtgttttaacggaagaat
vir	ggtagtgtagcagcgaatgggctcaagaagcattatggaaaacgccaagatgagctcaaaacc	tur	ggtttgatgtaacagctcctgggttttgaatcctaattgtgtgtgtgttttaacggaagaat
*****	*****	*****	*****
brya	ttactggagcatggtcctcaaggtgaatccatggacatcgatgaagaagaagtctcgggg	brya	ggtttaattcatttttactgaggatgttagcgtgtagctgtttacaacaagtctcggcg
vu	ttactggagcatggtcctcaaggtgaatccatggacatcgatgaagaagaagtctcgggg	vir	ggtttaattcatttttactgaggatgttagcgtgtagctgtttacaacaagtctcggcg
hav	ttactggagcatggtcctcaaggtgaatccatggacatcgatgaagaagaagtctcgggg	vu	ggtttaattcatttttactgaggatgttagcgtgtagctgtttacaacaagtctcggcg
ind	ttactggagcatggtcctcaaggtgaatccatggacatcgatgaagaagaagtctcgggg	hav	ggtttaattcatttttactgaggatgttagcgtgtagctgtttacaacaagtctcggcg
ori	ttactggagcatggtcctcaaggtgaatccatggacatcgatgaagaagaagtctcgggg	ind	ggtttaattcatttttactgaggatgttagcgtgtagctgtttacaacaagtctcggcg
ori	ttactggagcatggtcctcaaggtgaatccatggacatcgatgaagaagaagtctcgggg	ori	ggtttaattcatttttactgaggatgttagcgtgtagctgtttacaacaagtctcggcg
tur	ttactggagcatggtcctcaaggtgaatccatggacatcgatgaagaagaagtctcgggg	su	ggtttaattcatttttactgaggatgttagcgtgtagctgtttacaacaagtctcggcg
su	ttactggagcatggtcctcaaggtgaatccatggacatcgatgaagaagaagtctcgggg	tur	ggtttaattcatttttactgaggatgttagcgtgtagctgtttacaacaagtctcggcg
vir	ttactggagcatggtcctcaaggtgaatccatggacatcgatgaagaagaagtctcgggg	tur	ggtttaattcatttttactgaggatgttagcgtgtagctgtttacaacaagtctcggcg
*****	*****	*****	*****
brya	gccaaacgactactgtatctctatgtcgactgccgacaaatgatgcgctgctttttggg	brya	cttctgtggcaagcaggaaaatccggctcttgaattttttcaagcaacagcagatataacc
vu	gccaaacgactactgtatctctatgtcgactgccgacaaatgatgcgctgctttttggg	vir	cttctgtggcaagcaggaaaatccggctcttgaattttttcaagcaacagcagatataacc
hav	gccaaacgactactgtatctctatgtcgactgccgacaaatgatgcgctgctttttggg	vu	cttctgtggcaagcaggaaaatccggctcttgaattttttcaagcaacagcagatataacc
ind	gccaaacgactactgtatctctatgtcgactgccgacaaatgatgcgctgctttttggg	hav	cttctgtggcaagcaggaaaatccggctcttgaattttttcaagcaacagcagatataacc
ori	gccaaacgactactgtatctctatgtcgactgccgacaaatgatgcgctgctttttggg	ind	cttctgtggcaagcaggaaaatccggctcttgaattttttcaagcaacagcagatataacc
ori	gccaaacgactactgtatctctatgtcgactgccgacaaatgatgcgctgctttttggg	ori	cttctgtggcaagcaggaaaatccggctcttgaattttttcaagcaacagcagatataacc
tur	gccaaacgactactgtatctctatgtcgactgccgacaaatgatgcgctgctttttggg	su	cttctgtggcaagcaggaaaatccggctcttgaattttttcaagcaacagcagatataacc
su	gccaaacgactactgtatctctatgtcgactgccgacaaatgatgcgctgctttttggg	tur	cttctgtggcaagcaggaaaatccggctcttgaattttttcaagcaacagcagatataacc
vir	gccaaacgactactgtatctctatgtcgactgccgacaaatgatgcgctgctttttggg	tur	cttctgtggcaagcaggaaaatccggctcttgaattttttcaagcaacagcagatataacc
*****	*****	*****	*****
brya	agatccatgccctataatgtagggaggcaccctccttacgacctcccccgatccat	brya	ggacctacgcccgaatgtttaatttctaatttctctgctgcttttgaagaacg
vu	agatccatgccctataatgtagggaggcaccctccttacgacctcccccgatccat	vir	ggacctacgcccgaatgtttaatttctaatttctctgctgcttttgaagaacg
hav	agatccatgccctataatgtagggaggcaccctccttacgacctcccccgatccat	vu	ggacctacgcccgaatgtttaatttctaatttctctgctgcttttgaagaacg
ind	agatccatgccctataatgtagggaggcaccctccttacgacctcccccgatccat	hav	ggacctacgcccgaatgtttaatttctaatttctctgctgcttttgaagaacg
ori	agatccatgccctataatgtagggaggcaccctccttacgacctcccccgatccat	ind	ggacctacgcccgaatgtttaatttctaatttctctgctgcttttgaagaacg
ori	agatccatgccctataatgtagggaggcaccctccttacgacctcccccgatccat	ori	ggacctacgcccgaatgtttaatttctaatttctctgctgcttttgaagaacg
tur	agatccatgccctataatgtagggaggcaccctccttacgacctcccccgatccat	su	ggacctacgcccgaatgtttaatttctaatttctctgctgcttttgaagaacg
su	agatccatgccctataatgtagggaggcaccctccttacgacctcccccgatccat	tur	ggacctacgcccgaatgtttaatttctaatttctctgctgcttttgaagaacg
vir	agatccatgccctataatgtagggaggcaccctccttacgacctcccccgatccat	tur	ggacctacgcccgaatgtttaatttctaatttctctgctgcttttgaagaacg
*****	*****	*****	*****
brya	gctgatggtgcgcccctagaggtgtctcgtggactcrggaaggcaacogatttctcagc	brya	cogacatgcttccacaagcagcgaagctgctgctgcttgaaaaacogttaaattatgac
vu	gctgatggtgcgcccctagaggtgtctcgtggactcrggaaggcaacogatttctcagc	vir	cogacatgcttccacaagcagcgaagctgctgctgcttgaaaaacogttaaattatgac
hav	gctgatggtgcgcccctagaggtgtctcgtggactcrggaaggcaacogatttctcagc	vu	cogacatgcttccacaagcagcgaagctgctgctgcttgaaaaacogttaaattatgac
ind	gctgatggtgcgcccctagaggtgtctcgtggactcrggaaggcaacogatttctcagc	hav	cogacatgcttccacaagcagcgaagctgctgctgcttgaaaaacogttaaattatgac
ori	gctgatggtgcgcccctagaggtgtctcgtggactcrggaaggcaacogatttctcagc	ind	cogacatgcttccacaagcagcgaagctgctgctgcttgaaaaacogttaaattatgac
ori	gctgatggtgcgcccctagaggtgtctcgtggactcrggaaggcaacogatttctcagc	ori	cogacatgcttccacaagcagcgaagctgctgctgcttgaaaaacogttaaattatgac
tur	gctgatggtgcgcccctagaggtgtctcgtggactcrggaaggcaacogatttctcagc	su	cogacatgcttccacaagcagcgaagctgctgctgcttgaaaaacogttaaattatgac
su	gctgatggtgcgcccctagaggtgtctcgtggactcrggaaggcaacogatttctcagc	tur	cogacatgcttccacaagcagcgaagctgctgctgcttgaaaaacogttaaattatgac
vir	gctgatggtgcgcccctagaggtgtctcgtggactcrggaaggcaacogatttctcagc	tur	cogacatgcttccacaagcagcgaagctgctgctgcttgaaaaacogttaaattatgac
*****	*****	*****	*****
brya	tatgaggtgctatctgtagtgctactctcgtgctctttcttccagcgggccaagtca	brya	cctaggacaacaacaacagaactatgctcctaattagcgggtgtggcaaaagatcaacgaa
vu	tatgaggtgctatctgtagtgctactctcgtgctctttcttccagcgggccaagtca	vir	cctaggacaacaacaacagaactatgctcctaattagcgggtgtggcaaaagatcaacgaa
hav	tatgaggtgctatctgtagtgctactctcgtgctctttcttccagcgggccaagtca	vu	cctaggacaacaacaacagaactatgctcctaattagcgggtgtggcaaaagatcaacgaa
ind	tatgaggtgctatctgtagtgctactctcgtgctctttcttccagcgggccaagtca	hav	cctaggacaacaacaacagaactatgctcctaattagcgggtgtggcaaaagatcaacgaa
ori	tatgaggtgctatctgtagtgctactctcgtgctctttcttccagcgggccaagtca	ind	cctaggacaacaacaacagaactatgctcctaattagcgggtgtggcaaaagatcaacgaa
ori	tatgaggtgctatctgtagtgctactctcgtgctctttcttccagcgggccaagtca	ori	cctaggacaacaacaacagaactatgctcctaattagcgggtgtggcaaaagatcaacgaa
tur	tatgaggtgctatctgtagtgctactctcgtgctctttcttccagcgggccaagtca	su	cctaggacaacaacaacagaactatgctcctaattagcgggtgtggcaaaagatcaacgaa
su	tatgaggtgctatctgtagtgctactctcgtgctctttcttccagcgggccaagtca	tur	cctaggacaacaacaacagaactatgctcctaattagcgggtgtggcaaaagatcaacgaa
vir	tatgaggtgctatctgtagtgctactctcgtgctctttcttccagcgggccaagtca	tur	cctaggacaacaacaacagaactatgctcctaattagcgggtgtggcaaaagatcaacgaa
*****	*****	*****	*****

Рис. 21. Множественное выравнивание фрагментов ТА.1 (слева) и ТА.2 (справа), полученных на матрицах 8 сортов *N. tabacum*, звездочками отмечены совпадающие нуклеотиды.

Таким образом, можно сделать вывод, что низкий уровень генетического разнообразия внутри вида *N. tabacum* показан как для трансгенных последовательностей, приобретенных в результате горизонтального переноса, так и для последовательностей «растительного происхождения». Полученные результаты согласуются с показанными ранее другими методами (Lewis, Nicholson, 2007).

hav	gcgctcaactgtgcccagcgataaactcaccggtcttctcgaacaacttgcgggagcggaa	brya	togatcgccaataaaggttgtatagagtgccctagcgttccggactacaagaggaacacgc
ind	gcgctcaactgtgcccagcgataaactcaccggtcttctcgaacaacttgcgggagcggaa	hav	togatcgccaataaaggttgtatagagtgccctagcgttccggactacaagaggaacacgc
ori	gcgctcaactgtgcccagcgataaactcaccggtcttctcgaacaacttgcgggagcggaa	ind	togatcgccaataaaggttgtatagagtgccctagcgttccggactacaagaggaacacgc
tur	gcgctcaactgtgcccagcgataaactcaccggtcttctcgaacaacttgcgggagcggaa	ori	togatcgccaataaaggttgtatagagtgccctagcgttccggactacaagaggaacacgc
vir	gcgctcaactgtgcccagcgataaactcaccggtcttctcgaacaacttgcgggagcggaa	su	togatcgccaataaaggttgtatagagtgccctagcgttccggactacaagaggaacacgc
su	gcgctcaactgtgcccagcgataaactcaccggtcttctcgaacaacttgcgggagcggaa	tur	togatcgccaataaaggttgtatagagtgccctagcgttccggactacaagaggaacacgc
vu	gcgctcaactgtgcccagcgataaactcaccggtcttctcgaacaacttgcgggagcggaa	vir	togatcgccaataaaggttgtatagagtgccctagcgttccggactacaagaggaacacgc
brya	gcgctcaactgtgcccagcgataaactcaccggtcttctcgaacaacttgcgggagcggaa	vu	togatcgccaataaaggttgtatagagtgccctagcgttccggactacaagaggaacacgc
*****	*****	*****	*****
hav	cgccacaaaatagttgtaacaagcaggctcaagagatccgggtgcagcaccoccatcaac	brya	gggggcaccacacactatcatcttctgtacctctatcttaatttaagcattaaagttcg
ind	cgccacaaaatagttgtaacaagcaggctcaagagatccgggtgcagcaccoccatcaac	hav	gggggcaccacacactatcatcttctgtacctctatcttaatttaagcattaaagttcg
ori	cgccacaaaatagttgtaacaagcaggctcaagagatccgggtgcagcaccoccatcaac	ind	gggggcaccacacactatcatcttctgtacctctatcttaatttaagcattaaagttcg
tur	cgccacaaaatagttgtaacaagcaggctcaagagatccgggtgcagcaccoccatcaac	ori	gggggcaccacacactatcatcttctgtacctctatcttaatttaagcattaaagttcg
vir	cgccacaaaatagttgtaacaagcaggctcaagagatccgggtgcagcaccoccatcaac	su	gggggcaccacacactatcatcttctgtacctctatcttaatttaagcattaaagttcg
su	cgccacaaaatagttgtaacaagcaggctcaagagatccgggtgcagcaccoccatcaac	tur	gggggcaccacacactatcatcttctgtacctctatcttaatttaagcattaaagttcg
vu	cgccacaaaatagttgtaacaagcaggctcaagagatccgggtgcagcaccoccatcaac	vir	gggggcaccacacactatcatcttctgtacctctatcttaatttaagcattaaagttcg
brya	cgccacaaaatagttgtaacaagcaggctcaagagatccgggtgcagcaccoccatcaac	vu	gggggcaccacacactatcatcttctgtacctctatcttaatttaagcattaaagttcg
*****	*****	*****	*****
hav	aatggagttgagtttctcaaacgctcgggtcttgcacgctccaggtatggtggcag	brya	acatattccgctatcttctgtatggtcttattagttctgatacaactcgcggaacacaca
ind	aatggagttgagtttctcaaacgctcgggtcttgcacgctccaggtatggtggcag	hav	acatattccgctatcttctgtatggtcttattagttctgatacaactcgcggaacacaca
ori	aatggagttgagtttctcaaacgctcgggtcttgcacgctccaggtatggtggcag	ind	acatattccgctatcttctgtatggtcttattagttctgatacaactcgcggaacacaca
tur	aatggagttgagtttctcaaacgctcgggtcttgcacgctccaggtatggtggcag	ori	acatattccgctatcttctgtatggtcttattagttctgatacaactcgcggaacacaca
vir	aatggagttgagtttctcaaacgctcgggtcttgcacgctccaggtatggtggcag	su	acatattccgctatcttctgtatggtcttattagttctgatacaactcgcggaacacaca
su	aatggagttgagtttctcaaacgctcgggtcttgcacgctccaggtatggtggcag	tur	acatattccgctatcttctgtatggtcttattagttctgatacaactcgcggaacacaca
vu	aatggagttgagtttctcaaacgctcgggtcttgcacgctccaggtatggtggcag	vir	acatattccgctatcttctgtatggtcttattagttctgatacaactcgcggaacacaca
brya	aatggagttgagtttctcaaacgctcgggtcttgcacgctccaggtatggtggcag	vu	acatattccgctatcttctgtatggtcttattagttctgatacaactcgcggaacacaca
*****	*****	*****	*****
hav	cgaggttacttttgccgctgaattgctactgttgcaaacagggcagcgtctccacataca	brya	tagtccacactcatcttctgacttcaacatcttctgtgcccgtccggcaagataccgaaaa
ind	cgaggttacttttgccgctgaattgctactgttgcaaacagggcagcgtctccacataca	hav	tagtccacactcatcttctgacttcaacatcttctgtgcccgtccggcaagataccgaaaa
ori	cgaggttacttttgccgctgaattgctactgttgcaaacagggcagcgtctccacataca	ind	tagtccacactcatcttctgacttcaacatcttctgtgcccgtccggcaagataccgaaaa
tur	cgaggttacttttgccgctgaattgctactgttgcaaacagggcagcgtctccacataca	ori	tagtccacactcatcttctgacttcaacatcttctgtgcccgtccggcaagataccgaaaa
vir	cgaggttacttttgccgctgaattgctactgttgcaaacagggcagcgtctccacataca	su	tagtccacactcatcttctgacttcaacatcttctgtgcccgtccggcaagataccgaaaa
su	cgaggttacttttgccgctgaattgctactgttgcaaacagggcagcgtctccacataca	tur	tagtccacactcatcttctgacttcaacatcttctgtgcccgtccggcaagataccgaaaa
vu	cgaggttacttttgccgctgaattgctactgttgcaaacagggcagcgtctccacataca	vir	tagtccacactcatcttctgacttcaacatcttctgtgcccgtccggcaagataccgaaaa
brya	cgaggttacttttgccgctgaattgctactgttgcaaacagggcagcgtctccacataca	vu	tagtccacactcatcttctgacttcaacatcttctgtgcccgtccggcaagataccgaaaa
*****	*****	*****	*****
hav	ccagctctgctgttgatcgatccatatacatataaagaagcttgtcgacgtcacagctctg	brya	atcaacttttctctgtgtctgacatttttaccacatatttttaacgattttttgttta
ind	ccagctctgctgttgatcgatccatatacatataaagaagcttgtcgacgtcacagctctg	hav	atcaacttttctctgtgtctgacatttttaccacatatttttaacgattttttgttta
ori	ccagctctgctgttgatcgatccatatacatataaagaagcttgtcgacgtcacagctctg	ind	atcaacttttctctgtgtctgacatttttaccacatatttttaacgattttttgttta
tur	ccagctctgctgttgatcgatccatatacatataaagaagcttgtcgacgtcacagctctg	ori	atcaacttttctctgtgtctgacatttttaccacatatttttaacgattttttgttta
vir	ccagctctgctgttgatcgatccatatacatataaagaagcttgtcgacgtcacagctctg	su	atcaacttttctctgtgtctgacatttttaccacatatttttaacgattttttgttta
su	ccagctctgctgttgatcgatccatatacatataaagaagcttgtcgacgtcacagctctg	tur	atcaacttttctctgtgtctgacatttttaccacatatttttaacgattttttgttta
vu	ccagctctgctgttgatcgatccatatacatataaagaagcttgtcgacgtcacagctctg	vir	atcaacttttctctgtgtctgacatttttaccacatatttttaacgattttttgttta
brya	ccagctctgctgttgatcgatccatatacatataaagaagcttgtcgacgtcacagctctg	vu	atcaacttttctctgtgtctgacatttttaccacatatttttaacgattttttgttta
*****	*****	*****	*****
hav	atcgaggtccagcaccgggtaaccgagcttttccagaggtcgagatcaccgcggtagtc	brya	ttggttttggccaagaagaagcaataaaaggcggataaaaaacgtacatgtagtagagatt
ind	atcgaggtccagcaccgggtaaccgagcttttccagaggtcgagatcaccgcggtagtc	hav	ttggttttggccaagaagaagcaataaaaggcggataaaaaacgtacatgtagtagagatt
ori	atcgaggtccagcaccgggtaaccgagcttttccagaggtcgagatcaccgcggtagtc	ind	ttggttttggccaagaagaagcaataaaaggcggataaaaaacgtacatgtagtagagatt
tur	atcgaggtccagcaccgggtaaccgagcttttccagaggtcgagatcaccgcggtagtc	ori	ttggttttggccaagaagaagcaataaaaggcggataaaaaacgtacatgtagtagagatt
vir	atcgaggtccagcaccgggtaaccgagcttttccagaggtcgagatcaccgcggtagtc	su	ttggttttggccaagaagaagcaataaaaggcggataaaaaacgtacatgtagtagagatt
su	atcgaggtccagcaccgggtaaccgagcttttccagaggtcgagatcaccgcggtagtc	tur	ttggttttggccaagaagaagcaataaaaggcggataaaaaacgtacatgtagtagagatt
vu	atcgaggtccagcaccgggtaaccgagcttttccagaggtcgagatcaccgcggtagtc	vir	ttggttttggccaagaagaagcaataaaaggcggataaaaaacgtacatgtagtagagatt
brya	atcgaggtccagcaccgggtaaccgagcttttccagaggtcgagatcaccgcggtagtc	vu	ttggttttggccaagaagaagcaataaaaggcggataaaaaacgtacatgtagtagagatt
*****	*****	*****	*****

Рис. 22. Множественное выравнивание фрагментов ТВ.1 (слева) и ТВ.2 (справа), полученных на матрицах 8 сортов *N. tabacum*, звездочками отмечены совпадающие нуклеотиды

Рассматривают несколько возможных причин для формирования подобного уровня полиморфизма. Во-первых, недавнее время образования вида *N. tabacum*. Действительно, культурный табак относится к молодым полиплоидам рода *Nicotiana*, чей возраст не превышает 200 000 лет (Leitch et al., 2008), в связи с чем его геном мог не успеть накопить должное количество мутаций.

```

PTM2_BasmaXanth atggaagtcataatctaccaacacaaaatggctctaccatcttcaagagtggtgccattccc
PTM2_K326 atggaagtcataatctaccaacacaaaatggctctaccatcttcaagagtggtgccattccc
PMT2_TN90 atggaagtcataatctaccaacacaaaatggctctaccatcttcaagagtggtgccattccc
*****

PTM2_BasmaXanth atgaatggccaccaggaatggcacttccaacacaaaaacggccacaagaatgggacttcc
PTM2_K326 atgaatggccaccataatggcacttccaacacaaaaacggccacaagaatgggacttcc
PMT2_TN90 atgaatggccaccataatggcacttccaacacaaaaacggccacaagaatgggacttcc
*****

PTM2_BasmaXanth gaacaacagaacgggacaatcagccttgataatggcaacgagctactgggaaactccaat
PTM2_K326 gaacaacagaacgggacaatcagccttgataatggcaacgagctactgggaaactccaat
PMT2_TN90 gaacaacagaacgggacaatcagccttgataatggcaacgagctactgggaaactccaat
*****

PTM2_BasmaXanth tgtattaagcctggttggtttcagagtttagcgcattatggccagggttagtactgagaa
PTM2_K326 tgtattaagcctggttggtttcagagtttagcgcattatggccagggttagtactgagaa
PMT2_TN90 tgtattaagcctggttggtttcagagtttagcgcattatggccagggttagtactgagaa
*****

.....

PTM2_BasmaXanth cgtcaagtcctttaagggtctctgtcaactatgcttggactactgttccaacatatccBacg
PTM2_K326 cgtcaagtcctttaagggtctctgtcaactatgcttggactactgttccaacatatccaacg
PMT2_TN90 cgtcaagtcctttaagggtctctgtcaactatgcttggactactgttccaacatatccaacg
*****

PTM2_BasmaXanth tattttctctctctctctctctctataaaaatggaagtttgattctataattgtca
PTM2_K326 tatttt--tctctctctctctctctataaaaatggaagtttgattctataattgtca
PMT2_TN90 tatttt--tctctctctctctctctataaaaatggaagtttgattctataattgtca
*****

PTM2_BasmaXanth agaatggagaatcagttccaagaaaaaccaacttcttttcttttactcttcaaggtatt
PTM2_K326 agaatggagaatcagttccaagaaaaaccaacttcttttcttttactcttcaaggtatt
PMT2_TN90 agaatggagaatcagttccaagaaaaaccaacttcttttcttttactcttcaaggtatt
*****

PTM2_BasmaXanth gtgtttaattttttcaactgatatgatcaattattttgatttcagcgggtgtgattggt
PTM2_K326 gtgtttaattttttcaactgatatgatcaattattttgatttcagcgggtgtgattggt
PMT2_TN90 gtgtttaattttttcaactgatatgatcaattattttgatttcagcgggtgtgattggt
*****

PTM2_BasmaXanth tataatgctctgctctactgaaggaccagaaattgacttcaagaatccagtaaatccaatt
PTM2_K326 tataatgctctgctctactgaaggaccagaaattgacttcaagaatccagtaaatccaatt
PMT2_TN90 tataatgctctgctctactgaaggaccagaaattgacttcaagaatccagtaaatccaatt
*****

PTM2_BasmaXanth gacaagagacagctcaagtcagtcctcaaatatgacacctctcaagttctacaactctgat
PTM2_K326 gacaagagacagctcaagtcagtcctcaaatatgacacctctcaagttctacaactctgat
PMT2_TN90 gacaagagacagctcaagtcagtcctcaaatatgacacctctcaagttctacaactctgat
*****

.....

PTM2_BasmaXanth gtaacttcataatctcacaatttctttttcctattgtactttatgttcttctcgtcaaattt
PTM2_K326 gtaacttcataatctcacaatttctttttcctattgtactttatgttcttctcgtcaaattt
PMT2_TN90 gtaacttcataatctcacaatttctttttcctattgtactttatgttcttctcgtcaaattt
*****

PTM2_BasmaXanth tataattaactcttttcaaattgtcttttttttttccagattcacaagcagcattca
PTM2_K326 tataattaactcttttcaaattgtc--tttttttttccagattcacaagcagcattca
PMT2_TN90 tataattaactcttttcaaattgtc--tttttttttccagattcacaagcagcattca
*****

PTM2_BasmaXanth ttttgccatctttcgccagaagtatgatcgagtcctaa
PTM2_K326 ttttgccatctttcgccagaagtatgatcgagtcctaa
PMT2_TN90 ttttgccatctttcgccagaagtatgatcgagtcctaa

```

Рис. 23. Множественное выравнивание гена *PMT2* в геномах *Nicotiana tabacum* сортов K326 (NCAA00000000.1, AWOJ00000000.1), TN90 (AYMY00000000.1) и Basma Xanthi (AWOK00000000.1), показаны фрагменты с различиями.

В качестве второй причины называют небольшое разнообразие исходных видов-предшественников, сформировавших вид *N. tabacum* (Lewis, Nicholson, 2007). В связи с отсутствием данных об уровнях полиморфизма ДНК в генофондах *N. tomentosiformis* и *N. sylvestris* в настоящий момент невозможно подтвердить или опровергнуть данную гипотезу. Направленная селекция также могла оказать влияние на формирование современного генетического разнообразия культурного табака. Снижение уровня разнообразия было показано на Коллекции зародышевой плазмы *Nicotiana* США, что свидетельствует в пользу данной гипотезы (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 разъяснению данного вопроса.

3.2.2. Анализ крупных перестроек в клТ-ДНК на примере ТА

Геном *Nicotiana tabacum* содержит три различающиеся по составу и возрасту клТ-ДНК (ТА, ТВ, TD) (рис. 4, 5, 7). При изучении клТ-ДНК *N. tabacum* Chen с коллегами отметила, что ряд сортов несет в центральной части ТА делецию длиной 9684 п.н., расположенную между двумя неполными последовательностями *orf13*. Вместо протяженного фрагмента агробактериального происхождения у сортов *Basma Drama*, *Samsoun*, и *Xanthi* в центральном районе ТА находится последовательность неизвестного происхождения длиной в 42 п.н., в то время как геномы *N. tabacum* сортов *Wisconsin 38* и *Havana 425* содержат полноразмерную ТА, как и предковый для культурного табака вид - *N. tomentosiformis* (см. рис. 13, 24). Таким образом, для ТА *N. tabacum* описано два структурных варианта - с делецией и без нее. В связи с тем, что для сортов культурного табака описан низкий уровень внутривидового

разнообразия, для проведения филогенетических исследований 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.

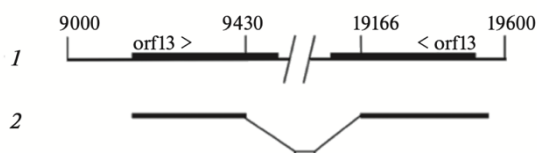


Рис. 24. схема последовательности ТА из *N. tomentosiformis* (KJ599826), цифрами отмечены координаты в контиге, черными прямоугольниками отмечен ген *orf13*, центральная часть не показана. (сверху) клТ-ДНК ТА без делеции, описана в сортах Wisconsin 38, Havana 425; (снизу) клТ-ДНК ТА с делецией, белым прямоугольником отмечена последовательность неизвестного происхождения длиной 42 пн, расположенная между двумя фрагментами *orf13*, описана у сортов Basma Drama, Samsoun, и Xanthi (Chen et al., 2014)

Ранее была описана структура ТА в 5 сортах *N. tabacum* (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. 14). В ходе работы для каждого сорта были получены фрагменты только с одним из сочетаний праймеров (см. рис. 25), 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). 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 появились 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.

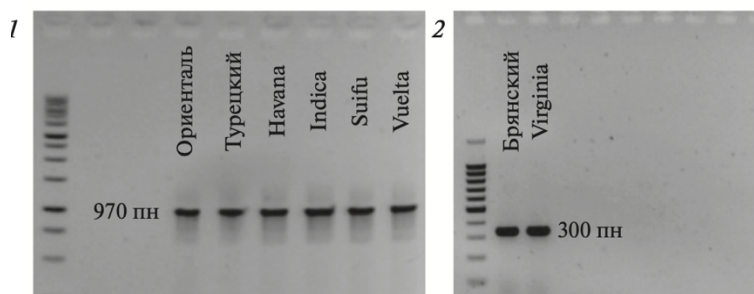


Рис. 25. Фрагменты, полученные на матрицах сортов *N. tabacum*: (1) – фрагменты, полученные с использованием праймеров F и innR, (2) – фрагменты, полученные с использованием праймеров F и R

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 также относят к классу наполнителя для сигар, а по Псаревой он является гибридом Виргинской и Мерилендской разновидностей Американского подвида. Таким образом, пока невозможно однозначно определить принадлежность 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. Сложно сказать, насколько стоит опираться на уже имеющиеся распределения сортов по группам при составлении новой классификации, поскольку непонятно, насколько корректно существующие группы отражают взаимоотношения сортов. Наличие данных о происхождении сортов было бы полезно для установления родительских форм сортов и для восстановления их родословных. 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, таких как, например, 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 - выявили принадлежность сортов К326, TN90 и сорта *Vasma Xanthi* к разным группам, с делецией и без делеции в центральном участке клТ-ДНК ТА. 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.

В результате анализа ряда сортов культурного табака, *N. tabacum*, был выявлен низкий уровень полиморфизма как для последовательностей «растительного происхождения», так и для фрагментов клТ-ДНК ТВ и ТА. При этом в ТА были отмечены структурные варианты, позволяющие разделить исследованные сорта *N. tabacum* на 2 группы по происхождению.

ЗАКЛЮЧЕНИЕ

Собранные в данной работе геномы видов *N. glauca* и *N. noctiflora*, относящихся к эволюционной ветви *Petunioides*, позволили провести детальный анализ клТ-ДНК этих видов. 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 же время присутствие в геноме единственной клТ-ДНК гТ свидетельствует об однократном акте агробактериальной трансформации *N. glauca*, 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* с секвенированными геномами, 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, Nicholson, 2007).

Молекулярный маркер на основе делеции в центральной части ТА, разработанный для *Nicotiana tabacum*, позволяет разделить сорта на две группы, и может быть использован как для изучения родственных 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 NnТ-DNA1 в геноме *N. noctiflora* и гТ 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. КлТ-ДНК gT является единственной в геноме *N. glauca*.
2. Геном *N. noctiflora* содержит две разные по составу клТ-ДНК (NnТ-ДНК1, и NnТ-ДНК2), вероятными источниками которых могут быть штаммы сходные с *Agrobacterium rhizogenes* 1724 и *Agrobacterium tumefaciens* CFBR1935.
3. В составе NnТ-ДНК1 интактными являются гены *iaaM*, *iaaH*, *acs*, в составе NnТ-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.

Список литературы

1. Вавилов Н. И. Мировые ресурсы сортов хлебных злаков, зерновых бобовых, льна и их использование в селекции/ Вавилов Н. И. - Издательство Академии наук СССР. 1957. - 464 стр.
2. Псарева Е. Н. Классификация *Nicotiana tabacum* L./ Псарева Е. Н.// Сборник НИР ВНИИ табака и махорки, Краснодар. - 1969. - Вып. 154.- С. 25–91
3. Хафизова Г. В. Изучение сайтов интеграции клеточной Т-ДНК у представителей различных секций рода *Nicotiana*/ Хафизова Г. В., Матвеева Т. В.// Молодежный научный форум: естественные и медицинские науки. Электронный сборник статей по материалам XV–XVI студенческой международной заочной научно-практической конференции. М.:Изд.«МЦНО». - 2014. - Т. - 8. N. 9. С. 15. URL:[http://www.nauchforum.ru/archive/MNF_nature/8-9\(15\).pdf](http://www.nauchforum.ru/archive/MNF_nature/8-9(15).pdf)
4. Akiyoshi D.E. T-DNA of *Agrobacterium tumefaciens* encodes an enzyme of cytokinin biosynthesis/ Akiyoshi D.E., Klee H., Amasino R.M., Nester E.W., Gordon M.P.// Proc. Natl Acad. Sci. USA. - 1984. - No 81. - P. 5994–5998
5. Aoki S. Sequence of the cellular T-DNA in the untransformed genome of *Nicotiana glauca* that is homologous to ORFs 13 and 14 of the Ri plasmid and analysis of its expression in genetic tumors of *N.glauca* x *N.langsdorffii* / Aoki S., Kawaoka A., Sekine M., Ichikawa T., Fujita T., Shinmyo A. and Syono K.// Mol. Gen. Genet. - 1994. - No 243- P.- 706 — 710
6. Aoki S. Function of Ngrol genes in the evolution of *Nicotiana glauca*: conservation of the function of NgORF13 and NgORF14 after ancient infection by an *Agrobacterium rhizogenes*-like ancestor/ Aoki, S., and Syono, K.// *Plant Cell Physiol.* 1999. - N. 40 - P. 222–230. 10.1093/oxfordjournals.pcp.a029531
7. Aoki S. Molecular phylogeny of *Nicotiana* (Solanaceae) based on nucleotide sequence of the matK gene/ Aoki S., Ito M.// *Plant Biol.* - 2000. - N. 2. - P. 316–324

8. Bally J. The extremophile *Nicotiana benthamiana* has traded viral defence for early vigour/ Bally J., Nakasugi, K., Jia, Jung F. H., Ho S. Y.W., Wong M., Paul C. M., Naim F., Wood C. C., Crowhurst R. N., Hellens R.P., Dale J. L., Waterhouse P. M.// *Nature Plants* 1. - 2015. - N. 15165 (2015). 10.1038/nplants.2015.165
9. Baranska M. Tobacco alkaloids analyzed by Raman spectroscopy and DFT calculations/ Baranska M., Dobrowolski J. C., Kaczor A., Chruszcz-Lipska K., Gorz K., & Rygula A.// *Journal of Raman Spectroscopy*. 2012. - V. 43. - N. 8. - P. 1065-1073. doi.org/10.1002/jrs.3127
10. Bland M.M. Comparison of the mitochondrial genome of *Nicotiana tabacum* with its progenitor species/ Bland M.M., Matzinger DF, Levings, CS, III// *Theor Appl Genet*. - 1985. - V. 69. -P. 535–541. 10.1007/BF00251100
11. Bombarely A. Deciphering the complex leaf transcriptome of the allotetraploid species *Nicotiana tabacum*: a phylogenomic perspective/ Bombarely A., Edwards K.D., Sanchez-Tamburrino J., Mueller L.A.// *BMC Genomics*. - 2012. - V. 13. P. 406. 10.1186/1471-2164-13-406
12. Bogani P. A Physiological and Molecular Analysis of the Genus *Nicotiana*/ Bogani, P., Liò, P., Intrieri, M. C., & Buiatti, M.// *Molecular Phylogenetics and Evolution*. - 1997. - V. 7- N. 1. - P. 62-70. 10.1006/mpev.1996.0356
13. Brandle J. Biotechnology: uses and applications in tobacco improvement/ Brandle J., Bai D. // *Tobacco: Production, Chemistry and Technology*. – 1999. – C. 49-65.
14. Brooks J.E. The mighty leaf / Brooks J.E. - Little Brown and Co., Boston MA. 1952. - 395ср.
15. Bulgakov V. Functions of rol genes in plant secondary metabolism/ Bulgakov V.// *Biotechnology advances*. - 2008. - V. 26. - N. 4. - P. 318-324.
16. Burr T. J. Crown gall of grape: biology of *Agrobacterium vitis* and the development of disease control strategies/ Burr T. J., Bazzi, C., Süle, S., & Otten, L.// *Plant disease*. - 1998. - N. 82. - I. 12. - P. 1288-1297

17. Byrne M. Birth of a biome: insights into the assembly and maintenance of the Australian arid zone biota/ Byrne M, Yeates DK, Joseph L, Kearney M, Bowler J, Williams MAJ, Cooper S, Donnellan SC, Keogh JS, Leys R, Melville J, Murphy DJ, Proch N, Wyrwoll K- H.// *Mol Ecol.* - 2008. - V. 17 - P. 4398–4417
18. Cameron D. R. Inheritance in *Nicotiana tabacum*. XXIV. Intraspecific differences in chromosome structure/ Cameron D. R. // *Genetics.* – 1952. – T. 37. – №. 3. – C. 288.
19. Cardarelli M. *Agrobacterium rhizogenes* T-DNA genes capable of inducing hairy root phenotype/ Cardarelli M., Mariotti D., Pomponi M., Spanò L., Capone I., Costantino P.// *Mol Gen Genet.* - 1987. - No 209. - P. 475 – 480
20. Charest P.J. Virulence of *Agrobacterium tumefaciens* strains with *Brassica napus* and *Brassica juncea*/ Charest, P.J., Iyer, V.N. & Miki, B.L.// *Plant Cell Reports.* 1989. - V. 8. - P. 303–306. 10.1007/BF00274136
21. Chase M.W. Molecular Systematics, GISH and the Origin of Hybrid Taxa in *Nicotiana* (Solanaceae)/ Chase M.W., Knapp S., Cox A.V., Clarkson J.J., Butsko Y., Joseph J., Savolainen V., Parokony A.S.// *Annals of Botany.* 2003. - V. 92. - N. 1. - P. 107–127. 10.1093/aob/mcg087
22. Chase M.W. Unexpected diversity of Australian tobacco species (*Nicotiana* section *Suaveolentes*, Solanaceae)/ Chase MW, Christenhusz MJM, Conran JG, Dodsworth S, Medeiros Nollet, de Assis F, Felix LP, Fay MF// *Curtis' Bot Mag.* - 2018 - V. 35 - P. 212–227
23. Chen K. Root-specific expression of opine genes and opine accumulation in some cultivars of the naturally occurring genetically modified organism *Nicotiana tabacum*/ Chen, K., de Borne, F. D., Julio, E., Obszynski, J., Pale, P., & Otten, L.// *The Plant Journal.* - 2016. - V. 87. - N. 3 - N. 258-269. 10.1111/tpj.13196
24. Chen K. Deep sequencing of the ancestral tobacco species *Nicotiana tomentosiformis* reveals multiple T-DNA inserts and a complex evolutionary history of natural

- transformation in the genus *Nicotiana*/ Chen K., Dorlhac de Borne F., Szegedi E., Otten L.// *Plant Journal*. - 2014. - V. 80. - No. 4. - P. 669–82. 10.1111/tpj.12661
- 25.Chen K. Natural *Agrobacterium* transformants: recent results and some theoretical considerations/ Chen, K., Otten, L. // *Frontiers in plant science*. 2017 - V. 8. - 1600. doi.org/10.3389/fpls.2017.01600
- 26.Chen K. Organization of the TC and TE cellular T-DNA regions in *Nicotiana otophora* and functional analysis of three diverged TE-6b genes/ Chen K., Dorlhac de Borne F., Siervo N., Ivanov N. V., Alouia M., Koechler S., & Otten L.// *The Plant Journal*. - 2018. - V. 94 - N. 2. - P. 274-287. 10.1111/tpj.13853
- 27.Clarkson J. J. *Phylogenetic relationships in Nicotiana (Solanaceae) inferred from multiple plastid DNA regions*/ Clarkson, J. J., Knapp, S., Garcia, V. F., Olmstead, R. G., Leitch, A. R., & Chase, M. W.// *Molecular Phylogenetics and Evolution*. 2004. - V. 33. - N. 1. - P. 75–90. 10.1016/j.ympev.2004.05.002
- 28.Clarkson J. J. Long-term genome diploidization in allopolyploid *Nicotiana* section *Repandae* (Solanaceae)/ Clarkson JJ, Lim KY, Kovarik A, Chase MW, Knapp S, Leitch AR.// *New Phytol*. 2005 - N. 168 - P. 241–252
- 29.Clarkson JJ. Time- calibrated phylogenetic trees establish a lag between polyploidisation and diversification in *Nicotiana* (Solanaceae)/ Clarkson JJ, Dodsworth S, Chase MW.// *Plant Syst Evol*. 2017 - V. 303. - P. 1001–1012
- 30.Clayton E. E. The genetics and breeding progress in tobacco during the last 50 years/ Clayton E. E.// *Agron J*. 1958 - V. 50. - P. 352–356
31. Crisp M.D. Phylogenetic biome conservatism on a global scale/ Crisp MD, Arroyo MTK, Cook LG, Gandolfo MA, Jordan GJ, McGlone MS, Weston PH, Westoby M, Wilf P, Linder HP.// *Nature*. 2009 - V. 458. - P. 754–756
- 32.Darvishzadeh R. *Genetic variation in oriental tobacco (Nicotiana tabacum L.) by agro-morphological traits and simple sequence repeat markers*/ Darvishzadeh, R.,

- Mirzaei, L., Maleki, H. H., Laurentin, H., & Alavi, S. R.// *Revista Ciência Agronômica*. - 2013 - V. 44. -N. 2. - P. 347–355. 10.1590/S1806-66902013000200018
33. Davis D. L. Tobacco: production, chemistry and technology/ Davis DL, Nielsen MT. - Blackwell Science Ltd., 1999 - 467 p
34. Dodsworth S. Is post- polyploidization diploidization the key to the evolutionary success of angiosperms? / Dodsworth S, Chase MW, Leitch AR// *Bot J Linn Soc*. 2016. - N. 180 - P. 1–5
35. Edwards K. D. A reference genome for *Nicotiana tabacum* enables map-based cloning of homeologous loci implicated in nitrogen utilization efficiency/ Edwards, K. D., Fernandez-Pozo, N., Drake-Stowe, K., Humphry, M., Evans, A. D., Bombarely, A., Allen, Hurst R., White B., Kernodle S. P., Bromley J. R., Sanchez-Tamburrino J. P., Lewis R. S., Mueller, L. A.// *BMC genomics* 2017. - V. 18. - N. 1. - P. 1-14. 10.1186/s12864-017-3791-6
36. Elliott P. E. Evaluation of tobacco germplasm for seedling resistance to stem rot and target spot caused by *Thanatephorus cucumeris*/ Elliott, P. E., Lewis, R. S., Shew, H. D., Gutierrez, W. A., and Nicholson, J. S.// *Plant Dis*. 2008. - V. 92. - P. 425–430. 10.1094/PDIS-92-3-0425
37. Escobar M. A. RNAi-mediated oncogene silencing confers resistance to crown gall tumorigenesis/ Escobar, M. A., Civerolo, E. L., Summerfelt, K. R., and Dandekar, A. M.// *Proc. Natl. Acad. Sci. U.S.A.* - 2001. - N. 98. - P. 13437–13442. 10.1073/pnas.241276898
38. Furner I.J. An *Agrobacterium* transformation in the evolution of the genus *Nicotiana*/ Furner I.J., Huffman G.A., Amasino R.M. // *Nature*. - 1986. - V. 319. - P. 422 – 427
39. Garner W.W. Superior germplasm in tobacco/ Garner WW, Allard H, Clayton EE// In: 1936 Yearbook of agriculture. USDA, Washington, DC. 1936. - pp 785–830
40. Gavilano L. B. *Genetic Engineering of Nicotiana tabacum for Reduced Nicotine Content*/ Gavilano, L. B., Coleman, N. P., Burnley, L.-E., Bowman, M. L.,

- Kalengamaliro, N. E., Hayes, A., Bush, L., Siminszky, B.// *Journal of Agricultural and Food Chemistry*. 2006 - V. 54. - N. 24. - P. 9071–9078. 10.1021/jf0610458
41. Gavrilova V. A. Genealogy of the Sunflower Lines Created on the Basis of Russian Varieties/ Gavrilova, V. A. and Anisimova, I. N.// *Helia*. - 2017 - v. 40. - N. 67. - pp. 133-146. 10.1515/helia-2017-0025
42. Gerstel D.U. Tobacco/ Gerstel DU, Sisson VA. - In: Smartt J, Simmonds NW (eds) *Evolution of crop plants*, 2nd edn. John Wiley & Sons, Inc., New York. 1995. - pp 458–463
43. Gerstel D. U. Segregation in new allopolyploids of *Nicotiana*. II. Discordant ratios from individual loci in $6\times$ (*N. tabacum* \times *N. sylvestris*)/ Gerstel D. U.// *Genetics*. 1963 - V. 48 - P. 677– 689. PMID: PMC1210503
44. Goodspeed T.H. The Genus *Nicotiana*: Origins, Relationships and Evolution of its Species in the Light of Their Distribution, Morphology and Cytogenetics/ Goodspeed T.H. - Chronica Botanica Co., Waltham, MA. - 1954. - 536 pp.
45. Gray J. C. Origin of *Nicotiana tabacum* L. detected by polypeptide composition of Fraction I protein/ Gray, J. C., Kung, S. D., Wildman, S. G., & Sheen, S. J.// *Nature*. - 1974. - V. 252. - N. 5480 - P. 226-227. 10.1038/252226a0
46. Guivarch A. Tissue-specific expression of the rolA gene mediates morphological changes in transgenic tobacco/ Guivarch, A., Carneiro, M., Vilaine, F., Pautot, V., and Criqui, D.// *Plant Mol. Biol.* 1996. - N. 30. - P. 125–134. 10.1007/BF00017807
47. Haoyu C. Haplotype-resolved de novo assembly with phased assembly graphs/ Haoyu Cheng, Gregory T Concepcion, Xiaowen Feng, Haowen Zhang, Heng Li// *Nature Methods*. - 2021. - N. 18. - p. 170–175. 10.1038/s41592-020-01056-5.
48. Hughes C.E. Island radiation on a continental scale: exceptional rates of plant diversification after uplift of the Andes/ Hughes CE, Eastwood R *Proc Nat Acad Sci USA*. - 2006. - N. 103. - P. 10334–10339. 10.1073/ pnas.0601928103

49. Intriери M.C. The horizontal transfer of *Agrobacterium rhizogenes* genes and the evolution of the genus *Nicotiana*/ Intriери M.C., Buiatti M.// *Molecular Phylogenetics and evolution*. - 2001. - V. 20. - P. 100 – 110
50. Iskandar H. M. Comparison of reference genes for quantitative real-time polymerase chain reaction analysis of gene expression in sugarcane/ Iskandar, H. M., Simpson, R. S., Casu, R. E., Bonnett, G. D., Maclean, D. J., & Manners, J. M.// *Plant Molecular Biology Reporter*. - 2004. - V. 22. - N. 4. - P. 325-337.
51. Jakob S. S. The considerable genome size variation of *Hordeum* species (Poaceae) is linked to phylogeny, life form, ecology, and speciation rates/ Jakob S. S., Meister A., Blattner F. R. // *Molecular Biology and Evolution*. – 2004. – T. 21. – №. 5. – C. 860-869
52. Jokat W. Timing and geometry of early Gondwana breakup/ Jokat W., Boebel, T., König, M., & Meyer, U. // *Journal of Geophysical Research: Solid Earth*. – 2003. – T. 108. – №. B9. 10.1029/2002JB001802
53. Katoh MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization/ Katoh, Rozewicki, Yamada// *Briefings in Bioinformatics*. - 2019. - N. 20. - P. 1160-1166.
54. Kelly L. J. Intragenic recombination events and evidence for hybrid speciation in *Nicotiana* (Solanaceae)/ Kelly LJ, Leitch AR, Clarkson JJ, Hunter RB, Knapp S, Chase MW// *Mol Biol Evol*. - 2010. - N. 27. - P. 781–799
55. Kelly L.J. Reconstructing the complex evolutionary origin of wild allopolyploid tobaccos (*Nicotiana* section *Suaveolentes*)/ Kelly LJ, Leitch AR, Clarkson JJ, Knapp S, Chase MW// *Evolution*. - 2013 - N. 76 - P. 80–94
56. Kenton A. Characterization of the *Nicotiana tabacum* L. genome by molecular cytogenetics/ Kenton, A., Parokony, A. S., Gleba, Y. Y., & Bennett, M. D.// *MGG Molecular & General Genetics*. - 1993 - V. 240 - N. 2 - P. 159–169. 10.1007/bf00277053

57. Kishore K. Monograph of tobacco (*Nicotiana tabacum*)/ Kishore K. // *Indian Journal of Drugs*. – 2014. – T. 2. – №. 1. – C. 5-23.
58. Knapp S. Biodiversity of *Nicotiana* (Solanaceae)/ Knapp S. // *The Tobacco Plant Genome*. – Springer, Cham. - 2020. – C. 21-41. 10.1007/978-3-030-29493-9_2
59. Knapp S. *Nomenclatural Changes and a New Sectional Classification in Nicotiana (Solanaceae)*/ Knapp, S., Chase, M. W., & Clarkson, J. J. // *Taxon*. - 2004. - V. 53 - N. 1. - 73. 10.2307/4135490
60. Koltunow A. M. Expression of *rolB* in apomictic *Hieracium piloselloides* Vill. causes ectopic meristems in planta and changes in ovule formation, where apomixis initiates at higher frequency/ Koltunow, A. M., Johnson, S. D., Lynch, M., Yoshihari, T., and Costantino, P. // *Planta*. - 2001. - N. 214. P. 196–205. 10.1007/s004250100612
61. Kovarik A. Concerted evolution of 18-5.8-26S rDNA repeats in *Nicotiana* allotetraploids/ Kovarik A, Matyasek R, Lim KY, Skalická K, Koukalová B, Knapp S, Chase M, Leitch AR // *Biol J Linn Soc*. - 2004. - V.82. - N. 4. - P. 615–625
62. Koukalova B. Family of highly repeated DNA sequences of *Nicotiana tabacum*/ Koukalova, B., Reich, J., Matyášek, R., Kuhrova V., Bezděk M. A Bam H.I. // *Theoret. Appl. Genetics*. - 1989. - N. 78. - p. 77–80. 10.1007/BF00299757
63. Kyndt T. The genome of cultivated sweet potato contains *Agrobacterium* T-DNAs with expressed genes: An example of a naturally transgenic food crop/ Kyndt T., Quispe D., Hong Zhai, Jarret R., Ghislain M., Liu Q., Gheysen G., and Jan F. Kreuze // *PNAS*. - 2015. - V.112. - No 18. - P. 5844 – 5849
64. Lagomarsino L. The abiotic and biotic drivers of rapid diversification in Andean bellflowers/ Lagomarsino L, Condamine FL, Antonelli A, Mulch A, Davis CC // *New Phytol*. - 2016. - N. 210 - P. 1430–1442. 10.1111/nph.13920
65. Lee C. B. *Post-pollination hybridization barriers in Nicotiana section Alatae*/ Lee C. B., Page, L. E., McClure, B. A., & Holtsford, T. P. // *Sexual Plant Reproduction*. 2008.

- V. 21 - N. 3. - P. 183–195. 10.1007/s00497-008-0077-9 doi.org/10.1007/s00497-008-0077-9
66. Leitch I. L. The ups and downs of genome size evolution in polyploid species of *Nicotiana* (Solanaceae)/ Leitch IL, Hanson L, Lim YK, Kovarik A, Clarkson JJ, Chase MW, Leitch AR.// *Ann Bot.* - 2008. - N. 101 - P. 805–814
67. Lemcke K. Gain of function assays identify non-rol genes from *Agrobacterium rhizogenes* TL-DNA that alter plant morphogenesis or hormone sensitivity/ Lemcke, K., and Schmulling, T.// *Plant J.* - 1998. - N. 15. - P. 423–433. 10.1046/j.1365-313X.1998.00223.x
68. Lewis R. Aspects of the evolution of *Nicotiana tabacum* L. and the status of the United States *Nicotiana* Germplasm Collection/ Lewis R., Nicholson J.// *Genet. Resour. Crop Evol.* - 2007. - V. 54. - P. 727–740. 10.1007/s10722-006-0024-2
69. Li H. Minimap2: pairwise alignment for nucleotide sequences/ Li H.// *Bioinformatics.* - 2018. - N. 34. - P. 3094–3100.
70. Li H. Minimap and miniasm: fast mapping and de novo assembly for noisy long sequences/ Li H.// *Bioinformatics.* - 2016. - N. 32. - P. 2103-2110. 10.1093/bioinformatics/btw152
71. Lim K. Y. Molecular and cytogenetic analyses and phylogenetic studies in the *Nicotiana* section *Tomentosae*/ Lim KY, Matyasek R, Lichtenstein CP, Leitch AR.// *Chromosoma.* - 2000. - N. 109. - P. 245–258
72. Lim K.Y. Sequence of events leading to near-complete genome turnover in allopolyploid *Nicotiana* within five million years/ Lim KY, Kovarik A, Matyasek R, Chase MW, Clarkson JJ, Grandbastien MA, Leitch AR.// *New Phytol.* - 2007. - N. 175. - P. 756–763
73. Lim K.Y. Comparative genomics and repetitive sequence divergence in the species of diploid *Nicotiana* section *Alatae*/ Lim KY, Kovarik A, Matyasek R, Chase MW,

- Knapp S, McCarthy E, Clarkson JJ, Leitch AR.// *Plant J.* - 2006. - N. 48. - P. 907–919
- 74.Lim K. Y. Gene conversion of ribosomal DNA in *Nicotiana tabacum* is associated with undermethylated, decondensed and probably active gene units/ Lim K.Y., Matyasek R., M. Bezděek, C.P. Lichtenstein & A.R. Leitch// *Chromosoma.* - 2000. - N. 109. - P. 161-172. 10.1007/s004120050424
- 75.Lim K.Y. Genome evolution in allotetraploid *Nicotiana*/ Lim K.Y., Matyasek R., Kovarik A., Leitch A. R.// *Biological Journal of the Linnean Society* - 2004. - V. 82. - I. 4. - P. 599–606. 10.1111/j.1095-8312.2004.00344.x
- 76.Litton C.C. Outcrossing in burley tobacco/ Litton CC, Stokes GW// *Tob Sci.* 1964. - N. 8. - P. 113–115
- 77.Long N. Sequencing and characterization of leaf transcriptomes of six diploid *Nicotiana* species/ Long, N., Ren, X., Xiang, Z.// *J of Biol Res-Thessaloniki.* - 2016. - N. 23. - 6. 10.1186/s40709-016-0048-5
- 78.Ma L. S. *Agrobacterium tumefaciens* deploys a superfamily of type VI secretion DNase effectors as weapons for interbacterial competition in planta/ Ma, L. S., Hachani, A., Lin, J. S., Filloux, A., & Lai, E. M.// *Cell host & microbe.* - 2014. - V. 16. - N. 1. - P. 94-104. 10.1016/j.chom.2014.06.002
- 79.Mallah G. S. Inheritance in *Nicotiana tabacum*. XVI. Structural differences among the chromosomes of a selected group of varieties/ Mallah G. S. // *Genetics.* – 1943. – T. 28. – №. 6. – C. 525.
- 80.Matveeva T.V. Horizontal Gene Transfer from Genus *Agrobacterium* to the Plant *Linaria* in Nature/ Matveeva T.V., Bogomaz D.I., Pavlova O.A., Nester E.W., Lutova L.A.// *MPMI.* - 2012. - V. 25. - No 12. - P. 1542 — 1551
- 81.Matveeva T. New naturally transgenic plants: 2020 update/ Matveeva T. // *Biological Communications.* – 2021. – T. 66. – №. 1. – C. 36-46. doi.org/10.21638/spbu03.2021.105

82. Matveeva T. V. Widespread occurrence of natural genetic transformation of plants by *Agrobacterium*/ Matveeva T. V., Otten L. // *Plant Molecular Biology*. – 2019. – T. 101. – №. 4. – C. 415-437. doi.org/10.1007/s11103-019-00913-y
83. McCarthy E. W. *Transgressive phenotypes and generalist pollination in the floral evolution of Nicotiana polyploids*/ McCarthy, E. W., Chase, M. W., Knapp, S., Litt, A., Leitch, A. R., & Le Comber, S. C. // *Nature Plants*. - 2016. - V. 2. - N. 9. - P. 16119. doi:10.1038/nplants.2016.119
84. McClintock B. The significance of responses of the genome to challenge/ McClintock B. // *Science*. - 1984. - V. 226 - N. 4676. - P. 792–801. doi:10.1126/science.15739260
85. Melayah D. Distribution of the Tnt1 retrotransposon family in the amphidiploid tobacco (*Nicotiana tabacum*) and its wild *Nicotiana* relatives/ Melayah D., Lim K.Y., Bonnivard E., Chalhoub B., Dorlhac De Borne F., Mhiri C., Leitch A. R., Grandbastien M.-A. // *Biological Journal of the Linnean Society*. - 2004. - V. 82. - I. 4. - P. 639–649. doi.org/10.1111/j.1095-8312.2004.00348.x
86. Merxmüller H. *Nicotiana in der Afrikanischen Namibein Pflanzengeographisches und Phylogenetisches Ratsel*/ Merxmüller H., Butler K.P. // *Mitteilungen aus der Botanischen Staatssammlung München*. - 1975. - N. 12. - P. 91–104
87. Milla S. R. Taxonomic relationships among *Arachis* sect. *Arachis* species as revealed by AFLP markers/ Milla SR, Isleib TG, Stalker HT. // *Genome*. - 2005. - N. 48. - P. 1–11
88. Mohajjel-Shoja, H., Clément, B., Perot, J., Alioua, M., & Otten, L. (2011). Biological activity of the *Agrobacterium* rhizogenes–derived *trnC* Gene of *Nicotiana tabacum* and its functional relation to other plast genes/ Mohajjel-Shoja, H., Clément, B., Perot, J., Alioua, M., & Otten, L. // *Molecular plant-microbe interactions*. - 2011. - V. 24. - N. 1. - P. 44-53.

89. Moon H.S. Use of transferable *Nicotiana tabacum* L. microsatellite markers for investigating genetic diversity in the genus *Nicotiana*/ Moon, H.S., Nicholson, J.S. and Lewis, R.S.// *Genome*. - 2008. - No 51. - P. 547 – 559
90. Moon H. S. Microsatellite-based analysis of tobacco (*Nicotiana tabacum* L.) genetic resources/ Moon H. S., Nifong, J. M., Nicholson, J. S., Heineman, A., Lion, K., Van der Hoeven, R., Lewis, R. S.// *Crop Science*. - 2009. - V. 49. - N. 6. - P. 2149-2159. 10.2135/cropsci2009.01.0024
91. S.A. Mousavi Revised phylogeny of *Rhizobiaceae*: proposal of the delineation of *Pararhizobium* gen. nov., and 13 new species combinations/ S.A. Mousavi, A. Willems, X. Nesme, P. de Lajudie, K. Lindström// *Syst. Appl. Microbiol.* - 2014. - N. 38. - pp. 84-90, 10.1016/j.syapm.2014.12.003
92. Murad L. The origin of tobacco's T genome is traced to a particular lineage within *Nicotiana tomentosiformis* (*Solanaceae*)/ Murad, L., Lim, K.Y., Christopodulou, V., Matyasek, R., Lichtenstein, C., Kovarik, A. and Leitch, A.// *Am. J. Bot.* - 2002. - No 89. - P. 921 – 928. 10.3732/ajb.89.6.921
93. Nagata T. Tobacco BY-2 cells: the present and beyond/ Nagata T., Sakamoto K., Shimizu T. // *In Vitro Cellular & Developmental Biology-Plant*. – 2004. – T. 40. – №. 2. – C. 163-166. doi.org/10.1079/IVP2003526
94. Narayan, R. K. J. *Nuclear DNA changes, genome differentiation and evolution in Nicotiana (Solanaceae)*/ Narayan, R. K. J.// *Plant Systematics and Evolution*. - 1987. - N. 157 - P. 161–180. doi:10.1007/bf00936195
95. Okamuro J. Tobacco single-copy DNA is highly homologous to sequences present in the genomes of its diploid progenitors/ Okamuro, J., and B. Goldberg// *Molecular and General Genetics*. - 1985. - N. 198 - P. 290– 298. doi.org/10.1007/BF00383009
96. Okonechnikov K. Unipro UGENE: a unified bioinformatics toolkit/ Okonechnikov K., Golosova O., Fursov M., the UGENE team// *Bioinformatics*. - 2012. - V. 28. - I. 8. - P. 1166–1167. doi.org/10.1093/bioinformatics/bts091

97. Olmstead R. G. A molecular phylogeny of the Solanaceae/ Olmstead R. G., Bohs, L., Migid, H. A., Santiago-Valentin, E., Garcia, V. F., & Collier, S. M. Olmstead R. G., Bohs, L., Migid, H. A., Santiago-Valentin, E., Garcia, V. F., & Collier, S. M. //Taxon. – 2008. – T. 57. – №. 4. – C. 1159-1181. doi.org/10.1002/tax.574010
98. Olmstead RG. Chloroplast DNA and systematics of the Solanaceae. In: Hawkes JG, Lester RN, Nee M, Estrada N (eds) Solanaceae III: taxonomy, chemistry, evolution. 1992. Royal Botanic Gardens, Kew, London, pp 161–168 doi.org/10.2307/2399773
99. Olmstead R.G. Combining data in phylogenetic systematics: an empirical approach using three molecular data sets in Solanaceae. Syst Biol. - 1994. - N. 43. P. 467–481. doi.org/10.1093/sysbio/43.4.467
100. Olmstead R.G. A molecular phylogeny of the Solanaceae/ Olmstead RG, Bohs L, Migid HA, Santiago-Valentín E, Garcia VF, Collier SM// Taxon. - 2008. - N. 57. - P. 1159–1181. doi.org/10.1002/tax.574010
101. Otten L. Restoration of virulence of Vir region mutants of *Agrobacterium tumefaciens* strain B6S3 by coinfection with normal and mutant *Agrobacterium* strains/ Otten, L., De Greve, H., Leemans, J., Hain, R., Hooykaas, P., & Schell, J.// Molecular and General Genetics MGG. - 1984. - V. 195. - I. 1-2. - P. 159–163. doi:10.1007/bf00332739
102. Otten L. The *Agrobacterium* phenotypic plasticity (Plast) genes/ Otten L.// *Agrobacterium Biology. Current Topics in Microbiology and Immunology.* - 2018. - N. 418. - P. 375–419. DOI 10.1007/82_2018_93
103. Owens L. D. Transformation of soybean cells using mixed strains of *Agrobacterium tumefaciens* and phenolic compounds/ Owens, L. D., & Smigocki, A. C.// *Plant physiology.* - 1988. - V. 88. - N. 3. - P. 570-573. doi.org/10.1104/pp.88.3.570
104. Parokonny A. S. Comparative physical mapping and evolution of the *Nicotiana tabacum* L. karyotype/ Parokonny A. S., Kenton A. Y.// *Kew Chromosome*

- Conference IV. – Kew, UK : Royal Botanic Gardens, 1995. – C. 301-320.
doi.org/10.1007/BF00713050
105. Petit A. The role of opines in the ecology of the Ti-plasmids of *Agrobacterium*/ Petit A., Tempe J., Guyon P., Tepfer D.// Plasmids of medical, environmental and commercial importance. Elsevier/North Holland Biomedical Press. - 1979. - P. 353 – 363
106. Posada D. Evaluation of methods for detecting recombination from DNA sequences: empirical data/ Posada D. //Molecular biology and evolution. – 2002. – T. 19. – №. 5. – C. 708-717.
107. Reddy T.V. Identification of SNPs in nicotine biosynthesis related genes by targeted re-sequencing of TILLING population and germplasm with varying nicotine levels in tobacco/ Reddy, T.V., Saiprasad, G.V.S.// Euphytica. - 2015. - N. 203. - P. 659–671. doi.org/10.1007/s10681-014-1300-4
108. Ren N. AFLP analysis of genetic polymorphism and evolutionary relationships among cultivated and wild *Nicotiana* species/ Ren N., Timko M.P.// Genome. - 2001. - V. 44. - No 4. - P. 559 – 71
109. Renny-Byfield, S. Next generation sequencing reveals genome downsizing in allotetraploid *Nicotiana tabacum*, predominantly through the elimination of paternally derived repetitive DNAs/ Renny-Byfield, S., Chester, M., Kovařík, A., Le Comber, S. C., Grandbastien, M. A., Deloger, M., Richard A. Nichols, Jiri Macas, Petr Novák, Mark W. Chase, & Leitch, A. R.// *Molecular biology and evolution*. - 2011. - V. 28. - N.10. - P. 2843-2854. doi.org/10.1093/molbev/msr112
110. Riechers, D.E. Structure and expression of the gene family encoding putrescine N-methyltransferase in *Nicotiana tabacum*: new clues to the evolutionary origin of cultivated tobacco/ Riechers, D.E. and Timko, M.P.// *Plant Mol.Biol.* - 1999. - N. 41. - P. 387 – 401

111. Rossi, L., Bindler, G., Pijnenburg, H., Isaac, P. G., Giraud-Henry, I., Mahe, M., Gadani, F. Potential of molecular marker analysis for variety identification in processed tobacco/ Rossi, L., Bindler, G., Pijnenburg, H., Isaac, P. G., Giraud-Henry, I., Mahe, M., Gadani, F.// *Plant Varieties and Seeds*. - 2001. - V. 14. - N. 2. - P. 89-102.
112. Särkinen T. A phylogenetic framework for evolutionary study of the nightshades (Solanaceae): a dated 1000-tip tree/ Särkinen T, Bohs L, Olmstead RG, Knapp S.// *BMC Evol Biol*. - 2013. - N. 13. - P. 214. doi.org/10.1186/1471-2148-13-214
113. Schueller S. K. Self-pollination in island and mainland populations of the introduced hummingbird-pollinated plant, *Nicotiana glauca* (Solanaceae)/ Schueller S. K. // *American Journal of Botany*. – 2004. – T. 91. – №. 5. – C. 672-681 doi.org/10.3732/ajb.91.5.672
114. Seppey M. BUSCO: Assessing Genome Assembly and Annotation Completeness/ Seppey M., Manni M., Zdobnov E.M.// In: Kollmar M. (eds) *Gene Prediction. Methods in Molecular Biology*. 2019. vol 1962. Humana, New York, NY. doi.org/10.1007/978-1-4939-9173-0_14
115. Setchell WA. Aboriginal tobaccos/ Setchell W. A.// *Am Anthropol*. - 1921. - N. 23. - P. 397–414
116. Shamel A. D. Tobacco breeding/ Shamel AD, Cobey WW// *USDA Bureau Plant Ind. Bull*. - 1907 - No. 96
117. Shen W. SeqKit: a cross-platform and ultrafast toolkit for FASTA/Q file manipulation/ Shen W., Le S., Li Y., Hu F.// *PLOS ONE*. - 2016. - V. 11. - N.10. - e0163962 doi:10.1371/journal.pone.0163962
118. Sheen, S. J. Isozymic evidence bearing on the origin of *Nicotiana tabacum* L./ Sheen, S. J.// *Evolution*. - 1972. - P. 143-154.

119. Siervo N. Reference genomes and transcriptomes of *Nicotiana sylvestris* and *Nicotiana tomentosiformis*/ Siervo N., Battey, J.N., Ouadi, S.// *Genome Biol.* - 2013. - V. 14. - N. R60. doi.org/10.1186/gb-2013-14-6-r60
120. Siervo, N. The tobacco genome sequence and its comparison with those of tomato and potato/ Siervo, N., Battey, J.N., Ouadi, S., Bakaher, N., Bovet, L., Willig, A., Goepfert, S., Peitsch, M.C. and Ivanov, N.V.// *Nat. Commun.* - 2014. - N.5. - P. 3833.
121. Skalicka K. Preferential elimination of repeated DNA sequences from the paternal *N. tomentosiformis* genome donor of a synthetic allotetraploid tobacco/ Skalicka K, Lim Y, Matyasek R, Matzka M, Leitch A, Kovarik A.// *New Phytol.* - 2005. - N. 166. - P. 291–303
122. Spinden HJ. Tobacco is American/ Spinden H. J. //The New York Public Library, New York. - 1950
123. Stanfill, S. Comprehensive chemical characterization of Rapé tobacco products: nicotine, un-ionized nicotine, tobacco-specific N'-nitrosamines, polycyclic aromatic hydrocarbons, and flavor constituents/ Stanfill, S. B., da Silva, A. L. O., Lisko, J. G., Lawler, T. S., Kuklenyik, P., Tyx, R. E., Watson, C. H.// *Food and Chemical Toxicology.* - 2015. - N. 82. - P. 50-58. doi.org/10.1016/j.fct.2015.04.016
124. Suzuki K. Tobacco plants were transformed by *Agrobacterium rhizogenes* infection during their evolution/ Suzuki K, Yamashita I., Tanaka N.// *Plant J.* - 2002. - V. 32. - No 5. - P. 775 — 787
125. Suzuki K. Mikimopine synthase (mis) gene on pRi1724/ Suzuki K., Tanaka N., Kamada H., Yamashita I.// *Gene.* - 2001. - V. 263 - I. 1-2. - P. 49-58. doi.org/10.1016/S0378-1119(00)00578-3
126. Tilley NM. The bright-tobacco industry 1860–1929/ Tilley NM// The Univ. of North Carolina Press, Chapel Hill. 1948

127. Tso TC. Production, physiology, and biochemistry of tobacco plant/ Tso TC// Ideals, Inc., Beltsville Maryland. 1990
128. Untergasser A. Primer3 - new capabilities and interfaces/ Untergasser A., Cutcutache I., Koressaar T., Ye J., Faircloth B., Remm M., Rozen S.// *Nucleic Acids Research.* - 2012. - V. 40. - N.15. - e115
129. Vaser R. Fast and accurate de novo genome assembly from long uncorrected reads/ Vaser R., Sović I., Nagarajan N., Šikić M.// *Genome Res.*- 2017. - No 27. - P.737–746.
130. Volkov, R. A. Elimination and rearrangement of parental rDNA in the allotetraploid *Nicotiana tabacum*/ Volkov, R. A., Borisjuk, N. V., Panchuk, I. I., Schweizer, D., & Hemleben, V.// *Molecular Biology and Evolution.* - 1999. - V.16. - N.3. - P. 311-320. 10.1093/oxfordjournals.molbev.a026112
131. White F.F. Sequence homologous to *Agrobacterium rhizogenes* T-DNA in the genomes of uninfected plants/ White F.F., Garfinkel D.J., Huffman G.A., Gordon M.P., and Nester E.W.// *Nature.* - 1983. - No 301. P. 348 — 350
132. Wiuf C. A simulation study of the reliability of recombination detection methods/ Wiuf C., Christensen T., Hein J. // *Molecular Biology and Evolution.* – 2001. – T. 18. – №. 10. – C. 1929-1939.
133. Yaldiz G. Genetic diversity assessment in *Nicotiana tabacum* L. with iPBS-retrotransposons/ Yaldiz G., Çamlica M., Nadeem M. A., Nawaz M. A., Baloch F. S.// *Turkish Journal of Agriculture and Forestry.* - 2018. - N. 42. - P. 154-164. doi:10.3906/tar-1708-32.
134. Yi Ye-Hee. Randomly amplified polymorphic DNA (RAPD) analysis of cultivated tobacco (*Nicotiana tabacum*, L.)/ Yi Ye-Hee// North Carolina State University. 1997.
135. Young J.M. A revision of *Rhizobium* Frank 1889, with an emended description of the genus, and the inclusion of all species of *Agrobacterium* Conn 1942

- and *Allorhizobium undicola* de Lajudie et al. 1998 as new combinations: *Rhizobium radiobacter*, *R. rhizogenes*, *R. rubi*, *R. undicola* and *R. vitis*/ Young J.M., Kuykendall L.D., Martinez-Romero E., Kerr A., Sawada H.// Int. J. Syst. Evol. Microbiol. - 2001. - N. 51. - pp. 89-103. 10.1099/00207713-51-1-89
136. Yukawa M. The 2005 version of the chloroplast DNA sequence from tobacco (*Nicotiana tabacum*)/ Yukawa M., Tsudzuki T., Sugiura M. //Plant Molecular Biology Reporter. – 2005. – T. 23. – №. 4. – C. 359-365. doi.org/10.1007/s00438-005-0092-6
137. Zhang Y. Parasitic plant dodder (*Cuscuta* spp.): A new natural Agrobacterium-to-plant horizontal gene transfer species/ Zhang Y. Wang D., Wang Y., Dong H., Yuan Y., Yang W., Lai D., Zhang M., Jiang L., Li Z.// Science China Life Sciences. - 2020. - T.63. №.2. C.312-316. doi.org/10.1007/s11427-019-1588-x
138. Zhang H. Y. Genetic diversity among flue-cured tobacco (*Nicotiana tabacum* L.) revealed by amplified fragment length polymorphism/ Zhang H. Y., Liu, X. Z., Li, T. S., & Yang, Y. M.// *Bot Stud.* - 2006 - V. 47. - N.3. - P. 223-229.
139. Zimin A. Hybrid assembly of the large and highly repetitive genome of *Aegilops tauschii*, a progenitor of bread wheat, with the MaSuRCA mega-reads algorithm/ Zimin A., Puiu D., Luo M., Zhu T., Koren S., Marçais G., Yorke J., Dvořák J., Salzberg S.// *Genome Res.* - 2017- N. 27. - 5.
140. Zimmerman, J. L. DNA sequence organization in the genome of *Nicotiana tabacum*/ Zimmerman, J. L. & Goldberg, R. B.// *Chromosoma.* - 1977. - N. 59- P. 227–252.

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 РМІ – Николая Владимировича Иванова и Николаса Сьерро (Nicolas Sierra) 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. Работа выполнена при поддержке грантов: РФФ 16-16-10010, РФФ 21-14-00050, РФФИ 18-016-0011, грант Министерства науки и высшего образования Российской

Федерации в рамках соглашения №075–15-2020-922 от 16.11.2020 на создание и развитие Научного центра мирового уровня «Агротехнологии будущего». В работе использовано оборудование ресурсных центров научного парка СПбГУ «Развитие молекулярных и клеточных технологий» и «Биобанк».

Приложения

Приложение 1. Современный видовой состав рода *Nicotiana*.

Подрод	Секция (Knapp et al., 2004)	Вид	Хромосомный набор (n)	Ареал произрастания
<i>Tabacum</i>	<i>Tomentosae</i>	<i>Nicotiana kawakamii</i> Y. Ohashi	12	Боливия
		<i>Nicotiana leguiana</i> J.F. Macbr.	12	Боливия
		<i>Nicotiana otophora</i> Griseb.	12	Аргентина, Боливия
		<i>Nicotiana setchellii</i> Goodsp.	12	Перу
		<i>Nicotiana tomentosa</i> Ruiz and Pav.	12	Боливия, Эквадор, Перу
		<i>Nicotiana tomentosiformis</i> Goodsp.	12	Боливия
	<i>Nicotiana</i>	<i>Nicotiana tabacum</i> L.	24	Повсеместно (выращивают)
<i>Rustica</i>	<i>Rusticae</i>	<i>Nicotiana rustica</i> L.	24	Повсеместно (выращивают)
	<i>Paniculatae</i>	<i>Nicotiana benavidesii</i> Goodsp.	12	Перу
		<i>Nicotiana cordifolia</i> Phil.	12	Чили
		<i>Nicotiana cutleri</i> D'Arcy	неизвестно	Боливия
		<i>Nicotiana knightiana</i> Goodsp.	12	Перу

		<i>Nicotiana paniculata</i> L.	12	Перу
		<i>Nicotiana raimondii</i> J.F. Macbr.	12	Перу
		<i>Nicotiana solanifolia</i> Walp.	12	Чили (острова Хуан- Фернандес)
<i>Petunioides</i>	<i>Alatae</i>	<i>Nicotiana azambujae</i> L.B. Sm. and Downs	неизвестен	Бразилия
		<i>Nicotiana alata</i> Link and Otto	9	Аргентина, Бразилия, Парагвай, Уругвай
		<i>Nicotiana bonariensis</i> Lehm.	9	Аргентина, Бразилия, Уругвай
		<i>Nicotiana forgetiana</i> Hemsl.	9	Бразилия
		<i>Nicotiana langsdorffii</i> Weinm.	9	Аргентина, Бразилия, Парагвай
		<i>Nicotiana longiflora</i> Cav.	10	Аргентина, Бразилия, Боливия, Парагвай, Уругвай
		<i>Nicotiana mutabilis</i> Stehmann and Samir	9	Бразилия

	<i>Nicotiana plumbaginifolia</i> Viv.	10	Аргентина, Боливия, Парагвай
<i>Noctiflorae</i>	<i>Nicotiana acaulis</i> Speg.	12	Аргентина
	<i>Nicotiana ameghinoi</i> Speg.	12	Аргентина
	<i>Nicotiana glauca</i> Graham	12	Аргентина, Боливия
	<i>Nicotiana noctiflora</i> Hook.	12	Аргентина
	<i>Nicotiana paa</i> Mart.Crov.	неизвестно	Аргентина, Чили
	<i>Nicotiana petunioides</i> (Griseb.) Millán	12	Аргентина, Чили
<i>Petunioides</i>	<i>Nicotiana acuminata</i> (Graham) Hook.	12	Аргентина, Чили
	<i>Nicotiana attenuata</i> S. Wats.	12	США
	<i>Nicotiana corymbosa</i> J. Rémy	12	Аргентина, Чили
	<i>Nicotiana linearis</i> Phil.	12	Аргентина, Чили
	<i>Nicotiana longibracteata</i> Phil.	12	Аргентина, Чили
	<i>Nicotiana miersii</i> J. Rémy	12	Аргентина
	<i>Nicotiana pauciflora</i> J. Rémy	12	Чили
	<i>Nicotiana spegazzinii</i> Millán	12	Аргентина

<i>Polydichiae</i>	<i>Nicotiana clevelandii</i> A. Gray	24	США, Мексика
	<i>Nicotiana quadrivalvis</i> Pursh	24	Канада, США, Мексика (выращивают)
<i>Repandae</i>	<i>Nicotiana nesophila</i> I.M. Johnst.	24	Мексика (острова Ревилья- Хихедо)
	<i>Nicotiana nudicaulis</i> S. Wats.	24	Мексика
	<i>Nicotiana repanda</i> Willd.	24	Мексика, Куба
	<i>Nicotiana stocktonii</i> Brandeggee	24	Мексика (острова Ревилья- Хихедо)
<i>Suaveolentes</i>	<i>Nicotiana africana</i> Merxm.	23	Намибия
	<i>Nicotiana amplexicaulis</i> N.T. Burb.	18	Австралия
	<i>Nicotiana benthamiana</i> Domin.	19	Австралия
	<i>Nicotiana burbridgeae</i> Symon	21	Австралия
	<i>Nicotiana cavicola</i> N.T. Burb.	23; 20	Австралия
	<i>Nicotiana excelsior</i> (J.M. Black) J.M. Black	19	Австралия

	<i>Nicotiana exigua</i> H.-M. Wheeler	неизвестно	Австралия
	<i>Nicotiana fatuhivensis</i> F. Br.	неизвестно	Маркизские острова
	<i>Nicotiana faucicola</i> M.W. Chase and Christenh.	неизвестно	Австралия, Новая Каледония
	<i>Nicotiana forsteri</i> Roem. and Schult.	24	Австралия
	<i>Nicotiana fragrans</i> Hook.	24	Новая Каледония, Тонгатапу
	<i>Nicotiana gascoynica</i> M.W. Chase and Christenh.	20	Австралия
	<i>Nicotiana goodspeedii</i> H.-M. Wheeler	16	Австралия
	<i>Nicotiana gossei</i> Domin	18	Австралия
	<i>Nicotiana hesperis</i> N.T. Burb.	21	Австралия
	<i>Nicotiana heterantha</i> Kenneally and Symon	неизвестно	Австралия
	<i>Nicotiana ingulba</i> J.M. Black	20	Австралия
	<i>Nicotiana karijini</i> M.W. Chase and Christenh.	неизвестно	Австралия
	<i>Nicotiana maritima</i> H.-M. Wheeler	16	Австралия

	<i>Nicotiana megalosiphon</i> Van Huerck and Müll.-Arg.	20	Австралия
	<i>Nicotiana monoschizocarpa</i> (P.Horton) Symon and Lepschi	24	Австралия
	<i>Nicotiana occidentalis</i> H.-M. Wheeler	21	Австралия
	<i>Nicotiana rosulata</i> (S.Moore) Domin	20	Австралия
	<i>Nicotiana rotundifolia</i> Lindl.	22	Австралия
	<i>Nicotiana simulans</i> N.T. Burb.	20	Австралия
	<i>Nicotiana stenocarpa</i> H.-M. Wheeler	20	Австралия
	<i>Nicotiana truncata</i> Symon	неизвестно	Австралия
	<i>Nicotiana suaveolens</i> Lehm.	15	Австралия
	<i>Nicotiana umbratica</i> N.T. Burb.	23	Австралия
	<i>Nicotiana velutina</i> H.-M. Wheeler	16	Австралия
	<i>Nicotiana wuttkei</i> Clarkson and Symon	15	Австралия

	<i>Nicotiana yandinga</i> M.W. Chase and Christenh.	21	Австралия
<i>Sylvestres</i>	<i>Nicotiana sylvestris</i> Speg.	12	Аргентина, Боливия
<i>Trigonophyllae</i>	<i>Nicotiana obtusifolia</i> M. Martens and Galeottii	12	Мексика, США
<i>Undulatae</i>	<i>Nicotiana arentsii</i> Goodsp.	24	Боливия
	<i>Nicotiana glutinosa</i> L.	12	Боливия, Эквадор, Перу
	<i>Nicotiana thrysiflora</i> Goodsp.	12	Перу
	<i>Nicotiana undulata</i> Ruiz and Pav.	12	Аргентина, Боливия, Перу
	<i>Nicotiana wigandioides</i> Koch and Fintelm.	12	Боливия