

Федеральное государственное автономное образовательное учреждение высшего
образования

«ЮЖНЫЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»

Академия биологии и биотехнологии им. Д.И. Ивановского

На правах рукописи

Макаренко Максим Станиславович

**ИЗМЕНЧИВОСТЬ ХЛОРОПЛАСТНОГО И
МИТОХОНДРИАЛЬНОГО ГЕНОМОВ У ОДНОЛЕТНИХ И
МНОГОЛЕТНИХ ВИДОВ ПОДСОЛНЕЧНИКА (*HELIANTHUS L.*)**

03.02.07 – Генетика

Диссертация на соискание ученой степени

кандидата биологических наук

Научный руководитель: доктор биологических наук,

профессор Усатов Александр Вячеславович

Ростов-на-Дону

2018

Содержание

ВВЕДЕНИЕ	4
ГЛАВА 1. Обзор литературы	9
1.1. Систематика рода <i>Helianthus</i> L.....	9
1.2. Особенности структурно-функциональной организации хлоропластной ДНК цветковых растений.....	12
1.3. Особенности структурно-функциональной организации митохондриальной ДНК цветковых растений.....	18
ГЛАВА 2. Материалы и методы исследования	35
2.1. Объекты исследования.....	35
2.2. Выделение хлоропластов и митохондрий из листьев подсолнечника.....	36
2.3. Молекулярно-генетические исследования.....	36
2.3.1. Выделение ДНК.....	37
2.3.2. Измерение концентрации ДНК.....	38
2.3.3. Секвенирование фрагментов ДНК методом Сэнгера.....	39
2.3.4. Анализ полиморфизма митохондриальных SSR локусов.....	40
2.3.5. Подготовка библиотек для высокопроизводительного секвенирования.....	41
2.3.6. Высокопроизводительное секвенирование.....	44
2.3.7. Анализ данных высокопроизводительного секвенирования. Сборка и аннотация геномов.....	44
2.3.8. Выделение РНК.....	45
2.3.9. Анализ транскрипционной активности митохондриальных генов.....	46
ГЛАВА 3. Результаты и обсуждение	49
3.1. Изменчивость генов и некодирующих локусов хлоропластной и митохондриальной ДНК у однолетних и многолетних видов подсолнечника.....	49
3.2. Сравнительный анализ хлоропластного и митохондриального геномов у культурного и дикорастущего подсолнечника (<i>Helianthus annuus</i> L.).....	57

3.3. Особенности структурно-функциональной организации геномов хлоропластов и митохондрий у фертильного и стерильных аналогов линии НА89 подсолнечника с различными типами ЦМС	68
3.3.1. Изменчивость хлоропластной ДНК у фертильного и стерильных аналогов линии НА89 подсолнечника с различными типами ЦМС	68
3.3.2. Изменчивость митохондриальной ДНК у фертильного и стерильного аналогов линии НА89 подсолнечника с различными типами ЦМС	80
3.3.2.1. Митохондриальный геном ЦМС линии НА89(РЕТ1).....	80
3.3.2.2. Митохондриальный геном ЦМС линии НА89(РЕТ2).....	83
3.3.2.3. Митохондриальный геном ЦМС линии НА89(МАХ1).....	93
3.3.2.4. Митохондриальный геном ЦМС линии НА89(АНН2).....	102
ЗАКЛЮЧЕНИЕ	113
ВЫВОДЫ	116
Список сокращений	118
Список литературы	119
Приложение 1	134
Приложение 2	150
Приложение 3	162

ВВЕДЕНИЕ

Актуальность темы исследования. Доминирующая роль ядра в растительной клетке неопровержимо доказана многочисленными генетическими исследованиями. Однако в пластидах и митохондриях растительных клеток локализованы еще и собственные генетические системы (Stanley et al., 2015). Эти системы функционируют различно, но взаимосвязано. Взаимодействие генетических систем приводит к тому, что каждая из них влияет на другие, что в свою очередь приводит к направленным или канализированным изменениям в эволюционных и селекционных процессах (Woodson, Chory, 2008).

В основном, при рассмотрении механизмов видообразования во внимание принимаются только ядерные гены, в то время как роль цитоплазматических генов, часто недооценивают. Тем не менее, изучение структурно-функциональной организации геномов органелл также актуально с точки зрения современной систематики и филогенетики (Zhang et al., 2012).

Следует заметить, что к настоящему времени в базу данных NCBI включены полные последовательности хлоропластной ДНК (хлДНК) только 3 однолетних и 8 многолетних видов подсолнечника, а митохондриальной ДНК (мтДНК) только одного вида - *H. annuus*, что явно недостаточно для корректного анализа межвидовой внеядерной изменчивости в роде *Helianthus* L. В связи с этим исследование изменчивости хлоропластных и митохондриальных генов у однолетних и многолетних видов подсолнечника достаточно актуально.

В настоящее время в коммерческом производстве семян подсолнечника доминируют гетерозисные гибриды, полученные на основе цитоплазматической мужской стерильности (ЦМС) только одного типа РЕТ1, открытого П. Леклерком еще в 1966 году у межвидового гибрида *H. petiolaris* Nutt. × *H. annuus* L. (Leclercq, 1969). Такая генетическая унификация цитоплазмы возделываемых гибридов может привести к различным негативным явлениям, например, массовым эпифитотиям (Miller, 1996).

Поскольку ЦМС растений представляет собой не только прекрасную модель, для изучения взаимодействия ядерных и цитоплазматических генетических систем клетки, но и эффективную мутацию для производства гибридных гетерозисных семян (Chen, Liu, 2014; Hu et al., 2014), дальнейший перевод лучших селекционных линий на новые типы ЦМС позволит избежать в

будущем рисков, связанных с унификацией цитоплазмона у коммерческих гибридов подсолнечника.

Степень разработанности темы. Полиморфный род *Helianthus*, по оценкам разных систематиков, включает от 10 до 254 видов (Анащенко, 1979; Schilling, Heiser, 1981; Rieseberg, 1991; Гаврилова, Анисимова, 2003). Эти классификации в основном базируются на анализе морфологических признаков растений, ареалах видов, кариотипах и результатах скрещиваний. В последнее время для вопросов систематики, включая и род *Helianthus*, все шире внедряются молекулярные методы анализа, в том числе высокопроизводительное секвенирование ДНК органелл (Bock et al., 2014; Stephens et al. 2015). Однако полные нуклеотидные последовательности хлДНК и мтДНК определены только у небольшого числа видов подсолнечника, что явно недостаточно для корректного анализа межвидовой внеядерной изменчивости в данном таксоне.

Цитоплазматическая мужская стерильность растений, как правило, вызвана нарушениями в структуре мтДНК (Horn et al., 2014). На сегодняшний день у подсолнечника выделено более 70 типов ЦМС (Ardila et al., 2010; Reddemann, Horn, 2018). Однако только для двух из них: PET1 (Kohler et al. 1991, Horn et al., 1996) и PET2 (Reddemann, Horn, 2018; Makarenko et al., 2018) выявлены специфичные реорганизации структуры митохондриальных геномов, связанные с ЦМС фенотипом.

Цель и задачи исследования. Целью работы является определение уровня изменчивости хлДНК и мтДНК у однолетних и многолетних видов подсолнечника (*Helianthus* L) и изучение особенностей структурно-функциональной организации хлоропластных и митохондриальных геномов у фертильной линии HA89 и ее ЦМС аналогов на основе однолетних (PET1, PET2 - *H. petiolaris*, ANN2 - *H. annuus*) и многолетнего (MAX1 - *H. maximilliani*) видов рода *Helianthus* L.

Исходя из поставленной цели, были сформулированы следующие задачи:

1. Исследовать полиморфизм хлоропластных и митохондриальных генов у образцов 5-и однолетних и 16-и многолетних видов подсолнечника из коллекции ВИР.

2. Провести высокопроизводительное секвенирование и определить уровень изменчивости хлДНК и мтДНК у линий культурной и дикорастущей форм *H annuus* L.

3. Провести секвенирование, сборку и аннотацию хлоропластных и митохондриальных геномов у фертильной линии НА89 и ее стерильных аналогов НА89(РЕТ1), НА89(РЕТ2), НА89(АНН2) и НА89(МАХ1).

4. Определить структурные и функциональные особенности митохондриальных геномов у стерильных аналогов линии НА89 с различными типами ЦМС: РЕТ1, РЕТ2, АНН2 и МАХ1.

Научная новизна исследования. Впервые определен уровень полиморфизма хлоропластных (*atpB*, *matK*, *rbcL*) и митохондриальных (*atp1*, *matR*, *nad4*) генов у образцов 5 однолетних и 16 многолетних видов рода *Helianthus* L. из коллекции ВИР. Проведено полногеномное секвенирование, сборка, аннотация и сравнительный анализ хлоропластных и митохондриальных геномов культурной и дикорастущей форм *H. annuus* L. Определены специфичные реорганизации митохондриальных геномов у ЦМС линий НА89(РЕТ1), НА89(РЕТ2), НА89(МАХ1), НА89(АНН2).

Теоретическая и практическая значимость работы. Полученные данные о полиморфизме хлоропластных и митохондриальных генов образцов 21 вида подсолнечника вносят вклад в решение проблемы микроэволюции рода *Helianthus* L. Полиморфные сайты хлДНК и мтДНК, локализованные у культурного и дикорастущего подсолнечника *H. annuus* L. в результате анализа полных нуклеотидных последовательностей внеядерных геномов, являются информативными мишенями для генотипирования селекционных линий и образцов.

Данные об уровне мутационной изменчивости внеядерных геномов у аллоплазматических ЦМС линий подсолнечника вносят вклад в решение фундаментальной проблемы взаимодействия ядерных и цитоплазматических генетических систем растительной клетки, а также в изучение механизмов возникновения ЦМС у цветковых растений. Дальнейший перевод лучших селекционных линий на новые типы ЦМС позволит избежать в будущем рисков, связанных с унификацией цитоплазмона у гибридов подсолнечника.

Аннотированные полногеномные последовательности мтДНК ЦМС линий НА89(РЕТ1), НА89(РЕТ2), НА89(МАХ1) депонированы в международную базу данных NCBI GenBank под номерами MG735191.1, MG770607.2, MH704580.1.

Результаты диссертационной работы включены в учебные курсы для студентов, специализирующихся на кафедре генетики Южного федерального университета.

Методология и методы исследования. В работе были использованы три уникальные генетические модели. Первая представляет собой образцы однолетних и многолетних видов рода *Helianthus* L из коллекции ВИР. Вторая включает 4 линии культурного и 1 дикорастущего подсолнечника *H. annuus*. Третья представлена фертильным и стерильными аналогами линии НА89 подсолнечника с различными типами ЦМС (PET1, PET2, ANN2, MAX), полученными на основе однолетних и многолетнего видов рода *Helianthus* L. Используя современные методы молекулярной генетики, сравнительной геномики и биоинформатики были определены и проанализированы последовательности хлоропластных и митохондриальных генов, а также исследован уровень изменчивости внеядерных геномов, что позволило в полной мере реализовать цель и задачи работы.

Положения, выносимые на защиту:

1. Частота мутаций в хлоропластных генах у представителей однолетних и многолетних видов рода *Helianthus* L. многократно превышает этот показатель в митохондриальных генах.

2. Дикорастущий подсолнечник *H. annuus* L. значительно отличается от селекционных линий по уровню изменчивости как хлДНК, так и мтДНК.

3. У стерильных аналогов линии НА89 с различными типами ЦМС частота точковых мутаций в хлДНК значительно выше, чем в мтДНК, однако только для митохондриальных геномов характерно наличие крупных реорганизаций структуры, уникальных для каждого типа ЦМС.

Апробация работы. Материалы диссертации были представлены на II и III Международной научной конференции «Генетика и биотехнология XXI века: проблемы, достижения, перспективы» (Минск, 2015, 2016); II Международной научно-практической конференции «Наука и образование сегодня: теория, практика и инновации – 2015» (Темрюк, 2015); Всероссийской конференции с международным участием «50 лет ВОГиС: успехи и перспективы» (Москва, 2016); Международной научной конференции, посвященная 90-летию Ботанического сада Южного федерального университета «Современные технологии в изучении биоразнообразия и интродукции растений» (Ростов-на-

Дону, 2017); Научно-практической конференции с международным участием «Генетика - фундаментальная основа инноваций в медицине и селекции» (Ростов-на-Дону, 2017); Международной школе-конференции молодых ученых «Наука и молодежь: фундаментальные и прикладные проблемы в области селекции и генетики сельскохозяйственных культур» (Зерноград, 2017); XXV Международной научной конференции студентов, аспирантов и молодых ученых «Ломоносов-2018» (Москва, 2018)

Публикации. По теме диссертационной работы опубликовано 17 работ, из них 5 работ в изданиях, индексируемых в библиографических базах данных Scopus и Web of Science; 2 базы данных.

Связь с другими проектами.

Исследование выполнено в рамках государственного задания Минобрнауки России, проект № 6.929.2017/4.6.

Объем и структура диссертации. Работа состоит из введения, трех глав, заключения, выводов, списка литературы, приложений. Диссертация изложена на 163 страницах, включает 30 таблиц и 17 рисунков. Список цитируемой литературы содержит 132 источника, из них 122 на иностранных языках.

ГЛАВА 1. Обзор литературы

1.1. Систематика рода *Helianthus* L.

Род *Helianthus* по оценкам разных систематиков включает в себя от 10 до 254 видов (Гаврилова, Анисимова, 2003). Карл Линней в 1753 году впервые описал 9 видов подсолнечника, А. Грей в 1889 определила 42 вида, Уотсон – 108 видов, Сацперов – 254 вида (Гаврилова, Анисимова, 2003). Данные классификации в основном базировались на морфологических признаках, в связи с чем в ранг вида были включены разнообразные декоративные формы подсолнечника или подвиды. В середине XX века более детально вопросом систематики рода *Helianthus* занялся Ч. Хейзер с коллегами, которые предложили включить в род *Helianthus* 14 однолетних и 36 многолетних видов, произрастающих в Северной Америке и 17 видов – в Южной Америке (Heiser et al., 1969). В более поздней классификации (Schilling, Heiser, 1981) был выделен 51 вид подсолнечника, включая 14 однолетних и 37 многолетних видов. Эта классификация базируется не только на данных морфофизиологического анализа, но также учитывает, ареал произрастания, число хромосом и возможность скрещиваемости видов. Среди отечественных работ, можно выделить классификацию А.В. Анащенко (1977, 1979), согласно которой род *Helianthus* включает один однолетний вид – *H. annuus* (с тремя подвидами и шестью формами) и 9 многолетних видов (с тринадцатью подвидами). Данная классификация основывается на скрещиваемости и анализе геномного состава образцов дикорастущего подсолнечника мировой коллекции Всероссийского института растениеводства (ВИР) (Анисимова и др., 1982).

Виды рода *Helianthus* также подвергались разделению по секциям. Торрей и Грэй в 1843 году выделили шесть секций: *Annui*, *Angustifolii*, *Atrorubentes*, *Laetiflori*, *Corono-Solis* и *Microcephali*. Позже Грэй (1884 год) признал только две секции: *Annuals* и *Perennials* (Гаврилова, Анисимова, 2003). Уотсон (1929 год) предложил разделение по признаку окраски венчика на две секции: *Rubri* и *Flavi* (Гаврилова, Анисимова, 2003). Очевидно, что это разделение является

искусственным. С тех пор, как стало ясно, что существуют виды, у которых окраска венчика может варьировать, возможность использования этой классификации стала ограниченной. Ридберг в 1932 году определил только одну секцию среди однолетних и шесть секций среди многолетних видов, произрастающих в Северной Америке (Гаврилова, Анисимова, 2003). Однако, два вида – *H. chartaceus* и *H. hirsutus* – по его классификации принадлежат как к секции *Divaricati*, так и к секции *Tuberosi*.

Хейзер, чья классификация является наиболее распространенной среди исследователей, разделил род *Helianthus* на 4 секции, основываясь главным образом на способности видов к гибридизации (Heiser, 1965).

1. Секция *Annui* включает 14 однолетних видов, среди которых несколько имеют стержневую корневую систему, а остальные – мочковатую. Набор хромосом – диплоидный. Согласно данным филогении и способности к скрещиванию друг с другом все виды данной секции (кроме *H. similis*, *H. ludens* и *H. agrestis*) являются близкородственными. Позднее *H. similis* и *H. ludens* были отнесены к роду *Viguiera* (Schilling, Heiser, 1981).

2. Секция *Ciliares* включает 6 видов, произрастающих в западной части Северной Америки. Пять из шести видов имеют способность скрещиваться с любым видом из секции *Divaricati*. Эти гибриды морфологически неотличимы от представителей секции *Divaricati* (Heiser, 1965).

3. Секция *Divaricati* включает 30 видов, произрастающих на востоке Северной Америки и представляющих большую сложность для их идентификации. Набор хромосом у видов этой секции может быть диплоидным, тетраплоидным и гексаплоидным. Гибридизация возможна практически между всеми видами внутри секции, имеющими одинаковую ploidy набора хромосом.

4. Секция *Fruticosi* включает многолетние виды, произрастающие в Южной Америке. По всей вероятности, эти виды не близкородственны североамериканским и произошли непосредственно в Южной Америке от

представителей рода *Viguiera*. Позднее южноамериканские виды были отнесены к роду *Helianthopsis*.

Род *Helianthus* включает в себя большое количество видов, которые в процессе эволюции предположительно скрещивались между собой, в связи с чем реконструкция филогенетических взаимоотношений между ними является сложной задачей, в особенности для многолетних видов (Schilling, Heiser 1981; Timme et al. 2007). Объединение обширной информации о скрещиваемости видов (Heiser et al., 1969) с морфологическими признаками для 49 видов привело к важным открытиям в области установления родства видов и их филогении. Однако, данные о скрещиваемости имеют ограниченное применение для многолетних видов по причине их полиплоидности (Schilling, Heiser, 1981). Проведённое сравнение рестрикционных карт рибосомальных генов прояснило взаимоотношения между большей частью однолетних видов подсолнечника (Riesenberg, 1991). Анализ участка хлДНК позволило установить уникальный ДНК-штрих код для 4-х видов подсолнечника - *H. annuus*, *H. argophyllus*, *H. debilis*, *H. tuberosus* (Vischi et al., 2006).

Основательная работа по выяснению филогении рода *Helianthus* на основе молекулярных маркеров опубликована в 2007 году (Timme et al., 2007). Было проведено секвенирование обширного транскрибируемого спейсера, расположенного внутри региона 18S-25S ядерной ДНК, у 47 видов подсолнечника, что позволило установить взаимоотношения между многолетними видами, идентифицировать происхождение некоторых гибридов и полиплоидных групп, а также утверждать, что однолетние виды не составляют монофилетическую группу (Timme et al., 2007).

Данные секвенирования фрагментов хлДНК и ядерных внутренних транскрибируемых спейсеров в значительной степени способствовали пониманию степени родства как между основными группами внутри рода *Helianthus*, а так и между другими родами семейства *Asteraceae* (Kane et al., 2013). Важно отметить, что в разных популяциях один и тот же вид рода *Helianthus* может иметь разную

плоидность ядерного генома (Bock et al., 2014), что осложняет использование ядерных маркеров при анализе степени родства.

В 2014 году Дэн Бок с коллегами (Bock et al., 2014) впервые провел секвенирование полных нуклеотидных последовательностей хлДНК и частичных – мтДНК у 8 многолетних видов подсолнечника: *H. decapetalus*, *H. divaricatus*, *H. giganteus*, *H. grosseserratus*, *H. hirsutus*, *H. maximiliani*, *H. strumosus*, *H. tuberosus*. Данное исследование позволило установить филогенетические взаимоотношения между ними, а также предоставило генетические доказательства полифилетического происхождения топинамбура (*H. tuberosus*). Наиболее масштабное молекулярное-филогенетическое исследование, основанное на данных секвенирования более 170 генов (ядерных и хлоропластных), у 37 диплоидных видов и подвидов рода *Helianthus* проведено Джесикой Стефанс с соавторами (Stephens et al., 2015).

1.2. Особенности структурно-функциональной организации хлоропластной ДНК цветковых растений

Еще в начале XX века были сделаны предположения о том, что хлоропласты растений содержат свой собственный генетический материал. Спустя более 50 лет наличие ДНК в пластидах было неопровержимо доказано целым рядом работ (Chun et al., 1963; Sager, Ishida 1963; Tewari, Wildman 1966).

Хлоропластная ДНК в значительной степени отличается от ядерной ДНК структурно-функциональной организацией и типом наследования (Bock, 2007). Хлоропластный геном имеет многочисленные прокариотические особенности, включая: структуру бактериального типа (кольцевая хромосома), упаковку генома в нуклеоиды, оперонную организацию генов (Даниленко, Давыденко, 2003). С другой стороны, пластидная ДНК обладает гораздо большей структурной сложностью, чем бактериальные геномы. Исследования показали, что помимо кольцевой формы хлДНК могут встречаться также линейные формы, в том числе в виде разветвленных мономеров и мультимеров (Krupinska et al., 2013)

Одной из основных особенностей пластидной ДНК является многокопийность. Каждая растительная клетка содержит много пластид, и каждая пластида содержит многочисленные копии ДНК. Так в одной растительной клетке может быть более 10 000 идентичных копий хлДНК (Bendich, 1987).

Количество копий пластидной ДНК зависит от ряда факторов: вида растения, типа ткани, стадии развития, условий окружающей среды и т.д. Показано что количество копий ДНК на клетку в процессе онтогенеза может изменяться многократно (Даниленко, Давыденко, 2003).

Полиплоидность пластидного генома, случайное распределение молекул хлДНК при митотических делениях клеток и, соответственно, пластидных генов приводит к неменделевскому типу наследования. Также полиплоидное состояние пластидных генов приводит к более низкой частоте закрепления в них мутаций по сравнению с ядерными генами (Drouin et al., 2008).

Более 40 лет назад были секвенированы первые пластидные геномы двух видов - *Marchantia polymorpha* и *Nicotiana tabacum*, что внесло огромный вклад в изучение молекулярно-генетических особенностей хлДНК (Воск, 2007). На сегодняшний день в банке данных NCBI находится информация о полных нуклеотидных последовательностях более 1000 видов цветковых растений (<https://www.ncbi.nlm.nih.gov/genome/browse/#!/organelles/>). Так у большинства цветковых растений размер хлоропластных геномов находится в диапазоне 130-160 т.п.н. GC-состав хлДНК - в пределах 35-40 % (табл. 1.1). При этом низкое содержание G и C нуклеотидов особенно заметно в некодирующих межгенных областях, где может составлять всего 20 % нуклеотидного состава (Воск, 2007).

Изучение структурной организации хлоропластных геномов цветковых растений показало, что для большинства из них можно отметить значительное сходство: наличие двух инвертированных, повторяющихся регионов (inverted repeats – IRa и IRb), которые делят кольцевую ДНК пластид на большой однокопийный регион (large single copy region (LSC)) и малый однокопийный регион (small single copy region (SSC)). Примером может служить генетическая карта хлДНК подсолнечника (рис. 1.1), имеющая типичную организацию.

Таблица 1.1. Размер и GC-состав хлоропластной ДНК у некоторых цветковых растений.

	Виды	Размер хлДНК, т.п.н.	GC-состав, %	Источник
Однодольные	<i>Avena sativa</i>	135,9	38,5	Saarela et al., 2015
	<i>Oryza sativa</i>	134,5	39	Tang et al., 2004
	<i>Sorghum bicolor</i>	140,8	38,5	Saski et al., 2007
	<i>Triticum aestivum</i>	134,5	38,3	Ogihara et al., 2002
	<i>Zea mays</i>	140,4	38,5	Maier et al., 1995
Двудольные	<i>Arabidopsis thaliana</i>	154,5	36,3	Sato et al., 1999
	<i>Glycine max</i>	152,2	35,4	Saski et al., 2005
	<i>Helianthus annuus</i>	151,1	37,6	Timme et al., 2007
	<i>Solanum lycopersicum</i>	155,5	37,9	Kahlau et al., 2006
	<i>Vitis vinifera</i>	160,9	37,4	Jansen et al., 2006

Размер регионов отличается даже у близкородственных видов (Воск, 2007). Интересно отметить, что у представителей рода *Pelargonium* описано значительное увеличение каждого IR региона до 75 т.п.н., что примерно в три раза превышает размер соответствующего участка у большинства цветковых растений (Chumley et al., 2006). Два инвертированных повтора имеют идентичные нуклеотидные последовательности, так что каждый ген, содержащийся в них, присутствует в молекуле хлДНК в двух копиях, которые отличаются только относительной ориентацией. Наличие двух крупных идентичных областей в хлДНК облегчает два типа генетических взаимодействий между гомологичными последовательностями, а именно: внутримолекулярную рекомбинацию и генную конверсию (Khakhlova, Воск 2006). Гомологичная рекомбинация между двумя IR регионами приводит к существованию двух изоформ пластидного генома, которые отличаются относительной ориентацией LSC и SSC (Marechal, Brisson, 2010). Частота мутаций в генах IR регионов ниже по сравнению с генами других

регионов (LSC, SSC) хлДНК. Это связано с тем, что процессы генной конверсии протекают активнее в IR чем в LSC и SSC (Perry, Wolfe, 2002; Marechal, Brisson, 2010).

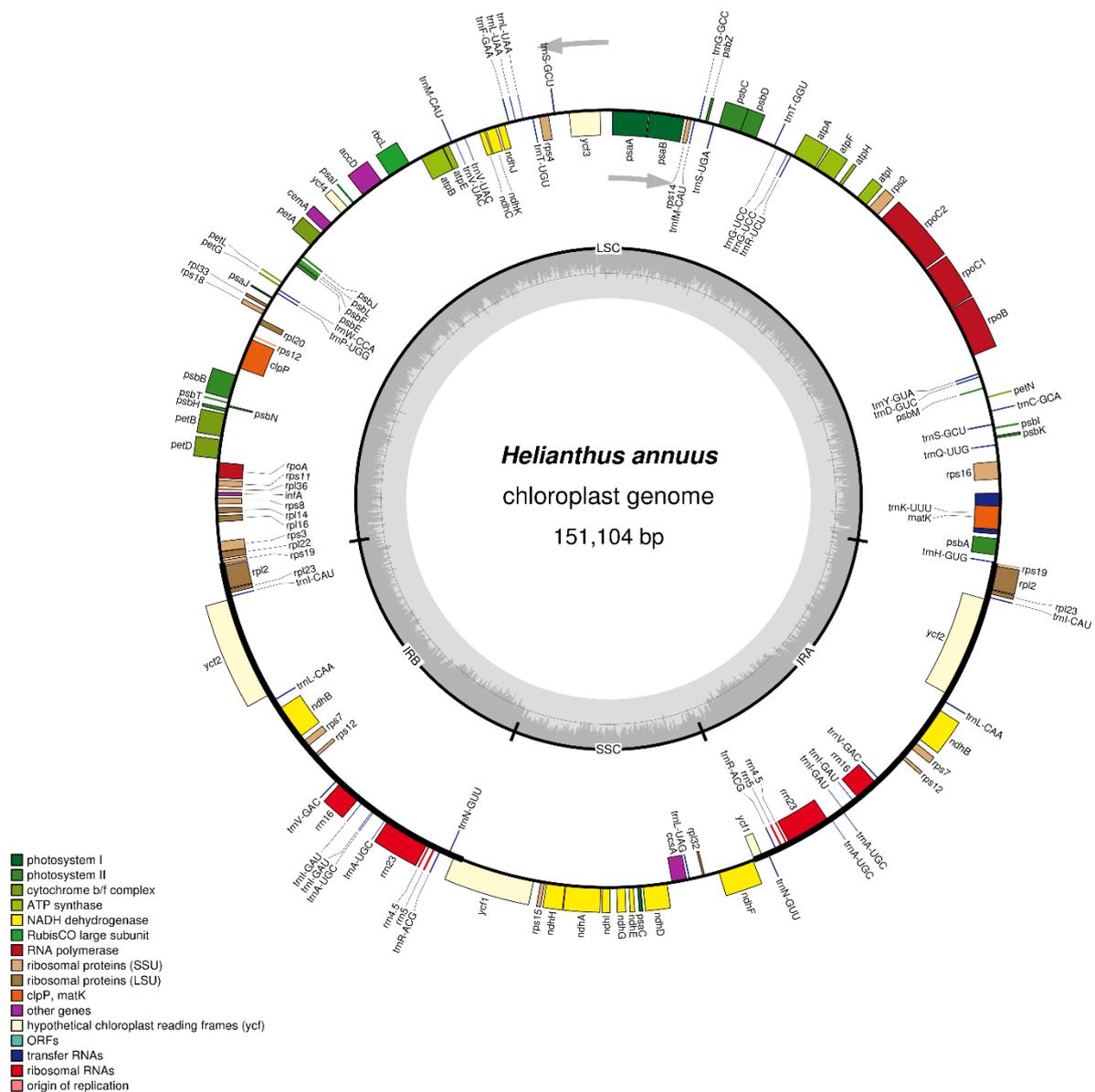


Рисунок 1.1. Генетическая карта хлоропластной ДНК подсолнечника (на основе данных Timme et al., 2007).

Среди трех геномов растительной клетки хлоропластный геном имеет наибольшую плотность генов, включая более 100 генов на 130-160 т.п.н. генома

(Воск, 2007). Хлоропластные гены цветковых растений, можно условно разделить на три основные группы (табл. 1.2) (Kahlau et al., 2006):

- 1) Гены (около 47) кодирующие белки, участвующие в фотосинтезе);
- 2) Гены (около 62), связанные с работой генетического аппарата пластид: гены рРНК и тРНК, гены, кодирующие рибосомальные белки и субъединицы РНК-полимеразы;
- 3) Гены биосинтеза (2 гена) и открытые рамки считывания (4 гена).

Таблица 1.2. Гены, кодируемые хлоропластной ДНК.

Фотосинтез	
Фотосистема I (ФСI)	<i>psaA, psaB, psaC, psaI, psaJ, ycf3, ycf4</i>
Фотосистема II (ФСII)	<i>psbA, psbB, psbC, psbD, psbE, psbF, psbH, psbI, psbJ, psbK, psbL, psbM, psbN, psbT, psbZ</i>
Цитохром b ₆ /f комплекс	<i>petA, petB, petD, petG, petL, petN</i>
АТФ-синтаза	<i>atpA, atpB, atpE, atpF, atpH, atpI</i>
Рибулозобисфосфаткарбоксилаза-оксигеназа	<i>rbcL</i>
НАДН-дегидрогеназа	<i>ndhA, ndhB, ndhC, ndhD, ndhE, ndhF, ndhG, ndhH, ndhI, ndhJ, ndhK</i>
Белок тилакоидной мембраны	<i>cemA</i>
Генетические функции пластид	
Инициация транскрипции	<i>infA</i>
РНК полимеразы	<i>rpoA, rpoB, rpoC1, rpoC2</i>
Матураза	<i>matK</i>
Протеаза	<i>clpP</i>
Рибосомальные РНК	<i>rrn4.5, rrn5, rrn16, rrn23</i>
Рибосомальные белки	<i>rps2, rps3, rps4, rps7, rps8, rps11, rps12, rps14, rps15, rps16, rps18, rps19, rpl2, rpl14, rpl16, rpl20, rpl22, rpl23, rpl32, rpl33, rpl36</i>
Транспортные РНК	<i>trnA-UGC, trnC-GCA, trnD-GUC, trnE-UUC, trnF-GAA, trnG-GCC, trnG-UCC, trnH-GUG,</i>

	<i>trnI-CAU, trnI-GAU, trnK-UUU, trnL-CAA, trnL-UAA, trnL-UAG, trnM-CAU, trnfM-CAU, trnN-GUU, trnP-UGG, trnQ-UUG, trnR-ACG, trnS-GCU, trnS-GGA, trnS-UGA, trnT-GGU, trnT-UGU, trnV-GAC, trnV-UAC, trnW-CCA, trnY-GUA</i>
Биосинтез и ОРС	
Ацетил-СоА-карбоксилаза	<i>accD</i>
Комплекс биосинтеза цитохрома С	<i>ccsA</i>
Неизвестная функция	<i>ycf1, ycf2, ycf15, sprA</i>

Для большинства цветковых растений число и состав генов соответствует описанному в таблице 1.2. Однако, у паразитических покрытосеменных растений, пластидные геномы радикально отличаются по числу генов. У паразитических растений отсутствует способность к автотрофному питанию, что отражается на числе активно транскрибирующихся хлоропластных генов, главным образом за счет потери или псевдогенизации фотосинтетических генов. Например, хлДНК паразитического растения *Epifagus virginiana* (семейство *Orobanchaceae*) состоит всего из 70 т.п.н., что в два раза меньше, чем у большинства фотосинтезирующих цветковых растений, и содержит всего 25 белок кодирующих генов, 4 гена рРНК и 23 тРНК (Wolfe et al., 1992). У другого паразита *Lathraea squamaria* (семейство *Orobanchaceae*) хлДНК имеет размер 150,5 т.п.н., а 35 генов (преимущественно связанных с фотосинтезом) превратились в псевдогены (Samigullin et al., 2016).

Транскрипция пластидных генов – сложный процесс, в регуляции которого задействован ряд молекул: РНК-полимеразы, сигма-факторы, транскрипционные регуляторы, белки пластидных нуклеоидов и различные сигнальные молекулы (Синявская и др., 2015). Транскрипция пластидного генома цветковых растений в значительной степени зависит от экспрессии генома ядра. Экспрессия пластидных генов осуществляется двумя разными типами полимераз: РНК-полимеразой бактериального типа (PEP), большинство субъединиц которой кодируются в хлДНК и фагоподобной РНК-полимеразой (NEP), кодируемой ядром (Borner et al.,

2015). РНК-полимеразы распознают разные типы промоторов и отличаются транскрипционной активностью в разных типах пластид. Важным отличием от ядерных генов является то, что большинство пластидных генов транскрибируются в виде полицистронных матриц. Гены, кодирующие белки с общими функциями, в ряде случаев экспрессируются в составе одного оперона, что определяет их скоординированную работу и стехиометрическую аккумуляцию. Примером является *proB* оперон, в состав которого входят три гена РНК полимеразы *proB*, *proC1*, *proC2*. У разных видов растений пластидные опероны консервативны (Синявская и др., 2015).

1.3. Особенности структурно-функциональной организации митохондриальной ДНК цветковых растений

Организация митохондриальной ДНК, как и хлоропластной ДНК, во многом схожа с таковой у прокариотических организмов. Тем не менее митохондриальный геном растений характеризуется большим количеством особенностей, отличающих его от ядерного и хлоропластного геномов. Митохондриальный геном высших растений отличается уникальностью его размеров, а также сложностью и особенностями эволюционной динамики. Он в десятки и даже сотни раз больше по величине, чем митохондриальные геномы животных, грибов и пластидные геномы растений и может составлять 150-2900 т.п.н. (Daniell, Chase, 2004, Alverson et al., 2010). По сравнению с хлДНК размер мтДНК сильно варьируется даже у близкородственных видов (Gualberto et al., 2014). Например, у некоторых представителей семейства *Cucurbitaceae* хлДНК варьируется в пределах 155-157 т.п.н., при этом размер мтДНК отличается многократно, находясь в пределах 380-2740 т.п.н. (табл. 1.3)

Большой вклад в увеличение размера митохондриального генома растений вносят повторяющиеся последовательности, АТ-богатые некодирующие регионы и протяженные интроны. Наличие протяженных повторов - уникальное свойство генома митохондрий растений. Количество таких повторов в митохондриальном

геноме растений сильно варьирует: *Nicotiana tabacum* содержит 3 пары больших повторов размером 18 т.п.н., 6,9 т.п.н. и 4,7 т.п.н. (Sugiyama et al., 2005), у *Helianthus annuus* обнаружено два повтора размером 12,9 п.н. и шесть повторов размером 200-750 п.н. (Makarenko et al., 2018), а у *Triticum aestivum* – 10 пар повторов (Ogihara et al., 2005).

Таблица 1.3. Размер хлоропластного и митохондриального геномов у некоторых видов рода *Cucurbitaceae*.

Вид	Размер хлДНК(т.п.н.)	Размер мтДНК(т.п.н.)	Источники
<i>Citrullus lanatus</i>	156.9	379.2	Alverson et al., 2010
<i>Cucurbita pepo</i>	156.3	982.8	Alverson et al., 2010, Zhang et al., 2018
<i>Cucumis sativus</i>	155.3	1685	Plader et al., 2007, Alverson et al., 2010
<i>Cucumis melo</i>	156	2740	Rodriguez-Moreno et al., 2011

Также размер мтДНК растений может сильно варьировать даже внутри одного вида из-за включения ядерных и хлоропластных последовательностей (Morley, Nielsen, 2017). Большая часть нуклеотидной последовательности мтДНК цветковых растений является некодирующей, поэтому различия в количестве митохондриальных генов между разными видами растений не зависят от размера митохондриального генома. Например, последовательность митохондриального генома огурца (*Cucumis sativus*) более чем на 1300 т.п.н. превышает размер мтДНК резуховидки Таля (*Arabidopsis thaliana*), но при этом содержит всего на 4 белок-кодирующих гена больше (Alverson et al., 2011).

Еще одной особенностью мтДНК цветковых растений является ее ploидность в клетке. Ploидность мтДНК растений значительно ниже, чем таковая животных, а также чем ploидность хлДНК растений. Результаты анализа ploидности нескольких митохондриальных генов у *Arabidopsis thaliana*, *Nicotiana tabacum* и *Hordeum vulgare* показали, что отдельные митохондрии могут содержать только

часть генома или не содержать ДНК вообще (Preuten et al., 2010). Тем не менее, плоидность мтДНК в растительной клетке выше, чем плоидность ядерной ДНК. Количество копий мтДНК на одну митохондрию варьирует от 50 до 500 единиц и зависит от вида растения, типа ткани, возраста растения, условий окружающей среды. На примере кукурузе было показано, что с увеличением возраста растения количество копий мтДНК снижается (Oldenburg et al., 2013). А клетки корневой системы растений, как правило, содержат значительно больше копий мтДНК, чем клетки листа (Kumar et al., 2014; Ma, Li, 2015).

Наиболее отличительной особенностью молекул мтДНК растений является динамичность структуры. В 70-80-х годах XX века было положено начало исследованиям структуры митохондриальной ДНК растений. Метод электронной микроскопии позволил выявить как кольцевые, так и линейные молекулы мтДНК в различных соотношениях (Даниленко, Давыденко, 2003). Считалось, что митохондриальный геном растений представляет собой кольцевую хромосому, а обнаруженные линейные молекулы являются результатом ее повреждения в процессе приготовления препарата для электронной микроскопии. Однако, в результате дальнейших исследований мтДНК были выявлены кольцевые молекулы меньшего размера, чем «главная хромосома», а также линейные молекулы с одним или несколькими разветвлениями, сложные розеткообразные молекулы. Позже было показано, что в митохондриях только малая доля молекул мтДНК представлены, так называемой «основной кольцевой хромосомой», которая содержит весь набор митохондриальных генов и соответствующая по размеру полному геному (Allen et al., 2007; Smith, Keeling 2013). Чаще молекулы ДНК в митохондриях представлены в виде большого числа субгеномных колец и линейных молекул (Liu et al., 2011; Wu et al., 2015), что несомненно, осложняет исследование мтДНК.

Наличие большого числа различных структурных изоформ мтДНК связано, в том числе, и с высокой частотой рекомбинации мтДНК (Wu et al., 2015). В последовательностях мтДНК цветковых растений присутствует большое число гомологичных повторов, которые подвергаются частым гомологичным

рекомбинациям, приводящим к образованию многочисленных конфигураций молекул мтДНК (рис. 1.2) (Sloan, 2013, Gualberto et al., 2014). Нарушение гомологичной рекомбинации или же ошибки при процессах репарации, приводят к накоплению альтернативных конфигураций мтДНК, которые вносят вклад в гетерогенность изоформ молекул мтДНК (Alverson et al., 2011; Sloan et al., 2012).

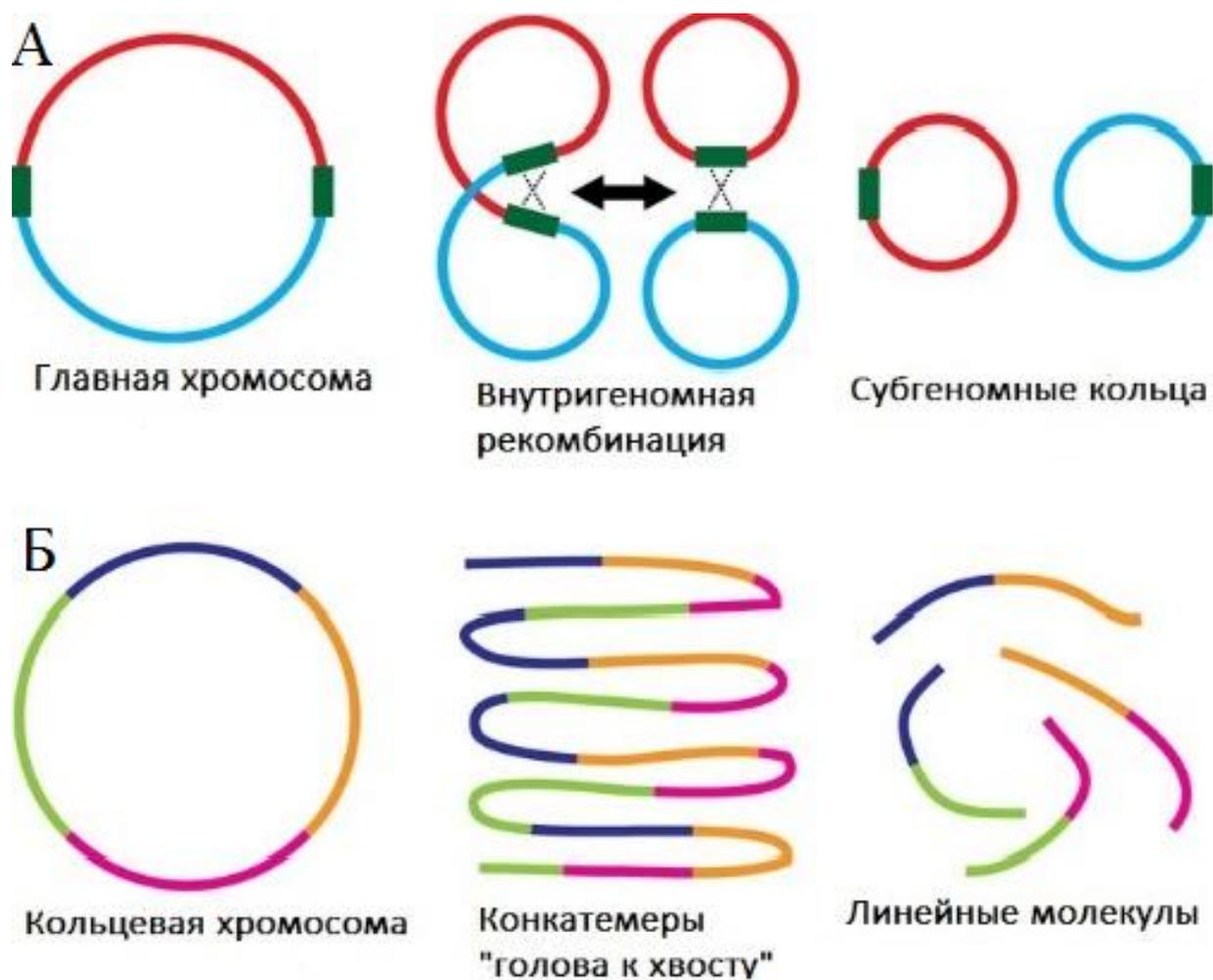


Рисунок 1.2. Модели преобразования конфигураций структуры митохондриального генома растений. А – Классическая модель, показывающая как при рекомбинации между гомологичными повторами (обозначены зеленым цветом) главной хромосомы могут образовываться субгеномные кольца. Б – Альтернативные модель, показывающая образование линейных форм мтДНК из кольцевой формы (Sloan, 2013).

Имеются также данные о том, что структура мтДНК может изменяться в зависимости от условий окружающей среды. Например, в условиях стресса

растения могут снижать экспрессию ядерных генов, которые способствуют подавлению рекомбинации между короткими повторами митохондриального генома (Davila et al., 2011). Так, например, можно объяснить наличие обширных структурных перестроек генома митохондрий и высокую частоту встречаемости суперспирализованных молекул мтДНК при выращивании растительных тканей в условия *in vitro* (Sun et al., 2012).

Благодаря увеличению доступности секвенирования растительных митохондриальных геномов стало возможным обнаружить существенные различия в ее структурной организации между видами. Например, в результате недавних исследований было установлено, что в процессе эволюции как минимум в двух независимых группах покрытосеменных появились мультихромосомные митохондриальные геномы. Впервые такой геном был обнаружен у огурца (*Cucumis sativus*). Митохондриальный геном этого вида имеет размер 1685 т.п.н., и большая его часть представлена в виде основной кольцевой хромосомы размером 1556 т.п.н. (Alverson et al., 2011). Кроме основной кольцевой хромосомы также выявлены еще две автономные кольцевые хромосомы размером 45 т.п.н. и 84 т.п.н., содержащие общую пару рекомбинационных повторов размером 3,6 т.п.н. Эти две хромосомы могут существовать и в объединенной конформации – кольцевой молекулы размером 129 т.п.н. Митохондриальные хромосомы размером 45 т.п.н. и 84 т.п.н. огурца не имеют больших гомологичных повторов в основной хромосомой, и до сих пор не доказана возможность рекомбинации между ними и основной митохондриальной хромосомой. Также не было обнаружено сходство хромосом с митохондриальными плазмидами (Alverson et al., 2011). Такая мультихромосомная организация митохондриального генома огурца отличается от типичной структуры мтДНК других видов растений.

Мультихромосомная организация митохондриального генома также была обнаружена у представителей рода *Silene* (*Caryophyllaceae*) (Sloan et al., 2012). Изучение нескольких популяций *S. vulgaris* показало, что митохондрии представителей данного вида содержат от одной до четырех небольших автономных хромосом, составляющих от 1 до 15 % от всего митохондриального

генома (Sloan et al., 2012). Интересно, что многие из мелких автономных хромосом *Cucumis* и *Silene* не содержат генов, в связи с чем неясно, какую функциональную роль они играют (Alverson et al., 2011; Sloan et al., 2012). Вопрос о том, как мультихромосомные митохондриальные геномы подвергаются репликации и как они наследуются при делении митохондрий, остается открытым.

Отличительной чертой организации митохондриального генома растений является наличие митохондриальных плазмид - кольцевых или линейных молекул ДНК размером 1-20 т.п.н, негомологичных геномной мтДНК (Handa, 2008, Gualberto et al., 2014). Митохондриальным плазмидам свойственна способность автономной репликации, в связи с чем они находятся в высокой стехиометрии относительно основного митохондриального генома. Паттерн митохондриальных плазмид видоспецифичен, но, тем не менее, могут наблюдаться его вариации внутри вида (Gualberto et al., 2014). Последовательности большинства плазмид растений не имеют сходства с основным митохондриальным геномом, а также не могут считаться субгеномными молекулами (Robison, Wolyn, 2005). На сегодняшний день не выявлено корреляций между наличием плазмид и какими-либо фенотипическими признаками, однако количество митохондриальных плазмид в клетках не постоянно и изменяется в течение развития растения (Handa, 2008). Несмотря на автономность митохондриальных плазмид их появление и количество копий частично зависят от влияния ядра растительной клетки (Gualberto et al., 2014).

Плазмиды митохондрий растений могут быть разделены на два класса: кольцевые и линейные. Кольцевые плазмиды были обнаружены у многих видов растений: кукурузы, свеклы, сорго, фасоли, подсолнечника, риса и др. (Даниленко, Давыденко, 2003). Размер кольцевых плазмид чаще всего колеблется в промежутке от 1 до 2,5 т.п.н., но иногда встречаются плазмиды размером 9 т.п.н. и более (Gualberto et al., 2014). Не было установлено гомологии кольцевых плазмид ни с мтДНК, ни с хлДНК, однако плазмиды могут включать последовательности нуклеотидов, имеющие сходство с ядерной ДНК (Даниленко,

Давыденко, 2003). Из-за небольшого размера кольцевых плазмид в них не было обнаружено протяженных (более 1 т.п.н.) открытых рамок считывания. Но в плазмидах выявлены небольшие (до 1 т.п.н.) транскрипционно активные ОРС, функция которых до сих пор не ясна (Gualberto et al., 2014). Примечательно, что была показана взаимосвязь между ЦМС фенотипом *Beta vulgaris* и ОРС, транскрибирующейся в кольцевой плазмиде (Thomas, 1986).

Линейные плазмиды были обнаружены у многих видов растений, в том числе: кукурузы, свеклы, сорго, капусты, моркови и др. (Handa, 2008). Линейные плазмиды по сравнению с кольцевыми плазмидами встречаются в митохондриях растений реже, но обычно состоят из большего числа нуклеотидов - от 2 до 20 т.п.н. (Handa, 2008, Gualberto et al., 2014). Также линейные плазмиды обычно содержат терминальные инвертированные повторы, и их 5'-концы ковалентно связаны со специфичными белками, предположительно участвующими в инициации репликации на обоих 5'-концах. Такие признаки делают линейные плазмиды схожими с митохондриальными плазмидами грибов, вирусными ДНК и транспозонами (Handa, 2008). На данный момент наиболее изученными являются S1, S2 и 2,3 т.п.н. плазмиды *Zea mays*, 10,4 т.п.н. плаزمиды *Beta vulgaris* и 11,6 т.п.н. плазмиды *Brassica napus* (Handa, 2008). Все линейные плазмиды (кроме 2,3 т.п.н. плазмиды *Z. mays*) содержат несколько открытых рамок считывания (до 6 в 11,6 т.п.н. плазмиде *B. napus*), по длине превышающих таковые в кольцевых плазмидах. Считается, что некоторые линейные плазмиды кодируют ДНК- или РНК-полимеразы фагового типа. Наличие генов ДНК- и РНК-полимераз в линейных плазмидах поддерживает идею о том, что они способны к автономной репликации и транскрипции (Smith, Keeling, 2013). Для 11,6 т.п.н. плазмиды *B. napus* была установлена экспрессия открытых рамок считывания с неизвестной функцией, которые скорее всего являются функционально активными генами (Gualberto et al., 2014).

В митохондриях растений также были идентифицированы РНК-плазмиды, которые являются наименее изученной группой. Они были обнаружены у сахарной свеклы, кукурузы, риса и представителей рода *Brassica* (Даниленко,

Давыденко, 2003). Для РНК-плазмид не было выявлено гомологии с ДНК-плазмидами. Происхождение данного типа плазмид неясно.

Митохондриальным плазмидам свойственно однородительское наследование по материнскому типу. Тем не менее, отклонение от строго однородительского наследования наблюдается у 10,4 т.п.н. плазмиды *B. vulgaris* и 11,6 т.п.н. плазмиды *B. napus* (Oshima, Handa, 2012).

Происхождение митохондриальных плазмид по большей части неизвестно. Сходство между нуклеотидными последовательностями некоторых кольцевых плазмид у разных видов растений предполагает наличие общего предка (Gualberto et al., 2014). Однако, в митохондриальных плазмидах часто встречаются последовательности, представленные в ядерном геноме растений, что может свидетельствовать о возникновении митохондриальных плазмид из ядерной ДНК, вследствие передачи генетической информации между двумя геномами. Последовательности открытых рамок считывания и структурная организация линейных плазмид отличаются у разных видов, что в свою очередь свидетельствует об их независимом происхождении (Handa, 2008). И, наконец, существуют предположения о том, что горизонтальный перенос ДНК, в частности, от грибов, мог сыграть роль в появлении митохондриальных плазмид (Handa, 2008).

Большой размер митохондриального генома растений не соответствует его кодирующей способности. Например, митохондриальный геном подсолнечника, в два раза больше по размерам, чем хлоропластный, но содержит почти вдвое меньше генов (Makarenko et al., 2016). В мтДНК цветковых растений содержится только небольшая часть генов (50-60) необходимых для функционирования митохондрий. Подавляющее большинство генов, обеспечивающих митохондриальные функции, детерминировано в ядре. Часть митохондриальных генов обеспечивают экспрессию и трансмиссию генетической системы мтДНК, в том числе: *matR*, гены рибосомальной РНК (*rrn5*, *rrn18*, 26) и тРНК, гены, кодирующие рибосомальные белки (Даниленко, Давыденко, 2003).

Гены тРНК обнаруженные в мтДНК цветковых растений могут иметь как митохондриальное так и пластидное происхождение. Набор митохондриальных тРНК отличается у однодольных и двудольных растений, например, ген *trnG* отсутствует в мтДНК большинства однодольных растений (Liao, et al., 2018). В таблице 1.4 отражено число генов тРНК выявленных у разных видов цветковых растений (Кноор, 2004, Daniell, Chase, 2004, Liao, et al., 2018). Следует заметить, что у цветковых растений число генов тРНК, кодируемых мтДНК, недостаточно для трансляции всех протеиногенных аминокислот, в связи с чем часть тРНК кодируется в ядре и затем импортируется в митохондрии (Даниленко, Давыденко, 2003).

Таблица 1.4. Гены тРНК в митохондриальных геномах цветковых растений

Ген	Однодольные виды					Двудольные виды												
	<i>Hordeum vulgare</i>	<i>Oryza sativa</i>	<i>Sorghum bicolor</i>	<i>Triticum aestivum</i>	<i>Zea mays</i>	<i>Arabidopsis thaliana</i>	<i>Brassica juncea</i>	<i>Brassica napus</i>	<i>Brassica oleracea</i>	<i>Cucumis melo</i>	<i>Cucurbita pepo</i>	<i>Glycine max</i>	<i>Nicotiana tabacum</i>	<i>Raphanus sativus</i>	<i>Ricinus communis</i>	<i>Vicia faba</i>	<i>Vigna radiate</i>	<i>Vitis vinifera</i>
тРНК митохондриального происхождения																		
trnC-GCA	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1
trnD-GUC	3	1	1	2	3	0	1	1	1	0	0	0	1	1	2	0	0	1
trnE-UUC	1	1	1	1	2	1	1	1	0	1	1	1	1	1	1	1	1	1
trnF-GAA	1	0	2	0	0	0	0	0	0	1	1	1	0	0	1	1	1	1
trnG-GCC	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1
trnH-GUG	0	0	0	1	0	0	0	0	1	0	0	0	1	0	1	1	0	1
trnI-CAU	1	2	1	0	2	1	1	1	1	0	1	1	1	1	0	0	1	0
trnK-UUU	3	1	1	4	1	5	1	1	1	0	1	1	1	1	1	2	1	2
trnL-UAA	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
trnM-CAU	2	1	1	3	1	1	0	1	2	3	0	0	0	1	3	0	0	0
trnM-CAU	3	1	1	2	1	1	1	1	0	0	1	4	2	1	0	1	2	2
trnN-GUU	0	1	0	0	1	1	0	0	0	0	0	1	2	0	0	0	0	0

trnP-UGG	1	2	1	2	2	1	1	1	1	1	1	0	1	1	2	1	1	1
trnQ-UUG	2	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1
trnR-UCU	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
trnS-AGA	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
trnS-GGA	1	0	0	0	0	0	1	1	1	0	0	0	0	0	0	1	0	
trnS-UGA	0	1	0	1	1	2	1	1	1	1	0	0	1	1	1	0	0	
trnS-GCU	1	1	1	1	1	2	1	1	1	0	0	1	1	2	1	0	1	
trnY-GUA	1	1	1	1	1	3	1	1	1	1	1	1	1	1	1	1	1	
тРНК пластидного происхождения																		
trnC-GCA	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
trnD-GUC	0	0	0	0	0	1	0	0	0	3	0	1	1	0	0	1	0	1
trnE-UUC	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1
trnF-GAA	0	1	1	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0
trnG-UCC	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
trnH-GUG	0	1	1	0	1	0	2	1	1	3	1	1	0	1	0	0	1	0
trnI-CAU	0	0	0	0	2	0	0	0	0	0	0	0	1	0	0	0	0	0
trnK-UUU	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
trnL-UAG	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
trnL-CAA	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0
trnM-CAU	0	1	1	0	0	0	0	0	0	2	1	1	1	0	0	1	1	0
trnM-CAU	2	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
trnN-GUU	1	0	1	1	1	0	1	1	1	2	1	0	0	1	1	1	1	3
trnP-UGG	0	0	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	2
trnR-ACG	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0
trnR-UCU	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
trnS-AGA	1	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
trnT-GGU	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
trnW-CCA	0	1	1	1	1	0	1	1	1	1	6	0	1	1	1	1	1	1
trnY-GUA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

Гомология и количество митохондриальных белков кодирующих генов растений еще малоизучены. Число белков кодирующих митохондриальных генов различается у разных видов цветковых растений, при этом наибольшие вариации наблюдаются в числе генов рибосомальных белков (табл. 1.5).

Таблица 1.5. Гены, кодирующих рибосомальные белки, в митохондриальных геномах цветковых растений.

Ген	Однолетние виды					Многолетние виды												
	<i>Hordeum vulgare</i>	<i>Oryza sativa</i>	<i>Sorghum bicolor</i>	<i>Triticum aestivum</i>	<i>Zea mays</i>	<i>Arabidopsis thaliana</i>	<i>Brassica juncea</i>	<i>Brassica napus</i>	<i>Brassica oleracea</i>	<i>Cucumis melo</i>	<i>Cucurbita pepo</i>	<i>Glycine max</i>	<i>Nicotiana tabacum</i>	<i>Raphanus sativus</i>	<i>Ricinus communis</i>	<i>Vicia faba</i>	<i>Vigna radiata</i>	<i>Vitis vinifera</i>
<i>rpl2</i>	-	+	-	-	-	+	+	+	+	+	+	-	+	+	-	-	-	+
<i>rpl5</i>	+	+	-	+	-	+	+	+	+	-	+	-	+	+	+	+	+	+
<i>rpl10</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-
<i>rpl14</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+
<i>rpl16</i>	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
<i>rpl20</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+
<i>rpl32</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+
<i>rpl33</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+
<i>rpl36</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+
<i>rps1</i>	+	+	+	+	+	-	-	-	-	+	+	+	-	-	+	-	+	+
<i>rps2</i>	+	+	-	+	+	-	-	-	-	-	-	-	-	-	-	+	-	-
<i>rps3</i>	+	+	+	+	+	+	+	+	+	+	+	-	+	-	+	+	+	+
<i>rps4</i>	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
<i>rps7</i>	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	-	-	-
<i>rps8</i>	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-
<i>rps10</i>	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
<i>rps12</i>	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
<i>rps13</i>	+	+	+	+	-	-	-	-	-	+	+	+	+	-	+	-	-	-

<i>rps14</i>	+	-	-	-	+	-	+	+	+	-	+	+	+	+	-	+	+	+
<i>rps19</i>	-	+	-	+	-	-	-	-	-	-	-	-	+	-	+	-	-	+

Известно, что митохондриальный геном растений кодирует ряд компонентов электрон транспортных комплексов. НАДН-дегидрогеназный комплекс (комплекс I), имеющий огромное значение для протекания процессов клеточного дыхания и окислительного фосфорилирования, у растений состоит из более чем 40 субъединиц (Klodmann et al., 2010, Sazanov 2012). В основном в мтДНК цветковых растений локализованы 9 генов, кодирующих субъединицы НАДН-дегидрогеназного комплекса - *nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, *nad6*, *nad7*, *nad9*.

Сукцинатдегидрогеназа (комплекс II) участвует как в цикле трикарбоновых кислот, так и дыхательной цепи переноса электронов и состоит из 8 субъединиц: 4-х субъединиц характерных для всех эукариот и 4-х субъединиц специфичных для растений (Millar, et al 2004; Popov et al., 2010) У растений только два гена (*sdh3*, *sdh4*) сукцинатдегидрогеназы локализованы в мтДНК, при этом у большинства цветковых растений эти гены транслоцированы в ядерный геном (Wendel et al., 2012, Liao, et al., 2018).

За исключением цитохрома-*b* (митохондриальный ген *cob*) все белки комплекса цепи переноса электронов цитохромов *b-c1* (комплекс III) кодируются в ядре.

Цитохром *c*-оксидаза (комплекс IV) катализирует перенос электронов с цитохрома-*c* на кислород и состоит из 14 субъединиц (Millar, et al 2004). Три основные субъединицы комплекса IV (гены *coxI*, *coxII*, *coxIII*) кодируются в мтДНК цветковых растений, за исключением некоторых представителей бобовых, у которых ген *coxII* может находиться в ядерной ДНК (Alverson et al., 2011, Wendel et al., 2012).

АТФ-синтаза (комплекс V) состоит из 11 субъединиц. У большинства растений 5 из 11 субъединиц кодируются митохондриальными генами: *atp1*, *atp4*, *atp6*, *atp8*, *atp9* (Wendel et al., 2012, Liao, et al., 2018).

Также в мтДНК цветковых растений локализованы, как правило, 4 гена биосинтеза цеохрома-*c* – *ccmB*, *ccmC*, *ccmFc* *ccmFn* (Wendel et al., 2012, Liao, et al., 2018). У многих растений митохондриальный геном включает еще и слабоизученный ген *mttB* (синонимы *orfX*, *tatC*, *ymf16*), белковым продуктом которого является транслоказа. Функция транслоказы сводится к межмембранному транспорту других белков (Wendel et al., 2012).

У некоторых видов митохондриальные гены могут быть дублированы, например, ген *atp6* у *Arabidopsis thaliana* и *Glycine max*, ген *atp1* у *Cucumis melo*, *Vicia faba*, *Vitis vinifera*, *Zea mays* (Liao, et al., 2018).

Кроме генов с известной функцией митохондриальный геном растений включает большое число открытых рамок считывания. В мтДНК *Arabidopsis thaliana* выявлено более 100 ОРС состоящих, преимущественно из 150-350 п.н. (Sloan, et al., 2018). Большинство ОРС не несут никакой функциональной нагрузки и являются случайными образованиями. Но некоторые ОРС активно транскрибируются и являются консервативными у разных видов, в связи с чем, можно говорить о них как о генах, функции которых пока неясны. Примером таких активно экспрессирующихся ОРС у подсолнечника являются *orf777*, *orf873* (Makarenko et al., 2018).

В состав некоторых ОРС входят последовательности идентичные митохондриальным генам, такие ОРС, можно назвать химерными генами (Даниленко, Давыденко, 2003). Химерные гены нарушают баланс взаимодействия между ядерным и митохондриальными геномами, и тем самым могут привести к возникновению фенотипа с цитоплазматической мужской стерильностью (ЦМС). Цитоплазматическая мужская стерильность (ЦМС) - наследуемая по материнской линии неспособность растений продуцировать жизнеспособную пыльцу, причиной которой чаще всего являются мутации в мтДНК (Анисимова, Гавриленко, 2016). Явление ЦМС широко распространено среди растений, описано у более чем 150 видов цветковых растений (Chen, Liu, 2014). В естественных популяциях растений ЦМС фенотип может возникнуть как из-за спонтанных мутаций в мтДНК, так и в результате межвидовой гибридизации при

несовместимости внутриклеточных геномов (Даниленко, Давыденко, 2003). ЦМС-Rf системы широко используются для получения гибридных семян F1 различных сельскохозяйственных культур: подсолнечника, кукурузы, свеклы, моркови, сорго, риса и др. (Liu et al., 2011; Kubo et al., 2011).

МтДНК ЦМС-типа обычно характеризуется множественными структурными перестройками, некоторые из них изменяют кодирующие части генов либо их локализацию и приводят к образованию химерных транскрибируемых рамок считывания (Брагин и др., 2011). Довольно редко ЦМС фенотип возникает из-за нарушений в генах НАДН-дегидрогеназного комплекса (Horn et al., 2014). Делеции экзонов гена *nad7* являются причиной двух ЦМС типов *Nicotiana sylvestris* (Gutierrez et al., 1997). В мтДНК ЦМС линии RT98A *Oryza rufipogon* возникает новая открытая рамка считывания - *orf113*. *Orf113* состоит из последовательности нуклеотидов идентичных гену *nad9*, и неизвестной последовательности (Igarashi et al., 2013). Более частой причиной ЦМС фенотипа цветковых растений являются химерные гены цитохром *c*-оксидазы. В мтДНК ЦМС линии *Sorghum bicolor* обнаружен измененный ген *coxI*: с удлинённой последовательностью на 3' и 5' концах и кодирующий белок, содержащий на 101 аминокислоту больше, чем нормальный белок (Bailey-Serres et al., 1986). Причиной ЦМС у *Raphanus sativus* является - *orf463*, состоящий из части гена *coxI* и 1261 п.н. неизвестного происхождения (Park et al., 2013). *Cox2*-химерные гены являются причиной ЦМС у кукурузы, риса, свеклы, петунии (Даниленко, Давыденко, 2003, Horn et al., 2014). Наиболее частой причиной ЦМС фенотипа цветковых растений становится наличие в митохондриальном геноме *atp6*- , *atp8*-, *atp9*-химерных генов (Hanson, Bentolila, 2004; Horn et al., 2014). Химерные гены АТФ-синтетазы обнаружены у кукурузы (Dewey et al., 1987), сои (Jiang, 2011), горчицы (Dieterich et al., 2003), подсолнечника (Reddemann, Horn, 2018) и др.

Митохондриальные гены растений могут быть представлены единичным экзоном или же иметь интрон-экзонную структуру. У покрытосеменных наибольшее количество интронов обнаружено в генах, кодирующих субъединицы

НАДН-дегидрогеназы - *nad1*, *nad2*, *nad4*, *nad5*, *nad7* (Liao, et al., 2018). Как правило, в генах *cob*, *coxI*, *coxII* у цветковых растений присутствуют интроны (Daniell, Chase, 2004, Wendel et al., 2012). В митохондриальных генах принято различать интроны I и II группы, которые имеют структурные отличия и механизмы сплайсинга (Кноор, 2004).

Полногеномное секвенирование значительно способствует пониманию генетических особенностей митохондрий. На сегодняшний день получены полные нуклеотидные последовательности мтДНК у более чем 1000 видов растений (http://megasun.bch.umontreal.ca/ogmp/projects/other/mt_list.html), однако только часть из них аннотирована. Примером митохондриального генома с локализованными генами является мтДНК *Arabidopsis thaliana* (рис. 1.3). Митохондриальный геном модельного организма *Arabidopsis thaliana* был полностью секвенирован и аннотирован еще в 1997 году (Unsold et al., 1997). При этом за последние 20 лет информация о его мтДНК неоднократно дополнялась и уточнялась (Sloan, et al., 2018).

Митохондриальная ДНК цветковых растений также имеет ряд особенностей реализации наследственной информации на уровне транскрипции, сплайсинге и процессинге мРНК. Обнаружено, что отличительной чертой транскрипции мтДНК является образование с одного участка генома большого числа варьирующих в размерах мРНК молекул (Даниленко, Давыденко, 2003). Многие митохондриальные гены имеют несколько промоторов, как правило, состоящих из 17-18 п.н. Митохондриальные гены, подобно хлоропластным, экспрессируются как в виде индивидуальных транскриптов, так и в составе полицистронных матриц. Некоторые кластеры генов консервативны среди высших растений, например, *rrn5-rrn18*, *nad1-matR*, *rps12-nad3*, *rps3-rpl16* (Liao, et al., 2018). Ключевую роль в регуляции транскрипции митохондриальных генов играет ядерный геном. В отличие от хлДНК в мтДНК нет генов РНК-полимеразы, а сам фермент кодируется ядерной ДНК.

Характерной посттранскрипционной модификацией мРНК в митохондриях является ее редактирование (Daniell, Chase, 2004). В мРНК цитозин может подвергаться превращению в уридин (C-U редактирование), что в свою очередь может приводить к изменению кодона CGG на UGG и замене в белках аргинин на триптофан (Gray, 2009). Также редактирование РНК может приводить к новым стартовым кодонам (замена ACG на AUG) или стоп кодонам (замена CGA на UGA). Например, у большинства покрытосеменных в мРНК гена *nad1* кодон инициации трансляции (AUG) возникает именно за счет редактирования первичного транскрипта (Daniell, Chase, 2004). Исследования показали, что у растений некоторые митохондриальные мРНК также подвергаются так называемому «обратному редактированию», то есть замене уридина на цитозин (Wendel et al., 2012).

Таким образом, геномы энергетических органелл растительной клетки, имеют как сходные, так и уникальные черты структурно-функциональной организации и отличаются рядом особенностей от ядерного генома. Хлоропластные геномы цветковых растений, как правило, консервативны в размерах и организации, в то время как структура митохондриальных геномов, наоборот, может сильно отличаться даже у близких видов. С другой стороны, следует отметить, что уровень нуклеотидных замен в хлоропластных генах в несколько раз выше, чем в митохондриальных (Drouin et al., 2008; Hoekstra et al., 2017).

ГЛАВА 2. Материалы и методы исследования

2.1. Объекты исследования

Объектом исследования служили образцы 5 однолетних и 16 многолетних видов подсолнечника (табл. 2.1) из Мировой коллекции ВИР. В качестве материала использовали листья подсолнечника, отобранные на Кубанской опытной станции Всероссийского института генетических ресурсов растений им. Н.И. Вавилова.

Таблица 2.1. Список исследованных образцов однолетних и многолетних видов подсолнечника.

Вид	Номер интродукции ВИР
Однолетние виды рода <i>Helianthus</i> L.	
<i>H. annuus</i> L.	398937
<i>H. argophyllus</i> Torr. & Gray	1000
<i>H. debilis</i> Nutt.	545666
<i>H. petiolaris</i> Nutt.	503232
<i>H. praecox</i> Engelm. & Gray	560400
Многолетние виды рода <i>Helianthus</i> L.	
<i>H. californicus</i> DC.	441063
<i>H. ciliaris</i> DC.	-
<i>H. decapetalus</i> L.	440439
<i>H. divaricatus</i> L.	545674
<i>H. giganteus</i> L.	441029
<i>H. grosseserratus</i> M. Martens	545711
<i>H. hirsutus</i> Raf.	560389
<i>H. laevigatus</i> Torr. & Gray	-
<i>H. maximiliani</i> Schrad.	440553
<i>H. mollis</i> Lam.	530453
<i>H. occidentalis</i> Riddell	441062

<i>H. rigidus</i> Nutt.	545660
<i>H. simulans</i> Watson	545659
<i>H. smithii</i> Heiser	-
<i>H. strumosus</i> L.	440679
<i>H. tuberosus</i> L.	441026

Объектом исследования также служили инбредная линия 3629 культурного и линии 398941 дикорастущего подсолнечника *Helianthus annuus* L. из коллекции ЮФУ, линия НА89 и ее ЦМС аналоги НА89(PET1), НА89(PET2), НА89(ANN2), НА89(MAX1) из коллекции ВИР. ЦМС линии были получены на основе однолетних *H. petiolaris* (PET1, PET2), *H. annuus* (ANN2) и многолетнего *H. maximillianii* (MAX1) видов рода *Helianthus* L (Leclercq, 1969; Serieys, 1996, Whelan, 1978; Whelan, 1980).

2.2. Выделение хлоропластов и митохондрий из листьев подсолнечника

Хлоропласты и митохондрии выделяли из первой пары настоящих листьев 14 дневных проростков подсолнечника (*Helianthus annuus*), выращенных в камере для роста растений KBWF720 (Binder, Германия) при температуре 26°C, 70 % влажности воздуха и 14 часовом суточном освещении. Все процедуры выделения клеточных органелл проводили при температуре 4-8°C. Листовую ткань (5 г) гомогенизировали в ступке в 25 мл STE буфера. Состав STE буфера: 50 мМ Трис-НСI (рН 7.8), 0,4 М сахараза, 5 мМ динатриевая соль этилендиаминтетрауксусной кислоты, 0,2 % бычий сывороточный альбумин и 0,2 % 2-меркаптоэтанол. Гомогенат фильтровали через один слой марли, двойной слой мираклоса (Calbiochem Behring, США) и центрифугировали в центрифуге с охлаждением Avanti J-30I (Beckman Coulter, США) по следующей схеме:

1. 5 минут при 200 g. Отбор супернатанта.
2. 5 минут при 1000 g. Отбор супернатанта.

3. 5 минут при 4000 g. Осадок промывали STE буфером, еще раз центрифугировали при тех же условиях и затем использовали для дальнейшего выделения хлоропластной ДНК. Супернатант использовали для еще одного центрифугирования.

4. 10 минут при 8000 g. Надосадочную жидкость аккуратно отбирали пипеткой, а осадок, содержащий как хлоропласты, так и митохондрии, использовали для дальнейшего выделения ДНК.

2.3. Молекулярно-генетические исследования

2.3.1. Выделение ДНК

Выделение ДНК из фракций органелл или гомогената листьев подсолнечника проводили с использованием коммерческого набора для выделения ДНК из растительных тканей «Сорб-ГМО-Б» (Синтол, Россия). Процедура выделения ДНК включала в себя следующие этапы.

1. К 50 мг образца добавляли 800 мкл лизирующего буфера и 10 мкл протеиназы К (2 мг/мл). Раствор перемешивали путем пипетирования.

2. Затем раствор инкубировали в термошейкере TS-100 (Biosan, Латвия) в течение 30 минут при температуре +60°C и режиме 1000 оборотов/минуту. Далее смесь центрифугировали 5 минут при 10000 g, и супернатант переносили в новые пробирки объемом 1,5 мл.

3. В каждую пробирку вносили по 500 мкл хлороформа. Смесь интенсивно перемешивали на вортексе в течение 15 секунд, а затем центрифугировали 10 минут при 10000 g.

4. Верхнюю водную фазу объемом 300 мкл аккуратно, не захватывая нижний слой хлороформа, переносили в новые пробирки объемом 1,5 мл, в которые предварительно вносили по 600 мкл осаждающего раствора и 25 мкл сорбента.

5. Пробирки интенсивно перемешивали на вортексе до полного ресуспензирования сорбента и смесь инкубировали при комнатной температуре в

течение 10 минут в ротаторе Multi Bio RS-24 (Biosan, Латвия) со скоростью вращения 50 оборотов/минуту. Суспензию осаждали при 8000 g 10 минут. Супернатант аккуратно декантировали.

6. К осадку добавляли 500 мкл промывающего раствора №1. Перемешивали смесь до полного ресуспензирования осадка. Центрифугировали суспензию 1 минуту при 8000 g. Не задевая осадок, удаляли супернатант.

7. К осадку добавляли 500 мкл промывающего раствора №2. Перемешивали смесь до полного ресуспензирования осадка. Центрифугировали суспензию 1 минуту при 8000 оборотов/мин. Не задевая осадок, удаляли супернатант.

8. К осадку добавляли 500 мкл промывающего раствора №3. Перемешивали смесь до полного ресуспензирования осадка. Центрифугировали суспензию 1 минуту при 8000 оборотов/мин. Не задевая осадок, удаляли супернатант.

9. Пробирки помешали в термостат при температуре +65°C, на 3-5 минуты до визуального высыхания осадка. Затем к осадку добавляли 100 мкл элюирующего раствора. Перемешивали смесь до полного ресуспензирования осадка. Затем смесь инкубировали при температуре +65°C в течение 5 минут, периодически встряхивая пробирки на вортексе. Проводили центрифугирование при 10000 g в течение 2 минут. Аккуратно отбирали супернатант объемом 75 мкл и переносили в чистые пробирки.

2.3.2. Измерение концентрации ДНК

Концентрацию выделенной ДНК измеряли с использованием флуориметра QuantiFluor ST (Promega, США) и набора реактивов QuantiFluor dsDNA System (Promega, США) согласно рекомендациям производителя. С этой целью в пробирки объемом 0,5 мл вносили по 98,5 мкл TE буфера, 0,5 мкл красителя и 1 мкл выделенной ДНК. Для калибровки использовали контрольный образец без добавления ДНК и стандарт с концентрацией ДНК 100 нг/мкл. Все измерения

проводили в трех повторностях и в дальнейшем использовали средние значения концентраций.

2.3.3. Секвенирование фрагментов ДНК методом Сэнгера

Для получения ампликонов и их дальнейшего секвенирования нами с помощью программы Primer BLAST (<https://www.ncbi.nlm.nih.gov/tools/primer-blast>) были разработаны праймеры комплементарные участкам хлоропластных генов - *atpB* (GeneID 4055708), *matK* (GeneID 4055637), *rbcL* (GeneID 4055709) и митохондриальных генов - *atp1* (GeneID 18250974), *matR* (GeneID 18250976), *nad5* (GeneID 18250966) подсолнечника.

Таблица 2.2. Праймеры, используемые для секвенирования хлоропластных и митохондриальных генов

Ген	Последовательность прямого праймера (5'-3')	Последовательность обратного праймера (5'-3')	Размер ПЦР продукта, п.н.
Хлоропластные гены			
<i>atpB</i>	AGGTTGATAACCCACAGCGG	AAAGAACCTGGGGCGTATCG	850
<i>matK</i>	AGCCAACGATCCAACCAGAG	TGGAGAAATTCCAAAGGCTAGA	1115
<i>rbcL</i>	GATTTGCGAATCCCGACTGC	GGGTGCCCTAAAGTTCCTCC	821
Митохондриальные гены			
<i>atp1</i>	TCTTTCGTTAGTTAAGCCACCT	CAGAAACGCTCAACTGTGGC	849
<i>matR</i>	ATGGCTTTTTAAGGGGGCCA	TCCAGACACATCGCACTTCC	832
<i>nad5</i>	ATCGTAAGATGTCTCGCCCG	TAGCTCCTGGGATTTCCGGGT	834

ПЦР проводили с помощью набора реактивов Tersus Plus PCR kit (Евроген, Россия) в термоциклере C1000 (Bio-Rad, США) по следующей программе: начальная денатурация при 94°C в течение 3 минут, затем 28 циклов, включающих денатурацию при 95°C – 10 секунд, отжиг при 62°C – 30 секунд, элонгация при 72°C – 60 секунд и один цикл финальной элонгации при 72°C – 5 минут. ПЦР продукты очищали с помощью набора спин-колонок Cleanup Mini

(Евроген, Россия) и анализировали методом электрофоретического разделения в 1 % агарозном геле. Затем с использованием набора флуоресцентно меченных дидезоксинуклеотидов BigDye Terminator v3.1 Cycle Sequencing Kit (Thermo Fisher Scientific, США) определяли нуклеотидные последовательности ампликонов на генетическом анализаторе ABIPrism 3130xl (Applied Biosystems, США). Полученные нуклеотидные последовательности анализировали с помощью программы BioEdit v. 7.0.5 (<http://www.mbio.ncsu.edu/BioEdit/bioedit.html>).

2.3.4. Анализ полиморфизма митохондриальных SSR локусов

Для каждого образца проводили 2 ПЦР реакции, при этом в каждой использовали пять пар праймеров для пяти SSR локусов, в результате чего получали 10 ампликонов. Праймеры были подобраны (табл. 2.3) на основе референсного митохондриального генома *H. annuus*, полученного из базы данных NCBI (NC_023337.1). Все используемые праймеры были мечены флуоресцентной меткой FAM для дальнейшего анализа размеров ПЦР продуктов. Проверку праймеров на наличие гетеро димеров проводили с использованием OligoAnalyzer 3.1 (<https://eu.idtdna.com/calc/analyzer>).

Таблица 2.3. Праймеры, используемые для анализа полиморфизма митохондриальных SSR локусов

SSR локус	Тип повтора	Последовательность прямого праймера (5'-3')	Последовательность обратного праймера (5'-3')
MT11	(A) _n	CGAACGAGATAAGGTTGTCAAATG	GGAAGAATGGATCCGAACGAA
MT4	(T) _n	GGACGCGAGAACCTAAGAAA	GAGCGCTCCTGCACTATAC
MT1	(A) _n	GCTCATCACCAGGTCCAATAG	GTTCCGGGACGGCTGTATTT
MT5	(A) _n	CCTTCTGAGCCAGGATCAATAC	CGAGATCAATTCAAAGAGGACTAGA
MT13	(C) _n	GGAGGCTTCACATACTCTTACTAC	TGTATGCCTGTTGGGTCTTT
MT14	(T) _n	GGACAGGAGAGACCCTCTTTA	GGATTGAGACTACGACTGGAATG
MT22	(T) _n	AAGATGCGAATCGGGAAGG	TCGGGTTACAGTCAACTCATAAA
MT23	(G) _n	СТCGTAGCTCCCGTTTGATAC	TGTCSTTGATGACCAACCATAA

MT20	(A) _n	CTTGTTCCTGCTATTGAGAGA	AATCGTAGGTCCCCTCGTC
MT15	(C) _n	AGTGGGAAGTAATGAGGCTAGT	CTGTCCTCGAAGGTCTCTC

ПЦР проводили с помощью набора реактивов iQ Multiplex Powermix (Bio-Rad, США). Реакционная смесь включала 12,5 мкл ПЦР-микса, 3 мкл ДНК (концентрацией 5 нг/мкл), 3 мкл праймеров (концентрацией 5 мкМ) и 6,5 мкл деионизированной воды. Реакцию амплификации проводили в термоциклере С1000 (Bio-Rad, США) по следующей программе: начальная денатурация при 95 °С в течение 3 мин., затем 27 циклов, включающих денатурацию при 95 °С – 15 сек., отжиг при 60°С – 25 сек. и элонгацию при 72 °С – 40 сек, и один цикл финальной элонгации при 72°С – 10 мин.

Размер продуктов амплификации анализировали методом капиллярного электрофореза на генетическом анализатора ABI PRISM 3130xl (Applied Biosystems, США) и программы Peak Scanner v1.0. Размер ПЦР продукта каждого SSR локуса учитывали в качестве аллельного варианта локуса. Для оценки информативности SSR локусов использовали величину информационного полиморфизма (Polymorphism Information Content (PIC)). Индекс PIC рассчитывали по формуле: $PIC=1 - \sum P_i^2$, где P_i - частота каждой аллели локуса (Чесноков, Артемьева, 2015). На основе выявленных аллельных вариантов определяли митотипы образцов. Для оценки внутри и межвидовой изменчивости митохондриальной ДНК использовали программу Treeson 1.3b.

2.3.5. Подготовка библиотек для высокопроизводительного секвенирования

Для подготовки библиотек использовали пулированную ДНК из 5 растений каждой исследуемого образца и набор реактивов Nextera XT DNA Library Prep Kit (Illumina, США). Процедура приготовления библиотек включала следующие этапы.

1) Тагментация ДНК. В микропробирки объемом 0,2 мл вносили по 5 мкл ДНК в концентрации 0,5 нг/мкл. Затем добавляли 10 мкл буфера (Tagment

DNA Buffer) и 5 мкл фермента (Amplicon Tagment Mix). Смесь перемешивали пипетированием, кратко центрифугировали и инкубировали 5 минут при температуре +55°C. Сразу после инкубации вносили нейтрализующий буфер (Neutralize Tagment Buffer) и перемешивали содержимое пипетированием.

2) Амплификация библиотек. К раствору тагментированной ДНК добавляли по 5 мкл праймера с индексом 1 (N70X) и 5 мкл праймера с индексом 2 (S5XX). Каждая библиотека имела свою уникальную комбинацию индексов 1 и 2 для дальнейшего баркодинга ридов. После внесения праймеров в раствор добавляли ПЦР-микс (Nextera PCR Master Mix). Содержимое пробирок перемешивали пипетированием, кратко центрифугировали и проводили ПЦР в термоциклере C1000 (Bio-Rad, США) по следующей программе: начальная денатурация при 72°C в течение 3 минут и 30 секунд при 95°C, затем 12 циклов, включающих денатурацию при 95°C – 10 секунд, отжиг при 55°C – 30 секунд, элонгация при 72°C – 30 секунд и один цикл финальной элонгации при 72°C – 5 минут.

3) Очистка библиотек. Очистку библиотек от побочных продуктов амплификации, таких как димеры адаптеров проводили с использованием магнитных частиц AMPure XP (Beckman Coulter, США). К полученным ранее 50 мкл ПЦР смеси добавляли 30 мкл частиц AMPure XP. Содержимое перемешивали на шейкере в течение 2 минут при 1800 оборотах/мин. Затем инкубировали 5 минут при комнатной температуре (+25°C) и помещали пробирки в магнитный штатив на 2 минуты. Пипеткой удаляли жидкость из каждой пробирки и добавляли 200 мкл 80 % этанола. Инкубировали 30 секунд и удаляли жидкость из каждой пробирки. Снова добавляли 200 мкл 80 % этанола, инкубировали 30 секунд и полностью удаляли жидкость. Содержимое пробирок сушили при комнатной температуре в течение 15 минут. Пробирки вынимали из магнитного штатива и добавляли в них по 52,5 мкл элюирующего буфера (Resuspension Buffer) с последующим перемешиванием содержимого путем пипетирования. Инкубировали 2 минуты при комнатной температуре и затем переносили на 2

минуты в магнитный штатив. Осторожно, не задевая магнитные частицы, переносили 50 мкл раствора в новые пробирки.

4) Качественный контроль и нормализация библиотек. Качество полученных библиотек (наличие димеров адаптеров, размер библиотек) проверяли с использованием высокоточного капиллярного электрофореза на приборе Bioanalyzer 2100 (Agilent, США). Пример такого анализа приведен на рисунке 2.1. Качественный анализ показал, что в приготовленных нами библиотеках содержание димеров адаптеров незначительно или полностью отсутствуют. Средний размер библиотек был в пределах 400-1000 п.н. Количественный анализ библиотек приводили с помощью флуориметра QuantiFluor ST и количественного ПЦР в режиме реального времени с помощью прибора Rotor-Gene 6000 (Corbett Research, Австралия). Проводили нормализацию всех библиотек до концентрации 8 пМ.

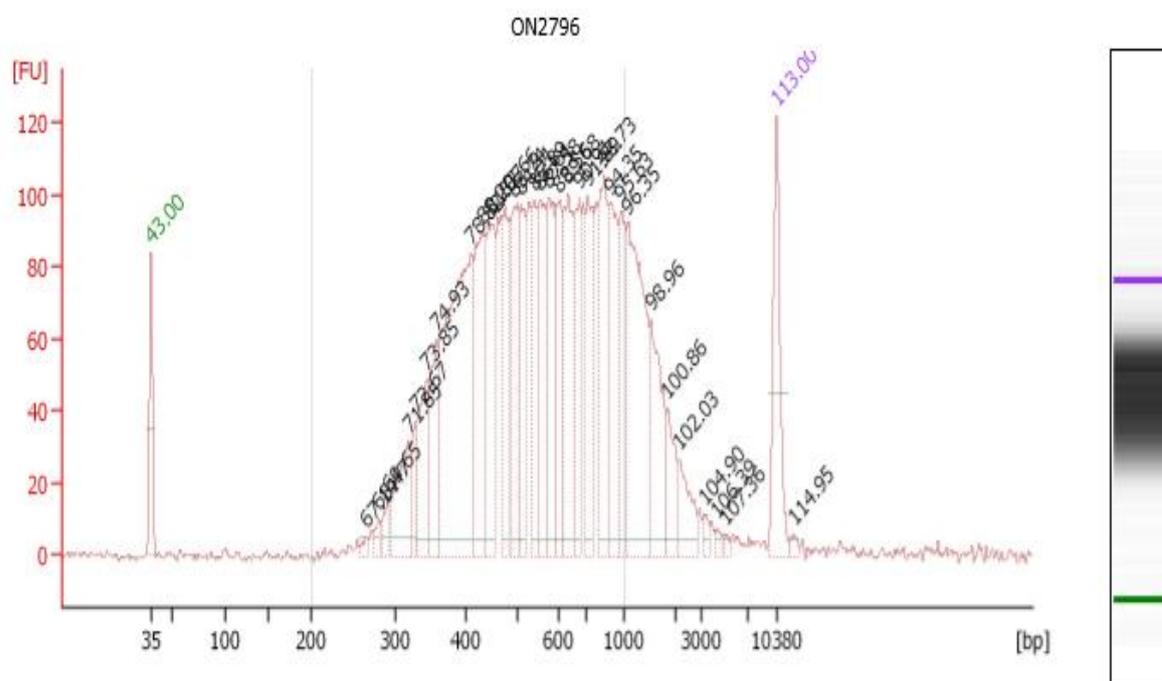


Рисунок 2.1. Результат качественного анализа библиотеки HA89(PET2), полученного на приборе Bioanalyzer 2100. По оси абсцисс размер фрагментов ДНК в п.н. По оси ординат относительные единицы флуоресценции, отображающие количественное содержание ДНК.

2.3.6. Высокопроизводительное секвенирование

Приготовленные библиотеки секвенировали на трех платформах фирмы Illumina (США): MiSeq, NextSeq 500 и HiSeq 2000 с использованием разных наборов для секвенирования – MiSeq Reagent Kit v2 500-cycles, High Output v2 kit и TruSeq SBS Kit v3-HS, соответственно.

Для каждой из библиотек были получены специфичные по числу и размеру парные короткие чтения (риды).

- 1) Линия 3629 - 5806412 ридов размером 100 п.н.
- 2) Линия 398941 - 4058566 ридов размером 100 п.н.
- 3) Фертильная линия НА89 - 13240056 ридов размером 150 п.н.
- 4) ЦМС линия НА89(РЕТ1) - 14758068 ридов размером 150 п.н.
- 5) ЦМС линия НА89(МАХ1) - 10174498 ридов размером 150 п.н.
- 6) ЦМС линия НА89(РЕТ2) - 4471774 и 4931318 ридов размером 125 и 250 п.н.
- 7) ЦМС линия НА89(АНН2) - 3063836 и 3305268 ридов размером 125 и 250 п.н.

2.3.7. Анализ данных высокопроизводительного секвенирования. Сборка и аннотация геномов

Качество полученных ридов оценивали с помощью программы FastQC (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Используя программу Trimmomatic (Bolger et al., 2014) проводили тримминг ридов, содержащих адаптерную последовательность или имеющих качество прочтения (Q-score) ниже 25. Выравнивание коротких чтений на референсные геномы проводили с использованием программы Bowtie 2 (Langmead, Salzberg, 2012). В качестве референсных нуклеотидных последовательностей использовали геномы органелл подсолнечника из базы данных NCBI – пластидный геном линии НА383 (NC_007977.1) и митохондриальный геном линии НА412 (NC_023337.1). Проводили выравнивание только парных, конкордантных, неперекрывающихся

ридов (опции --no-mixed, --no-discordant, --no-overlap). Поиск вариантных сайтов осуществляли с помощью набора программ samtools/bcftools (Li, 2011) с последующей валидацией полученных данных с помощью программы визуализации геномных данных IGV (Thorvaldsdottir et al., 2013). Сборку геномов *de novo* проводили с использованием геномного ассемблера SPAdes v 3.10.1 (Nurk et al., 2013). Для сборки контигов использовали разные значения размера k-мера (75, 85, 95, 109). Объединение контигов в скафолды проводили путем их выравнивания относительно друг друга, используя программу BLAST (<https://blast.ncbi.nlm.nih.gov/BlastAlign.cgi>). Аннотирование и поиск потенциальных рамок считывания осуществляли с помощью программ ORFfinder (<https://www.ncbi.nlm.nih.gov/orffinder/>), BLAST и MITOFY (Alverson et al., 2010). Для графического отображения геномных карт использовали программу OGDRAW (Lohse et al., 2007).

Верификацию геномных сборок проводили как биоинформатическим методом с помощью программ QUASt (Gurevich et al., 2013) и CONTIGuator (Galardini et al., 2011), так и методом ПЦР-анализа с последующим секвенированием по Сэнгеру. Также использовали картирование чтений на собранный геном с помощью программы Bowtie 2.

2.3.8. Выделение РНК

РНК выделяли из 3-4 пары листьев растений. В каждом образце исследовали по 10 растений. Выделение тотальной РНК проводили с использованием коммерческого набора Extract RNA (Евроген, Россия). Процедура выделения РНК включала в себя следующие этапы.

1. Навеску ткани массой 50 мг гомогенизировали в жидком азоте с помощью ступки и пестика.
2. Затем в каждую ступку добавляли по 1 мл лизирующего реагента и инкубировали 15 мин при комнатной температуре. Лизат переносили в пробирки объемом 1,5 мл и центрифугировали при 6°C в центрифуге с охлаждением 5417R

(Eppendorf, Германия) при 12 000 g в течение 10 минут. Супернатант объемом 500 мкл переносили в чистые пробирки.

3. В каждую из пробирок добавляли по 200 мкл хлороформа и интенсивно перемешивали в течение 15 секунд. Затем пробирки инкубировали 5 минут при комнатной температуре, периодически встряхивая образцы. Образцы центрифугировали при 12 000 g в течение 15 минут.

4. Верхнюю водную фазу аккуратно, не захватывая нижний слой хлороформа, переносили в новые пробирки объемом 1,5 мл, в которые предварительно вносили по 1 мл изопропанола и перемешивали 15 секунд. Затем пробирки центрифугировали при 12 000 g в течение 8 минут. Удаляли супернатант.

5. К осадку добавляли 1 мл 75 % этанола, встряхивали пробирки, центрифугировали при 12 000 g в течение 8 минут. Удаляли супернатант. Осадок подсушивали в пробирке с открытой крышкой 10 минут при 25 °С.

6. Осадок растворяли в 50 мкл воды свободной от РНКаз, перемешивая пипетированием.

2.3.9. Анализ транскрипционной активности митохондриальных генов

Для анализа транскрипционной активности генов выделенную РНК, сначала обрабатывали ДНКазой I (Thermo Fisher Scientific, США). Для реакции использовали 8 мкл РНК, 1 мкл ДНКазного буфера, 1 мкл ДНКазы. Смесь инкубировали 30 минут при 37 °С. Для инактивации фермента добавляли 1 мкл 50 мМ ЭДТА и затем инкубировали 10 минут при 65 °С. Затем проводили реакции обратной транскрипции с помощью набора реактивов MMLV RT kit (Евроген, Россия). Для этого в каждую пробирку добавляли 5 мкл обработанной ДНКазой РНК, 4 мкл буфера, 2 мкл dNTP (10 мМ), 2 мкл дитиотреитол (20 мМ), 2 мкл случайных декануклеотидных праймеров и 1 мкл MMLV ревертазы. Стерильной водой свободной от РНКаз доводили конечный объем реакционной смеси до 25 мкл. В качестве отрицательного контроля обратной транскрипции для каждого

образца использовали ту же реакционную смесь, однако, с добавлением воды вместо фермента (MMLV). Смесь инкубировали 60 минут при 39 °С и затем останавливали реакцию прогреванием смеси при 75°С в течение 10 минут. Полученную в ходе реакции обратной транскрипции кДНК использовали для дальнейшего ПЦР анализа.

ПЦР в режиме реального времени проводили с использованием разработанных нами праймеров (табл. 2.3) и набора реактивов с интеркалирующим красителем EVA Green (Синтол, Россия) на приборе Rotor-Gene 6000 (Corbett Research, Австралия) по следующей программе: начальная денатурация при 95 °С в течение 3 мин., затем 35 циклов, включающих денатурацию при 95 °С – 15 сек., отжиг при 60°С – 25 сек. и элонгацию при 72 °С – 30 сек.

Таблица 2.4. Праймеры, используемые для анализа транскрипционной активности генов.

Ген	Последовательность прямого праймера (5'-3')	Последовательность обратного праймера (3'-5')	Размер ПЦР продукта
<i>atp1</i>	CCCATGGCACAGCCAGAATA	CAGAAACGCTCAACTGTGGC	140
<i>atp6</i>	AGAACTGTAAGTACAAACGCA	CCTGAGTCCGAGTCTGCATC	106
<i>atp9</i>	CATTGGGGCAAACGATGCAA	TGAAAAAGAAAAGCGTGAGGAGA	233
<i>nad6</i>	TGGGAGGTCGGGGTATTCAA	CAAGCCTGGACCCGCTATAC	94
<i>orf285</i>	TCCCATCATGACCTACCCGT	GAGTTTCCGGGTCCACCAAC	101
<i>orf228</i> +285	CATTGGGGCAAACGATGCAA	GAGTTTCCGGGTCCACCAAC	343
<i>orf306</i>	AAGAAAGGCACCTCTGGACG	TCCGGGGGAAAGAAATCCAT	146
<i>orf324</i>	GAACCACCTTTCGCTAACCT	CATCGACTTGAGAATTTGCCTC	101
<i>orf327</i>	GGGGATCAGTACAAGACGGC	GCAAGCATTAAGCCGTCGAA	114
<i>orf480</i>	GGGCGATGACCCGGATAAG	GTGCAAACCCTCACGCAAG	136
<i>orf558</i>	CCCGAATCGCTCGGTAAGAG	TTTCACTGCTACGCCAGCTT	134
<i>orf645</i>	GCCTTCCACCTCTCGTTTGA	TCCGAAAGCCGGCCTAAAAT	162
<i>orf873</i>	GGGGAACATCGAGAGCTTGC	AATAGCAGGCGAGAACGAGG	146

<i>orf891</i>	CCTGGACCCGCTATACGATG	TGGAGAGTTTCCGTGTAGAAGA	147
<i>orf933</i>	AGCACTTTGGGCATTCCGT	ACAAGGGTTGGACCATCAGG	126
<i>orf1197</i>	AGTGGGCTAGGTACCACCAT	AAGCTTGACTCTATGCGCCA	140
<i>orf1287</i>	CGGATTCTGGGGGTGTCTTG	TGTTACCCCGCAAGTATCG	140
<i>orf2565</i>	TCAATCCATGTGTTCTCGCT	CGGAAAGAACAGGTCTCGGT	147

Эффективность ПЦР определяли путем построения калибровочной кривой с различными начальными концентрациями кДНК. Все ПЦР реакции проводили в 3-х повторностях.

ГЛАВА 3. Результаты и обсуждение

3.1. Изменчивость генов и некодирующих локусов хлоропластной и митохондриальной ДНК у однолетних и многолетних видов подсолнечника

Для определения уровня изменчивости хлДНК и мтДНК у однолетних и многолетних видов подсолнечника мы секвенировали участки трех хлоропластных – *atpB*, *rbcL*, *matK* (рис. 3.1) и трех митохондриальных – *atp1*, *nad5*, *matR* генов (рис. 3.2) у образцов 5 однолетних (*H. annuus*, *H. argophyllus*, *H. debilis*, *H. petiolatis*, *H. praecox*) и 16 многолетних (*H. californicus*, *H. ciliaris*, *H. decapetalus*, *H. divaricatus*, *H. giganteus*, *H. grosseserratus*, *H. hirsutus*, *H. laevigatus*, *H. maximiliani*, *H. mollis*, *H. occidentalis*, *H. rigidus*, *H. simulans*, *H. smithii*, *H. strumosus*, *H. tuberosus*) видов. Выбор данных цитогенов был сделан не случайно. Согласно многочисленным данным литературы (Qiu et al., 2006; Mower et al., 2007; Матвеева и др., 2011; Patwardhan et al., 2014) эти гены - *atpB*, *rbcL*, *matK*, *atp1*, *nad5* и *matR* представляют собой высокоинформативные мишени в хлДНК и мтДНК для проведения молекулярного штрих-кодирования и филогенетического анализа цветковых растений.

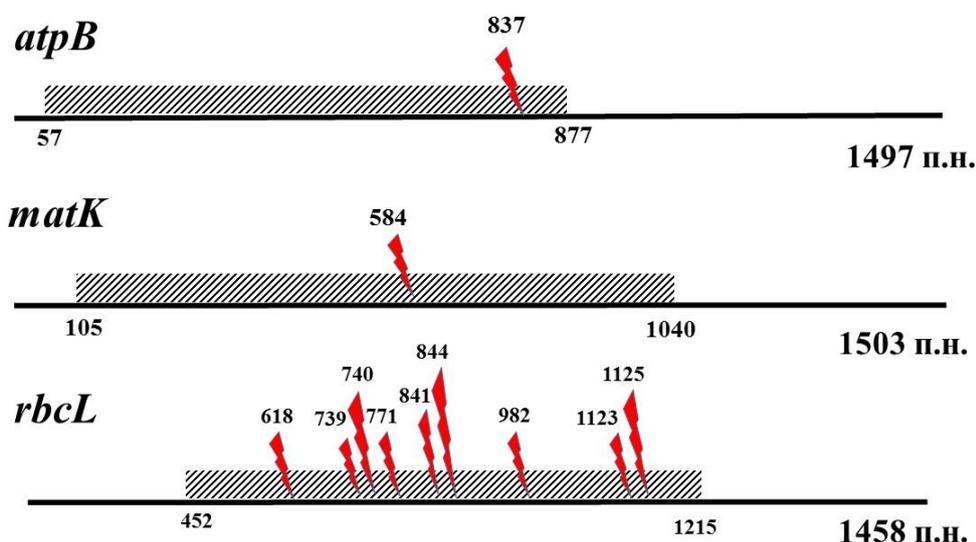


Рисунок 3.1. Схематическое изображение секвенированных участков (заштрихованная область) хлоропластных генов *atpB*, *matK*, *rbcL* и позиции локализованных в них точковых мутаций.

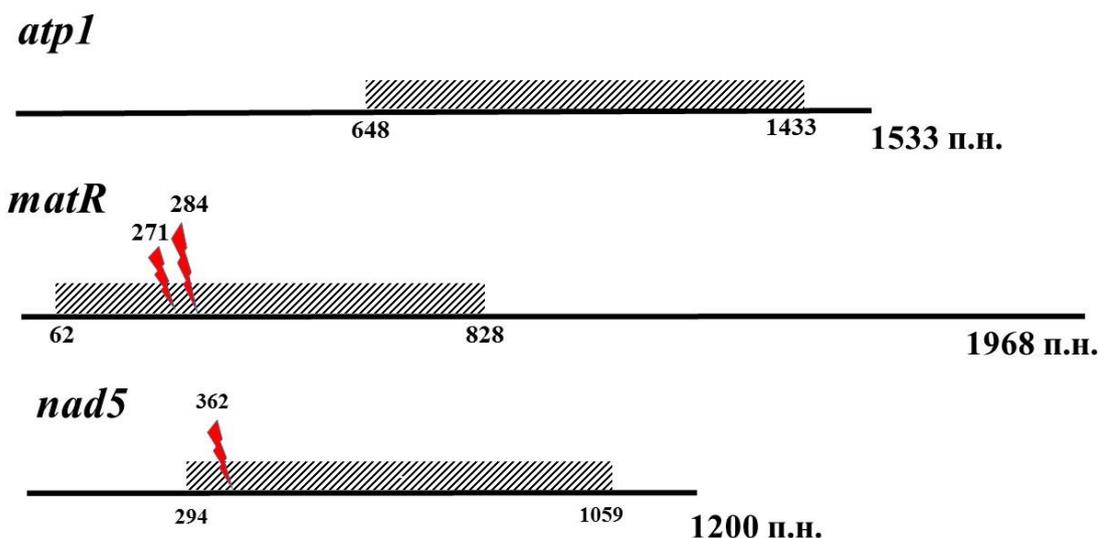


Рисунок 3.2. Схематическое изображение секвенированных участков (заштрихованная область) митохондриальных генов *atp1*, *matR*, *nad5* и позиции локализованных в них точковых мутаций.

В результате анализа полученных нуклеотидных последовательностей хлоропластных генов у представителей 5 однолетних и 16 многолетних видов подсолнечника было выявлено, что ген *atpB* у представителей рода *Helianthus* обладает крайне низкой изменчивостью. За исключением *H. strumosus*, у которого был локализован только один SNP в 837 позиции гена *atpB*, приводящий к замене цитозина на аденин и не изменяющий аминокислотную последовательность CF1 β -субъединицы АТФ-синтазы у других исследованных образцов данный участок гена был идентичен. Более информативной мишенью в хлДНК оказался ген *matK*. Обнаруженный SNP в 584 позиции четко кластеризовал исследуемые образцы на однолетние и многолетние виды. Данная мутация привела к замене тимина на цитозин у представителей многолетних видов подсолнечника, тем самым изменив 195 аминокислоту матуразы К с фениланина на серин. Наиболее информативным участком в хлДНК оказался ген *rbcL* (табл. 3.1).

Таблица 3.1 Полиморфные сайты в гене *rbcL* у образцов 5 однолетних и 16 многолетних видов рода *Helianthus L.*

Вид		Позиция в гене <i>rbcL</i>								
		618	739	740	771	841	844	982	1123	1125
Однолетние виды	<i>H. annuus</i>	G	T	G	T	T	C	T	C	G
	<i>H. argophyllus</i>	G	T	G	T	T	C	T	C	G
	<i>H. debilis</i>	G	T	G	T	T	C	T	C	G
	<i>H. petiolatis</i>	G	T	G	T	T	C	T	C	G
	<i>H. praecox</i>	G	T	G	T	T	C	T	C	G
Многолетние виды	<i>H. californicus</i>	G	A	A	C	G	A	G	A	A
	<i>H. ciliaris</i>	A	A	A	C	G	A	G	A	A
	<i>H. decapetalus</i>	G	A	A	C	G	A	G	A	A
	<i>H. divaricatus</i>	G	A	A	C	G	A	G	A	A
	<i>H. giganteus</i>	G	A	A	C	G	A	G	A	A
	<i>H. grosseserratus</i>	G	A	A	C	G	A	G	A	A
	<i>H. hirsutus</i>	G	A	A	C	G	A	G	A	A
	<i>H. laevigatus</i>	G	A	A	C	G	A	G	A	A
	<i>H. maximiliani</i>	G	A	A	C	G	A	G	A	A
	<i>H. mollis</i>	G	A	A	C	G	A	G	A	A
	<i>H. occidentalis</i>	G	A	A	C	G	A	G	A	A
	<i>H. rigidus</i>	G	A	A	C	G	A	G	A	A
	<i>H. simulans</i>	G	A	A	C	G	A	G	A	A
	<i>H. smithii</i>	G	A	A	C	G	A	G	A	A
	<i>H. strumosus</i>	G	A	A	C	G	A	G	A	A
<i>H. tuberosus</i>	G	A	A	C	G	A	G	A	A	

В секвенированной последовательности гена *rbcL* у исследованных образцов было локализовано 9 полиморфных сайтов. Один SNP в 618 позиции гена *rbcL* оказался уникальным для *H. ciliaris*. Данная мутация не приводит к

изменению транслируемой последовательности. Остальные 8 SNP четко кластеризуют исследуемые образцы на однолетние и многолетние виды (табл. 3.1). Среди них только 1 замена тимина на цитозин в 771 позиции гена *rbcL* оказалась синонимичной, а остальные 7 приводят к изменению аминокислотной последовательности большой субъединицы рибулозобисфосфаткарбоксилазы-оксигеназы (РБФК/О). В частности, SNP в 739 и 740 позициях гена *rbcL* приводят к замене 247-ой аминокислоты цистеина на аспаргин, SNP в 841 позиции *rbcL* – серина на аланин в 281 положении, SNP в 844 позиции *rbcL* – глутамина на лизин в 282 положении, SNP в 982 позиции *rbcL* – серина на аланин в 328 положении, SNP в 1123 и 1125 позициях *rbcL* - к замене 375-ой аминокислоты, а именно лейцина на изолейцин. Суммарно все 9 локализованных SNP в гене *rbcL* приводят к изменению только 5 аминокислот в транслируемом белке РБФК/О.

Таким образом, нами локализованы две специфичные мутации, одна в гене *atpB* у *H. strumosus* и одна в гене *rbcL* у *H. ciliaris*, а мутации в генах *matK* и *rbcL* кластеризующие представителей *Helianthus L.* на однолетние и многолетние. Следует добавить, что идентифицированные точечные мутации могут быть использованы в качестве молекулярных маркеров для баркодинга.

Анализ нуклеотидных последовательностей митохондриальных генов - *atp1*, *nad5* и *matR* у образцов 5 однолетних и 16 многолетних видов подсолнечника показал, что уровень их изменчивости значительно уступает хлоропластным генам. Так, например, нуклеотидный состав гена *atp1* был идентичным у всех исследуемых видов. Ген *nad5* оказался также консервативным. В нем обнаружена одна трансверсия гуанина на тимин в 362 позиции только у двух многолетних видов подсолнечника - *H. californicus* и *H. maximiliani*. Данный SNP приводит к замене аргинина на изолейцин в 121 положении субъединицы НАДН-дегидрогеназы. В гене *matR* были обнаружены два SNP, а именно замена гуанина на аденин в 271 позиции у *H. strumosus* и замена гуанин на цитозин в 284 позиции у *H. ciliaris*. При этом обе эти мутации в гене *matR* являются несинонимичными: в случае *H. strumosus* происходит замена глицин на аспаргиновую кислоту в 91

положении митохондриальной матуразы, а у *H. ciliaris* - глицина на аланин в 95 положении того же полипептида. Таким образом, всего в исследуемых участках митохондриальных генов *atp1*, *nad5* и *matR* у 5 однолетних и 16 многолетних видов подсолнечника обнаружены всего 3 точковые мутации.

В связи с тем, что митохондриальные гены продемонстрировали высокий уровень консерватизма у представителей рода *Helianthus L.*, мы провели анализ некодирующих последовательностей (SSR локусов) митохондриальной ДНК. Была разработана панель маркеров, позволяющая оценить уровень полиморфизма митохондриальной ДНК подсолнечника (раздел 2.3.4). Всего было использовано 10 молекулярных маркеров для идентификации аллельных вариантов 10 SSR локусов, соответственно. Информативность этих SSR маркеров зависела от числа и частоты встречаемости аллельных вариантов. Так в локусе MT5 было выявлено 2 аллеля, в локусах MT13, MT14, MT203 - 3 аллеля, в локусах MT1, MT11, MT15, MT22, MT23 - 4 аллеля и в локусе MT4 - 8 аллелей. Также для каждого SSR локуса был рассчитан индекс PIC - показатель, отражающий величину информационного полиморфизма. Значения PIC, а также число аллельных вариантов для каждого исследуемого SSR локуса представлены в таблице 3.2.

Таблица 3.2. Аллельные различия митохондриальных SSR локусов у подсолнечника.

SSR локус	Аллельные варианты, п.н.	PIC
MT11	125, 126, 127, 128	0,6
MT4	151, 153, 155, 156, 157, 159, 160, 161	0,8
MT1	185, 189, 190, 191	0,7
MT5	222, 223	0,2
MT13	264, 265, 266	0,6
MT14	290, 291, 292	0,4
MT22	308, 310, 311, 312	0,5
MT23	330, 331, 332, 333	0,6

MT20	364, 365, 368	0,6
MT15	401, 402, 403, 404	0,6

Значения PIC варьировали в зависимости от конкретного маркера в пределах 0,2–0,8. Наиболее информативным оказался маркер MT4, PIC которого составил 0,8. Маркер MT1 имел значение PIC равное 0,7. Для маркеров MT11, MT13, MT23, MT20, MT15 PIC составил 0,6, для MT22 - 0,5 и для MT14 - 0,4. Наименее информативным оказался маркер MT5, PIC которого составил только 0,2. Как видно из результатов расчета PIC, большая часть маркеров имеет значения выше 0,5, что свидетельствует о высокой степени их информационного полиморфизма и эффективности для SSR-анализа мтДНК различных видов подсолнечника.

На основе данных полиморфизма митохондриальных SSR локусов были определены митотипы для каждого вида подсолнечника и построена кладограмма (рис. 3.3), наглядно отражающая различия между ними.

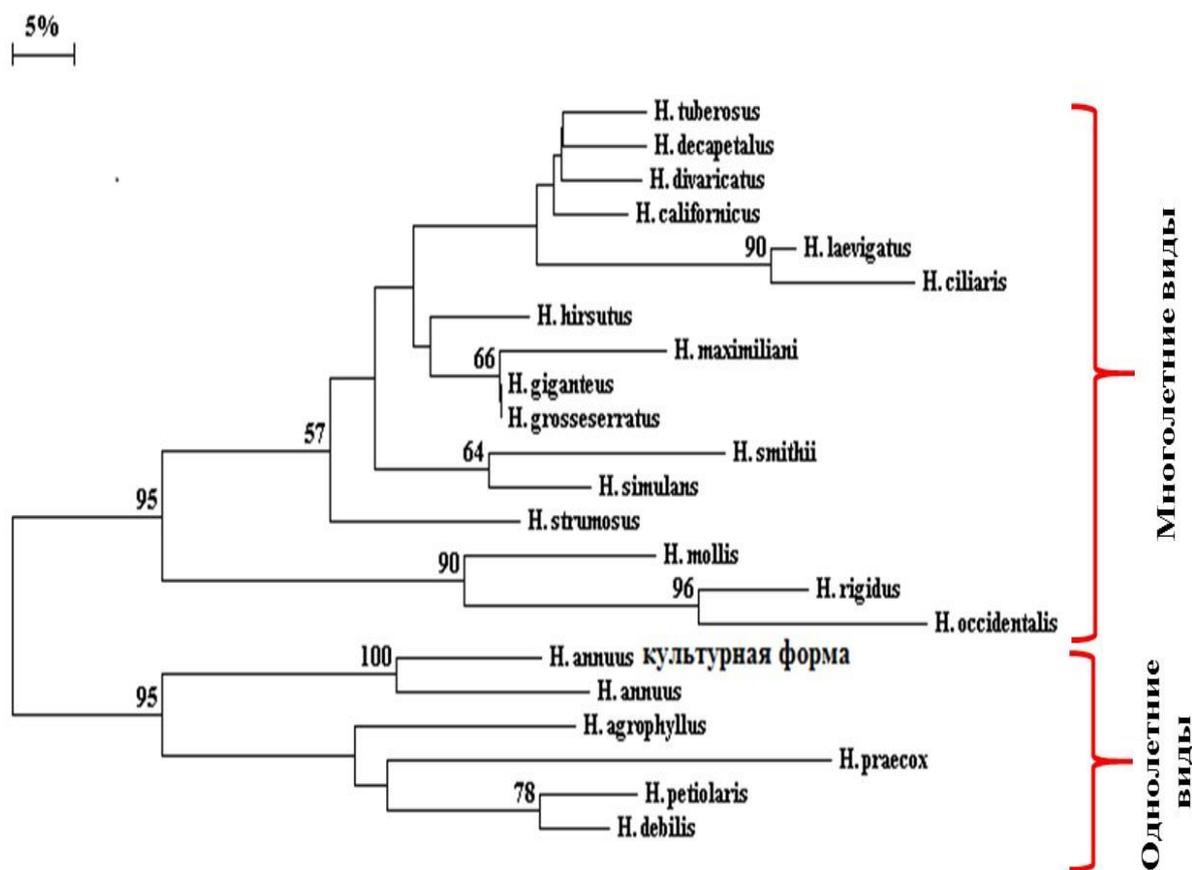


Рисунок 3.3. Кладограмма образцов однолетних и многолетних видов подсолнечника построенная на основе SSR-анализа мтДНК. Цифрами обозначены значения бутстреп-анализа; значения бутстрепа указаны для узлов с поддержкой более 50 %.

Из представленной кладограммы видно, что почти каждый образец подсолнечника имеет уникальный митотип. Исключением являются виды *H. giganteus* и *H. grosseserratus*, у которых идентичны аллели всех 10 исследуемых SSR локусов. Также отчетливо видны различия между однолетними и многолетними видами подсолнечника. Интересно, что митотип *H. annuus* значительно отличается от митотипов других однолетних видов. Многолетние виды можно кластеризировать на три группы. Первая включает 3 вида: *H. mollis*, *H. rigidus* и *H. occidentalis*, вторая – 2 вида: *H. ciliaris* и *H. laevigatus* и третья - 10

остальных видов: *H. californicus*, *H. decapetalus*, *H. divaricatus*, *H. giganteus*, *H. grosseserratus*, *H. hirsutus*, *H. maximiliani*, *H. simulans*, *H. smithii*, *H. strumosus*, *H. tuberosus*. Важно отметить, что такая кластеризация на основе митохондриальных митотипов хорошо согласуется с данными филогенетического анализа подсолнечника проведенного, как на основе нуклеотидных последовательностей ядерных генов 18S-26S рРНК (Timme et al., 2007), так и на анализе данных высокопроизводительного секвенирования 172 генов ядерного и хлоропластного происхождения (Stephens et al., 2015).

Таким образом, уровень изменчивости некодирующих последовательностей митохондриальной ДНК (SSR локусов) у подсолнечника выше, чем уровень изменчивости митохондриальных генов, и позволяет достоверно кластеризовать однолетние и многолетние виды подсолнечника.

В целом уровень изменчивости митохондриальных генов у представителей рода *Helianthus* значительно ниже, чем хлоропластных. Гены, кодирующие субъединицы АТФ-синтазы оказались наиболее консервативными как в хлоропластной (*atpB* – 1 SNP), так и митохондриальной ДНК (*atp1* нет изменчивости). В хлоропластных и митохондриальных генах выявлены видоспецифичные SNP, а именно: у вида *H. strumosus* в генах *rbcL* (хлДНК) и *matR* (мтДНК), у вида *H. ciliaris* в генах *atpB* (хлДНК) и *matR* (мтДНК). SNP в хлоропластных генах *matK* и *rbcL* позволяют различать однолетние и многолетние виды подсолнечника. В случае митохондриальных генов, таких мутаций не обнаружено, однако кластеризация видов возможна на основе данных о некодирующих последовательностях (SSR локусов) митохондриальной ДНК.

3.2. Сравнительный анализ хлоропластного и митохондриального геномов у культурного и дикорастущего подсолнечника (*Helianthus annuus* L.)

С целью определения спектра и уровня внутривидовой изменчивости внеядерных геномов *Helianthus annuus* L. мы провели полногеномное секвенирование хлоропластной и митохондриальной ДНК у линий НА89 и 3629 культурного и линии 398941 дикорастущего подсолнечника.

В качестве примера на рисунке 3.4 представлен результат сборки и аннотации хлоропластного генома линии НА89. Для повышения информативности результатов в сравнительный анализ хлоропластных геномов мы включили дополнительно данные из базы NCBI полных нуклеотидных последовательностей хлДНК еще двух линий культурного подсолнечника - НА383 (GenBank ID NC_007977.1) и SF193 (GenBank ID CM007907.1).

В результате сравнительного анализа хлоропластной ДНК 4-х линий культурного подсолнечника НА89, 3629, НА383, SF193 и линии 398941 дикорастущего подсолнечника всего было выявлено 62 вариантных сайта. Из них 29 были представлены мононуклеотидными повторами в микросателлитных локусах, т.е. в не кодирующих областях хлДНК (табл. 3.3). Наиболее распространённым оказался микросателлитный повтор (T)_n (51,7 %), далее следует - (A)_n (41,4 %) и наиболее редкий - (C)_n (6,9 %). Изменчивость в мононуклеотидных повторах SSR локусов (G)_n не обнаружена.

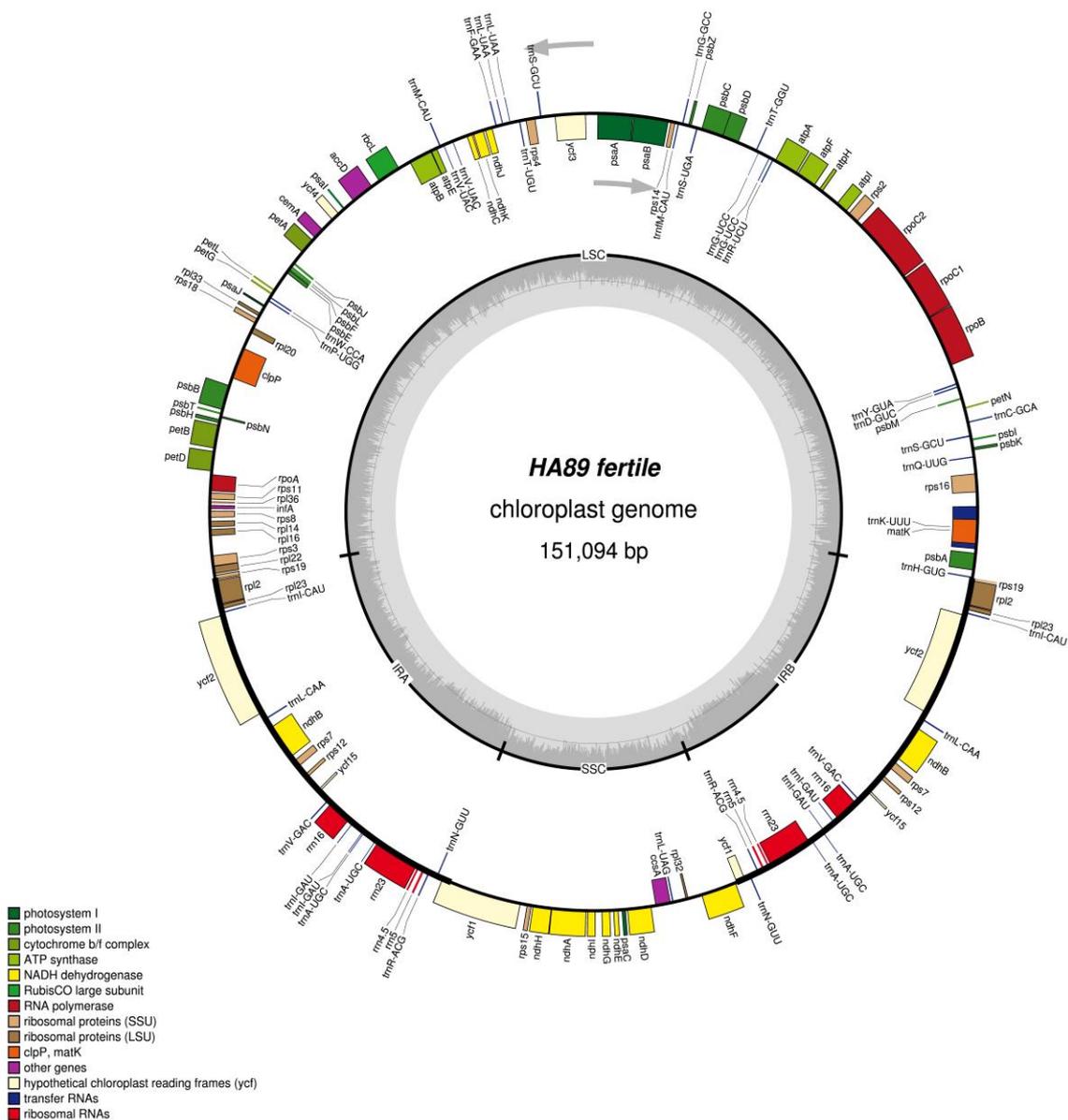


Рисунок 3.4. Генетическая карта хлоропластного генома линии HA89 подсолнечника. Внутреннее кольцо отражает содержание GC-оснований (темно-серая область) и AT-оснований (светло-серая область). LSC – большой однокопийный район, SSC – малый однокопийный район, IR – инвертированные повторы. На внешнем кольце жирными линиями обозначены инвертированные повторы IRa и IRb. Стрелками обозначены направления транскрипции генов.

В структуре исследуемых хлоропластных геномов также было локализовано 23 SNP и 10 INDEL (делеций и инсерций). Большая часть (60,9 %) точечных мутаций представлена транзициями, а именно: 8 замен A/G (34,8 %) и 6 замен C/T (26,1 %). На долю трансверсий приходится 9 мутаций (39,1 %), в том числе A/C – 5 (21,7 %), G/T – 2 (8,7 %) и A/T – 2 (8,7 %). Трансверсии типа C/G не обнаружены.

Среди 23 выявленных SNP 14 были локализованы в не кодирующих областях хлДНК (межгенные регионы и интроны генов), а оставшиеся 9 имели внутри генную локализацию. При этом 6 SNP представлены несинонимичными мутациями, а 3 SNP приводят к замене аминокислот в белках, кодируемыми генами *rpoC2*, *rps2*, *ndhF* (табл. 3.3).

Таблица 3.3. Полиморфные сайты хлоропластной ДНК, локализованные у 4-х линий HA383, SF191 HA89, 3629 культурной формы и линии 398941 дикорастущей формы подсолнечника.

Позиции референсного генома HA383	Тип полиморфизма	Линии					Локализация
		HA383	SF191	HA89	3629	ДФ* 398941	
206	SSR	(A) ₁₁	(A) ₁₁	(A) ₁₁	(A) ₁₁	(A) ₁₂	МГР <i>rpl2-psbA</i>
1991	SSR	(T) ₉	(T) ₉	(T) ₉	(T) ₉	(T) ₁₀	МГР <i>psbA-matK</i>
2032	SSR	(T) ₁₂	(T) ₁₂	(T) ₁₂	(T) ₁₂	(T) ₁₃	МГР <i>psbA-matK</i>
5450	SSR	(C) ₁₁	(C) ₉	(C) ₉	(C) ₉	(C) ₈	<i>rps16</i> (интрон)
5653	SNP	A	A	A	A	C	<i>rps16</i> (интрон)
5692	SSR	(T) ₁₃	(T) ₁₃	(T) ₁₃	(T) ₁₃	(T) ₁₁	<i>rps16</i> (интрон)
9883	SSR	(A) ₈	(A) ₈	(A) ₈	(A) ₈	(A) ₉	МГР <i>psbI-petN</i>
12984	SSR	(T) ₁₅	(T) ₁₅	(T) ₁₅	(T) ₁₅	(T) ₁₁	МГР <i>psbM-rpoB</i>
16887	SNP	C	C	C	C	A	<i>rpoC1</i> (интрон)
17424	INDEL	G	G	G	G	-	<i>rpoC1</i> (интрон)
20660	SNP	T	T	T	T	C	<i>rpoC2</i> несинонимичная мутация Leu490Pro
24141	SNP	A	A	A	A	C	<i>rps2</i> несинонимичная мутация Gln178His

25466	SSR	(A) ₁₀	(A) ₁₀	(A) ₁₀	(A) ₁₀	(A) ₁₃	МГР <i>atpI-atpH</i>
28373	SSR	(T) ₁₆	(T) ₁₅	(T) ₁₅	(T) ₁₅	(T) ₁₆	МГР <i>atpF-atpA</i>
29701	SNP	G	G	G	G	A	<i>atpA</i> синонимичная мутация
30166	SSR	(A) ₁₀	(A) ₁₀	(A) ₁₀	(A) ₁₀	(A) ₁₂	МГР <i>atpA-psbD</i>
35398	SSR	(A) ₁₈	(A) ₁₇	(A) ₁₈	(A) ₁₈	(A) ₁₈	МГР <i>psbC-psbZ</i>
35885	SSR	(A) ₉	(A) ₉	(A) ₉	(A) ₉	(A) ₈	МГР <i>psbZ-rps14</i>
39980	SNP	G	G	G	G	A	<i>psaA</i> синонимичная мутация
41995	SSR	(T) ₉	(T) ₉	(T) ₉	(T) ₉	(T) ₁₀	МГР <i>psaA-ycf3</i>
42006	SSR	(A) ₂₉	(A) ₂₈	(A) ₂₉	(A) ₂₉	(A) ₂₉	МГР <i>psaA-ycf3</i>
46980	SNP	A	A	A	A	G	МГР <i>rps14-ndhJ</i>
47912	SNP	G	T	T	T	T	МГР <i>rps14-ndhJ</i>
50032-41	INDEL	ATCTA CATAG	ATCT ACAT AG	ATCT ACAT AG	ATCT ACAT AG	-	МГР <i>ndhC-atpE</i>
50163	SSR	(T) ₁₀	(T) ₁₀	(T) ₁₀	(T) ₁₀	(T) ₁₂	МГР <i>ndhC-atpE</i>
50764	SSR	(T) ₁₁	(T) ₁₁	(T) ₁₁	(T) ₁₁	(T) ₈	МГР <i>ndhC-atpE</i>
51778	SSR	(T) ₁₀	(T) ₁₀	(T) ₁₀	(T) ₁₀	(T) ₉	МГР <i>ndhC-atpE</i>
54283	SNP	T	C	C	C	C	МГР <i>atpB-rbcL</i>
54286	SNP	A	A	A	A	T	МГР <i>atpB-rbcL</i>
54312-13	INDEL	CA	CGA	CGA	CGA	CGA	МГР <i>atpB-rbcL</i>
54313	SSR	(A) ₁₆	(A) ₁₅	(A) ₁₅	(A) ₁₅	(A) ₁₇	МГР <i>atpB-rbcL</i>
58654	SNP	A	A	A	A	G	МГР <i>accD-psaI</i>
60017	SSR	(C) ₆	(C) ₆	(C) ₆	(C) ₆	(C) ₇	МГР <i>ycf4-cemA</i>
62683	SSR	(T) ₂₂	(T) ₂₂	(T) ₂₂	(T) ₂₂	(T) ₁₅	МГР <i>petA-psbJ</i>
64145	SSR	(A) ₉	(A) ₉	(A) ₉	(A) ₉	(A) ₈	МГР <i>psbE-petL</i>
64668	SNP	C	C	C	C	T	МГР <i>psbE-petL</i>
65713	SSR	(T) ₁₆	(T) ₁₆	(T) ₁₆	(T) ₁₅	(T) ₁₆	МГР <i>petG-psaJ</i>
70584	SSR	(T) ₉	(T) ₉	(T) ₉	(T) ₉	(T) ₁₀	<i>clpP</i> (интрон)
73600-601	INDEL	AA	AA	AA	AA	-	МГР <i>psbT-psbN</i>
74157	SNP	T	T	T	T	A	МГР <i>psbH-petB</i>
76370	SNP	A	A	A	A	C	<i>petD</i> (интрон)

79932	SSR	(T) ₂₂	(T) ₂₂	(T) ₂₂	(T) ₂₂	(T) ₁₅	МГР <i>rps8-rpl14</i>
80456	SSR	(T) ₂₄	(T) ₂₃	(T) ₂₄	(T) ₂₄	(T) ₂₄	МГР <i>rpl14-rpl16</i>
80718	SNP	G	G	G	G	A	<i>rpl16</i> синонимичная мутация
81057	SNP	G	G	G	G	A	МГР <i>rpl16-rps3</i>
81555	SNP	A	A	A	A	G	МГР <i>rpl16-rps3</i>
102705	INDEL	G	-	-	-	-	МГР <i>rrn16-rrn23</i>
105130	INDEL	G	-	-	-	-	<i>rrn23</i>
106156	INDEL	C	-	-	-	-	МГР <i>rrn4.5-rrn5</i>
106393	SSR	(A) ₃₃	(A) ₃₁	(A) ₃₃	(A) ₃₃	(A) ₃₃	МГР <i>rrn5-ycf1</i>
108541	SNP	C	C	C	C	A	<i>ycf1</i> синонимичная мутация
108697	SNP	G	G	G	G	A	<i>ycf1</i> синонимичная мутация
115821	SSR	(A) ₂₇	(A) ₂₈	(A) ₂₇	(A) ₂₇	(A) ₂₇	<i>ndhA</i> (интрон)
117833	SNP	C	C	C	C	T	<i>ndhG</i> синонимичная мутация
118438	SNP	G	G	G	G	T	МГР <i>ndhG-ndhE</i>
120979	SSR	(A) ₁₆	(A) ₁₆	(A) ₁₆	(A) ₁₆	(A) ₁₃	МГР <i>ndhD-ccsA</i>
122970	SSR	(T) ₁₃	(T) ₁₃	(T) ₁₃	(T) ₁₃	(T) ₁₂	МГР <i>ccsA-rpl32</i>
123004	SNP	T	T	T	T	C	МГР <i>ccsA-rpl32</i>
125593	SNP	T	T	T	C	C	<i>ndhF</i> несинонимичная мутация Leu475Ser
128472	INDEL	G	-	-	-	-	МГР <i>rrn5-rrn4.5</i>
129498	INDEL	C	-	-	-	-	<i>rrn23</i>
131923	INDEL	C	-	-	-	-	МГР <i>rrn23-rrn16</i>

*ДФ – дикий тип.

**МГР – межгенный регион.

Четыре анализируемые линии культурного подсолнечника различаются между собой по 19 сайтам в хлДНК, а именно 9 SSR, 3 SNP и 7 INDEL. При этом наибольшая изменчивость хлДНК (3 SSR, 2 SNP, 7 INDEL) характерна для линии НА383. В свою очередь хлДНК линий НА89 и SF191 различаются по 5 SSR локусам, а линии 3629 и НА89 только по 2 сайтам – 1 SSR и 1 SNP.

Большая часть полиморфных сайтов хлДНК была обнаружена при сравнении культурной и дикорастущей форм подсолнечника. Так, например, хлоропластный геном линии 398941 дикорастущей формы отличается от референсного генома линии НА383 культурной формы 23 SSR, 21 SNP и 10 INDEL, то есть суммарно по 54 из 62 выявленных вариантных сайтов.

Особый интерес для оценки изменчивости ДНК представляют точечные мутации и их локализация в геноме (Yang et al., 2010; Doorduyn et al., 2011). Около половины SNP (9 из 21), обнаруженных в хлДНК дикорастущего подсолнечника, локализованы в кодирующих областях генома. В том числе 3 мутации приводят к изменению аминокислотного состава хлоропластных белков, а именно: замена лейцина на пролин в 490 положении β " субъединицы ДНК зависимой РНК полимеразы (ген *rpoC2*), замена глутамина на гистидин в 178 положении рибосомального белка S2 (ген *rps2*) и замена лейцина на серин в 475 положении F субъединицы НАДН-дегидрогеназы (ген *ndhF*). Известно, что кодирующие области хлДНК подсолнечника составляют около 77,6 т.п.н., что примерно составляет половину хлоропластного генома. У линии 398941 дикорастущей формы 9 SNP локализованы в генах, что в среднем составляет 0.11 SNP на 1 т.п.н кодирующей части хлоропластного генома. В случае 4-х линий культурной формы этот показатель существенно ниже - 0.013 мутаций на 1 т.п.н.

Следующим этапом нашего исследования является оценка внутривидовой изменчивости митохондриального генома *Helianthus annuus* L. На рисунке 3.5 представлен результат сборки и аннотации митохондриального генома линии НА89. Для повышения информативности результатов в сравнительный анализ митохондриальных геномов мы включили дополнительно данные из базы NCBI полных нуклеотидных последовательностей мтДНК еще двух линий культурного подсолнечника - НА412 (GenBank ID NC_023337.1) и SF193 (GenBank ID CM007908.1).

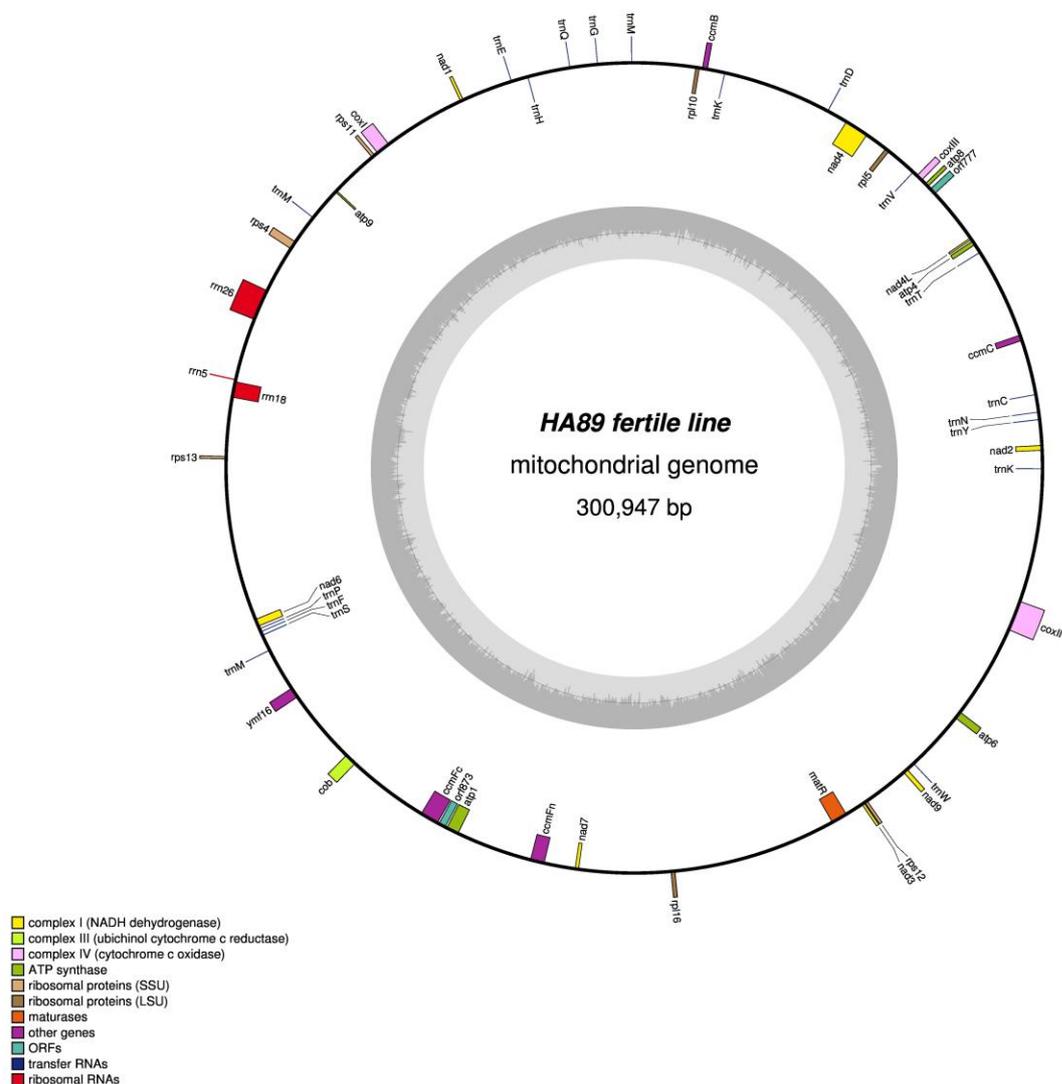


Рисунок 3.5. Генетическая карта митохондриального генома линии HA89 подсолнечника. Стрелками обозначены направления транскрипции генов мтДНК. Во внутреннем кольце серым и темно серым цветом отображено соотношение АТ и GC нуклеотидных пар.

Сравнительный анализ структуры мтДНК 4-х линий HA89, 3629, HA412, SF193 культурного и линии 398941 дикорастущего подсолнечника позволил нам локализовать 24 вариантных сайта, в том числе 11 SSR локусов представленные как в случае и с хлДНК, исключительно мононуклеотидными повторами. Однако

нуклеотидный состав полиморфных SSR локусов мтДНК существенно отличается от SSR локусов хлДНК; а именно: (T)_n – 36,4 %, (A)_n – 27,3 %, (G)_n – 27,3 %, (C)_n – 9 %. Очевидно, эти различия связаны с особенностями нуклеотидного состава анализируемых геномов, т.к. А-Т нуклеотидные пары в хлДНК составляют 62,4 %, а в мтДНК – только 54,9 %. Среди полиморфных сайтов мтДНК также было выявлено 9 SNP и 4 INDEL (табл. 3.4).

Таблица 3.4. Полиморфные сайты митохондриальной ДНК, локализованные у 4-х линий HA412, SF191, HA89, 3629 культурной формы и линии 398941 дикорастущей формы подсолнечника.

Позиции референсного генома HA412	Тип полиморфизма	Линии					Локализация
		HA412	SF191	HA89	3629	ДФ 398941	
6206	SSR	(A) ₈	(A) ₈	(A) ₈	(A) ₇	(A) ₇	МГР <i>nad2-ccmC</i>
7404	SSR	(G) ₁₀	(G) ₉	(G) ₁₀	(G) ₁₀	(G) ₁₀	МГР <i>nad2-ccmC</i>
35690-91	INDEL	CT	CTT	CTT	CTT	CTT	<i>orf777</i>
36360	SNP	T	T	T	T	G	МГР <i>orf777-atp8</i>
46039	INDEL	A	A	A	A	-	МГР <i>rpl5-nad4</i>
49272	SSR	(C) ₁₁	(C) ₁₀	(C) ₁₁	(C) ₁₁	(C) ₉	МГР <i>nad4-trnD</i>
51678	SSR	(G) ₁₀	(G) ₁₀	(G) ₁₀	(G) ₉	(G) ₉	МГР <i>nad4-ccmB</i>
105474	SSR	(T) ₃₅	(T) ₄₄	(T) ₃₅	(T) ₃₅	(T) ₃₅	МГР <i>nad1-coxI</i>
116777	SNP	G	G	G	G	T	МГР <i>atp9-trnM</i>
129368-69	INDEL	GG	GTG	GTG	GTG	GTG	<i>rrn26</i>
169028	SNP	G	G	G	G	T	<i>nad6</i> несинонимичная мутация (Ser232Phe)
170184	SSR	(T) ₁₄	(T) ₁₅	(T) ₁₅	(T) ₁₃	(T) ₁₂	МГР <i>nad6-ymf16</i>
178406	SSR	(T) ₉	(T) ₉	(T) ₉	(T) ₉	(T) ₈	<i>ymf16</i> интрон
184739	SSR	(A) ₁₀	(A) ₁₀	(A) ₁₀	(A) ₁₁	(A) ₁₁	МГР <i>ymf16-cob</i>
188363	SSR	(T) ₁₁	(T) ₁₁	(T) ₁₁	(T) ₁₀	(T) ₁₀	<i>cob</i> интрон
190813	SNP	G	G	G	G	A	МГР <i>cob-ccmFc</i>
202672	SNP	T	T	T	T	C	МГР <i>orf873-atp1</i>

209335-36	INDEL	AA	AA	AA	AA	-	МГР <i>atp1-ccmFn</i>
211809	SNP	C	G	G	G	G	МГР <i>atp1-ccmFn</i>
230112	SNP	A	A	A	A	C	<i>rpl16</i> несинонимичная мутация (Lys38Thr)
248266	SSR	(A) ₁₄	(A) ₁₄	(A) ₁₄	(A) ₁₁	(A) ₁₀	МГР <i>rpl16-matR</i>
269062	SNP	G	G	G	G	C	МГР <i>trnW-atp6</i>
284451	SSR	(G) ₉	(G) ₉	(G) ₉	(G) ₉	(G) ₁₀	МГР <i>coxII</i> интрон
293545	SNP	T	G	T	T	T	МГР <i>coxII-nad2</i>

Большая часть SNP (77,8 %) представлена трансверсиями, а именно: G/T – 4 (44,5 %), C/G – 2 (22,2 %), A/C – 1 (11,1%). Трансверсии типа A/T не обнаружены. Среди транзиций были выявлены 1 замена A/G (11,1 %) и 1 замена C/T (11,1 %). Количественные соотношения выявленных трансверсий и транзиций в мтДНК исследуемых образцов значительно отличаются от таковых в хлДНК. Данные различия, как и в случае SSR локусов, предположительно связаны с особенностями AT/GC-состава органельных геномов. 7 однонуклеотидных замен были локализованы в межгенных регионах мтДНК, а 2 имели внутригенную локализацию. При этом оба SNP приводят к замене аминокислот в белках, кодируемых генами *nad6* и *rpl16* (табл. 3.4).

Все четыре анализируемые линии культурного подсолнечника различаются по 13 сайтам: 9 SSR, 2 SNP и 2 INDEL. При этом только линия HA89 не имеет характерных вариантных сайтов мтДНК, которые бы отличали ее от трех остальных линий. Митохондриальный геном линии HA412 отличается от геномов других линий одним SNP и двумя инделями, линии SF193 – двумя SSR и одним SNP, линии 3629 – пятью SSR. Интересно отметить, что SSR локус, расположенный в межгенном регионе *nad6-ymf16* оказался наиболее полиморфным, его поли(T)₁₅ аллельный вариант был выявлен у HA89 и SF193, у HA412 – (T)₁₄, у 3629 – (T)₁₃, а у дикорастущей формы – (T)₁₂.

Большая часть полиморфных сайтов мтДНК, как и в хлДНК была обнаружена при сравнении культурной и дикорастущей форм подсолнечника.

Например, митохондриальный геном линии 398941 дикорастущей формы отличается от референсного генома линии HA412 по 9 SSR локусам, 7 SNP и 4 INDEL, то есть суммарно по 20 из 24 выявленных вариантных сайтов. Важно отметить, что два SNP, обнаруженных в митохондриальном геноме дикорастущей формы подсолнечника, приводят к изменению аминокислотных последовательностей белков, а именно: замена серина на фенилаланин в 232 положении шестой субъединицы НАДН-дегидрогеназы (ген *nad6*) и замена лизина на треонин в 38 положении рибосомального белка L16 (ген *rpl16*).

Результаты сравнительного анализа полиморфных сайтов внеядерной ДНК 4-х линий HA412, SF191, HA89, 3629 культурной формы и линии 398941 дикорастущей формы подсолнечника суммированы в таблице 3.5. Необходимо принять во внимание, что митохондриальный геном подсолнечника включает почти в два раза больше нуклеотидов, чем хлоропластный. В связи с этим мы ввели относительный показатель частоты полиморфизма (табл. 3.5), рассчитанный как отношение числа выявленных полиморфных сайтов к общему числу нуклеотидов (выраженное в т.п.н.) генома. Например, частота SNP в хлДНК составила 0,152 SNP, а в случае мтДНК – 0,03 SNP. То есть общая частота точечных мутаций в хлоропластном геноме в 5 раз выше. При этом частота мутаций в кодирующей части хлоропластного генома (около 77,6 т.п.н.) составляет 0,116 SNP, что почти в два раза выше, чем этот показатель в экзонах митохондриальных генов (29,8 т.п.н.) – 0,068 SNP.

Таблица 3.5. Количество SSR, SNP и INDEL, выявленных в хлоропластном и митохондриальном геномах у культурных (КФ) и дикорастущей (ДФ) форм *H. annuus* L.

Тип полиморф изма	Хлоропластный геном						Митохондриальный геном					
	КФ		ДФ		Общее		КФ		ДФ		Общее	
	Абс	Отн	Абс	Отн	Абс	Отн	Абс	Отн	Абс	Отн	Абс	Отн
Число полиморфных сайтов в геноме												
SSR	9	0,059	23	0,152	29	0,192	9	0,030	9	0,030	11	0,037
SNP	3	0,019	21	0,139	23	0,152	2	0,007	7	0,023	9	0,030
INDEL	7	0,046	10	0,066	10	0,066	2	0,007	4	0,013	4	0,013
Σ	19	0,126	54	0,357	62	0,411	13	0,043	20	0,066	24	0,080
Число полиморфных сайтов в некодирующих областях генома												
SSR	9	0,124	23	0,318	29	0,401	9	0,033	9	0,033	11	0,041
SNP	2	0,028	12	0,166	14	0,193	2	0,007	5	0,018	7	0,026
INDEL	5	0,069	8	0,110	8	0,110	0	0	2	0,007	2	0,007
Σ	16	0,221	43	0,594	51	0,704	11	0,041	16	0,059	20	0,074
Число полиморфных сайтов в кодирующих областях генома												
SNP	1	0,013	9	0,116	9	0,116	0	0	2	0,068	2	0,068
INDEL	2	0,026	2	0,026	2	0,026	2	0,068	2	0,068	2	0,068
Σ	3	0,039	11	0,142	11	0,142	2	0,068	4	0,134	4	0,134

Таким образом, из результатов, приведенных в таблице 3.5 можно сделать вывод, что у культурного и дикорастущего подсолнечника (*H. annuus* L.) изменчивость хлоропластного генома в 5,1 раз выше, чем митохондриального и дикорастущий подсолнечник значительно отличается от селекционных линий, как по уровню изменчивости хлДНК, так и мтДНК.

3.3. Особенности структурно-функциональной организации геномов хлоропластов и митохондрий у фертильного и стерильных аналогов линии НА89 подсолнечника с различными типами ЦМС

Целью исследования, изложенного в данной главе, был сравнительный анализ структурно-функциональной организации хлоропластных и митохондриальных геномов у фертильной линии НА 89 и ее ЦМС аналогов на основе однолетних (PET1, PET2 - *H. petiolaris*, ANN2 - *H. annuus*) и многолетнего (MAX1 - *H. maximilliani*) видов рода *Helianthus* L, с единым ядерным геномом и различными цитоплазмами. Полногеномное секвенирование органельных ДНК позволила нам провести сборку *de novo* и аннотацию хлоропластных и митохондриальных геномов 4-х линий аналогов: НА89(PET1), НА89(PET2), НА89(ANN2) и НА89(MAX1). Хлоропластные и митохондриальные геномы ЦМС линий сравнивали с ДНК органелл фертильной линии НА89, особенности структурной организации которых были рассмотрены ранее (см раздел 3.2).

3.3.1. Изменчивость хлоропластной ДНК у фертильного и стерильных аналогов линии НА89 подсолнечника с различными типами ЦМС

Количество нуклеотидов, входящих в состав хлоропластных геномов у фертильного и стерильных аналогов линии НА89 подсолнечника варьировало в довольно узких пределах: от 151094 п.н. (фертильный аналог НА89) до 151147 п.н. (НА89(PET2)). Структура хлоропластных геномов сходна с таковой для большинства высших растений и включает большой однокопийный район (LSC), малый однокопийный район (SSC) и два инвертированных повтора (IR). Размер LSC области варьирует от 83527 п.н. (фертильный аналог НА89) до 83605 п.н. (НА89(ANN2), SSC региона - от 19113 п.н. (НА89(MAX1)) до 19147 п.н. (фертильный аналог НА89) и IR региона – от 24210 п.н. (фертильный аналог

НА89, НА89(РЕТ1)) до 24212 п.н. (НА89(РЕТ2), НА89(АНН2), НА89(МАХ1)) (табл. 3.6). Содержание GC-оснований в исследуемых хлДНК аллоплазматических линий, имеет незначительные вариации от 37,60 до 37,62 % (табл. 3.6), что сравнимо с хлДНК других видов *Angiosperms* (Воск, 2007).

Таблица 3.6. Первичная структура хлДНК у фертильного и стерильных аналогов линии НА89 подсолнечника с различными типами ЦМС.

Линия	Размер хлДНК, п.н.	LSC , п.н.	SSC, п.н.	IR (2x), п.н.	GC-состав, %
Фертильная линия НА89	151094	83527	19147	24210x2	37,62
НА89(РЕТ1)	151110	83545	19145	24210x2	37,61
НА89(РЕТ2)	151127	83565	19138	24212x2	37,60
НА89(АНН2)	151147	83605	19118	24212x2	37,61
НА89(МАХ1)	151138	83601	19113	24212x2	37,61

В результате сравнительного анализа хлоропластных геномов всего выявлено 447 полиморфных сайта; 57 из которых представлены микросателлитными локусами в виде мононуклеотидных повторов, локализованными в некодирующих областях хлДНК (табл. 3.7). Соотношение нуклеотидного состава полиморфных SSR локусов составило: А – 47,4 %, Т – 45,6 %, С – 5,3 %, G – 1,7 %. Наибольшее число полиморфных SSR локусов нами выявлено при сравнении хлДНК фертильной линии НА89 с ЦМС аналогом НА89(МАХ1) – 40, далее следует НА89(АНН2) - 38, НА89(РЕТ2) – 36 и, наконец, НА89(РЕТ1) – только 21.

Таблица 3.7. Полиморфные микросателлитные локусы в хлДНК у стерильных аналогов линии НА89 подсолнечника с различными типами ЦМС.

Позиция в референсном геноме	Референсный геном - НА89 (ферг.)	PET1	PET2	ANN2	MAX1	Локализация	Регион
206	(A) ₁₁	(A) ₁₂	(A) ₉	(A) ₉	(A) ₉	МГР <i>rpl2-psbA</i>	LSC
370	(A) ₇	(A) ₇	(A) ₇	(A) ₇	(A) ₉	МГР <i>rpl2-psbA</i>	LSC
1667	(A) ₆	(A) ₆	(A) ₇	(A) ₈	(A) ₉	МГР <i>psbA-matK</i>	LSC
1991	(T) ₉	(T) ₁₀	(T) ₈	(T) ₈	(T) ₁₃	МГР <i>psbA-matK</i>	LSC
2032	(T) ₁₂	(T) ₁₃	(T) ₉	(T) ₁₄	(T) ₁₀	МГР <i>psbA-matK</i>	LSC
3953	(T) ₇	(T) ₇	(T) ₇	(T) ₇	(T) ₈	МГР <i>matk-rps16</i>	LSC
4437	(A) ₇	(A) ₇	(A) ₇	(A) ₆	(A) ₇	МГР <i>matk-rps16</i>	LSC
4788	(A) ₈	(A) ₈	(A) ₈	(A) ₈	(A) ₉	МГР <i>matk-rps16</i>	LSC
5450	(C) ₉	(C) ₈	(C) ₁₀	(C) ₉	(C) ₉	<i>rps16</i> интрон	LSC
5692	(T) ₁₃	(T) ₁₁	(T) ₇	(T) ₇	(T) ₇	<i>rps16</i> интрон	LSC
5718	(C) ₇	(C) ₇	(C) ₇	(C) ₁₁	(C) ₈	<i>rps16</i> интрон	LSC
7720	(A) ₈	(A) ₈	(A) ₈	(A) ₈	(A) ₉	МГР <i>rps16-psbK</i>	LSC
7944	(T) ₈	(T) ₈	(T) ₁₁	(T) ₉	(T) ₈	МГР <i>psbK-psbI</i>	LSC
8294	(T) ₈	(T) ₈	(T) ₈	(T) ₁₀	(T) ₉	МГР <i>psbK-psbI</i>	LSC
8552	(A) ₈	(A) ₈	(A) ₇	(A) ₇	(A) ₇	МГР <i>psbI-petN</i>	LSC
8880	(T) ₆	(T) ₆	(T) ₆	(T) ₆	(T) ₇	МГР <i>psbI-petN</i>	LSC
9882	(A) ₈	(A) ₉	(A) ₉	(A) ₉	(A) ₁₀	МГР <i>psbI-petN</i>	LSC
12759	(A) ₉	(A) ₉	(A) ₉	(A) ₉	(A) ₁₀	МГР <i>psbM-rpoB</i>	LSC
12984	(T) ₁₅	(T) ₁₁	(T) ₉	(T) ₁₀	(T) ₁₄	МГР <i>psbM-rpoB</i>	LSC
16964	(A) ₈	(A) ₈	(A) ₉	(A) ₉	(A) ₈	<i>rpoC1</i> интрон	LSC
17424	(G) ₉	(G) ₈	(G) ₇	(G) ₇	(G) ₉	<i>rpoC1</i> интрон	LSC
25296	(T) ₈	(T) ₈	(T) ₉	(T) ₈	(T) ₁₀	МГР <i>atpI-atpH</i>	LSC
25466	(A) ₁₀	(A) ₁₃	(A) ₂₁	(A) ₁₉	(A) ₂₀	МГР <i>atpI-atpH</i>	LSC
30166	(A) ₁₀	(A) ₁₂	(A) ₁₀	(A) ₁₀	(A) ₁₀	МГР <i>atpA-psbD</i>	LSC
35019	(A) ₉	(A) ₉	(A) ₁₆	(A) ₁₆	(A) ₈	МГР <i>psbC-psbZ</i>	LSC
35398	(A) ₁₈	(A) ₃₀	(A) ₂₄	(A) ₃₀	(A) ₂₂	МГР <i>psbC-psbZ</i>	LSC

35885	(A) ₉	(A) ₈	(A) ₇	(A) ₇	(A) ₇	МГР <i>psbZ-rps14</i>	LSC
44324	(A) ₈	(A) ₈	(A) ₁₀	(A) ₁₄	(A) ₁₀	МГР <i>ycf3-rps4</i>	LSC
46721	(A) ₁₅	(A) ₁₅	(A) ₁₃	(A) ₂₂	(A) ₁₁	МГР <i>rps4-ndhJ</i>	LSC
50049	(T) ₈	(T) ₈	(T) ₈	(T) ₉	(T) ₈	МГР <i>ndhC-atpE</i>	LSC
50163	(T) ₁₀	(T) ₁₂	(T) ₁₀	(T) ₁₅	(T) ₁₀	МГР <i>ndhC-atpE</i>	LSC
50764	(T) ₁₁	(T) ₈	(T) ₇	(T) ₇	(T) ₇	МГР <i>ndhC-atpE</i>	LSC
51778	(T) ₁₀	(T) ₉	(T) ₂₁	(T) ₂₂	(T) ₁₂	МГР <i>ndhC-atpE</i>	LSC
54313	(A) ₁₅	(A) ₁₅	(A) ₁₇	(A) ₂₁	(A) ₁₆	МГР <i>atpB-rbcL</i>	LSC
58257	(A) ₇	(A) ₇	(A) ₇	(A) ₇	(A) ₈	МГР <i>accD-psaI</i>	LSC
60017	(C) ₆	(C) ₇	(C) ₇	(C) ₇	(C) ₇	МГР <i>ycf4-cemA</i>	LSC
60434	(T) ₆	(T) ₆	(T) ₇	(T) ₈	(T) ₉	МГР <i>ycf4-cemA</i>	LSC
62683	(T) ₂₂	(T) ₂₀	(T) ₁₅	(T) ₁₆	(T) ₂₂	МГР <i>petA-psbJ</i>	LSC
64145	(A) ₉	(A) ₈	(A) ₁₄	(A) ₉	(A) ₉	МГР <i>psbE-petL</i>	LSC
64939	(A) ₂₃	(A) ₂₃	(A) ₂₃	(A) ₁₁	(A) ₂₃	МГР <i>psbE-petL</i>	LSC
65714	(T) ₁₆	(T) ₁₆	(T) ₁₆	(T) ₂₁	(T) ₁₃	МГР <i>petG-psaJ</i>	LSC
69743	(A) ₆	(A) ₆	(A) ₆	(A) ₆	(A) ₇	<i>clpP</i> интрон	LSC
69936	(T) ₁₀	(T) ₁₀	(T) ₁₃	(T) ₁₁	(T) ₉	<i>clpP</i> интрон	LSC
70545	(A) ₁₅	(A) ₂₀	(A) ₇	(A) ₇	(A) ₁₀	<i>clpP</i> интрон	LSC
70584	(T) ₉	(T) ₁₀	(T) ₁₀	(T) ₉	(T) ₉	<i>clpP</i> интрон	LSC
73405	(T) ₇	(T) ₇	(T) ₈	(T) ₈	(T) ₉	МГР <i>psbB-psbT</i>	LSC
77234	(A) ₁₀	(A) ₁₀	(A) ₁₁	(A) ₁₀	(A) ₉	МГР <i>petD-rpoA</i>	LSC
79932	(T) ₂₂	(T) ₁₅	(T) ₁₉	(T) ₁₇	(T) ₂₂	МГР <i>rps8-rpl14</i>	LSC
80456	(T) ₂₄	(T) ₂₄	(T) ₂₄	(T) ₂₄	(T) ₂₇	МГР <i>rpl14-rpl16</i>	LSC
82083	(T) ₈	(T) ₈	(T) ₈	(T) ₈	(T) ₉	МГР <i>rpl16-rps3</i>	LSC
83644	(T) ₈	(T) ₈	(T) ₉	(T) ₉	(T) ₈	МГР <i>rps19-rpl2</i>	IR _A
101481	(T) ₉	(T) ₉	(T) ₉	(T) ₉	(T) ₁₀	МГР <i>rrn16-rrn23</i>	IR _A
119344	(T) ₇	(T) ₇	(T) ₈	(T) ₈	(T) ₇	МГР <i>psaC-ndhD</i>	SSC
120979	(A) ₁₆	(A) ₁₃	(A) ₂₀	(A) ₁₉	(A) ₁₀	МГР <i>ndhD-ccsA</i>	SSC
122970	(T) ₁₃	(T) ₁₂	(T) ₁₈	(T) ₁₉	(T) ₁₃	МГР <i>ccsA-rpl32</i>	SSC
133146	(A) ₉	(A) ₉	(A) ₉	(A) ₉	(A) ₁₀	МГР <i>rrn23-rrn16</i>	IR _B

150984	(A) ₈	(A) ₈	(A) ₉	(A) ₉	(A) ₈	МГР <i>rpl2-rps19</i>	IR _B
--------	------------------	------------------	------------------	------------------	------------------	-----------------------	-----------------

В структуре исследуемых геномов также обнаружено 315 SNP. Локализованные SNP представлены всеми шестью возможными вариантами нуклеотидных замен, включая транзиции A/G (26,9 %), C/T (29,7 %) и трансверсии A/C (14,4 %), G/T (12,5 %), A/T (10,5 %), C/G (6,0 %). В целом наибольшее число полиморфных сайтов локализовано в большом однокопийном районе (LSC), наименьшее – в инвертированных повторах (IR), что характерно и для других цветковых растений (Ni et al., 2016; Shen et al., 2017). Из 315 SNP (приложение 1), 119 локализованы в экзонах белок кодирующих генов (табл. 3.8), причём 58 мутаций являются несинонимичными.

Известно, что хлоропластные гены цветковых растений характеризуются различной частотой точечных мутаций (Parks et al., 2009; Wicke et al., 2011). В нашем исследовании наиболее высокий уровень полиморфизма характерен для генов *ycf1*, *rps2*, *atpA*. При это в гене *ycf1* выявлено наибольшее число несинонимичных мутаций. Также большое количество мутаций было выявлено в гене *rbcL*, но только в случае ЦМС линии HA89(MAX1). Низкая изменчивость была отмечена в генах, кодирующих белки реакционных центров фотосистем – *psaB*, *psbB*, *psbD*, *psbC*, при этом все локализованные в них мутации синонимичные. Единичный SNP обнаружен и в гене *ycf2*, продукт которого также влияет на сборку реакционных центров фотосистем. Полученные данные согласуются с результатами других авторов о том, что хлоропластные гены, непосредственно контролирующие белки фотосистем, обладают низкой изменчивостью у большинства цветковых растений (Allen et al., 2011; Wicke et al., 2011; Hoekstra et al., 2017).

Таблица 3.8. SNP локализованные в экзонах хлоропластных генов стерильных аналогов линии НА89 подсолнечника с различными типами ЦМС.

Позиция в референсном геноме	Референсный геном - НА89 (ферт.)	PET1	PET2	ANN2	MAX1	Ген	Тип мутации
2984	A	A	A	A	G	<i>matK</i>	несинонимичная Phe195Ser
13459	T	T	T	T	C	<i>proB</i>	синонимичная
14470	A	A	A	A	G	<i>proB</i>	синонимичная
14496	A	A	G	A	A	<i>proB</i>	несинонимичная Gln481Arg
14992	C	C	C	C	T	<i>proB</i>	синонимичная
15565	T	T	C	T	T	<i>proB</i>	синонимичная
17643	A	A	A	A	C	<i>proC1</i>	несинонимичная Met210Leu
17783	A	A	C	C	A	<i>proC1</i>	синонимичная
17814	C	C	C	C	T	<i>proC1</i>	несинонимичная Arg267Cys
18307	G	G	G	A	G	<i>proC1</i>	несинонимичная Arg431Gln
18530	G	G	T	T	T	<i>proC1</i>	синонимичная
19839	C	C	T	T	C	<i>proC2</i>	синонимичная
20660	T	C	T	T	T	<i>proC2</i>	несинонимичная Leu490Pro
20750	C	C	C	T	C	<i>proC2</i>	несинонимичная Thr520Ile
20945	T	T	T	G	T	<i>proC2</i>	несинонимичная Leu585Arg
21054	C	C	C	C	T	<i>proC2</i>	синонимичная
22083	C	C	C	C	A	<i>proC2</i>	несинонимичная Asp964Glu
22539	C	C	C	C	T	<i>proC2</i>	синонимичная
23220	C	C	T	C	C	<i>proC2</i>	синонимичная
24138	T	T	C	T	T	<i>rps2</i>	синонимичная
24141	A	C	A	A	A	<i>rps2</i>	несинонимичная Gln178His
24198	C	C	T	C	C	<i>rps2</i>	синонимичная
24758	T	T	T	T	C	<i>atpA</i>	синонимичная
28592	A	A	A	A	G	<i>atpA</i>	несинонимичная Ile61Val
28747	G	G	G	G	A	<i>atpA</i>	синонимичная

29422	A	A	G	G	G	<i>atpA</i>	синонимичная
29458	C	C	A	A	A	<i>atpA</i>	синонимичная
29521	A	A	A	A	T	<i>atpA</i>	синонимичная
29701	G	A	G	G	G	<i>atpA</i>	синонимичная
29783	C	C	T	C	C	<i>atpA</i>	синонимичная
32669	A	A	G	A	A	<i>psbD</i>	синонимичная
34038	G	G	A	A	G	<i>psbC</i>	синонимичная
34218	G	G	A	A	A	<i>psbC</i>	синонимичная
38902	A	A	A	A	G	<i>psaB</i>	синонимичная
39980	G	A	A	A	A	<i>psaA</i>	синонимичная
40412	A	A	A	A	G	<i>psaA</i>	синонимичная
40748	C	C	C	T	C	<i>psaA</i>	синонимичная
40880	A	A	A	A	G	<i>psaA</i>	синонимичная
41098	G	G	G	G	T	<i>psaA</i>	несинонимичная Leu144Ile
42375	T	T	C	C	C	<i>ycf3</i>	синонимичная
48363	A	A	G	G	A	<i>ndhJ</i>	несинонимичная Val107Ala
49305	A	A	G	G	G	<i>ndhK</i>	синонимичная
49678	A	A	A	A	T	<i>ndhC</i>	синонимичная
52166	T	T	C	T	T	<i>atpE</i>	синонимичная
52330	C	C	C	C	T	<i>atpB</i>	синонимичная
52386	C	C	T	T	C	<i>atpB</i>	несинонимичная Gly468Ser
52717	A	A	G	G	G	<i>atpB</i>	синонимичная
52734	A	A	A	A	G	<i>atpB</i>	синонимичная
53110	A	A	A	G	A	<i>atpB</i>	синонимичная
54726	T	T	C	C	T	<i>rbcL</i>	синонимичная
55291	T	T	T	T	A	<i>rbcL</i>	несинонимичная Cys247Asn
55292	G	G	G	G	A	<i>rbcL</i>	несинонимичная Cys247Asn
55323	T	T	T	T	C	<i>rbcL</i>	синонимичная
55393	T	T	T	T	G	<i>rbcL</i>	несинонимичная Ser281Ala
55396	C	C	C	C	A	<i>rbcL</i>	несинонимичная Gln282Lys

55534	T	T	T	T	G	<i>rbcL</i>	несинонимичная Ser328Ala
55675	C	C	C	C	A	<i>rbcL</i>	несинонимичная Leu375Ile
55677	G	G	G	G	A	<i>rbcL</i>	несинонимичная Leu375Ile
55837	C	C	C	C	A	<i>rbcL</i>	несинонимичная Gln429Lys
55996	G	G	G	G	T	<i>rbcL</i>	несинонимичная Asp482Tyr
57148	A	A	A	A	C	<i>accD</i>	несинонимичная Glu191Ala
57593	A	A	G	G	G	<i>accD</i>	синонимичная
59445	C	C	C	C	T	<i>ycf4</i>	синонимичная
65328	C	C	C	C	T	<i>petL</i>	несинонимичная Pro2Leu
72744	G	G	C	C	G	<i>psbB</i>	синонимичная
75058	A	A	A	A	G	<i>petB</i>	синонимичная
75067	T	T	T	T	C	<i>petB</i>	синонимичная
77299	T	T	C	C	T	<i>rpoA</i>	синонимичная
78756	T	T	T	C	T	<i>rps11</i>	несинонимичная Lys8Arg
79664	G	G	T	T	T	<i>rps8</i>	несинонимичная Pro72Thr
79820	T	T	G	G	T	<i>rps8</i>	синонимичная
80718	G	A	G	G	G	<i>rpl16</i>	синонимичная
82328	G	G	G	G	A	<i>rps3</i>	синонимичная
82899	C	C	C	C	T	<i>rpl22</i>	несинонимичная Ala128Thr
89152	A	A	A	A	G	<i>ycf2</i>	Ile1110Val
95380	A	A	G	G	A	<i>ndhB</i>	синонимичная
108400	C	C	G	G	C	<i>ycf1</i>	несинонимичная Asn271Lys
108541	C	A	A	A	A	<i>ycf1</i>	синонимичная
108697	G	A	G	G	G	<i>ycf1</i>	синонимичная
108722	A	A	A	A	C	<i>ycf1</i>	несинонимичная Ile379Leu
109194	C	C	C	T	C	<i>ycf1</i>	несинонимичная Pro536Leu
109640	G	G	A	A	G	<i>ycf1</i>	несинонимичная Asp685Asn
109766	T	T	C	C	C	<i>ycf1</i>	синонимичная
109850	C	C	C	G	C	<i>ycf1</i>	несинонимичная Arg755Gly
110520	C	C	G	G	C	<i>ycf1</i>	несинонимичная Thr978Arg

110568	C	C	A	A	A	<i>ycfI</i>	несинонимичная Thr994Lys
110913	T	T	T	T	C	<i>ycfI</i>	несинонимичная Val1109Ala
110918	C	C	A	A	A	<i>ycfI</i>	несинонимичная His1111Asn
111048	C	C	C	C	A	<i>ycfI</i>	несинонимичная Ser1154Tyr
111059	C	C	C	G	C	<i>ycfI</i>	несинонимичная Pro1158Ala
111166	G	G	G	G	A	<i>ycfI</i>	синонимичная
111605	T	T	G	G	G	<i>ycfI</i>	несинонимичная Tyr1340Asp
111659	C	C	T	T	T	<i>ycfI</i>	несинонимичная Pro1358Ser
111983	A	A	C	C	C	<i>ycfI</i>	несинонимичная Ile1466Leu
114382	T	T	C	C	T	<i>ndhH</i>	несинонимичная Val298Ala
114585	C	C	T	C	C	<i>ndhH</i>	несинонимичная Pro366Ser
114863	C	C	T	C	C	<i>ndhA</i>	несинонимичная Pro64Leu
117052	C	C	C	C	T	<i>ndhI</i>	синонимичная
117833	C	T	T	T	T	<i>ndhG</i>	синонимичная
118566	T	T	T	C	T	<i>ndhE</i>	синонимичная
118728	A	A	G	G	A	<i>ndhE</i>	синонимичная
119751	T	T	G	T	T	<i>ndhD</i>	синонимичная
119835	A	A	A	T	A	<i>ndhD</i>	синонимичная
119952	C	C	T	T	C	<i>ndhD</i>	синонимичная
120314	T	T	C	C	C	<i>ndhD</i>	несинонимичная Leu296Pro
121676	A	A	A	A	C	<i>ccsA</i>	несинонимичная Asn167Lys
121775	C	C	C	C	T	<i>ccsA</i>	синонимичная
121813	G	G	G	G	A	<i>ccsA</i>	несинонимичная His122Tyr
122087	C	C	A	A	C	<i>ccsA</i>	несинонимичная Leu30Phe
124343	G	G	A	A	G	<i>ndhF</i>	синонимичная
125027	A	A	C	C	C	<i>ndhF</i>	синонимичная
125593	T	C	C	C	C	<i>ndhF</i>	несинонимичная Leu477Ser
125888	A	A	T	A	A	<i>ndhF</i>	несинонимичная Gln573His
125944	G	G	G	G	A	<i>ndhF</i>	несинонимичная Ser592Asn
125949	A	A	T	A	A	<i>ndhF</i>	несинонимичная Asn594Tyr

126087	G	G	T	T	T	<i>ndhF</i>	несинонимичная Val640Phe
126194	C	C	C	C	A	<i>ndhF</i>	несинонимичная Phe675Leu
139255	T	T	C	C	T	<i>ndhB</i>	синонимичная
145483	T	T	T	T	C	<i>ycf2</i>	несинонимичная Ile1110Val

Согласно полученным данным наибольшее количество SNP в кодирующих регионах хлоропластного генома было идентифицировано у ЦМС линии НА89(MAX1) в количестве 70 SNP, при этом 36 из них приводят к изменению аминокислотной последовательности белков, а наименьшее число - у ЦМС линии НА89(PET1) – 9 SNP, 3 из которых являются несинонимичными. У ЦМС линий НА89(PET2) и НА89(ANN2) выявлено примерно равное количество SNP в кодирующих областях генома – 54 и 51, а также число несинонимичных мутаций – 21 и 23, соответственно.

В результате сравнительного анализ хлоропластной ДНК у стерильных аналогов линии НА89 подсолнечника с различными типами ЦМС также выявлено 75 INDEL (приложение 1). Все INDEL локализованы в некодирующих областях хлоропластного генома, за исключением единичной мононуклеотидной делеции в гене *matK* у ЦМС линии НА89(PET2). Эта делеция приводит к сдвигу рамки считывания в 460 положении полипептида матуразы К и преждевременному стоп кодону.

Условно INDEL были разделены на короткие - размером до 10 п.н., и протяженные - размером более 10 п.н. Большая часть INDEL была представлена малыми инсерциями и делециями, около 1-5 п.н. В исследованных последовательностях хлДНК за исключением линии НА89(PET1) были обнаружены протяженные INDEL, размером 14-24 п.н (табл. 3.9). Данные вставки и делеции могут быть основой для создания эффективных тест-систем, идентифицирующих данные типы ЦМС у подсолнечника.

Таблица 3.9. INDEL размером (14-24 п.н.) локализованные в хлоропластных геномах стерильных аналогов линии НА89 подсолнечника с различными типами ЦМС.

Позиция в геноме фертильной линии НА89	Тип INDEL	НА89(PET2)	НА89(ANN2)	НА89(MAX1)	Локализация
23377-78	Инсерция	СТТТСТТТТТАТС ТТТТА			<i>rpoC2-rps2</i>
32066-67	Инсерция		TACATAAAAATT AGAATAGAAAG		<i>atpA-psbD</i>
49986-50004	Делеция			GTTCCTTTT GCGGGGCGT	<i>ndhC-atpE</i>
61429-30	Инсерция			TCCAGTAAA TAACAGAAT CGTGGA	<i>cemA-petA</i>
73364-65	Инсерция	AGGGTACCAG AGAA			<i>psbB-psbT</i>
78307-08	Инсерция		TCTCTTGTТТСТ СТТGC		<i>rpoA-rps11</i>
126449-50	Инсерция	ТТТААТТGAGT СТТТААТТСТ			<i>ndhF-rrn5</i>

Сравнительный анализ хлоропластных геномов ЦМС линий и фертильного аналога показал, что наибольшее число INDEL, также, как SSR и SNP, обнаружено у НА89(MAX1), а наименьшее - у НА89(PET1) (табл. 3.10).

Таблица 3.10. - Количество SSR, SNP и INDEL, идентифицированных у стерильных аналогов линии НА89 подсолнечника с различными типами ЦМС.

Линия	Тип последовательности										
	Кодирующая ДНК		Интроны			Межгены			Всего		
	SNP	Indel	SNP	Indel	SSR	SNP	Indel	SSR	SNP	Indel	SSR
НА89(PET1)	9	0	4	0	5	10	5	16	23	5	21
НА89(PET2)	54	1	13	5	7	87	33	29	154	38	36
НА89(ANN2)	52	0	10	4	6	80	32	32	142	36	38
НА89(MAX1)	70	0	11	2	5	114	38	35	196	40	40

Таким образом, структура и нуклеотидный состав хлоропластных геномов у фертильного и стерильных аналогов линии НА89 подсолнечника с различными типами ЦМС (РЕТ1, РЕТ2, АНН2 и МАХ1) в целом сходны. Однако сравнительный анализ геномов позволил нам локализовать 447 полиморфных сайта, в том числе 57 SSR, 315 SNP и 75 INDEL. Микросателлитные локусы в хлоропластной ДНК в основном представлены $(A)_n$ и $(T)_n$ повторами. Из 315 обнаруженных SNP, 119 локализованы в экзонах хлоропластных генов, 58 из которых являются несинонимичными. Только 1 из 75 INDEL был выявлен в кодирующей части хлоропластного генома у ЦМС линии НА89(РЕТ2). Среди аллоплазматических ЦМС линий наибольшее число полиморфных сайтов по сравнению с фертильным аналогом было идентифицировано у ЦМС линии НА89(МАХ1), а наименьшее – у ЦМС линии НА89(РЕТ1).

3.3.2. Изменчивость митохондриальной ДНК у фертильного и стерильного аналогов линии НА89 подсолнечника с различными типами ЦМС

В отличие от хлоропластной ДНК в митохондриальной ДНК у стерильных аналогов линии НА89 подсолнечника были обнаружены не только малые мутации, но и значительные реорганизации структуры. В связи с этим подробный анализ митохондриального генома у каждой ЦМС линии выделен в отдельный раздел.

3.3.2.1. Митохондриальный геном ЦМС линии НА89(РЕТ1)

Ранее, с помощью рестрикционного анализа у стерильной линии CMSBaso подсолнечника на основе ЦМС типа РЕТ1 были выявлены две крупные aberrации мтДНК, приводящие к стерильности растений - инверсия участка размером 12,1 т.п.н. между генами *cob* и *atp1* и инсерция размером 4,7 т.п.н. между генами *orf873* и *atp1* (Kohler et al., 1991). Нам с помощью полногеномного секвенирования удалось не только подтвердить наличие этих реорганизаций митохондриального генома, но и более точно определить размер данных перестроек. Размер инверсии составил 11852 п.н., а вставки - 4732 п.н. Следует также отметить, что в мтДНК линии НА89(РЕТ1) нами была обнаружена еще одна перестройка по сравнению с фертильным аналогом, а именно делеция участка размером 451 п.н. в регионе *cob-ccmFc* (позиция 190335-190786 п.н. в мтДНК фертильной линии НА89). То есть, в *cob-atp1* участке митохондриального генома у линии НА89(РЕТ1) возникли три крупные перестройки, представленные инверсией размером 11852 п.н., инсерцией размером 4732 п.н. и делецией размером 451 п.н. Для наглядности область перестроек обозначена красным овалом на генетической карте ЦМС линии РЕТ1 (рис. 3.6)

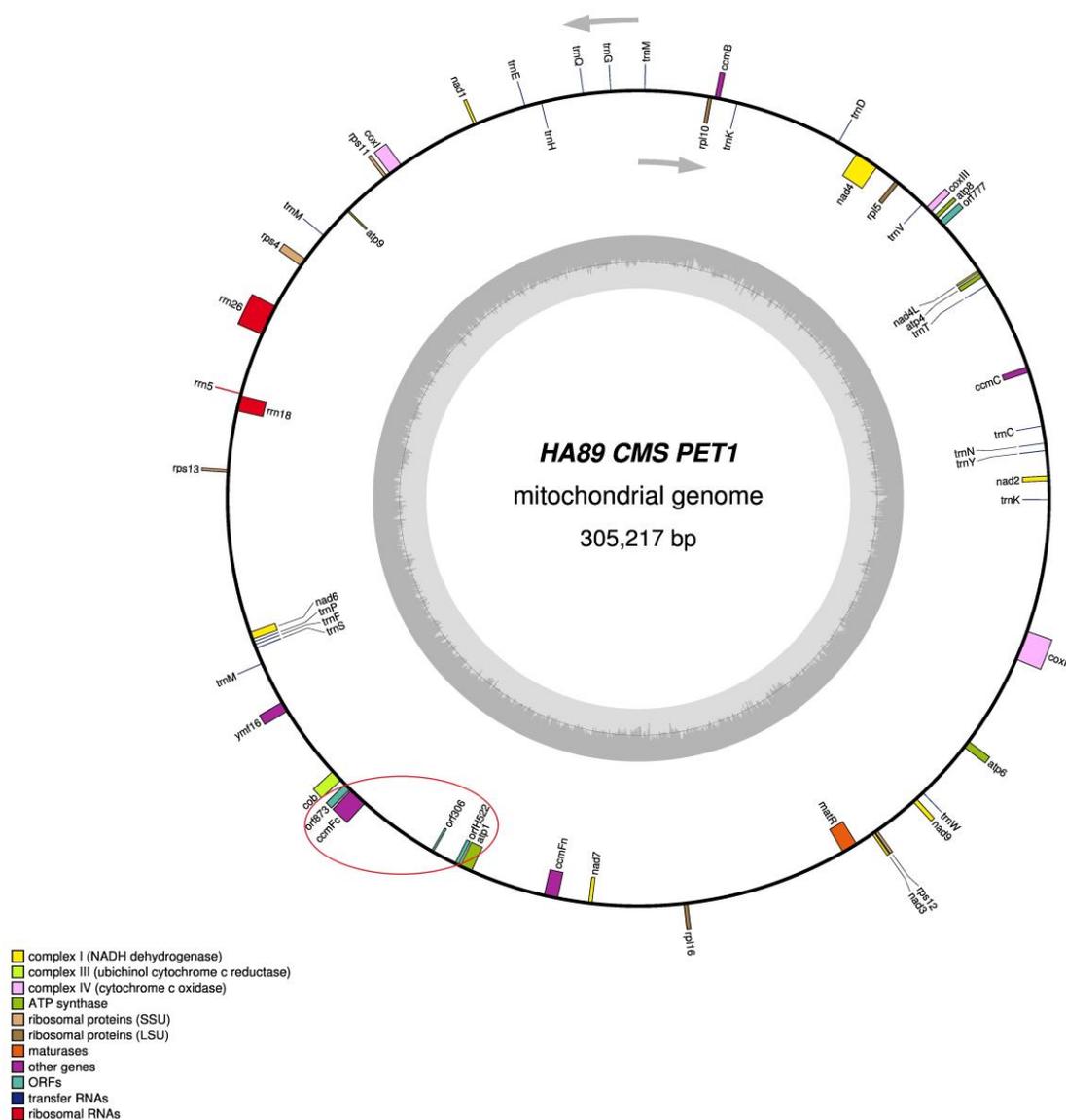


Рисунок 3.6. Генетическая карта митохондриального генома ЦМС линии HA89(PET1) подсолнечника. Стрелками обозначены направления транскрипции генов мтДНК. Во внутреннем кольце серым и темно серым цветом отображено соотношение АТ и GC нуклеотидных пар.

В результате инверсии участка мтДНК у ЦМС линии HA89(PET1) транскрипция генов *orf873* и *csmFc* протекает с противоположной цепи мтДНК, по сравнению с фертильным аналогом (рис 3.5 в разделе 3.2), а в результате инсерции возникает открытая рамка считывания *orfH522*. Авторами нескольких публикаций было показано, что именно экспрессия *orfH522* приводит к ЦМС типа

PET1 (Kohler et al., 1991; Horn et al., 1996; Horn, 2002). Следует также отметить, что в результате биоинформатического поиска новых открытых рамок считывания в нуклеотидной последовательности данной инсерции нами была обнаружена еще одна ОРС – *orf306*, кодирующая полипептид, состоящий из следующих 102 аминокислот – MTRKDRSAFLSFHQERHLWTSASTSSPSRVG IGTMDQDFWFPDFQLRSLLLLEGYGF LSPGYRFIGSATLSIKAITPLIHTHSAPLT PERKLAHLSDWIRS.

Кроме крупных перестроек в митохондриальном геноме HA89(PET1) также были выявлены и другие изменения структуры ДНК, представленные 17 вариантными сайтами: 8 SSR, 7 SNP и 2 INDEL (табл. 3.11).

Таблица 3.11. Полиморфные сайты митохондриальной ДНК ЦМС линии HA89(PET1).

Позиция в мтДНК фертильной линии HA89, п.н.	Тип	Фертильная линия HA89	HA89(PET1)	Локализация
6207	SSR	(A) ₈	(A) ₇	МГР <i>nad2-ccmC</i>
36361	SNP	T	G	<i>orf777</i> несинонимичная мутация Asp251Glu
46039	INDEL	A	-	МГР <i>rpl5-nad4</i>
49272	SSR	(C) ₁₁	(C) ₉	МГР <i>nad4-ccmB</i>
51678	SSR	(G) ₁₀	(G) ₉	МГР <i>nad4-ccmB</i>
75332	SNP	A	C	МГР <i>rpl10-nad1</i>
116777	SNP	G	T	МГР <i>atp9-rps4</i>
169028	SNP	G	T	<i>nad6</i> несинонимичная мутация Ser232Tyr
170185	SSR	(T) ₁₄	(T) ₁₂	МГР <i>nad6-ymf16</i>
178406	SSR	(T) ₉	(T) ₈	<i>ymf16</i> интрон
184739	SSR	(A) ₁₀	(A) ₁₁	МГР <i>ymf16-cob</i>
188363	SSR	(T) ₁₁	(T) ₁₀	<i>cob</i> интрон
202672	SNP	T	C	МГР <i>orf873-atp1</i>
209335-209336	INDEL	AA	-	МГР <i>atp1-ccmFn</i>
230112	SNP	A	C	<i>rpl16</i> несинонимичная мутация Lys32Gln

248266	SSR	(A) ₁₄	(A) ₁₀	МГР <i>rpl16-matR</i>
269062	SNP	G	C	<i>atp6</i> несинонимичная мутация Lys46Asn

Состав полиморфных SSR локусов был следующим: (A)_n – 37,5 %, (T)_n – 37,5 %, (C)_n – 12,5 %, (G)_n – 12,5 %. Большая часть (85,7 %) точечных мутаций представлена трансверсиями, а именно: G/T – 3 (42,8 %), A/C – 2 (28,6%), C/G – 1 (14,3 %). Трансверсии типа A/T не обнаружены. Была также выявлена 1 транзиция C/T (14,3 %). 3 из 7 обнаруженных нами SNP локализованы в межгенных регионах, а остальные 4 SNP затрагивают кодирующую часть генома, приводя к несинонимичным заменам в генах *orf777*, *nad6*, *rpl16* и *atp6*. INDEL представлены малыми делециями (1-2 п.н.), локализованные только в межгенных регионах. Важно отметить, что число SNP, SSR и INDEL в мтДНК НА89(PET1) в 2,5-3,3 раза меньше по сравнению с хлДНК (см. раздел 3.3.1).

Таким образом, в мтДНК ЦМС линии НА89(PET1) по сравнению с фертильным аналогом были выявлены следующие изменения структуры: 1 инверсия (11852 п.н.), 1 инсерция (4732 п.н.), 1 делеция (451 п.н.), 8 SSR, 7 SNP и 2 INDEL. В результате этих изменений структуры размер митохондриального генома НА89(PET1) составил 305217 п.н., что превышает на 4270 п.н. соответствующий показатель фертильной линии НА89. С функциональной точки зрения данные изменения структуры мтДНК привели к возникновению 2-х открытых рамок считывания *orf306* и *orfH522*.

Аннотированный митохондриальный геном ЦМС линии НА89(PET1) подсолнечника был размещен в банке данных NCBI, его идентификатор – MG735191.1.

3.3.2.2. Митохондриальный геном ЦМС линии НА89(PET2)

Результатом сборки и аннотации митохондриального генома ЦМС линии НА89(PET2) является кольцевая молекула ДНК размером 316582 п.н., представленная на рисунке 3.7.

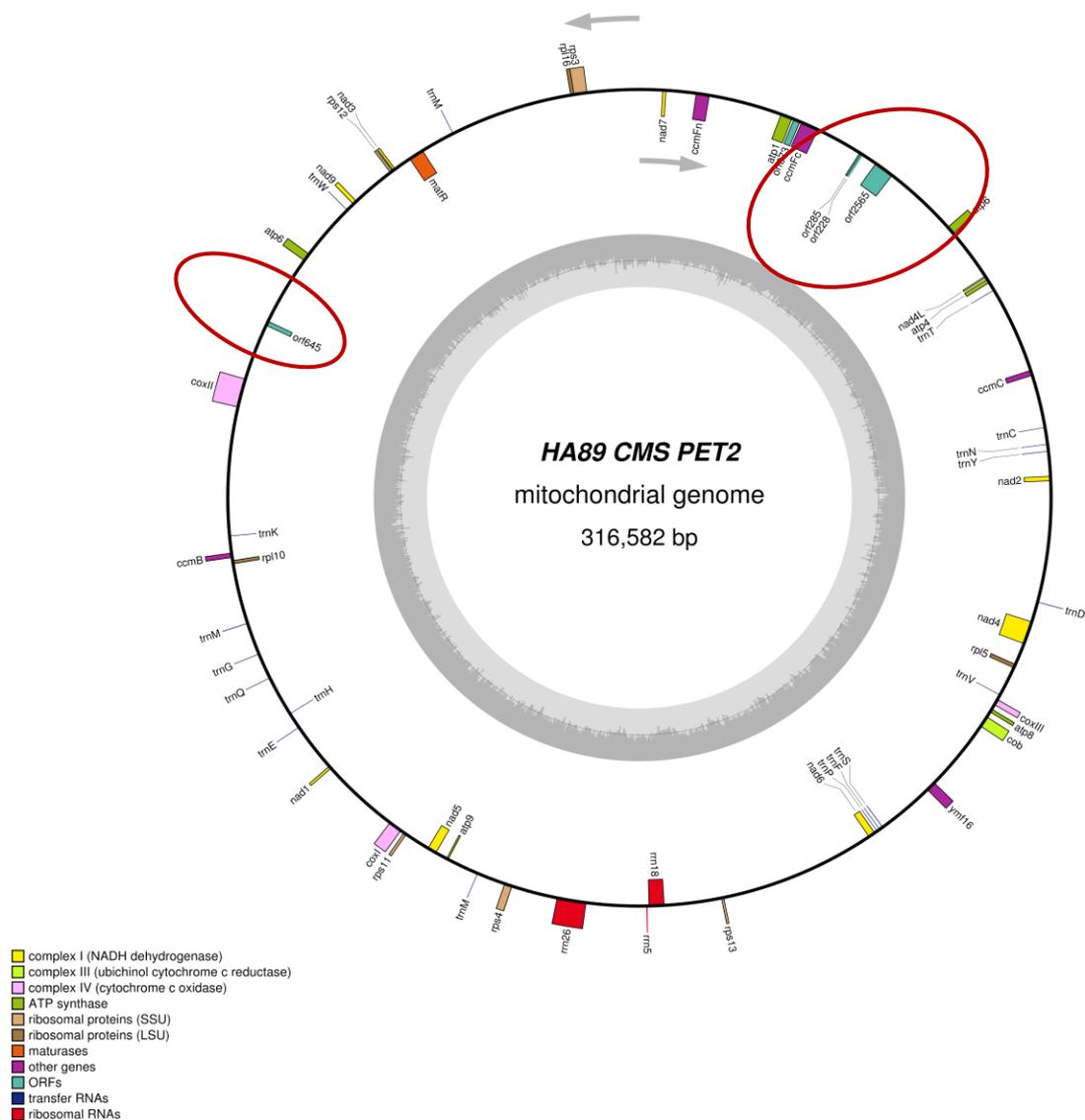


Рисунок 3.7. Генетическая карта митохондриального генома ЦМС линии HA89(PET2) подсолнечника. Стрелками обозначены направления транскрипции генов мтДНК. Во внутреннем кольце серым и темно серым цветом отображено соотношение АТ и GC нуклеотидных пар.

Митохондриальный геном ЦМС линии HA89(PET2) значительно отличается от фертильного аналога, как перестройками размером 100 и более пар нуклеотидов, так и менее протяженными (1-29 п.н.) изменениями структуры. Всего у ЦМС линии HA89(PET2) обнаружено 6 крупных реорганизаций митохондриального генома: 2 транспозиции, 2 делеции и 2 инсерции.

По сравнению с митохондриальным геномом фертильной линии НА89 в мтДНК НА89(РЕТ2) выявлены две транспозиции размером около 27,5 т.п.н. и 106,5 т.п.н. Транспозиция размером 27,5 т.п.н. соотносится с 36700–64200 нуклеотидными позициями мтДНК фертильной линии НА89 и на генетической карте фланкирована генами *atp8* и *trnD*. Транспозиция размером 106,5 т.п.н. соотносится с 194440–300945 нуклеотидными позициями мтДНК фертильной линии НА89 и на генетической карте фланкирована генами *scmFc* и *trnK*. Эти две транспозиции не затрагивают кодирующих областей мтДНК и не приводят к возникновению новых ОРС. Тем не менее, в геноме НА89(РЕТ2) можно отметить пространственное сближение генов *cob* и *atp8* (рис. 3.7), которые в митохондриальном геноме фертильной линии располагаются на значительном удалении (около 150 т.п.н.) друг от друга (рис. 3.5 в разделе 3.2).

В митохондриальном геноме НА89(РЕТ2) также идентифицированы две делеции размером 711 п.н. и 3780 п.н. В мтДНК НА89(РЕТ2) отсутствуют участки мтДНК фертильной линии, соответствующие позициям 35682–36393 (делеция 711 п.н.) и 190659–194439 (делеция 3780 п.н.). Делеция 3780 п.н. не затрагивает кодирующей части генома, а делеция 711 п.н. приводит к потере гена *orf777*. На сегодняшний день функция белка, кодируемого *orf777*, неизвестна. Более того *orf777* не имеет сходства с другими митохондриальными генами. Отсюда можно предположить, что элиминация гена *orf777* в мтДНК НА89(РЕТ2) не связана с ЦМС фенотипом.

Наиболее функционально значимые среди выявленных перестроек митохондриального генома являются две инсерции размером 5050 п.н. и 15885 п.н. На генетической карте мтДНК НА89(РЕТ2) инсерции обозначены красными овалами (рис. 3.7). Инсерция 5050 п.н. была локализована в межгенном регионе *atp6-coxII* (275230-275231 позиции мтДНК фертильной линии). Поиск потенциальных ОРС в данной инсерции выявил только одну открытую рамку считывания - *orf645*, кодирующую полипептид, состоящий из следующих 215 аминокислот – MARKGNPISVRLDLNRSSDSSRFVESTIHASIFFILLSLTFQNQLK KGGISLPVVKHGLLLIFILLSIFLMNELYNNAVVPFLAKSGDSGINPLPGPSDQNSI

FPLIEEGEVPSESRPKINRDPEFCISFQNACDLESEISEIMETLLHEKGVVDLSQREI
KREVEGFLSNVWDMЕPIPRNRRLTALKSLRSRGVNSRNFKDFEKHKIPVLL.

23 аминокислоты на N-конце этого полипептида комплементарны N-концу митохондриального рибосомального белка S3. Проведенный нами анализ транскрипционной активности *orf645* свидетельствует, что мРНК этого гена представлены только в тканях растений ЦМС линии НА89(РЕТ2) и отсутствует у фертильных растений линии НА89.

Следует отметить, что молекулярной причиной ЦМС у цветковых растений чаще всего является химерный белок, имеющий в своем составе трансмембранный домен, подобный субъединицам АТФ-синтазы или белкам ассоциированным с дыхательной цепью переноса электронов (Gillman et al., 2007; Yang et al., 2009). Исходя из этого, можно предположить, что возникновение открытой рамки считывания *orf645* в мтДНК НА89(РЕТ2) не связано в данном случае с ЦМС фенотипом.

Вторая инсерция размером 15885 п.н. локализована в некодирующей части генома между генами *nad4L* и *ccmFс*. Данная вставка условно была разделена нами на три области (рис. 3.8) Первая, включающая 9482 п.н., имеет полное сходство с другим участком митохондриального генома НА89(РЕТ2) (позициями 126260-135741). Эта область повтора не содержит кодирующих последовательностей, за исключением гена *atp6*. Следующая область размером 4849 п.н. представлена уникальной последовательностью митохондриального генома НА89(РЕТ2), в которой мы определили новую ОРС – *orf2565*. Третья область размером 1554 п.н. включает как повторяющиеся, так и уникальные последовательности. С функциональной точки зрения этот участок инсерции (1554 п.н.) представляет собой дубликацию гена *atp9* (позиции 114341-114640 мтДНК фертильной линии НА89) с последующей делецией размером 12 п.н. и инсерцией размером 271 п.н. В результате этих перестроек возникают две открытые рамки считывания – *orf228* и *orf285*. Таким образом, инсерция размером 15885 п.н. в митохондриальном геноме НА89(РЕТ2) приводит к дубликации гена *atp6* и возникновению трех открытых рамок считывания *orf2565*, *orf228* и *orf285*.

мтДНК НА89(РЕТ2)
35686-51570 п.н.

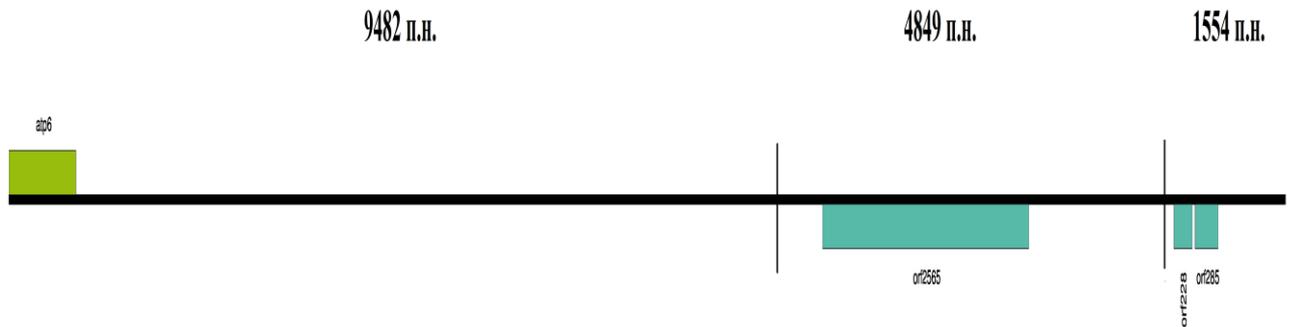


Рисунок 3.8. Схематическое изображение инсерции размером 15885 п.н., локализованной у ЦМС линии НА89(РЕТ2). Гены на схеме расположены согласно направлению их транскрипции.

Несмотря на дубликацию гена *atp6*, уровень мРНК данного гена не отличался у ЦМС линии и ее фертильного аналога. Маловероятно, что дубликация гена *atp6* может быть причиной ЦМС типа РЕТ2. Это можно предположить и в случае с *orf2565*. *Orf2565* кодирует белок, состоящий из 855 аминокислот (приложение 3). Поиск гомологичных белков в базе данных NCBI показал, что белок, кодируемый *orf2565*, относится к В-семейству ДНК полимераз и не содержит мембранно-связывающих мотивов.

Orf228 кодирует полипептид, состоящий из 76 аминокислот (приложение 3), среди которых 75 идентичны С-концу белка, кодируемого *atp9*. Другими словами, *orf228* – это часть гена *atp9*, у которого в результате инсерции размером 271 п.н. образовался новый стартовый кодон. IX субъединица АТФазы содержит на С-конце и N-конце два трансмембранных домена (рис. 3.9). Белок, кодируемый *orf228*, также содержит два потенциальных мембранно-связывающих мотива, однако по сравнению с IX субъединицей АТФазы в N-концевом мотиве у *orf228* отсутствуют 4 аминокислоты. Эти различия могут влиять на взаимодействие белка, кодируемого *orf228*, с митохондриальными мембранами. Другая ОРС - *orf285* кодирует полипептид, состоящий из 95 аминокислот (приложение 3) из

которых 18 комплементарны N-концу белка, кодируемого *atp9*. То есть *orf228* и *orf285* суммарно кодируют 93 из 99 аминокислот IX субъединицы АТФазы. В полипептиде, кодируемым *orf285*, также обнаружен трансмембранный домен (рис. 3.9).

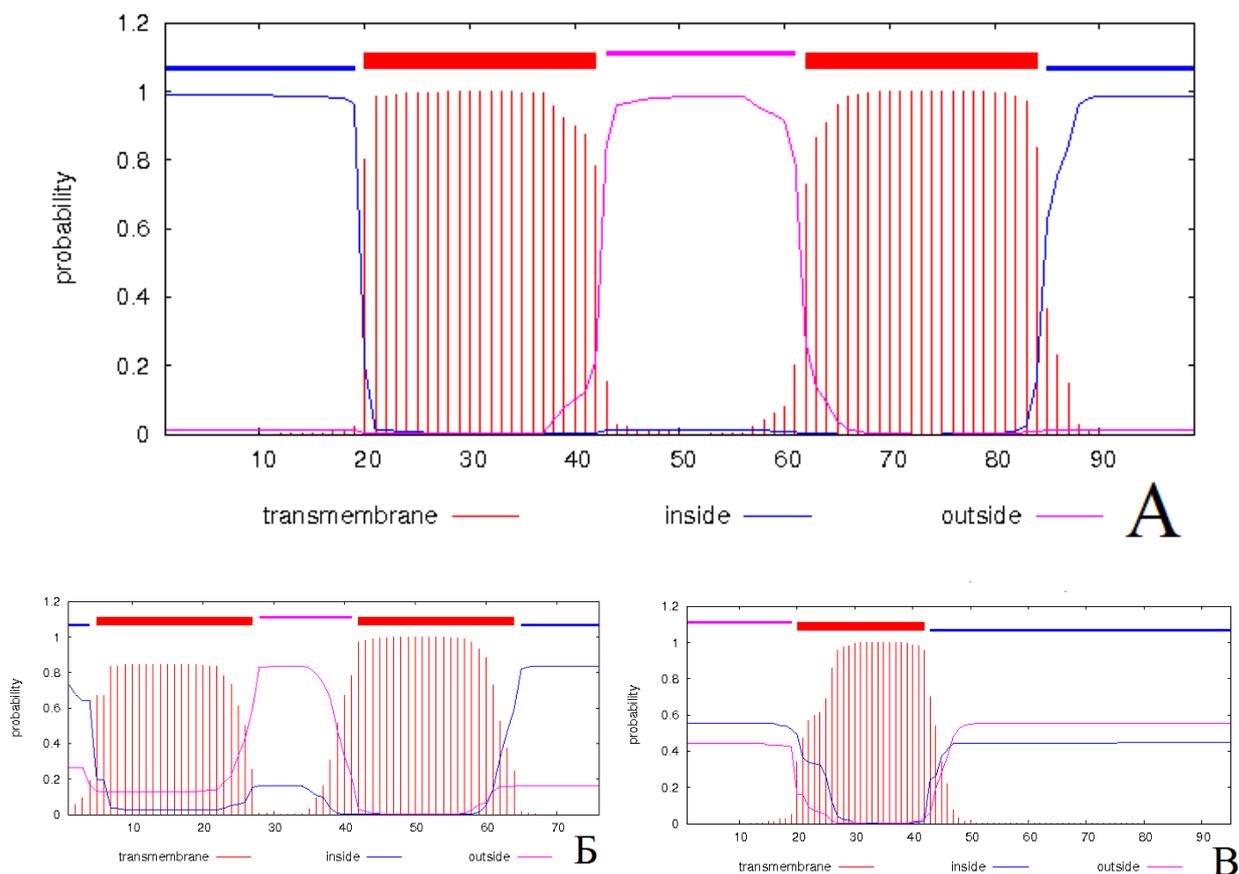


Рисунок 3.9. Расположение трансмембранных доменов в белках, кодируемых А - *atp9*; Б - *orf228*; В - *orf285*. Ось абсцисс отображает позиции аминокислот в белке.

Ось ординат - взаимодействие полипептида с внутренней (синяя линия) и внешней (сиреневая линия) митохондриальными мембранами и вероятность формирования трансмембранного домена (красные линии).

Определение уровня транскрипции *orf228* оказалось труднореализуемой задачей из-за практически полного сходства (>97 %) нуклеотидного состава *orf228* с *atp9*. Однако мы определили, уровень транскрипции *orf285*, которая была характерна только для растений ЦМС линии НА89(РЕТ2). Также мы обнаружили, что *orf228* и *orf285* транскрибируются в единой полицистронной матричной РНК.

Кроме крупных перестроек в мтДНК НА89(PET2) по сравнению с фертильным аналогом были обнаружены варианты сайты: 14 SSR, 55 SNP и 13 INDEL (табл. 3.12).

Таблица 3.12. Полиморфные сайты митохондриальной ДНК ЦМС линии НА89(PET2).

Позиция в мтДНК фертильной линии НА89, п.н.	Тип	Фертильная линия НА89	НА89(PET2)	Локализация
3031	SSR	(G) ₅	(G) ₆	МГР <i>nad2-ccmC</i>
3107	SSR	(T) ₅	(T) ₆	МГР <i>nad2-ccmC</i>
3275-3276	INDEL	TA	TTTA	МГР <i>nad2-ccmC</i>
3281-3281	INDEL	AT	ATT	МГР <i>nad2-ccmC</i>
4715	SSR	(T) ₈	(T) ₉	МГР <i>nad2-ccmC</i>
6207	SSR	(A) ₈	(A) ₇	МГР <i>nad2-ccmC</i>
6660	SNP	A	G	МГР <i>nad2-ccmC</i>
7404	SSR	(G) ₁₀	(G) ₉	МГР <i>nad2-ccmC</i>
7919	INDEL	A	-	МГР <i>nad2-ccmC</i>
9796	SNP	T	C	МГР <i>nad2-ccmC</i>
10467	SNP	A	C	МГР <i>nad2-ccmC</i>
10924	SNP	A	C	МГР <i>nad2-ccmC</i>
12314	SNP	T	C	МГР <i>nad2-ccmC</i>
19594	SNP	G	A	МГР <i>ccmC-atp4</i>
23917	SNP	G	T	МГР <i>ccmC-atp4</i>
31803	SNP	A	C	МГР <i>nad4L-orf777</i>
34099	SNP	A	C	МГР <i>nad4L-orf777</i>
34135-34136	INDEL	AT	ATGT	МГР <i>nad4L-orf777</i>
34162	SNP	T	C	МГР <i>nad4L-orf777</i>
35031	SNP	C	A	МГР <i>nad4L-orf777</i>
35114	SNP	C	A	МГР <i>nad4L-orf777</i>
35478	SNP	T	C	МГР <i>nad4L-orf777</i>

35511	SNP	G	A	МГР <i>nad4L-orf777</i>
35596	SNP	G	C	МГР <i>nad4L-orf777</i>
42296	SNP	C	A	МГР <i>coxIII-rpl5</i>
49273	SSR	(C) ₁₁	(C) ₁₀	МГР <i>nad4-ccmB</i>
50857	SNP	C	A	МГР <i>nad4-ccmB</i>
51679	SSR	(G) ₁₀	(G) ₉	МГР <i>nad4-ccmB</i>
62361	SNP	T	G	МГР <i>nad4-ccmB</i>
62404	SNP	G	A	МГР <i>nad4-ccmB</i>
63434-63435	INDEL	TC	TCC	МГР <i>nad4-ccmB</i>
71498-71499	INDEL	GT	GGGGCT	МГР <i>rpl10-nad1</i>
75333	SNP	A	C	МГР <i>rpl10-nad1</i>
91106	SNP	G	T	МГР <i>rpl10-nad1</i>
91107	SNP	A	C	МГР <i>rpl10-nad1</i>
105475	SSR	(T) ₃₅	(T) ₂₅	МГР <i>nad1-coxI</i>
108201	SNP	T	G	МГР <i>rps11-nad5</i>
115916	SNP	T	G	МГР <i>atp9-rps4</i>
119332	SNP	G	A	МГР <i>atp9-rps4</i>
121109-121110	INDEL	CC	CTTC	МГР <i>atp9-rps4</i>
122991	SNP	A	C	<i>rps4</i> синонимичная мутация
133548	SNP	T	A	МГР <i>rrn26-rrn5</i>
133549-133550	INDEL	AT	AGG	МГР <i>rrn26-rrn5</i>
133550	SNP	T	G	МГР <i>rrn26-rrn5</i>
133551	SNP	A	C	МГР <i>rrn26-rrn5</i>
156215	SNP	C	A	МГР <i>rps13-nad6</i>
156623-156624	INDEL	CC	CCTAC	МГР <i>rps13-nad6</i>
157461	SNP	T	G	МГР <i>rps13-nad6</i>
170187	SSR	(T) ₁₄	(T) ₁₂	МГР <i>nad6-ymf16</i>
174934-174935	INDEL	AC	ACTCGACTGAAA GGAAAGGTACG AAGTGGC	МГР <i>nad6-ymf16</i>
175181	SNP	G	T	МГР <i>nad6-ymf16</i>
188365	SSR	(T) ₁₁	(T) ₁₀	<i>cob</i> интрон

189982	SNP	G	T	МГР <i>cob-ccmFc</i>
195010	SNP	G	T	МГР <i>cob-ccmFc</i>
195017	SNP	C	A	МГР <i>cob-ccmFc</i>
200176	SNP	G	A	<i>ccmfc</i> интрон
200517	SNP	G	A	<i>ccmfc</i> интрон
204992	SNP	C	A	МГР <i>atp1-ccmFn</i>
204848-204849	INDEL	AA	ATA	МГР <i>atp1-ccmFn</i>
207967	SSR	(G) ₁₀	(G) ₁₂	МГР <i>atp1-ccmFn</i>
209460	SNP	G	A	МГР <i>atp1-ccmFn</i>
212640	SSR	(C) ₉	(C) ₁₂	МГР <i>atp1-ccmFn</i>
215918	SNP	C	T	МГР <i>ccmFn-rps3</i>
223919	SNP	A	C	МГР <i>ccmFn-rps3</i>
223927-223928	INDEL	GA	GAA	МГР <i>ccmFn-rps3</i>
226979-226980	INDEL	AC	ACGTTGTTTTC	МГР <i>ccmFn-rps3</i>
232828	SNP	G	T	МГР <i>rpl16-matR</i>
239882	SNP	G	A	МГР <i>rpl16-matR</i>
239990	SNP	A	C	МГР <i>rpl16-matR</i>
241037	SNP	G	A	МГР <i>rpl16-matR</i>
241477	SNP	A	C	МГР <i>rpl16-matR</i>
246055	SNP	C	T	МГР <i>rpl16-matR</i>
248268	SSR	(A) ₁₄	(A) ₉	МГР <i>rpl16-matR</i>
249349	SSR	(T) ₈	(T) ₉	МГР <i>rpl16-matR</i>
249363	SNP	C	A	МГР <i>rpl16-matR</i>
260903	SNP	G	T	МГР <i>nad9-atp6</i>
262082	SNP	G	A	МГР <i>matR-nad9</i>
269064	SNP	G	C	<i>atp6</i> несинонимичная мутация Lys46Asn
269136	SNP	A	C	<i>atp6</i> синонимичная мутация
270678	SNP	G	T	МГР <i>atp6-coxII</i>
273346	SNP	C	A	МГР <i>atp6-coxII</i>
276836	SNP	T	G	МГР <i>atp6-coxII</i>

Состав полиморфных SSR локусов был следующим: (T)_n – 42,8 %, (G)_n – 28,6 %, (C)_n – 14,3%, (A)_n – 14,3%. Большая часть (69,1 %) точечных мутаций представлена трансверсиями, а именно: A/C – 21 (38,2 %) G/T – 14 (25,5 %), C/G – 2 (3,6 %) A/T – (1,8 %). Транзиции (30,9 %) были представлены следующими заменами: A/G – 11 (20 %) C/T – 6 (10,9 %). Абсолютное большинство (94,5 %) выявленных SNP затрагивают некодирующие области митохондриального генома. Только 3 SNP были локализованы в кодирующей последовательности генов *rps3* и *atp6*, и только одна мутация привели к замене лизина на аспарагин в 46 положении VI субъединицы АТФ-синтазы. Все 13 обнаруженных INDEL локализованы в некодирующих областях мтДНК. По сравнению с хлДНК в мтДНК линии HA89(PET2) выявлено значительно меньше полиморфных сайтов: SSR в 2,7 раза, SNP и INDEL в 2,8 раз. При этом частота SNP в хлоропластных генах (0,696 SNP на 1 т.п.н.) в 7 раз выше, чем в митохондриальных (0,1 SNP на 1 т.п.н.).

Таким образом, в мтДНК ЦМС линии HA89(PET1) по сравнению с фертильным аналогом были выявлены следующие изменения структуры: 2 транспозиции (27,5 и 106,5 т.п.н.), 2 делеции (711 и 3780 п.н.), 2 инсерции (5050 и 15885 п.н.), 14 SSR, 55 SNP и 13 INDEL. В результате этих изменений структуры размер митохондриального генома HA89(PET2) составил 316582 п.н., что на 15635 п.н. превышает соответствующий показатель фертильной линии HA89. С функциональной точки зрения данные изменения структуры мтДНК привели к делеции гена *orf777*, дупликации гена *atp6* и возникновению 4-х открытых рамок считывания: *orf228*, *orf285*, *orf645* и *orf2565*. Мы предполагаем, что наиболее вероятной причиной ЦМС типа PET2 подсолнечника является возникновение в митохондриальном геноме двух открытых рамок считывания *orf228* и *orf285*, кодирующих *atp9*-химерные полипептиды.

Аннотированный митохондриальный геном ЦМС линии HA89(PET2) подсолнечника был размещен в банке данных NCBI, его идентификатор - MG770607.2.

3.3.2.3. Митохондриальный геном ЦМС линии HA89(MAX1)

Результатом сборки и аннотации митохондриального генома ЦМС линии HA89(MAX1) является кольцевая молекула ДНК размером 295586 п.н., представленная на рисунке 3.10.

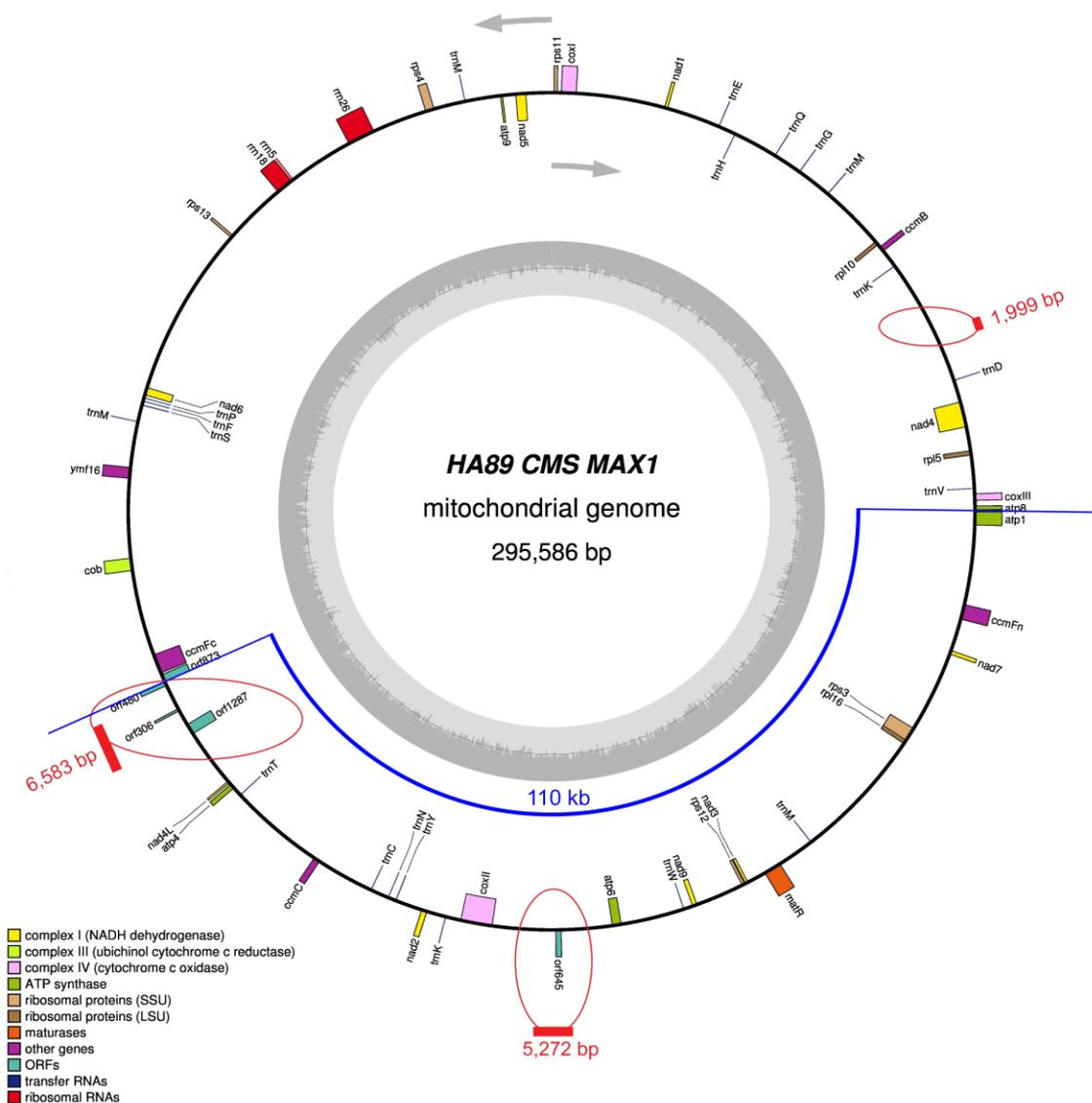


Рисунок 3.10. Генетическая карта митохондриального генома ЦМС линии HA89(MAX1) подсолнечника. Стрелками обозначены направления транскрипции генов мтДНК. Во внутреннем кольце серым и темно серым цветом в отображено соотношение АТ и GС нуклеотидных пар.

Митохондриальный геном ЦМС линии HA89(MAX1) существенно отличается от фертильного аналога, как крупными перестройками, так и

незначительными изменениями структуры. Всего было выявлено 8 крупных перестроек в мтДНК НА89(МАХ1), в том числе 1 инвертированная транспозиция, 4 делеции и 3 инсерции.

В митохондриальном геноме НА89(МАХ1) по сравнению с фертильным аналогом была обнаружена инвертированная транспозиция участка размером около 110 т.п.н., фланкированного генами *orf873* и *atp8*. Границы транспозиции обозначены синими линиями на генетической карте мтДНК НА89(МАХ1) (рис. 3.10). Интересно отметить, что транспозиция затрагивает более трети митохондриального генома, изменяя направления транскрипции целого ряда генов.

Также в митохондриальном геноме НА89(МАХ1) идентифицированы 4 делеции размером 439 п.н., 978 п.н., 3183 п.н. и 14296 п.н. (табл. 3.13). Несмотря на значительные размеры делетируемых участков данные изменения структуры мтДНК практически не затронули кодирующих областей генома, за исключением делеции размером 978 п.н., в результате которой в митохондриальном геноме НА89(МАХ1) полностью удален ген *orf777*. Данный ген также делетирован и у ЦМС линии НА89(РЕТ2).

Таблица 3.13. Делеции размером более 100 п.н., выявленные в митохондриальном геноме ЦМС линии НА89(МАХ1).

Размер делетируемого участка, п.н.	Позиции в мтДНК фертильной линии НА89, п.н.	Локализация делеции на генетической карте мтДНК фертильной линии НА89	Гены, входящие в делетируемый участок
439	56164-56602	<i>nad4-ccmB</i>	-
978	35559-36536	<i>nad4L-orf777-atp8</i>	<i>orf777</i>
3183	272049-275231	<i>nad9-coxII</i>	-
14296	286490-300785	<i>coxII-nad2</i>	-

Среди локализованных мутаций в мтДНК НА89(МАХ1) наибольший интерес представляют инсерции размером 1999 п.н., 5272 п.н. и 6583 п.н., обозначенные красными овалами на генетической карте мтДНК НА89(МАХ1) (рис. 3.10). Инсерция 1999 п.н. расположена в межгенном регионе *nad4-ccmB*, в

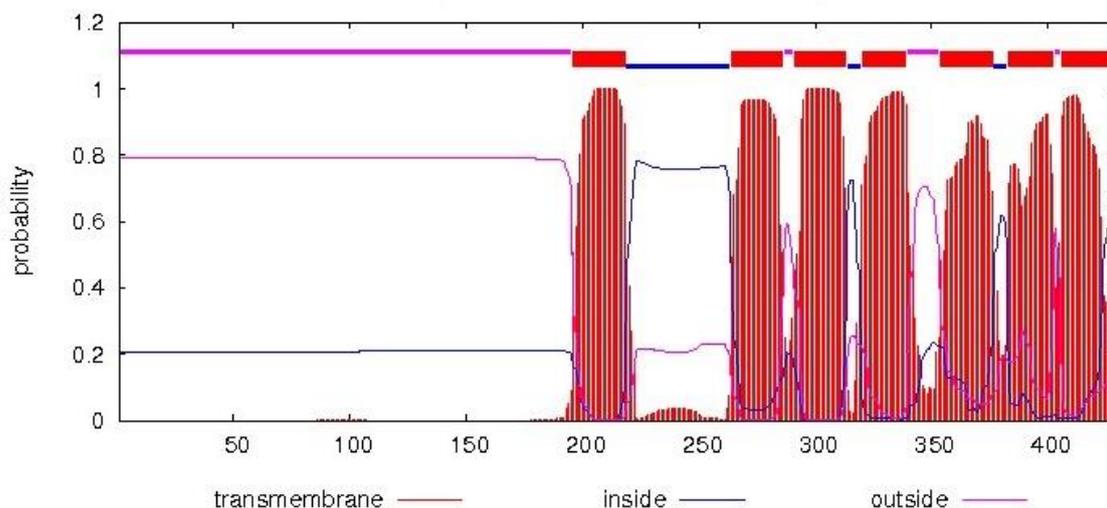
том же участке митохондриального генома, что и делеция размером 439 п.н. Эта инсерция интересна тем, что представляет собой вставку нуклеотидов идентичных области инвертированного повтора (IR) хлоропластного генома (позиции 102508-104507 или 130128-132127 хлДНК НА89(МАХ1)), включающую часть гена хлоропластной рибосомальной РНК – *rrn23*. Инсерция размером 5272 п.н. локализована в межгенном регионе *atp5-coxII*, в том же участке мтДНК, что и делеция размером 3183 п.н. Около 80 % нуклеотидов данной вставки комплементарны (99 % сходство) нуклеотидной последовательности другой инсерции размером 5050 п.н., обнаруженной в мтДНК ЦМС линии НА89(РЕТ2). Как у линии НА89(РЕТ2), так и НА89(МАХ1) в результате этих инсерций возникла активно транскрибируемая новая открытая рамка считывания *orf645*.

Наиболее протяженная инсерция (6583 п.н.) в мтДНК ЦМС линии НА89(МАХ1) расположена в межгенном регионе *orf873-nad4L* на 5' конце 110 т.п.н. инвертированной транспозиции (рис. 3.10). Данная вставка условно была поделена нами на две области. Первая область, включающая участок размером 4767 п.н. (позиции 167500-172267 в мтДНК НА89(МАХ1)), комплементарна (99 % сходство) другой вставке размером 4732 п.н., обнаруженной в мтДНК ЦМС линии НА89(РЕТ1). В данной области, были выявлены открытые рамки считывания *orf306*, *orf480*. Вторая область, состоящая из 1816 п.н. (позиции 172268-174083 в мтДНК НА89(МАХ1)), содержит, как повторяющуюся, так и уникальную последовательности. Последовательность размером 1227 п.н. идентична другому участку мтДНК НА89(МАХ1) (позиции 227962-229169), в результате чего этот участок генома оказался дублированным. Дублированный участок содержит в себе часть нуклеотидной последовательности гена *atp6* и в сочетании с уникальной последовательностью вставки образует новую рамку считывания - *orf1287*. Таким образом, в результате инсерции размером 6583 п.н. в митохондриальном геноме ЦМС линии НА89(МАХ1) возникли три новых ОРС - *orf306*, *orf480* и *orf1287*.

Открытая рамка считывания *orf306*, обнаруженная в мтДНК НА89(МАХ1), идентична *orf306* в мтДНК ЦМС линии НА89(РЕТ1) (см. раздел 3.3.2.1). *Orf480*

кодирует белок из следующих 160 аминокислот – MLFRLTARSLFLSFSFFMLVGRSVFMEQITPYKKGRSVSGPSSQKNLPLPGGSGDDPKRKKVPVSKDTANA AVSLLRQVILEILARARDPSLREGLHNPTTQAWNRAITTAIQERSGNYSISTLGAI QRTIEVAGELVFEGEQSAFFLRVLQLVRERYS. Последовательность из 128 аминокислот этого полипептида идентична последовательности 16,5 кДа белка, который кодируется *orfH522* и является причиной ЦМС типа PET1 (см. раздел 3.3.2.1). При этом важно отметить, что 16,5 кДа белок (*orfH522*) содержит мембрано-связывающий мотив, а в белке, кодируемом *orf480*, трансмембранный домен отсутствует. Отсюда можно предположить, что *orf480* не связана с возникновением стерильного фенотипа у линии НА89(MAX1) подсолнечника.

Orf1287 представляет собой *atp6*-химерный ген, состоящий из 1290 п.н., среди которых 760 п.н. комплементарны гену *atp6*, а остальные 530 п.н. уникальны. *Orf1287* транскрибируется только в тканях растений ЦМС линии НА89(MAX1) и не экспрессируется у фертильных растений линии НА89. Результатом трансляции *orf1287* является белок из 429 аминокислот, у которого 250 аминокислот на С-конце идентичны α -субъединице митохондриальной АТФ-синтазы. Эти 250 аминокислот формируют 7 трансмембранных доменов (рис. 3.11). Поиск в базе данных NCBI *orf1287* гомологичных белков выявил сходство *orf1287* с другим *atp6*-химерным белком, состоящим из 437 аминокислот (идентификатор NCBI CAA57790.1), обнаруженного у линии подсолнечника с ЦМС типа ANT1 (Spasova et al., 1994). Аминокислотные последовательности этих белков сходны приблизительно на 74 %.



MPNNSHAFLGRTKPTISDKSFLIFRARSQTTRKKIFTMTMKDFIQRFKRT
 VSPVLNESEEVPIQFGLSMDEIVEANLEKFTLYLEGDSVTPASIEALTKL
 NYLYVFMRHDLGTVKPAAIRSLQKELNKTPPESVVPKLESYQNELKSL
 DNFLEPFENVLSSKDFLNYCDEFYRSS**SPLEQFDILPLIPMNIGNLYFSF**
TNSSLFMLLTLSLVLLLIHFVTKKGGGNLVPNAWQSLVELIYDFVLNLVN
EQIGGLSGNVKQKFFPCILVTFTFLLFCNLOGMIPYSFTVTSHFLLITLGL
SFSIFIGITIVGFQRNGLHFLSFLLPAGVPLPLAPFLVLELISYCFRAL
SLGIRLFANMMAGHSLVKILSGFAWTMLCMNDLLYFIGDLGPLFIVLALT
GLELGVAILQAYVFTILICIYLNDAINLH

Рисунок 3.11. Расположение трансмембранных доменов в белке, кодируемом *orf1287*. Ось абсцисс отображает позиции аминокислот в белке. Ось ординат - взаимодействие полипептида с внутренней (синяя линия) и внешней (сиреневая линия) митохондриальными мембранами и вероятность формирования трансмембранного домена (красные линии). Жирным шрифтом выделены аминокислоты идентичные в белках, кодируемых *orf1287* и *atp6*. Красным подчеркиванием обозначены аминокислоты, формирующие трансмембранный домен.

Кроме крупных aberrаций в мтДНК ЦМС линии HA89(MAX1) были выявлены 18 SSR, 230 SNP и 29 INDEL (приложение 2). Состав полиморфных SSR локусов был следующим: (T)_n – 38,9 %, (G)_n – 33,3 %, (A)_n – 16,7%, (C)_n – 11,1% (табл. 3.14).

Таблица 3.14. Полиморфные SSR локусы митохондриальной ДНК ЦМС линии НА89(MAX1).

Позиция в мтДНК фертильной линии НА89, п.н.	Фертильная линия НА89	НА89(MAX1)	Локализация
1462	(A) ₈	(A) ₇	МГР <i>coxII-nad2</i>
3031	(G) ₅	(G) ₆	МГР <i>nad2-ccmC</i>
3107	(T) ₅	(T) ₄	МГР <i>nad2-ccmC</i>
4715	(T) ₈	(T) ₉	МГР <i>nad2-ccmC</i>
7404	(G) ₁₀	(G) ₈	МГР <i>nad2-ccmC</i>
44412	(G) ₆	(G) ₇	МГР <i>rpl5-nad5</i>
49273	(C) ₁₁	(C) ₉	МГР <i>nad4-ccmB</i>
51679	(G) ₁₀	(G) ₉	МГР <i>nad4-ccmB</i>
86188	(T) ₇	(T) ₆	МГР <i>rpl10-nad1</i>
170187	(T) ₁₄	(T) ₇	МГР <i>nad6-ymf16</i>
178408	(T) ₉	(T) ₈	<i>ymf16</i> интрон
179735	(A) ₁₀	(A) ₉	МГР <i>ymf16-cob</i>
188365	(T) ₁₁	(T) ₉	<i>cob</i> интрон
190823	(C) ₇	(C) ₁₀	МГР <i>cob-ccmFc</i>
212024	(G) ₉	(G) ₈	МГР <i>atp1-ccmFn</i>
236229	(T) ₈	(T) ₇	МГР <i>rpl16-matR</i>
248268	(A) ₁₄	(A) ₁₀	МГР <i>rpl16-matR</i>
284453	(G) ₉	(G) ₇	<i>coxII</i> интрон

Большинство (65,7 %) точечных мутаций в мтДНК ЦМС линии НА89(MAX1) представлены трансверсиями, а именно: А/С – 65 (28,3%), G/T – 60 (26,1 %), C/G – 15 (6,5 %), A/T – 11 (4,8 %). Среди транзиций (34,3 %) обнаружены следующие частоты нуклеотидных замен: A/G – 44 (19,1%), C/T – 35 (15,2 %). При этом большая часть (85,2 %) вариантных сайтов была локализована в некодирующих областях генома, тогда как 34 SNP, суммированные в таблице 3.15, локализованы в экзонах митохондриальных генов.

Таблица 3.15. SNP, выявленные в экзонах митохондриальных генов ЦМС линии НА89(MAX1).

Позиция в мтДНК фертильной линии НА89, п.н.	Фертильная линия НА89	НА89(MAX1)	Ген	Тип мутации
28717	T	C	<i>nad4L</i>	синонимичная мутация
28803	A	G	<i>nad4L</i>	несинонимичная мутация Ser50Pro
112693	T	G	<i>nad5</i>	синонимичная мутация
169209	A	G	<i>nad6</i>	синонимичная мутация
188296	G	T	<i>cob</i>	синонимичная мутация
188466	C	A	<i>cob</i>	несинонимичная мутация Asn42Lys
202282	G	C	<i>orf873</i>	несинонимичная мутация Ala129Gly
202308	C	T	<i>orf873</i>	синонимичная мутация
202311	T	C	<i>orf873</i>	синонимичная мутация
202324	G	T	<i>orf873</i>	несинонимичная мутация Ser115Glu
202325	A	C	<i>orf873</i>	несинонимичная мутация Ser115Glu
202332	C	T	<i>orf873</i>	несинонимичная мутация Glu112Gly
202333	T	C	<i>orf873</i>	несинонимичная мутация Glu112Gly
202350	T	C	<i>orf873</i>	синонимичная мутация
202353	T	C	<i>orf873</i>	синонимичная мутация
202380	C	G	<i>orf873</i>	несинонимичная мутация Trp96Cys
202441	C	G	<i>orf873</i>	несинонимичная мутация Cys76Ala
202442	A	C	<i>orf873</i>	несинонимичная мутация Cys76Ala
202445	A	G	<i>orf873</i>	синонимичная мутация
202446	C	T	<i>orf873</i>	синонимичная мутация
202455	G	T	<i>orf873</i>	синонимичная мутация
202520	T	C	<i>orf873</i>	несинонимичная мутация Ile50Val
202526	G	T	<i>orf873</i>	несинонимичная мутация Leu48Ile
202538	T	A	<i>orf873</i>	несинонимичная мутация Ile44Leu
202539	T	C	<i>orf873</i>	несинонимичная мутация Ile44Leu
202548	T	G	<i>orf873</i>	несинонимичная мутация Leu40Phe
202554	C	A	<i>orf873</i>	несинонимичная мутация Leu38Phe

202583	T	G	<i>orf873</i>	несинонимичная мутация Thr29Pro
202605	C	A	<i>orf873</i>	несинонимичная мутация Met21Ile
202616	T	G	<i>orf873</i>	несинонимичная мутация Ile18Leu
202620	C	T	<i>orf873</i>	синонимичная мутация
202628	T	C	<i>orf873</i>	несинонимичная мутация Trh14Ala
269064	G	C	<i>atp6</i>	несинонимичная мутация Lys46Asn
269155	T	C	<i>atp6</i>	несинонимичная мутация Tyr77His

Следует подчеркнуть, что 26 из 34 SNP, выявленных в кодирующих регионах мтДНК НА89(MAX1), локализованы всего лишь в одном гене - *orf873*. Учитывая консервативность митохондриальных генов, такая изменчивость в *orf873* на первый взгляд кажется парадоксальной. Однако именно в регионе мтДНК вблизи *orf873* и находится «горячая точка» рекомбинации митохондриального генома ЦМС линии НА89(MAX1), в которой мы обнаружили инвертированную транспозицию размером 110 т.н.п. и инсерцию размером 6583 п.н. Крупные перестройки мтДНК могли привести к увеличению числа точечных мутаций или к нарушению функции *orf873* с последующим накоплением точечных мутаций в данном участке генома. Анализ транскрипционной активности показал, что уровень мРНК *orf873* в тканях ЦМС линии НА89(MAX1) незначительно отличается от фертильного аналога, в связи с чем псевдогенизация *orf873* маловероятна. *Orf873* гомологи найдены в мтДНК многих других представителей цветковых растений: *Lactuca sativa* L (GeneBank ID PLY75385.1), *Beta vulgaris* L (GeneBank ID NP_064062.1), *Vicia faba* L (GeneBank ID AGC78928.1) и др., однако функция *orf873* и его гомологов в митохондриальном геноме цветковых растений до сих пор не ясна. Оставшиеся 8 из 34 SNP были выявлены в генах *nad4L*, *nad5*, *nad6*, *cob* и *atp6*. Мутации в генах *nad5*, *nad6* не вызывают изменение аминокислотной последовательности транслируемых белков, а в гене *atp6* наоборот две несинонимичные мутации, приводят к заменам 46-ой и 77-ой аминокислот в VI субъединице АТФ-синтазы. Интересно отметить, что несинонимичная мутация Lys46Asn в гене *atp6* характерна для мтДНК трех

ЦМС линий – НА89(РЕТ1), НА89(РЕТ2) и НА89(МАХ1), соответственно. В генах *nad4L* и *cob* было выявлено по одной синонимичной и одной несинонимичной мутации (табл. 3.15).

Также в мтДНК ЦМС линии НА89(МАХ1) были локализованы 29 INDEL. Среди них 15 делеций (1-5 п.н.) и 14 инсерций (1-10 п.н.) расположены исключительно в межгенных регионах митохондриального генома НА89(МАХ1) (приложение 2).

По сравнению с хлДНК в мтДНК НА89(МАХ1) выявлено меньшее число полиморфных сайтов: SSR (в 2,2 раза) и INDEL (в 1,4 раза). При этом абсолютное число SNP в мтДНК больше (в 1,2 раза), чем в хлДНК. Однако, учитывая размер геномов, мы рассчитали частоту SNP, которая составила 0,764 (мтДНК) и 1,298 (хлДНК) SNP на 1 т.п.н. То есть у ЦМС линии подсолнечника НА89(МАХ1) частота SNP в мтДНК (в 1,7 раз) ниже, чем в хлДНК.

Таким образом, в мтДНК ЦМС линии НА89(МАХ1) по сравнению с фертильным аналогом были выявлены следующие изменения структуры: 1 транспозиция (110 т.п.н.), 4 делеции (439, 978, 3183 и 14296 п.н.), 3 инсерции (1999, 5272 и 6583 п.н.), 18 SSR, 230 SNP и 29 INDEL. В результате этих изменений структуры размер митохондриального генома НА89(МАХ1) составил 295586 п.н., что на 5361 п.н. меньше, чем соответствующий показатель фертильной линии НА89. С функциональной точки зрения данные изменения структуры мтДНК привели к делеции гена *orf777* и возникновению 4-х открытых рамок считывания *orf306*, *orf480*, *orf645* и *orf1287*. Мы предполагаем, что наиболее вероятной причиной ЦМС типа МАХ1 подсолнечника является возникновение в митохондриальном геноме открытой рамки считывания *orf1287*, кодирующей *atp6*-химерный белок.

Аннотированный митохондриальный геном ЦМС линии НА89(МАХ1) подсолнечника был размещен в банке данных NCBI, его идентификатор - MN704580.1.

3.3.2.4. Митохондриальный геном ЦМС линии HA89(ANN2)

Наибольшее количество перестроек митохондриального генома среди исследованных нами ЦМС линий подсолнечника было обнаружено у HA89(ANN2). В мтДНК ЦМС линии HA89(ANN2) по сравнению с фертильным аналогом нами были выявлены, как крупные реорганизации структуры (более 100 п.н.), представленные транспозициями, делециями и инсерциями, так и незначительные изменения структуры, размером от 1 до 67 п.н. Результат сборки и аннотации митохондриального генома ЦМС линии HA89(ANN2) представлен на рисунке 3.12.

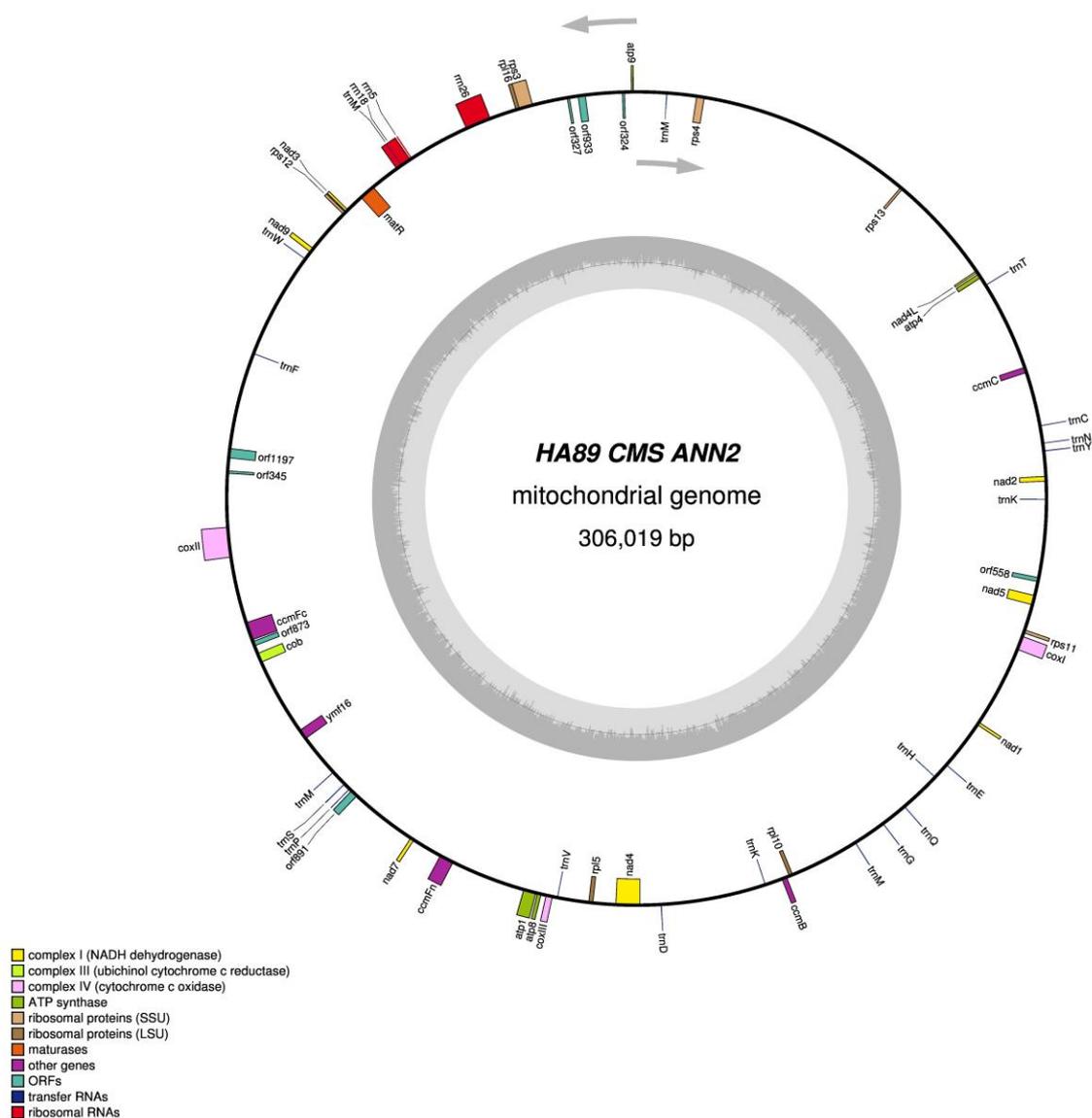


Рисунок 3.12 Генетическая карта митохондриального генома ЦМС линии НА89(ANN2) подсолнечника. Стрелками обозначены направления транскрипции генов мтДНК. Во внутреннем кольце серым и темно серым цветом отображено соотношение АТ и GC нуклеотидных пар

Сравнительный анализ нуклеотидных последовательностей мтДНК ЦМС линии НА89(ANN2) и ее фертильного аналога показал наличие 14 гомологичных областей (табл. 3.16).

Таблица 3.16. Гомологичные области митохондриальной ДНК фертильной линии НА89 и ЦМС линии НА89(ANN2).

№ участка	Размер участка, п.н.	Позиции в мтДНК фертильной линии НА89	Позиции в мтДНК ЦМС линии НА89(ANN2)	Ориентация	Сходство, %	Локализованные гены
1	29196	1-29196	1-29204	Прямая	99	<i>nad2, trnY, trnN, trnC, ccmC, trnT, atp4, nad4L</i>
2	557	33772-34328	78343-78899	Прямая	99	-
3	1245	34329-35573	148163-149411	Инверт.	98	-
4	77441	36739-114179	217575-295553	Прямая	95	<i>atp8, coxIII, trnV, rpl5, nad4, trnD, trnK, ccmB, rpl10, trnM, trnG, trnQ, trnH, trnE, nad1, coxI, nad5</i>
5	41702	114180-155882	35657-77315	Инверт.	99	<i>atp9, trnM, rps4, rrr26, rrr5, rrr18, rps13</i>
6	4150	155883-160032	300892-305041	Прямая	99	-
7	8584	160320-168903	129358-137946	Прямая	99	<i>nad6*</i>
8	21433	168906-190275	171388-192871	Инверт.	98	<i>nad6*, trnP, trnF, trnS, trnM, ymf16, cob</i>
9	8158	194543-202700	163232-171387	Прямая	99	<i>ccmFc, orf873</i>
10	24687	202701-227387	192915-217574	Инверт.	99	<i>atp1, ccmFn, nad7</i>
11	41505	227396-268900	87945-129446	Прямая	99	<i>rps3, rpl16, trnM, matR, nad3, rps12, nad9, trnW</i>

12	6029	269217-275245	141704-147732	Инверт.	99	<i>atp6*</i>
13	12520	275536-288055	150723-163231	Прямая	99	<i>coxII</i>
14	977	299971-300947	305042-306018	Прямая	99	<i>trnK</i>

*гены, у которых в результате транспозиции изменена структура по сравнению с фертильным аналогом

В результате многочисленных перестроек в мтДНК линии HA89(ANN2) за исключением участков № 1 и № 14, которые представляют фактически одну область кольцевой молекулы мтДНК, все другие участки (№ 2 - № 13) изменили свое расположение в митохондриальном геноме (рис. 3.13). В связи с этим участки № 2 - № 13 в мтДНК HA89(ANN2) мы определили, как транспозиции.

Фертильная линия HA89

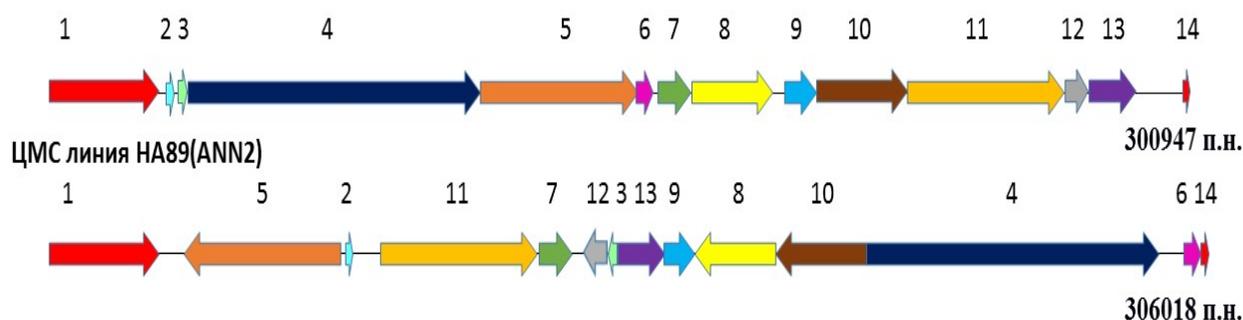


Рисунок 3.13. Схематичное изображение гомологичных участков мтДНК фертильной линии HA89 и ЦМС линии HA89(ANN2). 1 – 29196 п.н.; 2 – 557 п.н.; 3 – 1245 п.н.; 4 – 77441 п.н.; 5 – 41702 п.н.; 6 – 4150 п.н.; 7 – 8584 п.н.; 8 – 21433 п.н. 9 – 8158 п.н.; 10 – 24687 п.н.; 11 – 41505 п.н.; 12 – 6029 п.н.; 13 – 12520 п.н.; 14 – 977 п.н.

В большинстве случаев данные транспозиции приводят только к изменению направления транскрипции генов, или их расположения относительно друг друга. Однако, две транспозиции размером 8584 п.н. и 21433 п.н. изменяют структуру гена *nad6*, а транспозиция размером 6029 п.н. - гена *atp6*. Важно отметить, что большинство обнаруженных нами транспозиций в мтДНК HA89(ANN2) сопряжены с другими реорганизациями структуры – делециями и инсерциями.

Также в мтДНК ЦМС линии НА89(ANN2) мы обнаружили 9 крупных (более 100 п.н.) делеций размером 287 п.н., 290 п.н., 299 п.н., 316 п.н., 583 п.н., 1165 п.н., 4204 п.н., 4575 п.н. и 11901 п.н. (табл. 3.17). Несмотря на относительно большое число делеций и значительные размеры некоторых делетируемых участков, данные изменения структуры мтДНК практически не затронули кодирующих областей генома, за исключением 2-х делеций размером 316 п.н. и 1165 п.н., в результате которых элиминируется часть гена *atp6* и полностью ген *orf777* (табл. 3.17). Примечательно, что ген *orf777* также делетирован и у ЦМС линий НА89(РЕТ2) и НА89(МАХ1).

Таблица 3.17. Делеции размером более 100 п.н., выявленные в митохондриальной ДНК ЦМС линии НА89(ANN2).

Размер делетируемого участка, п.н.	Позиции в мтДНК фертильной линии НА89, п.н.	Локализация делеции на генетической карте мтДНК фертильной линии НА89	Гены, входящие в делетируемый участок
287	160032-160319	<i>rps13-nad6</i>	-
290	275246-275535	<i>atp6-coxII</i>	-
299	56701-56999	<i>nad4-ccmB</i>	-
316	268901-269216	<i>nad9-atp6</i>	<i>atp6</i> (часть гена)
583	70338-70920	<i>rpl10-nad1</i>	-
1165	35574-36738	<i>nad4L-orf777-atp8</i>	<i>orf777</i>
4204	190339-194542	<i>cob-ccmFc</i>	-
4575	29197-33771	<i>nad4L-orf777</i>	-
11901	288070-299970	<i>coxII-nad2</i>	-

Наряду с транспозициями и делециями в мтДНК ЦМС линии НА89(ANN2) были локализованы 7 крупных (более 100 п.н.) инсерций размером 430 п.н., 1027 п.н., 1310 п.н., 3757 п.н., 5338 п.н., 6452 п.н., 9045 п.н. (табл. 3.18).

Таблица 3.18. Инсерции более 100 п.н., выявленные в митохондриальной ДНК ЦМС линии HA89(ANN2).

Размер инсерции, п.н.	Позиции инсерции в мтДНК HA89(ANN2), п.н.	ОРС, возникшие в результате инсерции
430	147733-148162	<i>orf1197*</i>
1027	77316-78342	<i>orf324</i>
1310	149412-150722	<i>orf345</i>
3757	137947-141703	
5338	295554-300891	<i>orf558</i>
6452	29205-35656	
9045	78900-87944	<i>orf327, orf933</i>

* ОРС, возникающая в результате нескольких реорганизаций структуры мтДНК

В результате обнаруженных нами реорганизаций структуры мтДНК HA89(ANN2) возникли 7 новых ОРС (табл. 3.19), 5 из которых *orf324*, *orf327*, *orf345*, *orf558*, *orf933* являются следствием инсерций (табл. 3.18), еще одна *orf891* - двух транспозиций размером 21433 п.н. и 24687 п.н. и последняя *orf1197* - одновременно трех крупных перестроек, делеции размером 316 п.н., инсерции размером 430 п.н. и транспозиции размером 6029 п.н. Проведенный нами анализ транскрипционной активности этих 7-и ОРС свидетельствует, все они за исключением *orf891*, активно экспрессируются только в тканях растений HA89(ANN2).

Таблица 3.19. Открытые рамки считывания, обнаруженные в митохондриальной ДНК ЦМС линии HA89(ANN2).

ОРС	Локализация ОРС в мтДНК HA89(ANN2)	Аминокислотная последовательность, кодируемая ОРС
<i>orf324</i>	77915-78241	MKDQLIEGKSESIKMQLYEHRLLENLPREKARKRPNRHTNSLSRLPN RLAKGGSSYLYESRTRFAAEWNEIHVLGGTSIPWHVGVGRKHSHFID SLSPFLIGLQSFSP
<i>orf327</i>	84497-84826	MKSYGHIFILRAIELSLASIKPSNYRLYNKLLRAIFSSPNTFNLYLTQC RRLVLIPFGNSAGGKGLSVSSKFIRPLRMVGTGQDWDSQQLSFFSSN PYPPQGVPGGLSTSI
<i>orf345</i>	149669-150016	MCLRPLFLVYDYPCSSSDVDIFTWLK VEMCDWLRPGNSFPSRNSS RGVKS GGRSGCRGHSISGKEDGLDSL FMAWAKAALMTFSEDMLI RRSQSQLCSYAFTLSRPSK KDVHN

<i>orf558</i>	295955-296515	MIVLEWLFLTIAPCDAAEPRQLGSQDAATPMMQGIIDFLHPPLLIIL VFVSGVLLLMERRGPHRIASSSSLSVFAAGGEDPNRGRRRTPVTR ETITTTGISLLRQLILEILVDTPDQALREGLNNPSTQAWNRALESALT ERFGHSRYTWGHLWNIVNMVLELSEQGERSPFFLVRVIALVRARGT
<i>orf891</i>	192053-192946	MILSVLSSLALVSGLMVVRAKNPVHSVLFPIPVFRNTSGLLLLLLGL DFFAMIFPVVHIGAIASVFLFVMMFHIQIAEIHVEVLRVLPVSGIIG LIFWWEMFFILDNESIPLPTQRNTTSLRYMVYAGKVRSWTNLETL GNLLYTYYSVWFLVPSLILLVAMIGAIVLTMHRTTKVKRQDVFRR NAIDSRRTIMRGM TDLLKESSLILVRIQFVRWPSWSYRCLDTLLFFS HFRMTVPFHFVYNFKPGPAIRCISRTPGIRLLLEYHGIEYKDFKES LSYNLNLESFLLHGNSPN
<i>orf933</i>	82591-83526	MRFLFSFLPERFRVLLNAIHNQGSINDRPPGSVPSSSSVRIPLNTSVI VPPTRVGPSPVPLTNEIDLPLLELRRRLQLDVP LSNESLRNAQSALT IADSVPENQENVMGIRYRHDPGRTSFLNFEFEMTNELIRKLTNICNS IGIKSSSVPIYIPDLFVYSRFGDGLRSLVHETYDKVLVEILSFYKGRF GRFII VSLGVSCVWYTFVPGNPDIALLSDLFPRFETYQSFLNPDT YNKFFHKICYIERSYITGEVLKTIENEFPFSELNIPEESGGTRVAVGL GLMIGVFLAMGIVPFTDRVPVNLIE
<i>orf1197</i>	146967-148166	MINYVLQSMKVTASSSSSGISFPFVLEIVGTPLPGDSFTNRLTASTVS IPGYIVDPKISSLTHFNRLLDLKYDFLEGRMYPKYIFSLQQLDN TPIESLGK KLDSMRQDEL TGLSDYLQSLGYS DQAYLEYLKICDERY RDMVVPSPLEQFSILPLIPMKIENLYFSFTNSSLFMLLTSLVLLLIHF VTKKGGGNLVPNAWQSLVELIYDFVLNLVNEQIGGLSGNVKQKFF PCILVTFTLLFCNLQGMIPYSFTVTS HFLITLGLSFSIFIGITIVGFQR NGLHFLSFLLPAGVPLPLAPFLVLELISYCFRALSGLRIFANMMA GHSLVKILSGFAWTMLCMNDLLYFIGDLGPLFIVLALTGLELGVAI LQAYVFTILICIYLNDAINLH

С функциональной точки зрения, особый интерес представляют те ОРС, которые кодируют белки, включающие трансмембранные домены, а именно: *orf558*, *orf891*, *orf933*, *orf1197*. За исключением *orf933*, все остальные ОРС представляют собою химерные последовательности. Например, мембранно-связывающая область белка, кодируемого *orf558*, гомологична II субъединице цитохром с-оксидазы. Более 90 % последовательности *orf891* идентичны гену *nad6*, а более половины *orf1197* – гену *atp6*.

Как было описано ранее, *orf891* не экспрессируется в мтДНК ЦМС линии НА89(ANN2), и наиболее вероятно является псевдогеном. Вклад других ОРС (*orf558*, *orf933*, *orf1197*), кодирующих трансмембранные белки, в формирование ЦМС фенотипа довольно сложно оценить. Согласно литературным данным, *atp6* химерные гены или новые ОРС, которые котранскрибируются с геном *atp6*, чаще всего становятся причиной ЦМС фенотипа у цветковых растений (Yamamoto et

al., 2005; Kim, Kim, 2006; Jing et al., 2012; Horn R et al., 2014; Tan et al., 2017). Примечательно, что в случае ЦМС типа MAX1 подсолнечника нами также локализован *atp6* химерный ген - *orf1287* (раздел 3.3.2.3). В связи с этим мы предполагаем, что и в случае с ЦМС типа ANN2 ключевую роль играет *atp6* модифицированный ген - *orf1197*.

Кроме крупных реорганизаций в митохондриальном геноме ЦМС линии HA89(ANN2) нами были выявлено значительное число вариантных сайтов, а именно: 17 SSR, 288 SNP и 35 INDEL (1-67 п.н.) (приложение 2). Состав полиморфных SSR локусов был следующим: (T)_n – 58,8 %, (A)_n – 23,5%, (G)_n – 17,7 %. Полиморфные SSR локусы (C)_n в мтДНК HA89(ANN2) не обнаружены (табл. 3.20).

Таблица 3.20. Полиморфные SSR локусы митохондриальной ДНК ЦМС линии HA89(ANN2).

Позиция в мтДНК фертильной линии HA89, п.н.	Фертильная линия HA89	HA89(ANN2)	Локализация
3031	(G) ₅	(G) ₆	МГР <i>nad2-ccmC</i>
3107	(T) ₅	(T) ₄	МГР <i>nad2-ccmC</i>
7404	(G) ₁₀	(G) ₈	МГР <i>nad2-ccmC</i>
41462	(T) ₇	(T) ₈	МГР <i>coxIII-rpl5</i>
51679	(G) ₁₀	(G) ₈	МГР <i>nad4-ccmB</i>
56443	(A) ₆	(A) ₅	МГР <i>nad4-ccmB</i>
56550	(T) ₅	(T) ₆	МГР <i>nad4-ccmB</i>
85178	(T) ₈	(T) ₉	МГР <i>rpl10-nad1</i>
86188	(T) ₇	(T) ₆	МГР <i>rpl10-nad1</i>
149893	(T) ₁₀	(T) ₉	МГР <i>rps13-nad6</i>
152989	(T) ₉	(T) ₁₀	МГР <i>rps13-nad6</i>
170187	(T) ₁₄	(T) ₈	МГР <i>nad6-ymf16</i>
178408	(T) ₉	(T) ₈	<i>ymf16</i> интрон
179735	(A) ₁₀	(A) ₉	МГР <i>ymf16-cob</i>
248268	(A) ₁₄	(A) ₁₀	МГР <i>rpl16-matR</i>

260997	(T) ₇	(T) ₈	МГР <i>nad9-atp6</i>
261909	(A) ₈	(A) ₇	МГР <i>nad9-atp6</i>

Больше половины (55,8 %) точковых мутаций в мтДНК ЦМС линии HA89(ANN2) представлены трансверсиями, а именно: G/T – 62 (21,5 %), A/C – 47 (16,3%), C/G – 36 (12,5 %), A/T – 16 (5,5 %). Среди транзиций (44,2 %) обнаружены следующие частоты нуклеотидных замен: C/T – 70 (24,4 %), A/G – 57 (19,8%). 2 из 288 SNP были локализованы в гене рибосомальной РНК (*rrn26*), а 48 SNP – в экзонах белок кодирующих митохондриальных генов (табл. 3.21).

Таблица 3.21. SNP, локализованные в экзонах митохондриальных генов ЦМС линии HA89(ANN2).

Позиция в мтДНК фертильной линии HA89, п.н.	Фертильная линия HA89	HA89(ANN2)	Ген	Тип мутации
36947	A	G	<i>atp8</i>	несинонимичная мутация Asp59Gly
38134	C	A	<i>coxIII</i>	синонимичная мутация
43358	G	T	<i>rpl5</i>	синонимичная мутация
67110	C	T	<i>ccmB</i>	синонимичная мутация
107634	T	G	<i>coxI</i>	синонимичная мутация
122615	A	G	<i>rps4</i>	несинонимичная мутация Lys167Arg
178952	A	C	<i>ymf16</i>	несинонимичная мутация Met161Leu
188296	G	T	<i>cob</i>	синонимичная мутация
188443	T	G	<i>cob</i>	синонимичная мутация
188452	G	A	<i>cob</i>	синонимичная мутация
189084	A	G	<i>cob</i>	синонимичная мутация
189619	G	A	<i>cob</i>	несинонимичная мутация Glu427Lys
202282	G	C	<i>orf873</i>	несинонимичная мутация Ala129Gly
202308	C	T	<i>orf873</i>	синонимичная мутация
202311	T	C	<i>orf873</i>	синонимичная мутация
202324	G	T	<i>orf873</i>	несинонимичная мутация Ser115Glu
202325	A	C	<i>orf873</i>	несинонимичная мутация Ser115Glu

202332	C	T	<i>orf873</i>	несинонимичная мутация Glu112Gly
202333	T	C	<i>orf873</i>	несинонимичная мутация Glu112Gly
202350	T	C	<i>orf873</i>	синонимичная мутация
202353	T	C	<i>orf873</i>	синонимичная мутация
202380	C	G	<i>orf873</i>	несинонимичная мутация Trp96Cys
202441	C	G	<i>orf873</i>	несинонимичная мутация Cys76Ala
202442	A	C	<i>orf873</i>	несинонимичная мутация Cys76Ala
202445	A	G	<i>orf873</i>	синонимичная мутация
202446	C	T	<i>orf873</i>	синонимичная мутация
202455	G	T	<i>orf873</i>	синонимичная мутация
202520	T	C	<i>orf873</i>	несинонимичная мутация Ile50Val
202526	G	T	<i>orf873</i>	несинонимичная мутация Leu48Ile
202538	T	A	<i>orf873</i>	несинонимичная мутация Ile44Leu
202539	T	C	<i>orf873</i>	несинонимичная мутация Ile44Leu
202548	T	G	<i>orf873</i>	несинонимичная мутация Leu40Phe
202554	C	A	<i>orf873</i>	несинонимичная мутация Leu38Phe
202583	T	G	<i>orf873</i>	несинонимичная мутация Thr29Pro
202605	C	A	<i>orf873</i>	несинонимичная мутация Met21Ile
202616	T	G	<i>orf873</i>	несинонимичная мутация Ile18Leu
202620	C	T	<i>orf873</i>	синонимичная мутация
202628	T	C	<i>orf873</i>	несинонимичная мутация Trh14Ala
213951	C	T	<i>csmFn</i>	синонимичная мутация
214504	C	A	<i>csmFn</i>	несинонимичная мутация Arg193Leu
214655	C	T	<i>csmFn</i>	несинонимичная мутация Ala143Thr
214817	T	C	<i>csmFn</i>	несинонимичная мутация Ile91Val
228443	C	T	<i>rps3</i>	несинонимичная мутация Ala3Val
229032	G	A	<i>rps3</i>	синонимичная мутация
250521	C	T	<i>matR</i>	синонимичная мутация
251046	G	C	<i>matR</i>	синонимичная мутация
251052	C	T	<i>matR</i>	синонимичная мутация
251215	G	C	<i>matR</i>	несинонимичная мутация Ala227Gly

В гене *orf873* у HA89(ANN2) локализованы 26 SNP. Идентичные мутации мы обнаружили и у линии HA89(MAX1) (см. раздел 3.3.2.3). Остальные 22 SNP были локализованы в 11 генах. В генах *rpl5*, *csmB*, *coxI*, *coxIII* было выявлено только по одной синонимичной мутации, а в генах *atp8*, *rps4*, *ymf16* – по одной несинонимичной мутации. Более одного SNP мы обнаружили в генах *rps3*, *cob*, *csmFn* и *matR* (табл.3.21).

Также в митохондриальном геноме HA89(ANN2) было выявлено 35 INDEL (приложение 2) размером 1-67 п.н. Одна инсерция размером 5 нуклеотидов (AAGAA) возникла в гене *orf873*, в результате которой произошел сдвиг рамки считывания и образовался преждевременный стоп кодон. Все остальные 34 INDEL были локализованы в некодирующих участках митохондриального генома HA89(ANN2).

По сравнению с хлДНК в мтДНК HA89(ANN2) выявлено меньшее число SSR (в 2,4 раза) и почти равное - INDEL. При этом абсолютное число SNP в мтДНК больше (в 2 раза), чем в хлДНК. Однако, учитывая размер геномов, мы рассчитали частоту SNP, которая составила 0,96 (мтДНК) и 0,94 (хлДНК) SNP на 1 т.п.н. То есть у ЦМС линии подсолнечника HA89(ANN2) частота SNP в мтДНК равна - в хлДНК.

Таким образом, в мтДНК ЦМС линии HA89(ANN2) по сравнению с фертильным аналогом были выявлены следующие изменения структуры: 12 транспозиций (557, 1245, 4150, 6029, 8158, 8584, 12520, 21433, 24687, 41505, 41702 и 77441 п.н.), 9 делеций (287, 290, 299, 316, 583, 1165, 4204, 4575 и 11901 п.н.), 7 инсерции (430, 1027, 1310, 3757, 5338, 6452 и 9045 п.н.), 17 SSR, 288 SNP и 35 INDEL. В результате этих мутационных изменений размер митохондриального генома HA89(ANN2) составил 306018 п.н., что на 5071 п.н. больше, чем соответствующий показатель фертильной линии HA89. С функциональной точки зрения данные изменения структуры мтДНК привели к делеции генов *orf777*, *nad6*, *atp6* и возникновению 7-ми открытых рамок считывания: *orf324*, *orf327*, *orf345*, *orf558*, *orf891*, *orf933*, *orf1197*. Мы предполагаем, что наиболее вероятной

причиной ЦМС типа ANN2 подсолнечника является возникновение в митохондриальном геноме открытой рамки считывания *orf1197*, кодирующей *atp6*-химерный белок.

ЗАКЛЮЧЕНИЕ

Для реализации цели и задач настоящего исследования были использованы три уникальные генетические модели. Первая представляет собой однолетние и многолетние виды рода *Helianthus* L из коллекции ВИР. Вторая включает 4 линии культурного и 1 дикорастущего подсолнечника *H. annuus*. Третья представлена фертильным и стерильными аналогами линии НА 89 подсолнечника с различными типами ЦМС (PET1, PET2, ANN2, MAX), полученными на основе однолетних и многолетнего видов рода *Helianthus* L.

Для определения уровня межвидовой изменчивости хлДНК и мтДНК в роде *Helianthus* L. были определены полиморфные сайты в хлоропластных (*atpB*, *matK* и *rbcL*) и митохондриальных (*atp1*, *matR*, *nad5*) генов у 5 однолетних и 16 многолетних видов подсолнечника из коллекции ВИР. Количество точковых мутаций (SNP) в изученных генах хлДНК в среднем превышает в 3,6 раз соответствующий показатель в генах мтДНК. Гены, кодирующие субъединицы АТФ-синтазы оказались наиболее консервативными как в хлоропластном (*atpB*), так и митохондриальном (*atp1*) геномах. Полученные результаты согласуются с данными литературы о скорости эволюционной изменчивости хлоропластных и митохондриальных генов цветковых растений (Drouin et al., 2008; Hoekstra et al., 2017). Также были обнаружены видоспецифичные SNP у *H. strumosus* в генах *rbcL* (хлДНК) и *matR* (мтДНК), у *H. ciliaris* в генах *atpB* (хлДНК) и *matR* (мтДНК). Различия в нуклеотидных последовательностях хлоропластных генов *matK* и *rbcL* позволили кластеризовать однолетние и многолетние виды подсолнечника. В связи с тем, что митохондриальные гены продемонстрировали высокий уровень консерватизма у представителей рода *Helianthus* L., мы провели анализ некодирующих последовательностей (SSR локусов) митохондриальной ДНК. На основе данных о полиморфизме митохондриальных SSR локусов были определены митотипы для каждого вида подсолнечника и построена кладограмма, отражающая различия между ними.

Анализ нуклеотидных последовательностей хлоропластных и митохондриальных геномов у 4-х линий культурного и 1-ой линии дикорастущего подсолнечника (*H. annuus* L.) позволил оценить спектр и уровень внутривидовой изменчивости. Полиморфизм хлоропластного генома оказался в 5,1 раз выше по сравнению с митохондриальным. Также было отмечено, что дикорастущий подсолнечник значительно отличается от селекционных линий, как по уровню изменчивости хлДНК, так и мтДНК.

В результате сравнительного анализа хлДНК фертильного и стерильных аналогов линии НА89 подсолнечника с различными типами ЦМС (PET1, PET2, ANN2, MAX1) было показано, что размер и структура хлоропластных геномов в целом сходны. Тем не менее, нами локализовано 447 полиморфных сайта, в том числе 57 SSR, 315 SNP и 75 INDEL. 119 из 315 обнаруженных точковых мутаций локализованы в кодирующих участках генов, 58 из которых являются несинонимичными. Только 1 из 75 INDEL был выявлен в кодирующей части хлоропластного генома у ЦМС линии НА89(PET2). Среди аллоплазматических ЦМС линий наибольшее число полиморфных сайтов по сравнению с фертильным аналогом было идентифицировано у ЦМС линии НА89(MAX1), а наименьшее – у ЦМС линии НА89(PET1).

В отличие от хлДНК в мтДНК стерильных аналогов линии НА89 подсолнечника были обнаружены не только малые (<100 п.н.) изменения структуры, но и крупные (>100 п.н.) реорганизации структуры.

Например, в мтДНК ЦМС линии НА89(PET1) по сравнению с фертильным аналогом были выявлены следующие изменения: 1 инверсия (11852 п.н.), 1 инсерция (4732 п.н.), 1 делеция (451 п.н.), 8 SSR, 7 SNP и 2 INDEL. В результате этих изменений структуры размер митохондриального генома НА89(PET1) составил 305217 п.н., что превышает на 4270 п.н. соответствующий показатель фертильной линии. Реорганизации структуры мтДНК привели к возникновению у НА89(PET1) 2-х открытых рамок считывания *orf306* и *orfH522*.

В мтДНК ЦМС линии НА89(PET2) были выявлены 2 транспозиции (27,5 и 106,5 т.п.н.), 2 делеции (711 и 3780 п.н.), 2 инсерции (5050 и 15885 п.н.), 14 SSR,

55 SNP и 13 INDEL. В результате этих изменений размер митохондриального генома HA89(PET2) составил 316582 п.н., что на 15637 п.н. превышает соответствующий показатель фертильной линии. Реорганизации структуры мтДНК у HA89(PET2) привели к делеции гена *orf777*, дупликации гена *atp6* и возникновению 4-х открытых рамок считывания: *orf228*, *orf285*, *orf645* и *orf2565*.

В мтДНК ЦМС линии HA89(MAX1) были локализованы 1 транспозиция (110 т.п.н.), 4 делеции (439, 978, 3183 и 14296 п.н.), 3 инсерции (1999, 5272 и 6583 п.н.), 18 SSR, 230 SNP и 29 INDEL. В результате этих изменений размер митохондриального генома HA89(MAX1) составил 295586 п.н., что на 5631 п.н. меньше, чем соответствующий показатель фертильной линии. Реорганизации структуры мтДНК у HA89(MAX1) привели к делеции гена *orf777* и возникновению 4-х открытых рамок считывания *orf306*, *orf480*, *orf645* и *orf1287*.

В мтДНК ЦМС линии HA89(ANN2) локализованы 12 транспозиций (557, 1245, 4150, 6029, 8158, 8584, 12520, 21433, 24687, 41505, 41702 и 77441 п.н.), 9 делеций (287, 290, 299, 316, 583, 1165, 4204, 4575 и 11901 п.н.), 7 инсерции (430, 1027, 1310, 3757, 5338, 6452 и 9045 п.н.), 17 SSR, 288 SNP и 35 INDEL. В результате этих мутационных изменений размер митохондриального генома HA89(ANN2) составил 306018 п.н., что на 5071 п.н. больше, чем соответствующий показатель фертильной линии. Реорганизации структуры мтДНК у HA89(ANN2) привели к делеции генов *orf777*, *nad6*, *atp6* и возникновению 7-ми открытых рамок считывания: *orf324*, *orf327*, *orf345*, *orf558*, *orf891*, *orf933*, *orf1197*.

Проведенный нами биоинформатический анализ полипептидов, кодируемыми активно транскрибирующимися открытыми рамками считывания, возникшими в результате реорганизаций мтДНК позволяет предположить, что ЦМС типа PET2 непосредственно связана с *atp9*-химерными генами *orf228* и *orf285*, ЦМС типа MAX1 – с *atp6*-химерным геном *orf1287*, ЦМС типа ANN2 - с *atp6*-химерным геном *orf1197*.

ВЫВОДЫ

1. Изучен полиморфизм хлоропластных (*atpB*, *matK* и *rbcL*) и митохондриальных (*atp1*, *matR*, *nad5*) генов у представителей 5 однолетних и 16 многолетних видов рода *Helianthus* L. из коллекции ВИР. Количество точковых мутаций (SNP) в изученных генах хлДНК в среднем превышает в 3,6 раз соответствующий показатель в генах мтДНК. Выявлены специфичные SNP: у *H. strumosus* в генах *rbcL* (хлДНК) и *matR* (мтДНК), у *H. ciliaris* в генах *atpB* (хлДНК) и *matR* (мтДНК).

2. Получены и проанализированы полные нуклеотидные последовательности хлДНК и мтДНК линий HA89 и 3629 культурного и линии 398941 дикорастущего подсолнечника (*H. annuus*). Уровень изменчивости, детектируемый по числу SSR, SNP и INDEL полиморфных сайтов, оказался в 5,1 раз выше в хлДНК, чем в мтДНК. Дикорастущий подсолнечник (*H. annuus*) значительно отличается от селекционных линий по уровню изменчивости как хлДНК, так и мтДНК.

3. Сравнительный анализ хлоропластных геномов фертильной линии HA89 и ее стерильных аналогов HA89(PET1), HA89(PET2), HA89(ANN2), HA89(MAX1) позволил локализовать 447 полиморфных сайта, в том числе 57 SSR, 315 SNP и 75 INDEL. 119 из 315 SNP обнаружены в кодирующих участках генов, при этом 58 SNP приводят к изменению аминокислотных последовательностей кодируемых белков. Среди стерильных аналогов наибольшее число полиморфных сайтов (246) в сравнении с фертильной линией идентифицировано у HA89(MAX1), а наименьшее (49) – у HA89(PET1).

4. Сравнительный анализ митохондриальных геномов фертильной линии HA89 и ее стерильных аналогов HA89(PET1), HA89(PET2), HA89(ANN2), HA89(MAX1) позволил локализовать крупные реорганизации геномов, в том числе 1 инверсию, 12 инсерций, 15 транспозиций и 16 делеций, а также 560 полиморфных сайтов, в том числе 33 SSR, 463 SNP и 64 INDEL. Размер

митохондриальных геномов аллоплазматических линий с различными типами ЦМС значительно отличался по сравнению с фертильным аналогом. Изменения структуры мтДНК привели к возникновению новых открытых рамок считывания: *orf228*, *orf285*, *orf306*, *orf324*, *orf327*, *orf345*, *orf480*, *orfH522*, *orf558*, *orf645*, *orf891*, *orf933*, *orf1197*, *orf1287*, *orf2565*. Среди стерильных аналогов наибольшее число изменений структуры мтДНК идентифицировано у НА89(ANN2), а наименьшее – у НА89(PET1).

5. Биоинформатический анализ полипептидов, кодируемыми активно транскрибирующимися открытыми рамками считывания, возникшими в результате реорганизаций мтДНК позволяет предположить, что ЦМС типа PET2 непосредственно связана с *atp9*-химерными генами *orf228* и *orf285*, ЦМС типа MAX1 – с *atp6*-химерным геном *orf1287*, ЦМС типа ANN2 - с *atp6*-химерным геном *orf1197*.

Список сокращений

- ВИР – Всероссийский институт генетических ресурсов растений имени Н.И. Вавилова
- МГР – межгенный регион
- мтДНК – митохондриальная ДНК
- ОРС – открытая рамка считывания
- п.н. – пара нуклеотидов
- ПЦР – полимеразная цепная реакция
- РБФК/О – рибулозобисфосфаткарбоксилаза- оксигеназа
- т.н.п. – тысяча пар нуклеотидов
- хлДНК – хлоропластная ДНК
- ЦМС – цитоплазматическая мужская стерильность
- BLAST – basic local alignment search tool – программа для локального выравнивания нуклеотидных последовательностей
- GenBank – международная база данных нуклеотидных последовательностей (<http://www.ncbi.nlm.nih.gov>)
- INDEL – insertion or deletion – инсерции или делеция одного или нескольких нуклеотидов.
- IR – inverted repeat – инвертированный повторяющийся регион
- LSC – large single copy region – большой однокопийный регион
- PIC – polymorphism information content – величина информационного полиморфизма
- SNP – single nucleotide polymorphism – однонуклеотидный полиморфизм
- SSC – small single copy region – малый однокопийный регион
- SSR – simple sequence repeats - простые повторяющиеся последовательности

Список использованной литературы

1. Анащенко, А. В. Современные вопросы прикладной ботаники, генетики и селекции подсолнечника / А. В. Анащенко // Научно-технический бюллетень ВНИИ растениеводства. – 1977. – №. 69. – С. 47-51.
2. Анащенко, А. В. Филогенетические связи в роде *Helianthus* L / А. В. Анащенко // Труды по прикладной ботанике, генетике и селекции. – 1979. – Т. 64. – №. 2. – С. 146-156.
3. Анисимова, И. Н. Некоторые вопросы геномного анализа рода *Helianthus* L. / И. Н. Анисимова, А. В. Анащенко, И. П. Гаврилюк // Вестник с.-х. науки. – 1982. – № 3. – С. 98-104.
4. Анисимова, И. Н. Цитоплазматическая мужская стерильность и перспективы ее использования в селекционно-генетических исследованиях и семеноводстве картофеля / И.Н. Анисимова, Т.А. Гавриленко // Вавиловский журнал генетики и селекции. – 2017. –Т. 21. – № 1. – С. 83-95.
5. Брагин, А. Г. Анализ гетероплазматического состояния митохондриальной ДНК фертильных и мужскостерильных растений сахарной свеклы (*Beta vulgaris*) / А. Г. Брагин, М. К. Иванов, Л. А. Федосеева, Г. М. Дымщиц // Вавиловский журнал генетики и селекции. – 2011. – Т. 15. – № 3. – С. 524-530.
6. Гаврилова, В.А. Подсолнечник / В. А. Гаврилова, И. Н. Анисимова – СПб: ВИР, 2003. – 209 с.
7. Даниленко, Н. Г. Миры геномов органелл / Н. Г. Даниленко, О. Г. Давыденко – Минск: Тэхналогія, 2003. – 494 с.
8. Матвеева, Т. В. Молекулярные маркеры для видоидентификации и филогенетики растений / Т. В. Матвеева, О. А. Павлова, Д. И. Богомаз, А. Е. Демкович, Л. А. Лутова // Экологическая генетика. – 2011 – Том 9. – № 1. – С. 32-43.

9. Синявская, М. Г. Экспрессия хлоропластного генома: современные представления и экспериментальные пути изучения / М. Г. Синявская, Н. Г. Даниленко, Н. В. Луханина, А. М. Шимкевич, О. Г. Давыденко // Вавиловский журнал генетики и селекции. – 2015. – Том 19. – № 5. – С. 511-528.
10. Чесноков, Ю. В. Оценка меры информационного полиморфизма генетического разнообразия / Ю. В. Чесноков, А.М. Артемьева // Сельскохозяйственная биология. – 2015. – Том 50. – № 5. – С. 571-578.
11. Allen, J. O. Comparisons among two fertile and three male-sterile mitochondrial genomes of maize / J. O. Allen, C. M. Fauron, P. Minx et al. // Genetics. – 2007. – Vol. 177. – № 2. – P. 1173-1192.
12. Allen, J. F. A structural phylogenetic map for chloroplast photosynthesis / J. F. Allen, W. B. M. de Paula, S. Puthiyaveetil, J. Nield // Trends in plant science. – 2011. – Vol. 16. – № 12. – P. 645-655.
13. Alverson, A. J. Insights into the evolution of mitochondrial genome size from complete sequences of *Citrullus lanatus* and *Cucurbita pepo* (*Cucurbitaceae*) / A. J. Alverson, X. Wei, D. W. Rice, D. B. Stern, K. Barry, J. D. Palmer // Molecular biology and evolution. – 2010. – Vol. 27. – № 6. – P. 1436-1448.
14. Alverson, A. J. Origins and Recombination of the Bacterial-Sized Multichromosomal Mitochondrial Genome of Cucumber / A. J. Alverson, D. W. Rice, S. Dickinson, K. Barry, J. D. Palmer // The Plant Cell. – 2011. – Vol. 23. – № 7. – P. 2499-2513.
15. Ardila, F. Structural features of a cytoplasmic male sterility source from *Helianthus resinosus*, CMS RES1 / F. Ardila, M. M. Echeverria, R. Rios, R. H. Rodriguez // Journal of Plant Breeding and Crop Science. – 2010. – Vol. 2 – № 7. – P. 168-172.
16. Bailey-Serres, J. Mitochondrial genome rearrangement leads to extension and relocation of the cytochrome c oxidase subunit I gene in sorghum / J. Bailey-Serres, D. K. Hanson, T. D. Fox, C. J. Leaver // Cell. – 1986. – Vol. 47. – № 4. – P. 567-576.
17. Bendich, A. J. Why do chloroplasts and mitochondria contain so many copies of their genome? / A. J. Bendich // BioEssays. – 1987. – Vol. 6. – P. 279-282.

18. Bock, D. G. Genome skimming reveals the origin of the Jerusalem Artichoke tuber crop species: neither from Jerusalem nor an artichoke / D. G. Bock, N. C. Kane, D. P. Ebert, L. H. Rieseberg // *New Phytologist*. – 2014. – Vol. 201. – № 3. – P. 1021-1030.
19. Bock, R. *Cell and Molecular Biology of Plastids* / R. Bock – Heidelberg, Germany: Springer-Verlag, 2007. – 534 p.
20. Bolger, A. M. Trimmomatic: A flexible trimmer for Illumina sequence data / A. M. Bolger, M. Lohse, B. Usadel // *Bioinformatics*. – 2014. – Vol. 30. – № 15. – P. 2114-2120.
21. Borner, T. Chloroplast RNA polymerases: Role in chloroplast biogenesis / T. Borner, A. Y. Aleynikova, Y. O. Zubo, V. V. Kusnetsov // *Biochimica et biophysica acta*. – 2015. – Vol. 1847. – № 9. – P. 761-769.
22. Chumley, T. W. The complete chloroplast genome sequence of *Pelargonium x hortorum*: Organization and evolution of the largest and most highly rearranged chloroplast genome of land plants / T. W. Chumley, J. D. Palmer, J. P. Mower, H. M. Fourcade, P. J. Calie, J. L. Boore, R. K. // *Jansen Molecular Biology and Evolution*. – 2006. – Vol. 23. – № 11. – P. 2175-2219.
23. Daniel, H. *Molecular biology and biotechnology of plant organelles* / H. Daniel, C. Chase – Dordrecht, The Netherlands: Springer, 2004. – 659 p.
24. Davila, J. I. Double-strand break repair processes drive evolution of the mitochondrial genome in *Arabidopsis*. / J. I. Davila, M.P Arrieta-Montiel, Y. Wamboldt, J. Cao, J. Hagemann, V. Shedge, Y. Z. Xu, D. Weigel, S.A. Mackenzie // *BMC biology*. – 2011. – Vol. 9. – № 1. – Режим доступа: <http://dx.doi.org/10.1186/1741-7007-9-64>.
25. Dewey, R. E. A mitochondrial protein associated with cytoplasmic male sterility in the T cytoplasm of maize / R. E. Dewey, D. H. Timothy, C. S Levings // *Proceedings of the National Academy of Sciences of the United States of America*. – 1987. – Vol. 84. – № 15. – P. 5374–5378.
26. Dieterich, J. H. Alloplasmic male sterility in *Brassica napus* (CMS “*Tournefortii*-Stiewe”) is associated with a special gene arrangement around a novel

atp9 gene / J. H. Dieterich, H. P. Braun, U. K. Schmitz // Molecular genetics and genomics. – 2003. – Vol. 269. – № 6. – P. 723-731.

27. Doorduyn, L. The Complete Chloroplast Genome of 17 Individuals of Pest Species *Jacobaea vulgaris*: SNPs, Microsatellites and Barcoding Markers for Population and Phylogenetic Studies / L. Doorduyn, B. Gravendeel, Y. Lammers, Y. Ariyurek, T. Chin-A-Woeng, K. Vrieling // DNA Research. – 2011. – Vol. 18. – P. 93–105.

28. Drouin, G. Relative rates of synonymous substitutions in the mitochondrial, chloroplast and nuclear genomes of seed plants / G. Drouin, H. Daoud, J Xia // Molecular Phylogenetics and Evolution. – 2008. – Vol. 49. – № 3. – P. 827-831.

29. Galardini, M. CONTIGuator: a bacterial genome finishing tool for structural insights on draft genomes. / M. Galardini, E.G. Biondi, M. Bazzicalupo, A. Mengoni // Source Code for Biology and Medicine. – 2011. – Vol. 6. – № 11. – P. 11.

30. Gillman, J. D. The petunia restorer of fertility protein is part of a large mitochondrial complex that interacts with transcripts of the CMS-associated locus / J.D. Gillman, S. Bentolila, M. R. Hanson // The Plant Journal. – 2007. – Vol. 49. – № 2. – P. 217–227.

31. Gray, M. W. RNA Editing in Plant Mitochondria / M. W. Gray // IUBMB Life. – 2009. – Vol. 61. – № 12. – P. 1101-1104.

32. Gualberto, J.M. The plant mitochondrial genome: Dynamics and maintenance / J. M. Gualberto, D. Mileshina, C. Wallet, A. K. Niazi, F. Weber-Lotfi, A. Dietrich // Biochimie. – 2014. – Vol. 100. – № 1. – P. 107-120.

33. Gurevich, A. QUAST: quality assessment tool for genome assemblies / A. Gurevich, V. Saveliev, N. Vyahhi, G. Tesler // Bioinformatics. – 2013. – Vol. 29 – № 8. – P. 1072-1075.

34. Gutierrez, S. Lack of mitochondrial and nuclear-encoded subunits of complex I and alteration of the respiratory chain in *Nicotiana sylvestris* mitochondrial deletion mutants / S. Gutierrez, M. Sabar, C. Lelandais et al. // Proceedings of the National Academy of Sciences of the United States of America. – 1997. – Vol. 94. – № 7. – P. 3436-3441.

35. Handa, H. Linear plasmids in plant mitochondria: peaceful coexistences or malicious invasions? / H. Handa // *Mitochondrion*. – 2008. – Vol. 8. – № 1. – P. 15-25.
36. Hanson, M. R. Interactions of mitochondrial and nuclear genes that affect male gametophyte development / M. R. Hanson, S. Bentolila // *The Plant Cell*. – 2004. – Vol. 16. – P. 154-169.
37. Heiser, C.B. Sunflowers, weeds, and cultivated plants / C.B. Heiser // *The genetics of colonizing species*. Academic Press, New York. – 1965. – P. 375-386.
38. Heiser, C.B. The North American sunflowers (*Helianthus*) / C.B. Heiser, D.M. Smith, S.B. Clevenger, C. Martin // *Memoirs of the Torrey Botanical Club*. – 1969. – Vol. 22 – № 3. – P. 1-218.
39. Hoekstra, P. H. Correlated evolutionary rates across genomic compartments in *Annonaceae* / P. H. Hoekstra, J. J. Wieringa, E. Smets, R. D. Brandao, J. C. Lopes, R. H. J. Erkens, L. W. Chatrou // *Molecular phylogenetics and evolution*. – 2017 – Vol. 114. – P. 63-72.
40. Horn, R. The CMS-associated 16 kDa protein encoded by orfH522 in the PET1 cytoplasm is also present in other male-sterile cytoplasm of sunflower / R. Horn, J. E.G. Hustedt, A. Horstmeyer, J. Hahnen, K. Zetsche. W. Friedt // *Plant molecular biology*. – 1996. – Vol. 30. – № 3. – P. 523-538.
41. Horn, R. Molecular diversity of male sterility inducing and male-fertile cytoplasm in the genus *Helianthus* / R. Horn // *Theoretical and applied genetics*. – 2002. – Vol. 104. – № 4. – P. 562–570.
42. Horn, R. Mitochondrion role in molecular basis of cytoplasmic male sterility / R. Horn, K. J. Gupta, N. Colombo // *Mitochondrion*. – 2014. – Vol. 19 – P. 198-205.
43. Hu, J. Mitochondria and cytoplasmic male sterility in plants / J. Hu, W. Huang, Q. Huang X. Qin, C. Yu, L. Wang, S. Li, R. Zhu, Y. Zhu // *Mitochondrion*. – 2014. – Vol. 19. – P. 282–288.
44. Igarashi, K. Whole genomic sequencing of RT98 mitochondria derived from *Oryza rufipogon* and northern blot analysis to uncover a cytoplasmic male

sterility-associated gene / K. Igarashi, T. Kazama, K. Motomura, K. Toriyama // Plant and cell physiology. – 2013. – Vol. 54. – № 2. – P. 237-243.

45. Jansen, R.K. Phylogenetic analyses of *Vitis* (*Vitaceae*) based on complete chloroplast genome sequences: effects of taxon sampling and phylogenetic methods on resolving relationships among rosids / R. K. Jansen, C. Kaittanis, C. Saski, S.B Lee, J. Tomkins, A.J. Alverson, H. Daniell // BMC evolutionary biology. – 2006 – Vol 6. – Режим доступа: <http://dx.doi.org/10.1186/1471-2148-6-32>.

46. Jiang, W. A comparative study of ATPase subunit 9 (*atp9*) gene between cytoplasmic male sterile line and its maintainer line in soybeans / W. Jiang, S. Yang, D. Yu, J. Gai // The African Journal of Biotechnology. – 2011. – Vol. 10. – P. 10387-10392.

47. Jing, B. A male sterility-associated cytotoxic protein ORF288 in *Brassica juncea* causes aborted pollen development / B. Jing, S. Heng, D. Tong et al. // Journal of Experimental Botany. – 2012. – Vol. 63. – № 3. – P. 1285-1295.

48. Chen, L. Male Sterility and Fertility Restoration in Crops / L. Chen, Y.G. Liu // Annual Review of Plant Biology. – 2014. – Vol. 65. – P. 579-606.

49. Chun, E. H. L. The isolation and characterization of DNA associated with chloroplast preparations / E. H. L. Chun, M. H. Vaugham, A. Rich // Journal of Molecular Biology. – 1963. – Vol. 7. – P. 130-141.

50. Kahlau, S. Sequence of the tomato chloroplast DNA and evolutionary comparison of solanaceous plastid genomes / S. Kahlau, S. Aspinall, J. C. Gray, R. Bock // Journal of molecular evolution. – 2006. – Vol. 63. – № 2. – P. 194-207.

51. Kane, N.C. Sunflower genetic, genomic and ecological resources / N.C. Kane, J. M. Burke, L. Marek et al. // Molecular Ecology Resources. – 2013. – Vol. 13. – № 1. – P. 10-20.

52. Khakhlova, O. Elimination of deleterious mutations in plastid genomes by gene conversion / O. Khakhlova, R. Bock // Plant journal. – 2006. – Vol 46. – № 1. – P. 85-94.

53. Kim, D. H. The organization of mitochondrial *atp6* gene region in male fertile and CMS lines of pepper (*Capsicum annuum* L.) / D. H. Kim, B. D. Kim // *Current Genetics*. – 2006. – Vol. 49. – P. 59–67.
54. Klodmann, J. Internal Architecture of Mitochondrial Complex I from *Arabidopsis thaliana* / J. Klodmann, S. Sunderhaus, M. Nimtz, L. Jansch, H.P. Braun // *Plant Cell*. – 2010. – Vol. 22. – № 3. – P. 797–810.
55. Knoop, V. The mitochondrial DNA of land plants: peculiarities in phylogenetic Perspective / V. Knoop // *Current Genetics*. – 2004. – Vol. 46. – P. 123-139.
56. Kohler, R. H. Cytoplasmic male sterility in sunflower is correlated with the co-transcription of a new open reading frame with the *atpA* gene / R. H. Kohler, R. Horn, A. Lossl, K. Zetsche // *Molecular genetics and genomics*. – 1991. – Vol. 227. – P. 369-376.
57. Krupinska, K. New insights into plastid nucleoid structure and functionality / K. Krupinska, J. Melonek, K. Krause // *Planta*. – 2013. – Vol. 237. – № 3. – P. 653-664.
58. Kubo, T. Male Sterility-Inducing Mitochondrial Genomes: How Do They Differ? / T. Kubo, K. Kitazaki, M. Matsunaga, H. Kagami, T. Mikami. *Critical Reviews in Plant Sciences*. – 2011. – Vol. 30. – № 4. – P. 378-400.
59. Kumar, R.A. Changes in DNA damage, molecular integrity, and copy number for plastid DNA and mitochondrial DNA during maize development / R.A. Kumar, D.J. Oldenburg, A.J. Bendich // *Journal of Experimental Botany*. – 2014. – Vol. 65. – № 22. – P. 6425-6439.
60. Langmead, B. Fast gapped-read alignment with Bowtie 2 / B. Langmead, S. Salzberg // *Nature methods*. – 2012. – Vol. 9. – № 4. – P.357-359.
61. Leclercq, P. Une sterilité cytoplasmique chez le tournesol / P. Leclercq // *Ann Amélior Plant*. – 1969. – Vol. 19. – P. 99–106.
62. Li, H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data / H. Li // *Bioinformatics*. – 2011. – Vol. 27. – № 21. – P. 2987-2993.

63. Liao, X. Complete sequence of kenaf (*Hibiscus cannabinus*) mitochondrial genome and comparative analysis with the mitochondrial genomes of other plants / X. Liao, Y. Zhao, X. Kong, et al. // Scientific reports. – 2018. – Vol. 8. – № 1 – Режим доступа: <http://dx.doi.org/10.1038/s41598-018-30297-w>.
64. Liu, H. Comparative analysis of mitochondrial genomes between a wheat K-type cytoplasmic male sterility (CMS) line and its maintainer line / H. Liu, P. Cui, K. Zhan, et al. // BMC Genomics. – 2011. – Vol. 12. – Режим доступа: <http://dx.doi.org/10.1186/1471-2164-12-163>.
65. Lohse, M. Organellar Genome Draw (OGDRAW): a tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. / M. Lohse, O. Drechsel, R. Bock. // Current Genetics. – 2007. – Vol. 52. – P. 267-274.
66. Ma, J. Organellar genome copy number variation and integrity during moderate maturation of roots and leaves of maize seedlings. / J. Ma, X.Q. Li // Current genetics. – 2015. – Vol. 61. – № 4. – P. 591-600.
67. Maier, R.M. Complete sequence of the maize chloroplast genome: gene content, hotspots of divergence and fine tuning of genetic information by transcript editing / R.M Maier, K. Neckermann, G. L. Igloi, H. Kössel // Journal of molecular biology. – 1995. – Vol. 251. – № 5. – P. 614-628.
68. Makarenko, M.S. Comparative Genomics of Domesticated and Wild Sunflower: Complete Chloroplast and Mitochondrial Genomes / M.S. Makarenko, A.V. Usatov, N.V. Markin, K.V. Azarin, O.F. Gorbachenko, N.A. Usatov // OnLine Journal of Biological Sciences. – 2016. – Vol. 16. – № 1. – P. 71-75.
69. Makarenko, M. S. Mitochondrial genomes organization in alloplasmic lines of sunflower (*Helianthus annuus* L.) with various types of cytoplasmic male sterility / M. S. Makarenko, I. V. Kornienko, K. V. Azarin, A. V. Usatov, M. D. Logacheva, N. V. Markin, V. A. Gavrilova // PeerJ. – 2018. – Vol. 6. – Режим доступа: <http://dx.doi.org/10.7717/peerj.5266>.

70. Marechal, A Recombination and the maintenance of plant organelle genome stability / A. Marechal, N. Brisson // *New Phytologist*. – 2010 – Vol. 186. – № 2. – P. 299-317.
71. Mason, C. M. How Old Are Sunflowers? A Molecular Clock Analysis of Key Divergences in the Origin and Diversification of *Helianthus* (*Asteraceae*) / C. M. Mason // *International Journal of Plant Sciences*. – Vol. 179. – № 3. – P. 182-191.
72. Millar, A. H. Mitochondrial cytochrome c oxidase and succinate dehydrogenase complexes contain plant specific subunits / A. H. Millar, H. Eubel, L. Jansch, V. Kruff, J. L. Heazlewood, H. P. Braun // *Plant molecular biology*. – 2004. – Vol. 56. – № 1. – P. 77-90.
73. Miller, J.F. Inheritance of restoration of *Helianthus petiolaris* sp. *fallax* (PEF1) cytoplasmic male sterility / J.F. Miller // *Crop Science*. – 1996. – Vol. 36. – P. 83-86.
74. Morley, S.A. Plant mitochondrial DNA / S.A. Morley, B.L. Nielsen // *Frontiers in Bioscience*. – 2017. – Vol. 22. – № 1. – P. 1023-1032.
75. Mower, J. P. Extensive variation in synonymous substitution rates in mitochondrial genes of seed plants / J. P Mower, P. Touzet, J. S. Gummow, L. F. Delph, J. D. Palmer // *BMC Evolutionary Biology*. – 2007. – Vol. 7. – Режим доступа: <http://dx.doi.org/10.1186/1471-2148-7-135>.
76. Ni, L. H. The complete chloroplast genome of *Gentiana straminea* (*Gentianaceae*), an endemic species to the Sino-Himalayan subregion / L. H. Ni, Z. L. Zhao, H. X. Xu, S. L. Chen, G. Dorje. // *Gene*. – 2016. - Vol. 577. – P. 281–288.
77. Nurk, S. Assembling Genomes and Mini-metagenomes from Highly Chimeric Reads. *Research in Computational Molecular Biology* / S. Nurk, A. Bankevich, D. Antipov et al. // *Lecture Notes in Computer Science*. – 2013. – Vol. 7821. – P. 158-170.
78. Ogihara, Y. Structural features of a wheat plastome as revealed by complete sequencing of chloroplast DNA / Y. Ogihara, K. Isono, T. Kojima, et al. // *Molecular genetics and genomics*. – 2002. – Vol. 266. – № 5. – P. 740-746.

79. Ogihara, Y. Structural dynamics of cereal mitochondrial genomes as revealed by complete nucleotide sequencing of the wheat mitochondrial genome / Y. Ogihara, Y. Yamazaki, K. Murai et al. // *Nucleic acids research*. – 2005. – Vol. 33. – № 19. – P. 6235-6250.
80. Oldenburg, D.J. The amount and integrity of mtDNA in maize decline with development / D.J. Oldenburg, R.A. Kumar, A.J. Bendich // *Planta*. – 2013. – Vol. 237. – № 2. – P. 603-617.
81. Oshima, M. The identification of quantitative trait loci that control the paternal inheritance of a mitochondrial plasmid in rapeseed (*Brassica napus* L.) / M. Oshima, H. Handa // *Genes and genetic systems*. – 2012. – Vol. 87. – № 1. – P. 19-27.
82. Park, J. Y. Complete mitochondrial genome sequence and identification of a candidate gene responsible for cytoplasmic male sterility in radish (*Raphanus sativus* L.) containing DCGMS cytoplasm / J. Y. Park, Y. P. Lee, J. Lee, B. S. Choi, S. Kim, T. J. Yang // *Theoretical and applied genetics*. – 2013 – Vol. 126. – № 7. – P. 1763-1774.
83. Parks, M. Increasing phylogenetic resolution at low taxonomic levels using massively parallel sequencing of chloroplast genomes / M. Parks, R. Cronn, A. Liston // *BMC Biology*. – 2009. – Vol. 7. – Режим доступа: <http://dx.doi.org/10.1186/1741-7007-7-84>.
84. Patwardhan, A. Molecular Markers in Phylogenetic Studies-A Review / A. Patwardhan, S. Ray, A. Roy // *Journal of Phylogenetics and Evolutionary Biology*. – 2014. – Vol. 2. – Режим доступа: <http://dx.doi.org/10.4172/2329-9002.1000131>.
85. Perry, A. S. Nucleotide substitution rates in legume chloroplast DNA depend on the presence of the inverted repeat / A. S. Perry, K. H. Wolfe // *Journal of molecular evolution*. – 2002. – Vol. 55. – № 5. – P. 501-508.
86. Plader, W. The complete structure of the cucumber (*Cucumis sativus* L.) chloroplast genome: its composition and comparative analysis / W. Plader, Y. Yukawa, M. Sugiura, S. Malepszy // *Cellular and molecular biology letters*. – 2007. – Vol. 12. № 4. – P. 584-594.

87. Popov, V.N. Succinate dehydrogenase in *Arabidopsis thaliana* is regulated by light via phytochrome A. / V. N. Popov, A. T. Eprintsev, D. N. Fedorin, A. U. Igamberdiev // FEBS Letters. – 2010. – Vol. 584. – № 1. – P. 199–202.
88. Preuten, T. Fewer genes than organelles: extremely low and variable gene copy numbers in mitochondria of somatic plant cells / T. Preuten, E. Cincu, J. Fuchs, R. Zoschke, K. Liere, T. Borner // Plant journal. – 2010. – Vol. 64. – № 6 – P. 948-959.
89. Qiu, Y. L. Reconstructing the Basal Angiosperm Phylogeny: Evaluating Information Content of Mitochondrial Genes / Y. L. Qiu, L. Li, T. A. Hendry et al. // Taxon. – Vol. 55. – No. 4. – P. 837-856.
90. Reddemann, A. Recombination Events Involving the *atp9* Gene Are Associated with Male Sterility of CMS PET2 in Sunflower / A. Reddemann, R. Horn // International journal of molecular sciences. – 2018. – Vol. 19. – № 3. – Режим доступа: <http://dx.doi.org/10.3390/ijms19030806>.
91. Rieseberg, L.H. Differential cytoplasmic versus nuclear introgression in *Helianthus* / L.H. Rieseberg, H.C. Choi, D. Ham // Journal of Heredity. – 1991. – Vol. 82(6). – P. 489-493.
92. Robison, M. M. A mitochondrial plasmid and plasmid-like RNA and DNA polymerases encoded within the mitochondrial genome of carrot (*Daucus carota* L.) / M. M. Robison, D. J. Wolyn // Current genetics. – 2005. – Vol. 47. – № 1. – P. 57-66.
93. Rodriguez-Moreno, L. Determination of the melon chloroplast and mitochondrial genome sequences reveals that the largest reported mitochondrial genome in plants contains a significant amount of DNA having a nuclear origin. / L. Rodriguez-Moreno, V. M. Gonzalez, A. Benjak, M. C. Marti, P. Puigdomenech, M. A. Aranda, J. Garcia-Mas // BMC Genomics. – 2011. – Vol. 7. – Режим доступа: <http://dx.doi.org/10.1186/1471-2164-12-424>.
94. Saarela, J.M. Plastid phylogenomics of the cool-season grass subfamily: clarification of relationships among early-diverging tribes / J. M. Saarela, W. P. Wysocki, C. F. Barrett, R. J. Soreng, J. I. Davis, L. G. Clark, S. A. Kelchner, J. C. Pires, P. P. Edger, D. R. Mayfield, M. R. Duvall // AoB Plants. – 2015 – Vol. 7. – Режим доступа: <http://dx.doi.org/10.1093/aobpla/plv046>.

95. Sager, R. Chloroplast genetics of *Chlamydomonas*. II. Mapping by cosegregation frequency analysis / R. Sager, Z. Ramanis // *Genetics*. – 1976 – Vol. 83 – P. 323-340.
96. Samigullin, T. H. Complete Plastid Genome of the Recent Holoparasite *Lathraea squamaria* Reveals Earliest Stages of Plastome Reduction in Orobanchaceae / T. H. Samigullin, M. D. Logacheva, A. A. Penin, C. M. Vallejo-Roman // *PLoS One*. – 2016. – Vol. 11. – Режим доступа: <http://dx.doi.org/10.1371/journal.pone.0150718>.
97. Sasaki, C. Complete chloroplast genome sequences of *Hordeum vulgare*, *Sorghum bicolor* and *Agrostis stolonifera*, and comparative analyses with other grass genomes / C. Sasaki, S. B. Lee, S. Fjellheim, C. Guda, R. K. Jansen, H. Luo, J. Tomkins, O. A. Rognli, H. Daniell, J. L. Clarke. // *Theoretical and applied genetics*. – 2007. – Vol. 115. – № 4. – P. 571-590.
98. Sato, S. Complete structure of the chloroplast genome of *Arabidopsis thaliana* / S. Sato, Y. Nakamura, T. Kaneko, E. Asamizu, S. Tabata // *DNA research*. – 1999. – Vol. 6 – № 5. – P. 283-290.
99. Sasaki, C. Complete chloroplast genome sequence of *Glycine max* and comparative analyses with other legume genomes / C. Sasaki, S. B. Lee, H. Daniell, T. C. Wood, J. Tomkins, H. G. Kim, R. K. Jansen // *Plant molecular biology*. – 2005. – Vol 59. – № 2. – P. 309-322.
100. Sazanov, L. A Structural Perspective on Respiratory Complex I: Structure and Function of NADH:ubiquinone oxidoreductase / L. Sazanov – Dordrecht, The Netherlands: Springer, 2012. – 286 p.
101. Schilling, E.E. Infrageneric classification of *Helianthus (Compositae)* / E.E. Schilling, C.B. Heiser // *Taxon*. – 1981. – Vol. 30 – № 2. – P. 393-403.
102. Serieys, H. A. Identification, study and utilization in breeding programs of new CMS / H. A. Serieys // *Helia*. – 1996 – Vol. 19. – P.144-160.
103. Shen, X. Complete chloroplast genome sequence and phylogenetic analysis of the medicinal plant *Artemisia annua* / X. Shen, M. Wu, B. Liao et al. // *Molecules*. – 2017. – Vol. 22. – Режим доступа: <http://dx.doi.org/10.3390/molecules22081330>.

104. Sloan, D. B. Rapid evolution of enormous, multichromosomal genomes in flowering plant mitochondria with exceptionally high mutation rates / D. B. Sloan, A. J. Alverson, J. P. Chuckalovcak, M. Wu, D. E. McCauley, J. D. Palmer, D. R. Taylor // *PLoS biology*. – 2012. – Vol. 10. – № 1. – Режим доступа: <http://dx.doi.org/10.1371/journal.pbio.1001241>.
105. Sloan, D. B. One ring to rule them all? Genome sequencing provides new insights into the «master circle» model of plant mitochondrial DNA structure / D. B. Sloan // *New Phytologist*. – 2013. – Vol. 200. – № 4. – P. 978-985.
106. Sloan, D. B. Correction of Persistent Errors in Arabidopsis Reference Mitochondrial Genomes / D. B. Sloan, Z. Wu, J. Sharbrough // *Plant Cell*. – 2018 – Vol. 30. – № 3. – P. 525-527.
107. Smith, D.R. Gene conversion shapes linear mitochondrial genome architecture / D.R. Smith, P.J. Keeling // *Genome biology and evolution*. – 2013. – Vol. 5. – № 5. – P. 905-912.
108. Spassova, M. Characterisation and expression of the mitochondrial genome of a new type of cytoplasmic male-sterile sunflower / M. Spassova, F. Moneger, C. J. Leaver, P. Petrov, A. Atanassov, H. J. Nijkamp, J. Hille // *Plant Molecular Biology*. – 1994 – Vol. 26. – № 6. – P. 1819-1831.
109. Stanley, L. Geographic patterns of genetic variation in three genomes of North American diploid strawberries with special reference to *Fragaria vesca subsp. Bracteata* / L. Stanley, N. J. Forrester, R. Govindarajulu, A. Liston, T. L. Ashman // *Botany*. – 2015. – Vol. 93. – № 9. – P. 573-588.
110. Stephens, J.D. Species tree estimation of diploid *Helianthus (Asteraceae)* using target enrichment / J.D. Stephens, W. L. Rogers, C. M. Mason, L. A. Donovan, R. L. Malmberg // *American journal of botany*. – 2015. – Vol. 102. – № 6. – P. 910-920.
111. Sugiyama, Y. The complete nucleotide sequence and multipartite organization of the tobacco mitochondrial genome: comparative analysis of mitochondrial genomes in higher plants / Y. Sugiyama, Y. Watase, M. Nagase, N. Makita, S. Yagura, A. Hirai, M. Sugiura // *Molecular Genetics and Genomics*. – 2005. – Vol. 272. – № 6. – P. 603-615.

112. Sun, P. Utility of in vitro culture to the study of plant mitochondrial genome configuration and its dynamic features. / P. Sun, M. P. Arrieta-Montiel, S. A. Mackenzie // Theoretical and applied genetics. – 2012. – Vol. 125. – № 3. – P. 449-454.
113. Tan, G. F. Different lengths, copies and expression levels of the mitochondrial *atp6* gene in male sterile and fertile lines of carrot (*Daucus carota* L.) / G. F. Tan, F. Wang, X. Y. Zhang, A. I. Xiong // Mitochondrial DNA. Part A, DNA Mapping, Sequencing, and Analysis. – 2017. – Vol. 29. – № 3. – P. 446-454.
114. Tang, J. A comparison of rice chloroplast genomes / J. Tang, H. Xia, M. Cao, et al. // Plant physiology. – 2004. – Vol. 135. – № 1. – P. 412-420.
115. Tewari, K. K. Chloroplast DNA from tobacco leaves / K. K. Tewari, S. G. Wildman // Science. – 1966. – Vol. 153. – P. 1269-1271.
116. Thomas, C. M. The nucleotide sequence and transcription of minicircular mitochondrial DNA's associated with male-fertile and cytoplasmic male-sterile lines of sugarbeet / C. M. Thomas // Nucleic Acids Research. – 1986. – Vol. 14. – P. 9353–9370.
117. Thorvaldsdottir, H. Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration / H. Thorvaldsdottir, J. T. Robinson, J. P. Mesirov. // Briefings in bioinformatics. – 2013. – Vol. 14. – № 2. – P. 178-192.
118. Timme, R.E. A comparative analysis of the *Lactuca* and *Helianthus* (*Asteraceae*) plastid genomes: identification of divergent regions and categorization of shared repeats / R.E. Timme, J.V. Kuehl, J.L. Boore, R.K. Jansen // American Journal of Botany. – 2007. – Vol. 94. – № 3. – P. 302-312.
119. Unseld, M. The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366,924 nucleotides / M. Unseld, J. R. Marienfeld, P. Brandt, A. Brennicke // Nature Genetics. – 1997 – Vol. 15. – № 1. – P.57-61.
120. Vischi, M. Identification of wild species of sunflower by a specific plastid DNA sequence / M. Vischi, F. Arzenton, E. De Paoli, S. Paselli, E. Tomat, A. M. Olivieri // Helia. – 2006. – Vol. 29. – P 11-18

121. Wendel, J. F. Plant Genome Diversity Volume 1: Plant Genomes, Their Residents, and Their Evolutionary Dynamics / J. F. Wendel, J. Greilhuber, J. Dolezel, I. J. Leitch. – Wien, Austria: Springer-Verlag, 2012 – 279 p.
122. Whelan, E. D. P. Hybridization between annual and perennial diploid species of *Helianthus* / E. D. P. Whelan // Canadian Journal of Genetics and Cytology – 1978 – Vol. 20. – № 4. – P. 523-530.
123. Whelan, E. D. P. Registration of sunflower germplasm composite crosses CMG-1, CMG-2, and CMG-3 / E. D. P. Whelan, W. Dedio // Crop Science. – 1980 – Vol. 20. – № 6. – P. 832.
124. Wicke, S. The evolution of the plastid chromosome in land plants: gene content, gene order, gene function / S. Wicke, G. M. Schneeweiss, C. W. de Pamphilis, K. F. Muller, D. Quandt // Plant Molecular Biology. – 2011. – Vol. 76. – P. 273–297.
125. Wolfe, K.H. Function and evolution of a minimal plastid genome from a nonphotosynthetic parasitic plant / K. H. Wolfe, C. W. Morden, J. D. Palmer // Proceedings of the National Academy of Sciences of the United States of America. – 1992. – Vol. 89. – № 22. – P. 10648-10652.
126. Woodson J. D. Coordination of gene expression between organellar and nuclear genomes / J. D. Woodson, J. Chory // Nature Reviews Genetics. – 2008. – Vol. 9. – P. 383-395.
127. Wu, B. Extensive horizontal transfer and homologous recombination generate highly chimeric mitochondrial genomes in yeast / B. Wu, A. Buljic, W. Hao // Molecular biology and evolution. – 2015 – Vol 32. – № 10. – P. 2559–2570.
128. Yamamoto, M. P. The 5' -leader sequence of sugar beet mitochondrial atp6 encodes a novel polypeptide that is characteristic of Owen cytoplasmic male sterility / M. P. Yamamoto, T. Kubo, T. Mikami // Molecular genetics and genomics. – 2005. – Vol. 273. – № 4. – P. 342-349.
129. Yang, J.H. Mitochondrial atpA gene is altered in a new orf220-type cytoplasmic male-sterile line of stem mustard (*Brassica juncea*) / J. H. Yang, Y. Huai, M. F. Zhang // Molecular Biology Reports. – 2009. – Vol. 36. – № 2. – P. 273–280.

130. Yang, M. The Complete Chloroplast Genome Sequence of Date Palm (*Phoenix dactylifera* L.) / M. Yang, X. Zhang, G. Liu // PLoS One. – 2010. – Vol. 5. – Режим доступа: <http://dx.doi.org/10.1371/journal.pone.0012762>.

131. Zhang, C. The complete chloroplast genome sequence of the *Cucurbita pepo* L. (*Cucurbitaceae*) / C. Zhang, Q. Zhu, S. Liu, P. Gao, Z. Zhu, X. Wang, F. Luan // Mitochondrial DNA Part B. – 2018. – Vol 3. – № 2. – P. 717-718.

132. Zhang, T. The Complete Chloroplast and Mitochondrial Genome Sequences of *Boea hygrometrica*: Insights into the Evolution of Plant Organellar Genomes. / T. Zhang, Y. Fang, X. Wang, X Deng, X Zhang, S Hu, J Yu // PLoS One. – 2012. – Vol. 7. – Режим доступа: <http://dx.doi.org/10.1371/journal.pone.0030531>.

Приложение 1.

Полиморфные сайты хлоропластной ДНК, локализованные у стерильных аналогов линии НА89 подсолнечника с различными типами ЦМС.

Позиция в геноме НА89 ферт.	Тип	НА89 фертильная	НА89 (PET1)	НА89 (PET2)	НА89 (ANN2)	НА89(MAX1)	Локализация	Регион
124	SNP	G		A	A		МГР <i>rpl2-psbA</i>	LSC
191-192	INDEL	AG				ATTTT CAATT G	МГР <i>rpl2-psbA</i>	LSC
205	INDEL	T		-	-	-	МГР <i>rpl2-psbA</i>	LSC
206	SSR	A11	A12	A9	A9	A9	МГР <i>rpl2-psbA</i>	LSC
307	SNP	T				C	МГР <i>rpl2-psbA</i>	LSC
370	SSR	A7				A9	МГР <i>rpl2-psbA</i>	LSC
1649	SNP	G				T	МГР <i>psbA-matK</i>	LSC
1667	SSR	A6		A7	A8	A9	МГР <i>psbA-matK</i>	LSC
1991	SSR	T9	T10	T8	T8	T13	МГР <i>psbA-matK</i>	LSC
2032	SSR	T12	T13	T9	T14	T10	МГР <i>psbA-matK</i>	LSC
2310	INDEL	T		-			<i>matK</i> сдвиг рамки считывания	LSC
2984	SNP	A				G	<i>matK</i> Phe195Ser	LSC
3730	SNP	G				T	МГР <i>matk-rps16</i>	LSC
3739	SNP	C				A	МГР <i>matk-rps16</i>	LSC
3953	SSR	T7				T8	МГР <i>matk-rps16</i>	LSC
4135	SNP	T				G	МГР <i>matk-rps16</i>	LSC
4402	SNP	G				A	МГР <i>matk-rps16</i>	LSC
4437	SSR	A7			A6		МГР <i>matk-rps16</i>	LSC
4570	SNP	T				C	МГР <i>matk-rps16</i>	LSC
4684	SNP	G			C		МГР <i>matk-rps16</i>	LSC
4788	SSR	A8				A9	МГР <i>matk-rps16</i>	LSC
4917	SNP	A		G	G	G	МГР <i>matk-rps16</i>	LSC
5002-03	INDEL	AT			ATTGC AGGT		МГР <i>matk-rps16</i>	LSC

5450	SSR	C9	C8	C10	C9		<i>rps16</i> интрон	LSC
5621	INDEL	A				-	<i>rps16</i> интрон	LSC
5653	SNP	A	C	C	C	C	<i>rps16</i> интрон	LSC
5692	SSR	T13	T11	T7	T7	T7	<i>rps16</i> интрон	LSC
5718	SSR	C7			C11	C8	<i>rps16</i> интрон	LSC
5725	INDEL	T				-	<i>rps16</i> интрон	LSC
5969	SNP	A		T			<i>rps16</i> интрон	LSC
6024	SNP	G		T		T	<i>rps16</i> интрон	LSC
6025	SNP	A		T			<i>rps16</i> интрон	LSC
6026	SNP	T		C			<i>rps16</i> интрон	LSC
6153	SNP	T		C	C		<i>rps16</i> интрон	LSC
6508	SNP	G				T	МГР <i>rps16-psbK</i>	LSC
6550	INDEL	A		-			МГР <i>rps16-psbK</i>	LSC
6583	SNP	G				A	МГР <i>rps16-psbK</i>	LSC
6773	SNP	C		T	T		МГР <i>rps16-psbK</i>	LSC
6963-64	INDEL	AA				ATA	МГР <i>rps16-psbK</i>	LSC
7091-92	INDEL	AG				AAAA GG	МГР <i>rps16-psbK</i>	LSC
7611	SNP	G		A			МГР <i>rps16-psbK</i>	LSC
7720	SSR	A8				A9	МГР <i>rps16-psbK</i>	LSC
7944	SSR	T8		T11	T9		МГР <i>psbK-psbI</i>	LSC
7952-53	INDEL	GA	GATA AAAA TGAT AACA ATA				МГР <i>psbK-psbI</i>	LSC
8294	SSR	T8			T10	T9	МГР <i>psbK-psbI</i>	LSC
8516	SNP	T				G	МГР <i>psbI-petN</i>	LSC
8552	SSR	A8		A7	A7	A7	МГР <i>psbI-petN</i>	LSC
8880	SSR	T6				T7	МГР <i>psbI-petN</i>	LSC
8924	SNP	C				G	МГР <i>psbI-petN</i>	LSC
9019	SNP	A			G		МГР <i>psbI-petN</i>	LSC
9080	SNP	T				C	МГР <i>psbI-petN</i>	LSC
9359	INDEL	T				-	МГР <i>psbI-petN</i>	LSC

9384-85	INDEL	CA				CAA	МГР <i>psbI-petN</i>	LSC
9398	SNP	T				G	МГР <i>psbI-petN</i>	LSC
9781	SNP	T		G			МГР <i>psbI-petN</i>	LSC
9831-35	INDEL	ATTA G				-	МГР <i>psbI-petN</i>	LSC
9881	SNP	T		C	C		МГР <i>psbI-petN</i>	LSC
9882	SSR	A8	A9	A9	A9	A10	МГР <i>psbI-petN</i>	LSC
9890-91	INDEL	AT			ATT	ATT	МГР <i>psbI-petN</i>	LSC
10099	SNP	A		C	C		МГР <i>psbI-petN</i>	LSC
10749	SNP	T		A	A	A	МГР <i>petN-psbM</i>	LSC
10750	SNP	T		A	A	A	МГР <i>psbM-rpoB</i>	LSC
11132	SNP	T		C	C	C	МГР <i>psbM-rpoB</i>	LSC
11339	SNP	G				C	МГР <i>psbM-rpoB</i>	LSC
11530	SNP	G				T	МГР <i>psbM-rpoB</i>	LSC
11616-17	INDEL	CA				CACA GAAA A	МГР <i>psbM-rpoB</i>	LSC
12015-16	INDEL	AG				AGTA GACT ATTA GG	МГР <i>psbM-rpoB</i>	LSC
12204	SNP	G			A		МГР <i>psbM-rpoB</i>	LSC
12249	SNP	A			C		МГР <i>psbM-rpoB</i>	LSC
12667	SNP	A				G	МГР <i>psbM-rpoB</i>	LSC
12759	SSR	A9				A10	МГР <i>psbM-rpoB</i>	
12883	SNP	T		C	C		МГР <i>psbM-rpoB</i>	LSC
12939-41	INDEL	AGG		-	-		МГР <i>psbM-rpoB</i>	LSC
12984	SSR	T15	T11	T9	T10	T14	МГР <i>psbM-rpoB</i>	LSC
13459	SNP	T				C	<i>rpoB</i> синоним.	LSC
14470	SNP	A				G	<i>rpoB</i> синоним.	LSC
14496	SNP	A		G			<i>rpoB</i> Gln481Arg	LSC
14992	SNP	C				T	<i>rpoB</i> синоним.	LSC
15565	SNP	T		C			<i>rpoB</i> синоним.	LSC
16887	SNP	C	A	A	A	A	<i>rpoC1</i> интрон	LSC

16964	SSR	A8		A9	A9		<i>rpoC1</i> интрон	LSC
17273	SNP	G		A	A		<i>rpoC1</i> интрон	LSC
17424	SSR	G9	G8	G7	G7		<i>rpoC1</i> интрон	LSC
17643	SNP	A				C	<i>rpoC1</i> Met210Leu	LSC
17783	SNP	A		C	C		<i>rpoC1</i> синоним.	LSC
17814	SNP	C				T	<i>rpoC1</i> Arg267Cys	LSC
18307	SNP	G			A		<i>rpoC1</i> Arg431Gln	LSC
18530	SNP	G		T	T	T	<i>rpoC1</i> синоним.	LSC
19839	SNP	C		T	T		<i>rpoC2</i> синоним.	LSC
20660	SNP	T	C				<i>rpoC2</i> Leu490Pro	LSC
20750	SNP	C			T		<i>rpoC2</i> Thr520Ile	LSC
20945	SNP	T			G		<i>rpoC2</i> Leu585Arg	LSC
21054	SNP	C				T	<i>rpoC2</i> синоним.	LSC
22083	SNP	C				A	<i>rpoC2</i> Asp964Glu	LSC
22539	SNP	C				T	<i>rpoC2</i> синоним.	LSC
23220	SNP	C		T			<i>rpoC2</i> синоним.	LSC
23355	SNP	A		G	G	G	МГР <i>rpoC2-rps2</i>	LSC
23377-78	INDEL	CC		ССТТ СТТТ АТСТ ТТАС			МГР <i>rpoC2-rps2</i>	LSC
23432	SNP	G		A	A		МГР <i>rpoC2-rps2</i>	LSC
24138	SNP	T		C			<i>rps2</i> синоним.	LSC
24141	SNP	A	C				<i>rps2</i> Gln178His	LSC
24198	SNP	C		T			<i>rps2</i> синоним.	LSC
24758	SNP	T				C	<i>atpI</i> синоним.	LSC
25296	SSR	T8		T9		T10	МГР <i>atpI-atpH</i>	LSC
25321	SNP	T			C		МГР <i>atpI-atpH</i>	LSC
25442	SNP	T		G	G		МГР <i>atpI-atpH</i>	LSC
25466	SSR	A10	A13	A21	A19	A20	МГР <i>atpI-atpH</i>	LSC
25581	SNP	A				G	МГР <i>atpI-atpH</i>	LSC
25779	SNP	T		C	C	C	МГР <i>atpI-atpH</i>	LSC

25996	SNP	C		G	G	G	МГР <i>atpI-atpH</i>	LSC
27501	SNP	A				C	<i>atpF</i> интрон	LSC
27779-80	INDEL	TC		TCATT AC	TCATT AC		<i>atpF</i> интрон	LSC
28367-69	INDEL	TTA		0			МГР <i>atpF-atpA</i>	LSC
28369-70	INDEL	AT			ATTTT		МГР <i>atpF-atpA</i>	LSC
28372	SNP	A			T		МГР <i>atpF-atpA</i>	LSC
28372-73	INDEL	AT	ATT				МГР <i>atpF-atpA</i>	LSC
28592	SNP	A				G	<i>atpA</i> Ile61Val	LSC
28747	SNP	G				A	<i>atpA</i> синоним.	LSC
29422	SNP	A		G	G	G	<i>atpA</i> синоним.	LSC
29458	SNP	C		A	A	A	<i>atpA</i> синоним.	LSC
29521	SNP	A				T	<i>atpA</i> синоним.	LSC
29701	SNP	G	A				<i>atpA</i> синоним.	LSC
29783	SNP	C		T			<i>atpA</i> синоним.	LSC
30166	SSR	A10	A12				МГР <i>atpA-psbD</i>	LSC
30176	SNP	T		G			МГР <i>atpA-psbD</i>	LSC
30568	SNP	C				T	МГР <i>atpA-psbD</i>	LSC
30751	SNP	T				C	МГР <i>atpA-psbD</i>	LSC
30987-88	INDEL	TA		TAAA A	TAAA A		МГР <i>atpA-psbD</i>	LSC
31310	INDEL	T		-	-		МГР <i>atpA-psbD</i>	LSC
31376	SNP	G		T	T	T	МГР <i>atpA-psbD</i>	LSC
31434	SNP	A		C			МГР <i>atpA-psbD</i>	LSC
31659	SNP	T			C		МГР <i>atpA-psbD</i>	LSC
32066-67	INDEL	AT			АТАС АТАА АААТ ТАГА АТАГ АААГ Т		МГР <i>atpA-psbD</i>	LSC
32092	SNP	C		A	A	A	МГР <i>atpA-psbD</i>	LSC
32158	SNP	C				T	МГР <i>atpA-psbD</i>	LSC
32377	INDEL	G				-	МГР <i>atpA-psbD</i>	LSC

32669	SNP	A		G			<i>psbD</i> синоним.	LSC
34038	SNP	G		A	A		<i>psbC</i> синоним.	LSC
34218	SNP	G		A	A	A	<i>psbC</i> синоним.	LSC
35019	SSR	A9		A16	A16	A8	МГР <i>psbC-psbZ</i>	LSC
35038-42	INDEL	TCAA C		-	-		МГР <i>psbC-psbZ</i>	LSC
35060-61	INDEL	AT				ATT	МГР <i>psbC-psbZ</i>	LSC
35398	SSR	A18	A30	A24	A30	A22	МГР <i>psbC-psbZ</i>	LSC
35495	SNP	T				C	МГР <i>psbC-psbZ</i>	LSC
35885	SSR	A9	A8	A7	A7	A7	МГР <i>psbZ-rps14</i>	LSC
35927	SNP	A				G	МГР <i>psbZ-rps14</i>	LSC
36044-45	INDEL	CA				CATA	МГР <i>psbZ-rps14</i>	LSC
36456	SNP	G		C	C	C	МГР <i>psbZ-rps14</i>	LSC
36937	SNP	T		A	A	A	МГР <i>rps14-psaB</i>	LSC
38902	SNP	A				G	<i>psaB</i> синоним.	LSC
39980	SNP	G	A	A	A	A	<i>psaA</i> синоним.	LSC
40412	SNP	A				G	<i>psaA</i> синоним.	LSC
40748	SNP	C			T		<i>psaA</i> синоним.	LSC
40880	SNP	A				G	<i>psaA</i> синоним.	LSC
41098	SNP	G				T	<i>psaA</i> Leu144Ile	LSC
41604	SNP	A			T		МГР <i>psaA-ycf3</i>	LSC
42375	SNP	T		C	C	C	<i>ycf3</i> синоним.	LSC
42780	SNP	T				G	<i>ycf3</i> интрон	LSC
43668	SNP	T	A	A	A	A	<i>ycf3</i> интрон	LSC
43798	SNP	A		G	G	G	<i>ycf3</i> интрон	LSC
43894	SNP	C				T	<i>ycf3</i> интрон	LSC
44285	SNP	C		A			МГР <i>ycf3-rps4</i>	LSC
44324	SSR	A8		A10	A14	A10	МГР <i>ycf3-rps4</i>	LSC
44539	SNP	G		A			МГР <i>ycf3-rps4</i>	LSC
44686	SNP	A				C	МГР <i>ycf3-rps4</i>	LSC
45331	SNP	T		C	C		МГР <i>ycf3-rps4</i>	LSC

45355	SNP	C				T	МГР <i>ycf3-rps4</i>	LSC
46330	SNP	G		A			МГР <i>rps4-ndhJ</i>	LSC
46716	SNP	A				G	МГР <i>rps4-ndhJ</i>	LSC
46721	SSR	A15		A13	A22	A11	МГР <i>rps4-ndhJ</i>	LSC
46980	SNP	A	G				МГР <i>rps4-ndhJ</i>	LSC
47317	INDEL	T		0			МГР <i>rps4-ndhJ</i>	LSC
47349-59	INDEL	ATTG AAGA AAG		0	0		МГР <i>rps4-ndhJ</i>	LSC
47553	SNP	C				T	МГР <i>rps4-ndhJ</i>	LSC
47687	SNP	G			T		МГР <i>rps4-ndhJ</i>	LSC
47763	SNP	A		G	G	G	МГР <i>rps4-ndhJ</i>	LSC
47781	SNP	A				C	МГР <i>rps4-ndhJ</i>	LSC
47875	INDEL	T		0	0		МГР <i>rps4-ndhJ</i>	LSC
48193	SNP	T				A	МГР <i>rps4-ndhJ</i>	LSC
48363	SNP	A		G	G		<i>ndhJ</i> Val107Ala	LSC
49305	SNP	A		G	G	G	<i>ndhK</i> синоним.	LSC
49678	SNP	A				T	<i>ndhC</i> синоним.	LSC
49986-50004	INDEL	GTTC CTTT GCGG GGCG T				-	МГР <i>ndhC-atpE</i>	LSC
50032-41	INDEL	ATCTA CATA G	-				МГР <i>ndhC-atpE</i>	LSC
50049	SSR	T8			T9		МГР <i>ndhC-atpE</i>	LSC
50095	SNP	C		T	T		МГР <i>ndhC-atpE</i>	LSC
50163	SSR	T10	T12		T15		МГР <i>ndhC-atpE</i>	LSC
50261-62	INDEL	TA		TAAT ATATT GA	TAAT ATATT GA		МГР <i>ndhC-atpE</i>	LSC
50617	INDEL	A		-	-		МГР <i>ndhC-atpE</i>	LSC
50706	SNP	C				A	МГР <i>ndhC-atpE</i>	LSC
50739	SNP	A				C	МГР <i>ndhC-atpE</i>	LSC
50764	SSR	T11	T8	T7	T7	T7	МГР <i>ndhC-atpE</i>	LSC

50770-71	INDEL +SNP	TT				TGATT TTCA	МГР <i>ndhC-atpE</i>	LSC
50775	SNP	G				T	МГР <i>ndhC-atpE</i>	LSC
51294	SNP	C				G	МГР <i>ndhC-atpE</i>	LSC
51778	SSR	T10	T9	T21	T22	T12	МГР <i>ndhC-atpE</i>	LSC
51738	SNP	C		T	T		МГР <i>ndhC-atpE</i>	LSC
51753	SNP	A				C	МГР <i>ndhC-atpE</i>	LSC
52166	SNP	T		C			<i>atpE</i> синоним.	LSC
52330	SNP	C				T	<i>atpB</i> синоним.	LSC
52386	SNP	C		T	T		<i>atpB</i> Gly468Ser	LSC
52717	SNP	A		G	G	G	<i>atpB</i> синоним.	LSC
52734	SNP	A				G	<i>atpB</i> синоним.	LSC
53110	SNP	A			G		<i>atpB</i> синоним.	LSC
54161	SNP	G		T	T	T	МГР <i>atpB-rbcL</i>	LSC
54286	SNP	A	T				МГР <i>atpB-rbcL</i>	LSC
54287	SNP	A		G	G	G	МГР <i>atpB-rbcL</i>	LSC
54313	SSR	A15	A15	A17	A21	A16	МГР <i>atpB-rbcL</i>	LSC
54385	SNP	T				G	МГР <i>atpB-rbcL</i>	LSC
54399	SNP	A				G	МГР <i>atpB-rbcL</i>	LSC
54726	SNP	T		C	C		<i>rbcL</i> синоним.	LSC
55291	SNP	T				A	<i>rbcL</i> Cys247Asn совместно с 55292	LSC
55292	SNP	G				A	<i>rbcL</i> Cys247Asn совместно с 55291	LSC
55323	SNP	T				C	<i>rbcL</i> синоним.	LSC
55393	SNP	T				G	<i>rbcL</i> Ser281Ala	LSC
55396	SNP	C				A	<i>rbcL</i> Gln282Lys	LSC
55534	SNP	T				G	<i>rbcL</i> Ser328Ala	LSC
55675	SNP	C				A	<i>rbcL</i> Leu375Ile совместно с 55677	LSC
55677	SNP	G				A	<i>rbcL</i> Leu375Ile совместно с 55675	LSC
55837	SNP	C				A	<i>rbcL</i> Gln429Lys	LSC
55996	SNP	G				T	<i>rbcL</i> Asp482Tyr	LSC
56024-25	INDEL	AA				AATT GAAT	МГР <i>rbcL-accD</i>	LSC

						TTA		
56245	SNP	C		T	T		МГР <i>rbcL-accD</i>	LSC
56514	SNP	A		G	G		МГР <i>rbcL-accD</i>	LSC
56559	SNP	C				A	МГР <i>rbcL-</i>	LSC
57148	SNP	A				C	<i>accD</i> Glu191Ala	LSC
57593	SNP	A		G	G	G	<i>accD</i> синоним.	LSC
58051	SNP	G		A	A		МГР <i>accD-psaI</i>	LSC
58257	SSR	A7				A8	МГР <i>accD-psaI</i>	LSC
58362	SNP	T				A	МГР <i>accD-psaI</i>	LSC
58406-09	INDEL	CTAA		-	-		МГР <i>accD-psaI</i>	LSC
58509	SNP	C		T			МГР <i>accD-psaI</i>	LSC
58548-49	INDEL	TA		TAG	TAG		МГР <i>accD-psaI</i>	LSC
58639-40	INDEL	CC				ССТТА ТАС	МГР <i>accD-psaI</i>	LSC
58654	SNP	A	G	G	G	G	МГР <i>accD-psaI</i>	LSC
58854-55	INDEL	AT		ATCTC TTTT T	ATCTC TTTT TTTT	ATCTC TTTT TT	МГР <i>psaI-ycf4</i>	LSC
58899-00	INDEL	AA		AAAA TATG GTAT A	AAAA TATG GTAT A	AAAA TATG GTAT A	МГР <i>psaI-ycf4</i>	LSC
58978	SNP	A				C	МГР <i>psaI-ycf4</i>	LSC
59126-27	INDEL	GT		GTT			МГР <i>psaI-ycf4</i>	LSC
59133	SNP	C		T			МГР <i>psaI-ycf4</i>	
59445	SNP	C				T	<i>ycf4</i> синоним.	LSC
60017	SSR	C6	C7	C7			МГР <i>ycf4-cemA</i>	
60063	SNP	T		C	C	C	МГР <i>ycf4-cemA</i>	LSC
60256	SNP	G				T	МГР <i>ycf4-cemA</i>	LSC
60296-97	INDEL	AC		ATC	ATC	ATC	МГР <i>ycf4-cemA</i>	LSC
60337	SNP	G		A	A		МГР <i>ycf4-cemA</i>	LSC
60434	SSR	T6		T7	T8	T9	МГР <i>ycf4-cemA</i>	LSC
60534	SNP	T			C		МГР <i>ycf4-cemA</i>	LSC
61309-	INDEL	GT		GAT	GAT	GAT	МГР <i>cemA-petA</i>	LSC

10								
61429-30	INDEL	TT				TTCCA GTAA ATAA CAGA ATCGT GGAT	MFP <i>cemA-petA</i>	LSC
62567-68	INDEL	CT		CTT			MFP <i>petA-psbJ</i>	LSC
62574	SNP INDEL	G		T	-		MFP <i>petA-psbJ</i>	LSC
62578	SNP	A		T			MFP <i>petA-psbJ</i>	LSC
62683	SSR	T22	T20	T15	T16		MFP <i>petA-psbJ</i>	LSC
62981	SNP	A			C		MFP <i>petA-psbJ</i>	LSC
63061	SNP	A				C	MFP <i>petA-psbJ</i>	LSC
63095	SNP	A		G	G	G	MFP <i>petA-psbJ</i>	LSC
63114	SNP	G		A	A		MFP <i>petA-psbJ</i>	LSC
63123	SNP	T		A	A	A	MFP <i>petA-psbJ</i>	LSC
63199	SNP	C				T	MFP <i>petA-psbJ</i>	LSC
63214-15	INDEL	CA				CTA	MFP <i>petA-psbJ</i>	LSC
63681	SNP	T		A	A	A	MFP <i>psbL-psbF</i>	LSC
64145	SSR	A9	A8	A14			MFP <i>psbE-petL</i>	LSC
64603	SNP	T		C	C	C	MFP <i>psbE-petL</i>	LSC
64668	SNP	C	T	T	T	T	MFP <i>psbE-petL</i>	LSC
64939	SSR	A23			A11		MFP <i>psbE-petL</i>	LSC
64962	INDEL	G			0		MFP <i>psbE-petL</i>	LSC
65016	SNP	G		T	T		MFP <i>psbE-petL</i>	LSC
65328	SNP	C				T	<i>petL</i> Pro2Leu	LSC
65438	SNP	G		T	T	T	MFP <i>petL-petG</i>	LSC
65714	SSR	T16			T21	T13	MFP <i>petG-psaJ</i>	LSC
65930	SNP	C				G	MFP <i>petG-psaJ</i>	LSC
65941	SNP	G		A	A	A	MFP <i>petG-psaJ</i>	LSC
65948	SNP	A		G			MFP <i>petG-psaJ</i>	LSC
66020	SNP	C				T	MFP <i>petG-psaJ</i>	LSC
66305	SNP	G				A	MFP <i>petG-psaJ</i>	LSC

66341	SNP	T		C	C		МГР <i>petG-psaJ</i>	LSC
67224	SNP	G			T		МГР <i>rpl33-rps18</i>	LSC
67257-58	INDEL	TT		TTAT			МГР <i>rpl33-rps18</i>	LSC
67344	SNP	C				T	МГР <i>rpl33-rps18</i>	LSC
67858-67871	INDEL	TTTAT TTTGA CTAC		-	-		МГР <i>rps18-rpl20</i>	LSC
67903-04	INDEL	AT				ATTCT T	МГР <i>rps18-rpl20</i>	LSC
68357	SNP	C		T	T		МГР <i>rpl20-clpP</i>	LSC
68917	SNP	C		T	T		МГР <i>rpl20-clpP</i>	LSC
69192	SNP	C				A	МГР <i>rpl20-clpP</i>	LSC
69743	SSR	A6				A7	<i>clpP</i> интрон	LSC
69936	SSR	T10		T13	T11	T9	<i>clpP</i> интрон	LSC
70545	SSR	A15	A20	A7	A7	A10	<i>clpP</i> интрон	LSC
70560	INDEL	G		-	-		<i>clpP</i> интрон	LSC
70584	SSR	T9	T10	T10			<i>clpP</i> интрон	LSC
71373	SNP	C				A	МГР <i>clpP-psbB</i>	LSC
72744	SNP	G		C	C		<i>psbB</i> синоним.	LSC
73356-62	INDEL	TTTTT AT				0	МГР <i>psbB-psbT</i>	LSC
73363	SNP	T		A	A		МГР <i>psbB-psbT</i>	LSC
73364-65	INDEL	TT		TAGG GTAC CAGA GAAT			МГР <i>psbB-psbT</i>	LSC
73365	SNP	T		A	A		МГР <i>psbB-psbT</i>	LSC
73367	SNP	T				G	МГР <i>psbB-psbT</i>	LSC
73405	SSR	T7		T8	T8	T9	МГР <i>psbB-psbT</i>	LSC
73602-603	INDEL	AA	-	-	-	-	МГР <i>psbT-psbN</i>	LSC
74157	SNP	T	A	A	A	A	МГР <i>psbH-petB</i>	LSC
74221	SNP	C		T			МГР <i>psbH-petB</i>	LSC
74527	SNP	T		C	C	C	<i>petB</i> интрон	
74549	SNP	A				G	<i>petB</i> интрон	LSC
74682	SNP	G				A	<i>petB</i> интрон	LSC

74908	SNP	C		A	A		<i>petB</i> интрон	LSC
75058	SNP	A				G	<i>petB</i> синоним.	LSC
75067	SNP	T				C	<i>petB</i> синоним.	LSC
75820	SNP	A				T	МГР <i>petB-petD</i>	LSC
76370	SNP	A	C				<i>petD</i> интрон	LSC
77205	SNP	G		A			МГР <i>petD-rpoA</i>	LSC
77216	SNP	G		A	A		МГР <i>petD-rpoA</i>	LSC
77234	SSR	A10		A11		A9	МГР <i>petD-rpoA</i>	LSC
77299	SNP	T		C	C		<i>rpoA</i> синоним.	LSC
78307-08	INDEL	TT			TTCTC TTGTT TCTCT TGCT		МГР <i>rpoA-rps11</i>	LSC
78756	SNP	T			C		<i>rps11</i> Lys8Arg	LSC
79110	INDEL	A				-	МГР <i>rpl36-infA</i>	LSC
79380	SNP	C		T	T		МГР <i>infA-rps8</i>	LSC
79664	SNP	G		T	T	T	<i>rps8</i> Pro72Thr	LSC
79820	SNP	T		G	G		<i>rps8</i> синоним.	LSC
79932	SSR	T22	T15	T19	T17		МГР <i>rps8-rpl14</i>	LSC
80456	SSR	T24				T27	МГР <i>rpl14-rpl16</i>	LSC
80718	SNP	G	A				<i>rpl16</i> синоним.	LSC
81057	SNP	G	A	C	A	A	МГР <i>rpl16-rps3</i>	LSC
8123334	INDEL	AT		-			МГР <i>rpl16-rps3</i>	LSC
81347	SNP	G		A	A		МГР <i>rpl16-rps3</i>	LSC
81555	SNP	A	G				МГР <i>rpl16-rps3</i>	LSC
81762	SNP	C				T	МГР <i>rpl16-rps3</i>	LSC
82164	SNP	A			C		МГР <i>rpl16-rps3</i>	LSC
82083	SSR	T8				AT9	МГР <i>rpl16-rps3</i>	LSC
82328	SNP	G				A	<i>rps3</i> синоним.	LSC
82899	SNP	C				T	<i>rpl22</i> A1Thr	LSC
83644	SSR	T8		T9	T9		МГР <i>rps19-rpl2</i>	LSC
89152	SNP	A				G	<i>ycf2</i> Ile1110Val	LSC

92346	SNP	G		A	A		МГР <i>ycf2-ndhB</i>	LSC
92490	SNP	C				T	МГР <i>ycf2-ndhB</i>	IRA
92960	SNP	A		G			МГР <i>ycf2-ndhB</i>	IRA
95380	SNP	A		G	G		<i>ndhB</i> синоним.	IRA
95585	SNP	G		A	A	A	МГР <i>ndhB-rps7</i>	IRA
97727	SNP	T		C	C	C	МГР <i>rps7-rrn16</i>	IRA
98298	SNP	A				C	МГР <i>rps7-rrn16</i>	IRA
98515	SNP	A		T	T		МГР <i>rps7-rrn16</i>	IRA
98838	SNP	C		G	G	G	МГР <i>rps7-rrn16</i>	IRA
101481	SSR	T9				T10	МГР <i>rrn16-rrn23</i>	IRA
101685	INDEL	AG		AGG	AGG	AGAT TTTCG	МГР <i>rrn16-rrn23</i>	IRA
102333	SNP	G		A	A	A	МГР <i>rrn16-rrn23</i>	IRA
106389	SNP	T				A	МГР <i>rrn5-ycf1</i>	IRA
106431	SNP	T		G		G	МГР <i>rrn5-ycf1</i>	IRA
106600	SNP	C				G	МГР <i>rrn5-ycf1</i>	IRA
108400	SNP	C		G	G		<i>ycf1</i> Asn271Lys	IRA
108541	SNP	C	A	A	A	A	<i>ycf1</i> синоним.	IRA
108697	SNP	G	A				<i>ycf1</i> синоним.	IRA
108722	SNP	A				C	<i>ycf1</i> Ile379Leu	SSC
109194	SNP	C			T		<i>ycf1</i> Pro536Leu	SSC
109640	SNP	G		A	A		<i>ycf</i> Asp685Asn	SSC
109766	SNP	T		C	C	C	<i>ycf1</i> синоним.	SSC
109850	SNP	C			G		<i>ycf1</i> Arg755Gly	SSC
110520	SNP	C		G	G		<i>ycf1</i> Thr978Arg	SSC
110568	SNP	C		A	A	A	<i>ycf1</i> Thr994Lys	SSC
110913	SNP	T				C	<i>ycf1</i> Val1109Ala	SSC
110918	SNP	C		A	A	A	<i>ycf1</i> His1111Asn	SSC
111048	SNP	C				A	<i>ycf1</i> Ser1154Tyr	SSC
111059	SNP	C			G		<i>ycf1</i> Pro1158Ala	SSC
111166	SNP	G				A	<i>ycf1</i> синоним.	SSC
111605	SNP	T		G	G	G	<i>ycf1</i> Tyr1340Asp	SSC

111659	SNP	C		T	T	T	<i>ycf1</i> Pro1358Ser	SSC
111983	SNP	A		C	C	C	<i>ycf1</i> Ile1466Leu	SSC
113410	SNP	G				A	МГР <i>rps15-ndhH</i>	SSC
114382	SNP	T		C	C		<i>ndhH</i> Val298Ala	SSC
114585	SNP	C		T			<i>ndhH</i> Pro366Ser	SSC
114863	SNP	C		T			<i>ndhA</i> Pro64Leu	SSC
115489	SNP	G			A		<i>ndhA</i> интрон	SSC
115703	INDEL	C		-	-		<i>ndhA</i> интрон	SSC
115793	INDEL	T		-	-		<i>ndhA</i> интрон	SSC
116179	SNP	C		T	T		<i>ndhA</i> интрон	SSC
116185-89	INDEL	AATA A		-	-		<i>ndhA</i> интрон	SSC
117052	SNP	C				T	<i>ndhI</i> синоним.	SSC
117506	SNP	G			T		МГР <i>ndhI-ndhG</i>	SSC
117529	SNP	G		T	T	T	МГР <i>ndhI-ndhG</i>	SSC
117833	SNP	C	T	T	T	T	<i>ndhG</i> синоним.	SSC
118438	SNP	G	T	T	T	T	МГР <i>ndhG-ndhE</i>	SSC
118566	SNP	T			C		<i>ndhE</i> синоним.	SSC
118728	SNP	A		G	G		<i>ndhE</i> синоним.	SSC
118881-82	INDEL	TC		TCATA AATT AATTC AAAT TAGC	TCATA AATT AATTC AAAT TAGC	TCATA AATT AATTC AAAT TAGC	МГР <i>ndhE-psaC</i>	SSC
119344	SSR	T7		T8	T8		МГР <i>psaC-ndhD</i>	SSC
119363	SNP	T				C	МГР <i>psaC-ndhD</i>	SSC
119377	SNP	T		A		A	МГР <i>psaC-ndhD</i>	SSC
119378	SNP	C		G		G	МГР <i>psaC-ndhD</i>	SSC
119379	SNP	T		A		A	МГР <i>psaC-ndhD</i>	SSC
119751	SNP	T		G			<i>ndhD</i> синоним.	SSC
119835	SNP	A			T		<i>ndhD</i> синоним.	SSC
119952	SNP	C		T	T		<i>ndhD</i> синоним.	SSC
120314	SNP	T		C	C	C	<i>ndhD</i> Leu296Pro	SSC
120937	INDEL	A		-	-	-	МГР <i>ndhD-ccsA</i>	SSC

120979	SSR	A16	A13	A20	A19	A10	МГР <i>ndhD-ccsA</i>	SSC
121092	SNP	T				C	МГР <i>ndhD-ccsA</i>	SSC
121100-01	INDEL	AA				ACA	МГР <i>ndhD-ccsA</i>	SSC
121101	SNP	A				T	МГР <i>ndhD-ccsA</i>	SSC
121102	SNP	G				A	МГР <i>ndhD-ccsA</i>	SSC
121105	SNP	G				C	МГР <i>ndhD-ccsA</i>	SSC
121108	INDEL	T				0	МГР <i>ndhD-ccsA</i>	SSC
121109	SNP	A				C	МГР <i>ndhD-ccsA</i>	SSC
121110	SNP	G				T	МГР <i>ndhD-ccsA</i>	SSC
121119	SNP	G				A	МГР <i>ndhD-ccsA</i>	SSC
121160-67	INDEL	TATTG CAT		-	-	-	МГР <i>ndhD-ccsA</i>	SSC
121676	SNP	A				C	<i>ccsA</i> Asn167Lys	SSC
121775	SNP	C				T	<i>ccsA</i> синоним.	SSC
121813	SNP	G				A	<i>ccsA</i> His122Tyr	SSC
122087	SNP	C		A	A		<i>ccsA</i> Leu30Phe	SSC
122442-63	INDEL	AGTA AAGA TTCAA ATTG ATAA		-	-	-	МГР <i>ccsA-rpl32</i>	SSC
122603	SNP	A		C	C		МГР <i>ccsA-rpl32</i>	SSC
122917-18	INDEL	GA				GAAA	МГР <i>ccsA-rpl32</i>	SSC
122970	SSR	T13	T12	T18	T19	T13	МГР <i>ccsA-rpl32</i>	SSC
122982-83	INDEL	TA				TAGA TTTTA	МГР <i>ccsA-rpl32</i>	SSC
123004	SNP	T	C				МГР <i>ccsA-rpl32</i>	SSC
123388	SNP	A				C	МГР <i>rpl32-ndhF</i>	SSC
123406	SNP	A				T	МГР <i>rpl32-ndhF</i>	SSC
123963	SNP	A		G	G	G	МГР <i>rpl32-ndhF</i>	SSC
124343	SNP	G		A	A		<i>ndhF</i> синоним.	SSC
125027	SNP	A		C	C	C	<i>ndhF</i> синоним.	SSC
125593	SNP	T	C	C	C	C	<i>ndhF</i> Leu477Ser	SSC
125888	SNP	A		T			<i>ndhF</i> Gln573His	SSC

125944	SNP	G				A	<i>ndhF</i> Ser592Asn	SSC
125949	SNP	A		T			<i>ndhF</i> Asn594Tyr	SSC
126087	SNP	G		T	T	T	<i>ndhF</i> Val640Phe	SSC
126194	SNP	C				A	<i>ndhF</i> Phe675Leu	SSC
126449-50	INDEL	CT		CTTTA ATTG AGTCT TTAAT TCTT			МГР <i>ndhF-rrn5</i>	SSC
128035	SNP	G				C	МГР <i>ndhF-rrn5</i>	SSC
128204	SNP	A		C		C	МГР <i>ndhF-rrn5</i>	SSC
128243	SNP	C				T	МГР <i>ndhF-rrn5</i>	IRB
128244	SNP	A				T	МГР <i>ndhF-rrn5</i>	IRB
128246	SNP	A				T	МГР <i>ndhF-rrn5</i>	IRB
132302	SNP	C		T	T	T	МГР <i>rrn23-rrn16</i>	IRB
132944-45	INDEL	TC		TCC	TCC	TCC	МГР <i>rrn23-rrn16</i>	IRB
133146	SSR	A9				A10	МГР <i>rrn23-rrn16</i>	IRB
135797	SNP	G		C	C	C	МГР <i>rrn16-rps7</i>	IRB
136120	SNP	T		A	A		МГР <i>rrn16-rps7</i>	IRB
136337	SNP	T				G	МГР <i>rrn16-rps7</i>	IRB
136908	SNP	A		G	G	G	МГР <i>rrn16-rps7</i>	IRB
139050	SNP	C		T	T	T	МГР <i>rps7-ndhB</i>	IRB
139255	SNP	T		C	C		<i>ndhB</i> синоним.	IRB
141675	SNP	T		C			МГР <i>ndhB-ycf2</i>	IRB
142145	SNP	G				A	МГР <i>ndhB-ycf2</i>	IRB
142289	SNP	C		T	T		МГР <i>ndhB-ycf2</i>	IRB
145483	SNP	T				C	<i>ycf2</i> Ile1110Val	IRB
150984	SSR	A8		A9	A9		МГР <i>rpl2-rps19</i>	IRB

Приложение 2.

Полиморфные сайты митохондриальной ДНК, локализованные у стерильных аналогов линии НА89 подсолнечника с различными типами ЦМС.

Позиция в геноме НА89 ферг.	Тип	НА89 ферг.	НА89 (PET1)	НА89 (PET2)	НА89 (ANN2)	НА89 (MAX1)	Локализация
345	INDEL	T			-		МГР <i>coxII-nad2</i>
697	SNP	A			G	G	МГР <i>coxII-nad2</i>
1477	SNP	A			C	C	МГР <i>coxII-nad2</i>
1769	SNP	A			C		МГР <i>coxII-nad2</i>
3031	SSR	G5		G6	G6	G6	МГР <i>nad2-ccmC</i>
3107	SSR	T5		T6	T4	T4	МГР <i>nad2-ccmC</i>
3114	SNP	A			T	T	МГР <i>nad2-ccmC</i>
3116	SNP	T			A	A	МГР <i>nad2-ccmC</i>
3117	SNP	T			G	G	МГР <i>nad2-ccmC</i>
3124	INDEL	A			-	-	МГР <i>nad2-ccmC</i>
3147	SNP	T			G		МГР <i>nad2-ccmC</i>
3153	SNP	G			T		МГР <i>nad2-ccmC</i>
3175	SNP	G			A		МГР <i>nad2-ccmC</i>
3176	SNP	T			A		МГР <i>nad2-ccmC</i>
3177	SNP	T			A		МГР <i>nad2-ccmC</i>
3178	SNP	T			A		МГР <i>nad2-ccmC</i>
3183	SNP	A			T		МГР <i>nad2-ccmC</i>
3190	SNP	A			C		МГР <i>nad2-ccmC</i>
3200	SNP	A			T	T	МГР <i>nad2-ccmC</i>
3213-14	INDEL	AG			-		МГР <i>nad2-ccmC</i>
3253-95	INDEL	TCTAT CTTTT CATTT TTATA TTTAT ATATT TTGAA AATCA			TATT ATTA AGAA TA		МГР <i>nad2-ccmC</i>

		AGA					
3312	SNP	A			T		MFP <i>nad2-ccmC</i>
3339	SNP	A			T		MFP <i>nad2-ccmC</i>
3360-62	INDEL	AAG			-		MFP <i>nad2-ccmC</i>
3376	SNP	C			A		MFP <i>nad2-ccmC</i>
3377	SNP	T			A		MFP <i>nad2-ccmC</i>
3385-86	INDEL	AT			ATCT	ATCT	MFP <i>nad2-ccmC</i>
3393	SNP	T			G	G	MFP <i>nad2-ccmC</i>
3394	SNP	G			T	T	MFP <i>nad2-ccmC</i>
3398	INDEL	G			-	-	MFP <i>nad2-ccmC</i>
3400-02	INDEL	CGA			CAGA A	CAGA A	MFP <i>nad2-ccmC</i>
3404	SNP	A			C	C	MFP <i>nad2-ccmC</i>
3408	SNP/ NDEL	T			G	-	MFP <i>nad2-ccmC</i>
3409	SNP	A			G	G	MFP <i>nad2-ccmC</i>
3605	SNP	C			T		MFP <i>nad2-ccmC</i>
3865-68	INDEL	TTTA			-		MFP <i>nad2-ccmC</i>
3884	SNP	T			G		MFP <i>nad2-ccmC</i>
3993	SNP	C			T		MFP <i>nad2-ccmC</i>
4043	SNP	C			G		MFP <i>nad2-ccmC</i>
4337	SNP	G			C		MFP <i>nad2-ccmC</i>
4422	SNP	G			C		MFP <i>nad2-ccmC</i>
5012	SNP	G			A		MFP <i>nad2-ccmC</i>
6177-6243	INDEL	AAGAA TAATG AAAAA CAAAA ACAAA AAAAC AAAAA AAACA AAAAG AAAGA AAAAA AGAAA			AACA AAAA AGAA		MFP <i>nad2-ccmC</i>

		TAAAG AA					
7267	SNP	G			A		MFP <i>nad2-ccmC</i>
7404	SSR	G10		G9	G8	G8	MFP <i>nad2-ccmC</i>
8023	SNP	G			T		MFP <i>nad2-ccmC</i>
8338	SNP	C			A		MFP <i>nad2-ccmC</i>
8769	SNP	C			G		MFP <i>nad2-ccmC</i>
9304	SNP	A			G	G	MFP <i>nad2-ccmC</i>
9335	SNP	G			T		MFP <i>nad2-ccmC</i>
10467	SNP	A		C	C	C	MFP <i>nad2-ccmC</i>
10823	SNP	C			G	G	MFP <i>nad2-ccmC</i>
10996-99	INDEL	TTTC			-		MFP <i>nad2-ccmC</i>
11005	SNP	T			C		MFP <i>nad2-ccmC</i>
11005	SNP	C			T		MFP <i>nad2-ccmC</i>
11248-49	INDEL	CT			-	-	MFP <i>nad2-ccmC</i>
11850	SNP	G			T	T	MFP <i>nad2-ccmC</i>
11976	SNP	C			G		MFP <i>nad2-ccmC</i>
18258	SNP	A			G		MFP <i>ccmC-atp4</i>
19594	SNP	G		A	A	A	MFP <i>ccmC-atp4</i>
20527	SNP	A			G		MFP <i>ccmC-atp4</i>
21892	SNP	A			G		MFP <i>ccmC-atp4</i>
23376	SNP	T			C		MFP <i>ccmC-atp4</i>
23595	SNP	G			A		MFP <i>ccmC-atp4</i>
23887	SNP	T			G		MFP <i>ccmC-atp4</i>
23917	SNP	G		T	T	T	MFP <i>ccmC-atp4</i>
23918	SNP	A			C	C	MFP <i>ccmC-atp4</i>
25987	SNP	T			G		MFP <i>ccmC-atp4</i>
29048	SNP	C			T		MFP <i>nad4L-orf777</i>
29049	SNP	C			T		MFP <i>nad4L-orf777</i>
29068	SNP	G			T		MFP <i>nad4L-orf777</i>
29121-24	INDEL	AATC			-		MFP <i>nad4L-orf777</i>

29128	SNP	T			G		МГР <i>nad4L-orf777</i>
29129	SNP	G			A		МГР <i>nad4L-orf777</i>
29130	SNP	T			C		МГР <i>nad4L-orf777</i>
29140	SNP	T			G	G	МГР <i>nad4L-orf777</i>
29149	SNP	G			T		МГР <i>nad4L-orf777</i>
29151	SNP	T			C		МГР <i>nad4L-orf777</i>
29153	SNP	G			T		МГР <i>nad4L-orf777</i>
33779	SNP	C			G		МГР <i>nad4L-orf777</i>
33823	SNP	C			G		МГР <i>nad4L-orf777</i>
34181	SNP	G			A		МГР <i>nad4L-orf777</i>
34333-34	INDEL	CG			CCGG		МГР <i>nad4L-orf777</i>
34337	SNP	A			G		МГР <i>nad4L-orf777</i>
34750	SNP	C			T		МГР <i>nad4L-orf777</i>
35031	SNP	C		A	A		МГР <i>nad4L-orf777</i>
35114	SNP	C		A	A		МГР <i>nad4L-orf777</i>
35478	SNP	T		C	C	C	МГР <i>nad4L-orf777</i>
35507-15	INDEL	TCCGG CAAG			TTCTT TCCG AAAG		МГР <i>nad4L-orf777</i>
35522	SNP	C			T		МГР <i>nad4L-orf777</i>
35523	SNP	T			G		МГР <i>nad4L-orf777</i>
35525	SNP	A			T		МГР <i>nad4L-orf777</i>
35531	SNP	C			A		МГР <i>nad4L-orf777</i>
35542	SNP	C			G		МГР <i>nad4L-orf777</i>
35550	SNP	C			T		МГР <i>nad4L-orf777</i>
36947	SNP	A			G		<i>atp8 Asp59Gly</i>
37409	SNP	G			A		<i>МГР atp8-coxIII</i>
38134	SNP	C			A		<i>coxIII синоним.</i>
39589	INDEL	GA			GAA	GAA	МГР <i>coxIII-rpl5</i>
39787	SNP	G			A		МГР <i>coxIII-rpl5</i>
41238-39	INDEL	GG			GTTG		МГР <i>coxIII-rpl5</i>

41243-44	INDEL	AA			-		МГР <i>coxIII-rpl5</i>
41462	SSR	T7			T8		МГР <i>coxIII-rpl5</i>
41746	SNP	G			T	T	МГР <i>coxIII-rpl5</i>
41747	SNP	A			C		МГР <i>coxIII-rpl5</i>
43358	SNP	G			T		<i>rpl5</i> синоним.
43533	SNP	G			A		МГР <i>rpl5-nad4</i>
44995	SNP	A			C		МГР <i>rpl5-nad4</i>
47939	SNP	T			G		<i>nad4</i> интрон
47940	SNP	C			A		<i>nad4</i> интрон
47954	SNP	T			G		<i>nad4</i> интрон
47955	SNP	G			T		<i>nad4</i> интрон
49509	SNP	C			T	T	МГР <i>nad4-ccmB</i>
49779-80	INDEL	TA			TGA		МГР <i>nad4-ccmB</i>
49796	SNP	G			A		МГР <i>nad4-ccmB</i>
50856	SNP	T			G	G	МГР <i>nad4-ccmB</i>
50857	SNP	C		A	A	A	МГР <i>nad4-ccmB</i>
50987	SNP	G			T		МГР <i>nad4-ccmB</i>
51679	SSR	G10	G9	G9	G8	G9	МГР <i>nad4-ccmB</i>
51702	SNP	A			C		МГР <i>nad4-ccmB</i>
52002	SNP	T			C	C	МГР <i>nad4-ccmB</i>
56443	SSR	A6			A5		МГР <i>nad4-ccmB</i>
56444	SNP	A			C		МГР <i>nad4-ccmB</i>
56476-77	INDEL	AT			-		МГР <i>nad4-ccmB</i>
56482	SNP	T			G		МГР <i>nad4-ccmB</i>
56483	SNP	C			A		МГР <i>nad4-ccmB</i>
56550	SSR	T5			T6		МГР <i>nad4-ccmB</i>
57336	SNP	G			T	T	МГР <i>nad4-ccmB</i>
57338	SNP	T			G	G	МГР <i>nad4-ccmB</i>
57339	SNP	C			A	A	МГР <i>nad4-ccmB</i>
60934	SNP	G			A		МГР <i>nad4-ccmB</i>
62299	SNP	T			C		МГР <i>nad4-ccmB</i>

62357	SNP	T			C		<i>MFP nad4-ccmB</i>
62361	SNP	T		G	G	G	<i>MFP nad4-ccmB</i>
62388	SNP	C			T		<i>MFP nad4-ccmB</i>
62394	SNP	G			A		<i>MFP nad4-ccmB</i>
62404	SNP	G		A	A	A	<i>MFP nad4-ccmB</i>
63464-65	INDEL	TC			TCC	TCC	<i>MFP nad4-ccmB</i>
64351	SNP	G			T		<i>MFP nad4-ccmB</i>
67110	SNP	C			T		<i>ccmB</i> синоним.
69004	SNP	C			T	T	<i>MFP rpl10-nad1</i>
70130	SNP	A			C	C	<i>MFP rpl10-nad1</i>
71087	SNP	G			C		<i>MFP rpl10-nad1</i>
71460	SNP	G			C		<i>MFP rpl10-nad1</i>
71498-99	INDEL	GT		GGGG CT	GGGG CT	GGGG CT	<i>MFP rpl10-nad1</i>
71529	SNP	T			C		<i>MFP rpl10-nad1</i>
72511	SNP	G			A		<i>MFP rpl10-nad1</i>
72939	SNP	G			A		<i>MFP rpl10-nad1</i>
73615	SNP	C			T		<i>MFP rpl10-nad1</i>
77437	SNP	C			T		<i>MFP rpl10-nad1</i>
77796	SNP	C			G		<i>MFP rpl10-nad1</i>
77915	SNP	G			T		<i>MFP rpl10-nad1</i>
81971-72	INDEL	GA			GACA		<i>MFP rpl10-nad1</i>
83865	SNP	C			T		<i>MFP rpl10-nad1</i>
84472	SNP	A			G	G	<i>MFP rpl10-nad1</i>
85178	SSR	T8			T9		<i>MFP rpl10-nad1</i>
85594	SNP	C			T		<i>MFP rpl10-nad1</i>
86188	SSR	T7			T6	T6	<i>MFP rpl10-nad1</i>
88615	SNP	G			C		<i>MFP rpl10-nad1</i>
88616	SNP	C			T		<i>MFP rpl10-nad1</i>
89633	SNP	C			G		<i>MFP rpl10-nad1</i>
90169	SNP	A			C	C	<i>MFP rpl10-nad1</i>

90173	SNP	A			C		<i>МГР rpl10-nad1</i>
91106	SNP	G		T	T	T	<i>МГР rpl10-nad1</i>
91107	SNP	A		C	C	C	<i>МГР rpl10-nad1</i>
96417	SNP	T			C		<i>МГР nad1-cox1</i>
98340	SNP	T			C		<i>МГР nad1-cox1</i>
99127	SNP	T			C		<i>МГР nad1-cox1</i>
102190	SNP	G			C		<i>МГР nad1-cox1</i>
103204	SNP	G			C		<i>МГР nad1-cox1</i>
104337	SNP	T			G		<i>МГР nad1-cox1</i>
105195	SNP	A			G		<i>МГР nad1-cox1</i>
107634	SNP	T			G		<i>cox1</i> синоним.
109233	SNP	C			T		<i>МГР rps11-nad5</i>
111296-97	INDEL	TG			TGG		<i>МГР rps11-nad5</i>
113470	SNP	A			G	G	<i>МГР nad5-atp9</i>
113934	SNP	C			A		<i>МГР nad5-atp9</i>
115237	SNP	G			A		<i>МГР atp9-rps4</i>
117811	SNP	A			G		<i>МГР atp9-rps4</i>
119108	SNP	C			A	A	<i>МГР atp9-rps4</i>
121109-10	INDEL	CC		CTTC	CTTC	CTTC	<i>МГР atp9-rps4</i>
121973	SNP	C			T		<i>МГР atp9-rps4</i>
122615	SNP	A			G		<i>rps4</i> Lys167Arg
123565	SNP	C			T		<i>МГР rps4-rrn26</i>
125075	SNP	C			T		<i>МГР rps4-rrn26</i>
125854	SNP	A			G		<i>МГР rps4-rrn26</i>
127017	SNP	T			C		<i>МГР rps4-rrn26</i>
128372	SNP	C			T	T	<i>МГР rps4-rrn26</i>
130827	SNP	C			T		<i>rrn26</i>
131787	SNP	G			C		<i>rrn26</i>
132832	SNP	T			G		<i>МГР rrn26-rrn5</i>
143537	SNP	G			T		<i>МГР rrn18-rps13</i>
145449	SNP	T			C		<i>МГР rrn18-rps13</i>

149518	SNP	T			G		МГР <i>rps13-nad6</i>
149893	SSR	T10			T9		МГР <i>rps13-nad6</i>
152989	SSR	T9			T10		МГР <i>rps13-nad6</i>
157009	SNP	T			C		МГР <i>rps13-nad6</i>
160332	SNP	T			G		МГР <i>rps13-nad6</i>
160333	SNP	T			A		МГР <i>rps13-nad6</i>
160344	SNP	T			C		МГР <i>rps13-nad6</i>
160347-48	INDEL	TG			TAAG		МГР <i>rps13-nad6</i>
160349	SNP	T			A		МГР <i>rps13-nad6</i>
160350	SNP	A			G		МГР <i>rps13-nad6</i>
160353	SNP	G			T		МГР <i>rps13-nad6</i>
160366	SNP	C			A		МГР <i>rps13-nad6</i>
160372	SNP	G			T		МГР <i>rps13-nad6</i>
161379	SNP	T			G		МГР <i>rps13-nad6</i>
170187	SSR	T14	T12	T12	T8	T7	МГР <i>nad6-ymf16</i>
170562	SNP	G			T		МГР <i>nad6-ymf16</i>
174932	SNP	C			T	T	МГР <i>nad6-ymf16</i>
177554	SNP	G			T	T	МГР <i>nad6-ymf16</i>
178408	SSR	T9	T8		T8	T8	<i>ymf16</i> интрон
178952	SNP	A			C		<i>ymf16</i> Met161Leu
179735	SSR	A10			A9	A9	МГР <i>ymf16-cob</i>
181673	SNP	T			G		МГР <i>ymf16-cob</i>
188296	SNP	G			T	T	<i>cob</i> синоним.
188443	SNP	T			G		<i>cob</i> синоним.
188452	SNP	G			A		<i>cob</i> синоним.
189084	SNP	A			G		<i>cob</i> синоним.
189619	SNP	G			A		<i>cob</i> Glu427Lys
195010	SNP	G		T	T	T	МГР <i>cob-ccmFc</i>
195017	SNP	C		A	A		МГР <i>cob-ccmFc</i>
196318	SNP	G			T	T	МГР <i>cob-ccmFc</i>
197773	SNP	T			C		МГР <i>cob-ccmFc</i>

198038-44	INDEL	CATC TC			-		МГР <i>cob-ccmFc</i>
200176	SNP	G		A	A	A	<i>ccmfC</i> (интрон)
200517	SNP	G		A	A	A	<i>ccmfC</i> (интрон)
202282	SNP	G			C	C	<i>orf873</i> Ala129Gly
202308	SNP	C			T	T	<i>orf873</i> синоним.
202311	SNP	T			C	C	<i>orf873</i> синоним.
202324	SNP	G			T	T	<i>orf873</i> Ser115Glu
202325	SNP	A			C	C	<i>orf873</i> Ser115Glu
202332	SNP	C			T	T	<i>orf873</i> Glu112Gly
202333	SNP	T			C	C	<i>orf873</i> Glu112Gly
202350	SNP	T			C	C	<i>orf873</i> синоним.
202353	SNP	T			C	C	<i>orf873</i> синоним.
202380	SNP	C			G	G	<i>orf873</i> Trp96Cys
202441	SNP	C			G	G	<i>orf873</i> Cys76Ala
202442	SNP	A			C	C	<i>orf873</i> Cys76Ala
202445	SNP	A			G	G	<i>orf873</i> синоним.
202446	SNP	C			T	T	<i>orf873</i> синоним.
202455	SNP	G			T	T	<i>orf873</i> синоним.
202470-71	INDEL	GA			GAAG AAA		<i>orf873</i> сдвиг рамки считывания
202520	SNP	T			C	C	<i>orf873</i> Ile50Val
202526	SNP	G			T	T	<i>orf873</i> Leu48Ile
202538	SNP	T			A	A	<i>orf873</i> Ile44Leu
202539	SNP	T			C	C	<i>orf873</i> Ile44Leu
202548	SNP	T			G	G	<i>orf873</i> Leu40Phe
202554	SNP	C			A	A	<i>orf873</i> Leu38Phe
202583	SNP	T			G	G	<i>orf873</i> Thr29Pro
202605	SNP	C			A	A	<i>orf873</i> Met21Ile
202616	SNP	T			G	G	<i>orf873</i> Ile18Leu
202620	SNP	C			T	T	<i>orf873</i> синоним.
202628	SNP	T			C	C	<i>orf873</i> Trh14Ala

204848-49	INDEL	AA		ATA	ATA	ATA	МГР <i>atp1-ccmFn</i>
208112	SNP	T			G		МГР <i>atp1-ccmFn</i>
208113	SNP	C			A		МГР <i>atp1-ccmFn</i>
212584	SNP	C			T		МГР <i>atp1-ccmFn</i>
213951	SNP	C			T		<i>ccmFn</i> синоним.
214504	SNP	C			A		<i>ccmFn</i> Arg193Leu
214655	SNP	C			T		<i>ccmFn</i> Ala143Thr
214817	SNP	T			C		<i>ccmFn</i> Ile91Val
215173	SNP	T			C		МГР <i>ccmFn-rps3</i>
216014	SNP	T			C		МГР <i>ccmFn-rps3</i>
216356	SNP	G			A		МГР <i>ccmFn-rps3</i>
217162	SNP	G			A		МГР <i>ccmFn-rps3</i>
217235	SNP	G			C	C	МГР <i>ccmFn-rps3</i>
217356	SNP	G			C	C	МГР <i>ccmFn-rps3</i>
217931	SNP	C			A		МГР <i>ccmFn-rps3</i>
218706	SNP	G			A		МГР <i>ccmFn-rps3</i>
219383	SNP	G			C		МГР <i>ccmFn-rps3</i>
220996	SNP	C			T		МГР <i>ccmFn-rps3</i>
221032	SNP	C			T		МГР <i>ccmFn-rps3</i>
221286	SNP	A			T		МГР <i>ccmFn-rps3</i>
221468	SNP	G			T		МГР <i>ccmFn-rps3</i>
221998	SNP	C			G		МГР <i>ccmFn-rps3</i>
222555	SNP	C			T		МГР <i>ccmFn-rps3</i>
223717	SNP	G			A	A	МГР <i>ccmFn-rps3</i>
223834	SNP	C			G		МГР <i>ccmFn-rps3</i>
223928-29	INDEL	AA			AATA		МГР <i>ccmFn-rps3</i>
223993	SNP	G			C	C	МГР <i>ccmFn-rps3</i>
226257-58	INDEL	CA			CAA	CAA	МГР <i>ccmFn-rps3</i>
226661	SNP	G			T		МГР <i>ccmFn-rps3</i>
226663	SNP	G			C		МГР <i>ccmFn-rps3</i>
226665	SNP	A			C		МГР <i>ccmFn-rps3</i>

227086	SNP	A			G	G	МГР <i>ccmFn-rps3</i>
227123	SNP	T			G		МГР <i>ccmFn-rps3</i>
227124	SNP	C			A		МГР <i>ccmFn-rps3</i>
228443	SNP	C			T		<i>rps3</i> Ala3Val
229032	SNP	G			A		<i>rps3</i> синоним.
233878-79	INDEL	AA			ATA		МГР <i>rpl16-matR</i>
235275	SNP	C			A		МГР <i>rpl16-matR</i>
239196	SNP	G			A		МГР <i>rpl16-matR</i>
239403	SNP	G			C		МГР <i>rpl16-matR</i>
239446	SNP	A			C		МГР <i>rpl16-matR</i>
239882	SNP	G		A	A	A	МГР <i>rpl16-matR</i>
241037	SNP	G		A	A	A	МГР <i>rpl16-matR</i>
241477	SNP	A		C	C	C	МГР <i>rpl16-matR</i>
246055	SNP	C		T	T	T	МГР <i>rpl16-matR</i>
246869	SNP	A			C	C	МГР <i>rpl16-matR</i>
247578	SNP	A			G		МГР <i>rpl16-matR</i>
248268	SSR	A14	A10	A9	A10	A10	МГР <i>rpl16-matR</i>
249363	SNP	C		A	A	A	МГР <i>rpl16-matR</i>
250521	SNP	C			T		<i>matR</i> синоним.
251046	SNP	G			C		<i>matR</i> синоним.
251052	SNP	C			T		<i>matR</i> синоним.
251215	SNP	G			C		<i>matR</i> Ala227Gly
256986	SNP	G			C		МГР <i>rps12-nad9</i>
259018	SNP	C			A		МГР <i>rps12-nad9</i>
259715	SNP	G			T		МГР <i>rps12-nad9</i>
260997	SSR	T7			T8		МГР <i>nad9-atp6</i>
261328	SNP	C			G		МГР <i>nad9-atp6</i>
261909	SSR	A8			A7		МГР <i>nad9-atp6</i>
262082	SNP	G		A	A	A	МГР <i>nad9-atp6</i>
263441	SNP	A			C		МГР <i>nad9-atp6</i>
266518	SNP	T			C		МГР <i>nad9-atp6</i>

266956	SNP	G			C		МГР <i>nad9-atp6</i>
268498	SNP	G			C		МГР <i>nad9-atp6</i>
270678	SNP	G		T	T	T	МГР <i>atp6-coxII</i>
271164	SNP	G			A		МГР <i>atp6-coxII</i>
273346	SNP	C		A	A		МГР <i>atp6-coxII</i>
275541	SNP	A			T		МГР <i>atp6-coxII</i>
277624-33	INDEL	ATACT CACGA			-		МГР <i>atp6-coxII</i>
277852	SNP	G			A		МГР <i>atp6-coxII</i>
281255	SNP	C			G		МГР <i>atp6-coxII</i>
282154	SNP	C			T		<i>coxII</i> (интрон)
282185	SNP	G			A		<i>coxII</i> (интрон)
282429	INDEL	A			-		<i>coxII</i> (интрон)
282542	SNP	T			C		<i>coxII</i> (интрон)
283593	SNP	G			A		<i>coxII</i> (интрон)
284113	SNP	C			G		<i>coxII</i> (интрон)
285912	SNP	G			A		МГР <i>coxII-nad2</i>
286666	SNP	T			G		МГР <i>coxII-nad2</i>
287286-89	INDEL	AGGC			-		МГР <i>coxII-nad2</i>

Приложение 3.

Открытые рамки считывания, локализованные у стерильных аналогов линии HA89 подсолнечника с различными типами ЦМС.

ОРС, линии	Кол-во ТМД* в кодируемом белке	Аминокислотная последовательность, кодируемая ОРС
<i>orf228</i> HA89(PET2)	1	MAATIASAGAAIGIGNVLSSSIHVARNP SLAKQSFQYAILGF ALTEAIASFAPMMAFLISSVFRSKNQRKKVEVS
<i>orf285</i> HA89(PET2)	1	MKKKKREENDQLEMLEGALITLIDNIFVKFLLCLLLILVSFLI YTYDRSFRVHQQTLLWAHQHNVTPGVSFLYKIRVGHDTT LDPIPLPVKEHC
<i>orf306</i> HA89(PET1) HA89(MAX1)	0	MTRKDRSAFLSFHQERHLWTSASTSSPSRVGIGTMDQDFWF PFDQFLRSLLLLEGYGF LSPGYRFIGSATLSIKAITPLIHTSA PLTPERKLAHLSDWIRS
<i>orf324</i> HA89(ANN2)	0	MKDQLIEGKSESIKMQLYEHRENLPREKARKRPNRHTNSL SRLPNRLAKGGSSYLYESRTRFAAEWNEIHV LGGTSPWHV GVRKHSHFIDSLSPFLIGLGQSFSP
<i>orf327</i> HA89(ANN2)	0	MKSYGHIFILRAIELSLASIKPSNYRLYNKLLRAIFSSPNTFNY LTQCRRLVLPFGNSAGGKGLSVSSKFI RPLRMVTGQDWDS QQLSFFSSNPYPPQGVPGGLSTSI
<i>orf345</i> HA89(ANN2)	0	MCLRPLFLVYDYPCSSSDVDIFTWLK VEMCDWLRPGNSFPS RNSSRGVKSGGRSGCRGHSISGKEDGLDSL FMAWAKAAL MTFSEDMLIRRSQSQSLCSYAFTLSRPSK KDVHN
<i>orf480</i> HA89(MAX1)	0	MLFRLTARSLFLSFSFFMLVGRSVFMEQITPYKGRSVSGPS SQKNLPLPGSGDDDPDKRKKVPVSKDTANA AVSLLRQVILE ILARARDPSLREGLHNPTT
<i>orfH522</i> HA89(PET1)	1	MPQLDKFTYFTQFFW SCLFDLRRTLFFCLFIFLILVARSVFM EQITPYKGRSISGPN AQSIPLPGSGDDPNRKKVPVSKDTA TAAVSLLRQLVLEILARA
<i>orf558</i> HA89(ANN2)	1	MIVLEWLFLTIAPCDAAEPRQLGSQDAATPMMQGIIDFLHPP LLIILVFVSGVLLMERRGPHRIASSSSLSVFAGGAGEDPNR GRRRTPVTRETITTTGISLLRQLILEILV DTPDQALREGLNPNST QAWNRALESALTERFGHSRYTWGHLWNIVNMVLELSEQGE RSPFFLRVIALVRARGT
<i>orf645</i> HA89(PET2) HA89(MAX1)	0	MARKGNPISVRLDLNRSSDSSRFVESTIHASIFFILLSLTFQNQ LKKGGISLPVVKHGLLLIFILLLSIFLMNELYN AVVPFLAKSG DSGINPLPGPSDQNSIFPLIEEGEVPSER/QPKINRDPEFCISF QNACDLESEISEIMETLLHEKGVVDLSQREIKREVEGF LSNV WDMPIPRNRRLTAILKSLRSRGVNSRNFKDFEKHIKNIPVL
<i>orf891</i> HA89(ANN2)	3	MILSVLSSLALVSGLMVVRANKNPVHSV LFPPIPVFRNTSGLLL LLGLDFFAMIFPVVHIGAIAVSFLFVMMMFHIQIAEIH EEVLR YLPVSGIIGLIFWWEMFFILDNESIPLLP TQRNTTSLRYMVYA GKVRSWTNLETGNLLYTYYSVWFLVPSLILLVAMIGAI VL TMHRTTKVKRQDVFRNAIDSRRTIMRGM TDLLKSSLILV RIQFVRWPSWSYRCLDTLLFFSHFRMTVPFHF WYNFKPGPA IRCISRTPGIRRLLEYHGIEYKDFKESLSYNLNLESFLLHGN SPN

<p><i>orf933</i> HA89(ANN2)</p>	<p>2</p>	<p>MRFLFSFLPERFRVLLNAIHNQGSSINDRPPGS VPSSSSVRIPL NTSVIVPPTRVGPSTVPLTNEEDLPLLESELRRRLQLDVPLSNES LRNAQSALTIADSVPENQENVMGIRYRHDPGRTSFLNFefe MTNELIRKLTNICNSIGIKSSSVPPYIPDLFVYSRFGDGLRSL VHETYDKVLVEILSFYKGRFGRFIIIVSLGVSCVWYTFVPG NPDIALLSDLFPRFETYQSFLNPDYTNKFFHKICYIERSYITGE VLKTIENEFPFSELNIPEESGGTRVAVGLGLMIGVFLAMGIVP</p>
<p><i>orf1197</i> HA89(ANN2)</p>	<p>7</p>	<p>MINYVLQSMKVTASSSSSGISFPFVLEIVGTPLPGDSFTNRLT ASTVSIPGYIVDPKISSLTHFNRLVLDLYDFLEGRMYPKYI FSLQQQLDNTPIESLGKLDSDMRQDEL TGLSDYLQSLGYS QAYLEYLKICDERYRDMVVPSPLEQFSILPLIPMKIENLYFSF TNSSLFMLLTLVLLLIHFVTKKGGGNLVPNAWQSLVELIY DFVLNLVNEQIGGLSGNVKQKFFPCILVTF TFLFCNLQGM PYSFTVTSHFLITLGLSFSIFIGITIVGFQRNGLHFLSFLLPAGV PLPLAPFLVLELISYCFRALS LGIRLFANMMAGHSLVKILSG FAWTMLCMNDLLYFIGDLGPLFIVLALTGLELGVAILQAYV FTILICIYLNDAINLH</p>
<p><i>orf1287</i> HA89(MAX1)</p>	<p>7</p>	<p>MPNNSHAFLGRTKPTISDKSFLIFRARS GTTRKKIFTMTMKD FIQRFKRTVSPVLNESEEVP IQFGLSMDEIVEANLEKFTLYLE GDSVTPASIEALTKLNYLVFMRHDLEGT VKPAAIRSLQKE LNKTPPESVVPKLESIYQNELKSLDNFLEPF EVNLSSKDFLN YCDEFYRSSSPLEQFDILPLIPMNIGNLYFSFTNSSLFMLLTL LVLLLIHFVTKKGGGNLVPNAWQSLVELIYDFVLNLVNEQI GGLSGNVKQKFFPCILVTF TFLFCNLQGMIPYSFTVTSHFLI TLGLSFSIFIGITIVGFQRNGLHFLSFLLPAGVPLPLAPFLVLL ELISYCFRALS LGIRLFANMMAGHSLVKILSGFAWTMLCMN DLLYFIGDLGPLFIVLALTGLELGVAILQAYVFTILICIYLNDA INLH</p>
<p><i>orf2565</i> HA89(PET2)</p>	<p>0</p>	<p>MNQSFGEIITYTLGSAIPLTSNDGALLPKNEIYARVKEAFLK NAELYNGSCLVQVIIRAYMDQVEKQDRPELKVS DRYQELLS IYQTELGELEAITARKIQHSKGKYKDYITALKKASGGRAF MVS DLETILIDNKHRPYAAGLMLVHPGKDVKESLIYTYFSE DYSRYIESFEERSKKVLFDLV NKIIALVKIDRN AKIVYFHNFS RFDGVLVLKHLVCHHDYK LKPLFRNRLYELSVYSGRKL FRLRDSLNLPGTLNNAKSLCPSLGSKGS LDYQDVRDLNL VSKKDELIEYMKQDILLGGIMQKAQEYFHL YQLDIVSKIT LSSLALSIYRLKY YDEENWPYIPNMNQDNFIRKAYYGHT DTYKPYGEDLYYD VNSLYPFVMKNYQMPGGKPVWHGN LDEKDLDSLYGFIEAYVVC PKTIKPFYPYRNKNNTLTFPTG EFVGVYYSEELKFARDLG YTVLPLSGYL YERMDSPFIEFVN TQSEKRIEAKKAGNEALS FVYKILMNSLYGRFGINPNSTTSEI CDH DRYRTLFKKDSFIYGS LLEKNKYIVSYHVNTGNTPETW NPPKNGAVQLAAAITACARIYMPYISREECYYTDTDSVVL GQPLPDELISSSVLGMLKLEARIVKGYFLAPKSYGFIEK DAD GKIVLKHKGAAKSLVTLEWFQS QYDDPSRKEQVS VTSNFKI DWKDLEIQKQESLYKLGLSMDSKRLPVYNS ENTWIDTEPIHI RDLSNVSPLLGDRILFYLRNEMNRLQTKSEILSEKLSQRDRE MIGIIEK DREISEMKS KIDHLQDEM KKFTEKTTKVVTKIKV VKSDNKTKTKVVKKTERKNQTDKGKPP</p>

* ТМД – трансмембранный домен