

САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

БИОЛОГИЧЕСКИЙ ФАКУЛЬТЕТ
КАФЕДРА ЦИТОЛОГИИ И ГИСТОЛОГИИ

Золотарев Андрей Владимирович

Выпускная квалификационная работа бакалавра

Сравнительная генетическая характеристика
апомиктичных линий растений из рода
Boechera

Научный руководитель:

д.б.н. Родионов А. В.

Научный консультант:

к.б.н., Брюхин В. Б.

Центр геномной биоинформатики

им. Ф. Г. Добржанского

Санкт-Петербург
2019

Оглавление

Введение.....	3
Глава 1. Обзор литературы	5
1.1 Гаметофитный апомиксис и его роль для интенсификации сельского хозяйства.....	5
1.2 Таксономия и среда обитания наиболее важных представителей рода <i>Boechea</i>	7
1.3 Преимущества рода <i>Boechea</i> для изучения апомиксиса	10
1.4 Цитоэмбриологические исследования в роде <i>Boechea</i>	11
1.5 Популяционно-генетические исследования апомиксиса в роде <i>Boechea</i>	16
1.6 Наследование и генетические аспекты апомиксиса в роду <i>Boechea</i>	19
1.7 Геномные ресурсы для рода <i>Boechea</i> и проблемы геномного анализа	25
Глава 2. Материалы и методы исследования	29
2.1 Материал	29
2.2 Методы работы с данными	29
2.2.1 Предобработка данных	29
2.2.2 Примененные стратегии сборки генома	30
2.2.3 Удаление контаминаций из данных Illumina	32
2.2.4 Подготовка к сборке хлоропластной ДНК.....	33
2.2.5 Сборка и аннотация хлоропластной ДНК	33
2.2.6 Методы филогенетического анализа с использованием хлоропластной ДНК.....	34
2.2.7 Методы филогенетического анализа с использованием последовательности гена APOLLO	35
Глава 3. Результаты и обсуждение	37
3.1 Анализ и коррекция исходных прочтений	37
3.2 Выбор наилучшей стратегии сборки высоко гетерозиготного генома <i>Boechea</i> <i>divaricarpa</i> M4B.....	40
3.2.1 De novo сборка генома <i>Boechea divaricarpa</i> M4B программой MaSuRCA	40
3.2.2 De novo сборка генома <i>Boechea divaricarpa</i> M4B программой Platanus	41
3.2.3 Сравнительная характеристика различных стратегий сборки высоко гетерозиготного генома <i>Boechea divaricarpa</i> M4B.....	41
3.2.4 Очистка исходных коротких прочтений генома <i>Boechea divaricarpa</i> M4B выравниванием на наиболее успешную сборку	42
3.3 Анализ хлоропластной ДНК <i>Boechea divaricarpa</i> линий M4B и ES517	46
3.3.1 Подготовка прочтений для сборки хлоропластной ДНК	46
3.3.2 Реконструкция хлоропластной ДНК и поиск молекулярных маркеров для филогенетического анализа	48
3.3.3 Определение донора хлоропластов для растений <i>Boechea divaricarpa</i> линий M4B и ES517	48
3.4 Анализ ядерной ДНК <i>Boechea divaricarpa</i> линий ES517 и M4B.....	51
Заключение	54
Выводы.....	55
Список литературы	56

Введение

Половое размножение является одним из важнейших инструментов формирования генетического разнообразия организмов. В дикой природе оно обуславливает постоянную смену генотипов популяций, тем самым обеспечивая материал для адаптивной эволюции. Образование различных комбинаций генов в ходе данного процесса способно привести к появлению более приспособленного к условиям среды фенотипа. Тем не менее, этот же способ размножения может привести к расщеплению полезных признаков в потомстве, ввиду чего последующие поколения не будут обладать особыми свойствами родителей. Данная проблема особенно актуальна в областях, связанных с селекцией растений и животных, например, в сельском хозяйстве. Специально выведенные сорта растений рано или поздно теряют характерные для них качества из-за размножения половым путем, что не позволяет культивировать их в долгосрочной перспективе. Бесполое размножение также встречается у представителей высших растений, но наиболее частой формой такового является вегетативное размножение, не способное стать решением поставленной проблемы, ввиду сложности его реализации в промышленных масштабах.

Альтернативой бесполому размножению через ткани организма для агропромышленности может стать апомиксис — партеногенез, наблюдаемый у ряда растений в дикой природе. При апомиксисе потомство остается генетически идентичным родителю, и при этом размножение происходит через формирование семян, а значит его внедрение в сельское хозяйство не потребует разработки специальных методик выращивания. К сожалению, на данный момент искусственная интеграция данного типа размножения в выращиваемые сорта невозможна, поскольку до конца не установлены молекулярные механизмы его регуляции. Не так давно исследователи обратили внимание

на растения рода *Boechea*, как на потенциальные модельные объекты для изучения генетических аспектов апомиксиса. Виды, принадлежащие роду *Boechea*, обладают рядом особенностей, значительно облегчающих изучение такого типа бесполого размножения. Особенный интерес представляют диплоидные гибриды данного рода, которые характеризуются высокой пенетрантностью апомиксиса, обладая при этом геномом, близкородственным генетическому материалу половых линий.

Цель работы: Сборка генома *Boechea divaricarpa* линии *M4B* и ее сравнительный анализ с геномом линии *ES517*.

Задачи:

- Реализация и сравнение различных стратегий *de novo* сборки ядерной ДНК с целью выявления наиболее эффективного подхода к работе с высоко гетерозиготными организмами.
- Сборка и аннотация хлоропластной ДНК *Boechea divaricarpa M4B* с ее последующим филогенетическим анализом.
- Поиск гена APOLLO в полученной сборке и последовательностях других представителей рода *Boechea* и дальнейший филогенетический анализ по этому маркеру.

Глава 1. Обзор литературы

1.1 Гаметофитный апомиксис и его роль для интенсификации сельского хозяйства

Апомиксис - процесс бесполого размножения растений с помощью семян, который приводит к формированию генетически идентичного потомства (R. A. Bicknell & Koltunow, 2004; Grossniklaus, Nogler, & van Dijk, 2001; Kotani et al., 2014; Nogler, 1984). Если во время полового развития яйцеклетка и центральная клетка оплодотворяются каждая одним спермием, в следствие чего образуется диплоидная зигота, позже развивающаяся в зародыш (клетки которого содержат один материнский и один отцовский геном 1m: 1p) и триплоидный эндосперм (клетки которого содержат один материнский и два отцовских генома 1m: 2p) соответственно (Dresselhaus, Sprunck, & Wessel, 2016), то в случае апомиксиса зародыш формируется как клон нередуцированной материнской клетки семязпочки (матроклинное наследование). В общем случае, для апомиксиса характерны следующие свойства:

- отсутствие редукционного деления (апомейоз);
- развитие зародыша из нередуцированной яйцеклетки без оплодотворения (партеногенез);
- формирование функционального эндосперма или автономно, или при оплодотворении центральной клетки (псевдогамия) (Hand & Koltunow, 2014; A M Koltunow, 1993; Anna M Koltunow & Grossniklaus, 2003; C Spillane, Steimer, & Grossniklaus, 2001).

Главным недостатком полового размножения для сельского хозяйства является расщепление полезных признаков родителя в последующих поколениях потомства, что выражается в потере выгодных комбинаций генов и их аллелей в последующих поколениях (Brukhin, 2017; Charles Spillane, Curtis, & Grossniklaus, 2004). Данная проблема стала одной из причин возросшего за последние два десятилетия интереса к изучению апомиксиса, благодаря его потенциальной возможности зафиксировать ценный, с позиции его сельскохозяйственных качеств, набор генов в неограниченном ряду поколений. Иными словами, внедрение признаков апомиксиса в культуры выращиваемых растений могло бы способствовать долгосрочному закреплению свойств, уникальных для выращиваемых сортов растений. Данный подход облегчил бы культивирование многих сортов, в особенности тех, что получены в результате гибридизации и проявляют свойства гетерозиса, и, как следствие, предоставил бы производствам облегченный доступ к высокоурожайным гибридам (Conner & Ozias-Akins, 2017; Charles Spillane et al., 2004). В литературе уже освещен ряд возможных вариантов использования апомиксиса в сельском хозяйстве, признана его важность не только для создания устойчивых культур растений, но и в аспекте продовольственной безопасности (Conner & Ozias-Akins, 2017; Jefferson, 1994; Anna M Koltunow & Grossniklaus, 2003; Charles Spillane et al., 2004). К сожалению, среди культивируемых растений было найдено крайне мало естественных апомиктов, а перенос способности к апомиксису при помощи интрогрессии от природных апомиктов пока не увенчался успехом (Savidan & others, 2000).

Для внедрения апомиксиса в сельскохозяйственные культуры необходимо понимание молекулярно-генетических основ его регуляции в природе, а также изучение геномов половых и апомиктичных линий с целью выявления генетических вариантов, отвечающих за реализацию данного типа размножения. В настоящее время для исследования генетической регуляции апомиксиса используют: 1) анализ мутантов половых линий, проявляющих некоторые свойства апомиксиса; 2) изоляцию релевантных генов из природных апомиктов (Barcaccia & Albertini, 2013; Conner & Ozias-Akins, 2017; Grimanelli, Leblanc, Perotti, & Grossniklaus, 2001; Pupilli & Barcaccia, 2012). Важным аспектом для изучения апомиксиса является выбор подходящего модельного растения, которое позволит облегчить расшифровку молекулярных механизмов, лежащих в основе апомиксиса.

Три рода растений, получившие наибольшую популярность среди исследователей апомиксиса, *Hieracium*, *Paspalum* и *Pennisetum* – обладают большим геномом и являются полиплоидами, что, в целом, справедливо и для большинства других природных апомиктов

(Asker & Jerling, 1992; Hojsgaard, Klatt, Baier, Carman, & Horandl, 2014). Несмотря на то, что оба качества значительно усложняют молекулярно-генетические исследования, изучение этих растений позволило получить огромное количество новых данных об апомиксисе и его механизмах (R. Bicknell, Catanach, Hand, & Koltunow, 2016; Conner & Ozias-Akins, 2017; Ortiz et al., 2013). В отличие от вышеперечисленных модельных объектов, *Boecheira spp.* обладает относительно небольшим геномом (приблизительно 170-230 млн. пар оснований) и является единственным известным родом, в котором апомиксис проявляется на диплоидном уровне в природе (BOCHER, 1947; C. Dobes, Mitchell-Olds, & Koch, 2004; T F Sharbel, Mitchell-Olds, Dobes, Kantama, & de Jong, 2005). К тому же, *Boecheira spp.* находится в близком родстве с *Arabidopsis thaliana* – популярным модельным объектом, хорошо изученным с позиций молекулярной генетики и аннотации генов и их функций. Тем не менее, апомиктам *Boecheira spp.* также присущ и ряд свойств, усложняющих проводимые на них исследования: крайне высокая гетерозиготность, склонность к аллоплоидии и анеуплоидии, что, очевидно, является следствием гибридизации, часто наблюдаемой у этих растений (Koch, Dobes, & Mitchell-Olds, 2003; Mandakova, Schranz, Sharbel, de Jong, & Lysak, 2015; Schranz, Dobes, Koch, & Mitchell-Olds, 2005). Эти особенности значительно усложняют задачи поэтапной сборки и фазирования геномов апомиктических линий *Boecheira*.

1.2 Таксономия и среда обитания наиболее важных представителей рода *Boecheira*

Род *Boecheira* относится к семейству крестоцветных, в большинстве своем представлен североамериканскими двухлетними или многолетними травами, с характерным для них гаплоидным набором хромосом $n=7$. Ранее эти виды были включены в род *Arabis*, но, поскольку для *Arabis spp.* свойственен гаплоидный набор хромосом $n=8$, были выделены из него (Love & Love, 1975). Молекулярно-генетические данные также подтверждают разницу между представителями двух родов: было показано, что их сходство носит конвергентный характер, демонстрирующий две эволюционно-независимые линии внутри семейства *Brassicaceae* (Al-Shehbaz, 2003). Недавно, благодаря использованию молекулярных маркеров, исследования в области систематики рода *Boecheira* получили дальнейшее развитие. На данный момент в пределах (Beck et al., 2012) рода описаны 110 видов, 71 из

них — диплоиды, предположительно размножающиеся половым путем, хотя диплоидные апомикты также были зафиксированы в литературе; еще 38 видов были определены как триплоидные апомикты гибридогенного происхождения (Windham & Al-Shehbaz, 2007). Таким образом, *Boechea* является пятым по величине родом семейства *Brassicaceae*.

Существующие исследования репродуктивной биологии *Boechea spp.* затрагивают большое количество видов, среди которых: широко распространенный половой диплоид *Boechea stricta* (Love & Love, 1975), половые и апомиктические линии растений, прежде известные как *Boechea holboellii s. l. (sensu lato)* (Bocher, 1947; Love & Love, 1975), и апомиктические гибриды, ранее относимые к виду *Boechea divaricarpa* (Love & Love, 1975; Windham & Al-Shehbaz, 2007). Виду *Boechea gunnisoniana* также было уделено внимание в работах некоторых исследователей, но он пользовался куда меньшей популярностью, чем вышеперечисленные представители рода (Kirioukhova et al., 2018; Schmidt et al., 2014; Taskin, Turgut, & Scott, 2009). Прочие виды были в большинстве своем использованы для изучения только отдельных аспектов апомиксиса (Aliyu, Schranz, & Sharbel, 2010; Corral et al., 2013; Lovell et al., 2013).

До недавнего времени плейстоценовый реликт *B. holboellii (Arabis holboellii)* рассматривался как вид с дизъюнктивным ареалом (т. е. состоящим из нескольких географически изолированных районов) (Bocher, 1951). Тем не менее, недавние исследования систематики рода с использованием молекулярных маркеров показали, что данный таксон является полифилетическим, включающим в себя множество обособленных видов (Alexander et al., 2013). В настоящее время под *B. holboellii* понимают исключительно линии, произрастающие в Гренландии. Они включают в себя половые и апомиктические диплоидные и триплоидные формы, и последние, в отличие от североамериканских видов, по-видимому являются аутотриплоидами.

Включенные ранее в группу *B. holboellii* образцы из Северной Америки были выделены, как ряд видов, образующий агамный комплекс (Stebbins, 1950). Основу этого комплекса составляют четыре диплоидных половых вида, в которых, тем не менее, не может быть исключено присутствие факультативного апомиксиса: *Boechea collinsii*, *Boechea pendulocarpa*, *Boechea polyantha*, *Boechea retrofracta*. Остальные виды являются триплоидными апомиктами гибридогенного происхождения, и их морфология в значительной мере соответствует морфологии родительских половых видов. Среди данных гибридов выделяют: *Boechea consanguinea (retrofracta x fendleri)*, *B. goodrichii (retrofracta x gracilipes)*, *B. grahamii (stricta x collinsii)*, *B. pauciflora (sparsiflora x retrofracta)*, *B.*

pinetorum (*rectissima* x *retrofracta* x *sparsiflora*), *B. quebecensis* (*holboellii* x *stricta*) и *B. tularensis* (*retrofracta* x *rectissima* x *stricta*) (Windham & Al-Shehbaz, 2007). Вид *B. quebecensis* распространен в изолированных районах Северо-Восточной Америки, что подразумевает присутствие в прошлом на североамериканском континенте его предполагаемого предка (Гренландской *B. holboellii* s. l.). С систематической точки зрения, апомиктичная *B. divaricarpa*, вероятно, наиболее проблематичный представитель рода для анализа. В классическом представлении большое разнообразие гибридов, одним из родителей которых является *B. stricta* (включая гибриды *B. stricta* и *B. holboellii* s. l.), было отнесено к виду *B. divaricarpa* во множестве статей, посвященных репродуктивной биологии рода *Boechera*. Такое небрежное использование видового названия способно привести к множеству проблем и вызвать недопонимание между исследователями, в связи с чем Windham и Al-Shehbaz (2007) постановили, что название *B. divaricarpa* следует употреблять исключительно в отношении растений, совмещающих геномы *B. stricta* и *B. sparsiflora*, тогда как гибриды *B. stricta* и *B. collinsii* были выделены в вид *B. grahamii*, а имя *B. quebecensis* закрепили за видами, являющимися гибридами *B. stricta* и *B. holboellii* s. l. В том случае если второй родитель не определен, следует именовать его как «гибрид *B. stricta*», избегая использования названия *B. divaricarpa*.

В качестве перспективной модели для изучения апомиксиса отдельного внимания заслуживает *B. gunnisoniana*. Данный вид представлен триплоидами предположительно гибридогенного происхождения, чьими родителями являются половые диплоидные виды *B. oxylobula* и *B. thompsonii* (последний вид также известен под названием *B. pallidifolia*) (de Arias, 2015). *B. gunnisoniana* обладает следующими выделяющимися свойствами: почти облигатный псевдогамный апомиксис (Roy, 1995; Schmidt et al., 2014), малые размеры растений, относительно быстрое развитие (примерно 4 месяца от посадки до получения семян). Несмотря на то, что подавляющее большинство растений рода *Boechera* произрастают в Северной Америке, два вида были описаны и в дальневосточном регионе России, что является примером дизъюнкции во флоре Восточной Азии и Северной Америки. Первый, *B. falcata*, согласно анализу молекулярных маркеров состоит в близком родстве с хорошо изученными североамериканскими апомиктами (*Boechera* s. s.) (Al-Shehbaz, 2005; Alexander et al., 2013; Kiefer, Dobeš, & Koch, 2009), и эта связь вызывает большой интерес к изучению данного вида, поскольку подразумевает возможное наличие апомиксиса внутри него. Второй вид, эндемик Байкальского региона и других регионов Дальнего Востока России, носит название *Borodinia* (вероятно, также допустимо

использование родового названия *Boechea macrophylla*. Проведенные молекулярно-генетические исследования показали его тесную связь с семью видами *Boechea*, произрастающими на востоке Соединенных Штатов Америки (Alexander et al., 2013).

1.3 Преимущества рода *Boechea* для изучения апомиксиса

Ранее различные виды растений рода *Boechea* часто использовались в исследованиях, посвященных эволюционной экологии (Rushworth, Song, Lee, & Mitchell-Olds, 2011), но за последнее десятилетие они также обрели популярность в качестве модельных объектов для изучения генетических основ апомиксиса. Этому поспособствовали следующие преимущества использования *Boechea spp.*:

- I. Близкое родство с *Arabidopsis thaliana* (Huang et al., 2016), модельным объектом, для которого собрано и проанализировано большое количество генетических данных: его геном полностью отсеквенирован и подробно проаннотирован, известно множество генов, задействованных в исполнении репродуктивной функции, что, очевидно, облегчает поиск генов, задействованных в реализации апомиксиса у представителей рода *Boechea*;
- II. Относительно небольшой размер генома: от 170 до 230 миллионов пар оснований;
- III. Наличие апомиксиса не только у полиплоидных форм, но также и у диплоидов ($2n = 14$), в отличие от множества других растений-апомиктов (Bocher, 1951; Lovell, Williamson, Wright, McKay, & Sharbel, 2017; Voigt-Zielinski, Piwczynski, & Sharbel, 2012);
- IV. Формирование нередуцированных спор в процессе диплоспории по *Taraxacum*-типу (Crane, 2001), очень похожей на мейотическое деление при половом типе размножения, но в ходе нее формируются две нередуцированные клетки-предшественницы спор, вместо четырех клеток с гаплоидным набором хромосом. При этом разные образцы демонстрируют различные варианты контакта между клетками — от полностью сохранившейся связи, до полной изоляции;
- V. Апомиктические линии рода *Boechea* являются псевдогамными, что редко встречается у растений с диплоспорией (Talent, 2009);
- VI. Апомиксис у представителей *Boechea spp.* всегда носит факультативный характер, что позволяет проводить гибридизацию и определенный генетический анализ даже у растений видов, с очень высокой пенетрантностью данного типа размножения;

- VII. Наличие в роду *Boechea* как половых (размножающихся преимущественно половым способом), так и апомиктных видов, представители которых обладают различной плоидностью и географическим распространением, но генетические различия между многими из них крайне малы, что значительно облегчает поиск гомологичных последовательностей в молекулярных исследованиях;
- VIII. Половые линии *Boechea spp.* являются самосовместимыми и самоопыляемыми, в отличие от половых видов, лежащих в основе образования большинства прочих апомиктов; как следствие возможности самоопыления, половые линии рода *Boechea* обладают крайне низкой гетерозиготностью;
- IX. Наличие разработанных методик для генетической трансформации через соматический эмбриогенез в культурах тканей апомиктов *B. gunnisoniana* и *B. holboellii s. l.* (Taskin et al., 2009).

1.4 Цитоэмбриологические исследования в роде *Boechea*

Первые подробные цито-эмбриологические исследования апомиксиса в роде *Boechea* были проведены датским ботаником Tyge W. Böcher (1951, 1947). Он обнаружил наличие апомиксиса у вида *Boechea holboellii s. l.* (именуемого им *Arabis holboellii*) среди диплоидных и триплоидных форм, а также описал мегаспорогенез и микроспорогенез у ряда половых и апомиктичных представителей *Boechea spp.* из Гренландии и Северной Америки. Особенно примечательно его описание форм с варьирующим углом конъюгации хромосом в профазе мейоза I. Он также отметил наличие растений с различными уровнями плоидности (в основном $2n$ и $3n$, реже $4n$, $5n$ и $6n$), присутствие анеуплоидов ($2n = 16, 22, 23, 30$) и предположил гибридогенное происхождение последних.

Почти 50 лет спустя голландско-российская научная группа провела эмбриологические исследования просветленных препаратов *B. holboellii s. l.* (образцы из Гренландии и штата Колорадо США) с использованием дифференциальной интерференционно-контрастной микроскопии (ДИК) и проточной цитометрии (Naumova et al., 2001). Исследователями было продемонстрировано наличие как мейоза, так и апомейоза в процессе мегаспорогенеза (**Рисунки 1-3**). При скрининге большого количества семязачатков было подтверждено образование нередуцированного зародышевого мешка посредством

диплоспории и партеногенетического развития зародыша в апомиктичных растениях рода *Boechera*. У половых видов развитие зародышевого мешка происходит по *Polygonum*-типу (Maheshwari, 1950) (**Рисунки 1, 3**): диплоидная материнская клетка мегаспор претерпевает мейоз с образованием тетрады гаплоидных мегаспор, три из которых дегенерируют, в то время как функциональная мегаспора подвергается трем митотическим делениям, вследствие чего образуется восьмиядерный семиклеточный зародышевый мешок, включающий в себя яйцеклетку, две синергиды, три клетки-антиподы, а также два полярных ядра, которые сливаются, формируя гомодиплоидное ядро центральной клетки. В пыльниках материнские клетки пыльцы дают начало пыльцевым зернам, подвергаясь мейозу, в следствие которого образуются тетрады микроспор. В дальнейшем каждая микроспора претерпевает асимметричное деление, из-за чего образуются большая вегетативная и меньшая генеративная клетка. После опыления вегетативные клетки прорастают, формируя пыльцевую трубку, предназначенную для транспорта спермиев в зародышевый мешок, а генеративная клетка делится еще раз, образуя два спермия, один из которых впоследствии оплодотворит яйцеклетку, а второй сольется с центральной клеткой; все это в результате приводит к образованию диплоидного эмбриона и триплоидного эндосперма (**Рисунки 1, 3**).

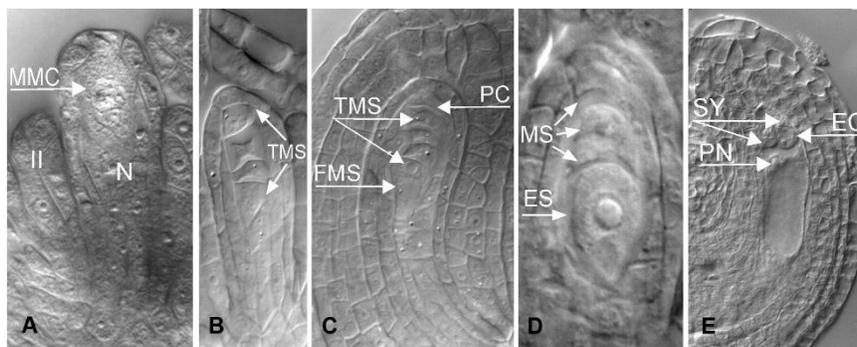


Рисунок 1. - Мейотический спорогенез в зародышевых мешках *B. holboellii s. l.* (**A**) материнская клетка мегаспоры (MMC), нуцеллус (N); (**B, C**) тетрада мегаспор (TMS), функциональная мегаспора (FMS) на халазальном конце; (**D**) одноядерный мейотический зародышевый мешок (ES) с остатками дегенерирующих нефункциональных мегаспор (MS); (**E**) зрелый семиклеточный зародышевый мешок *Polygonum*-типа с яйцеклеткой (EC), двумя синергидами (SY) и центральной клеткой, имеющей два полярных ядра (PN) (Naumova et al., 2001; Osadtchiy, Naumova, & Brukhin, 2017).

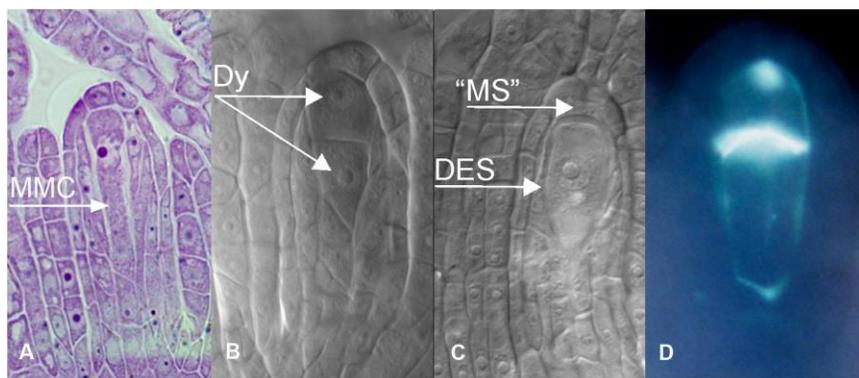


Рисунок 2. - Апомейотический спорогенез у *B. holboellii s. l.* (A) материнская клетка мегаспоры (MMC); (B) диплоспорическая диада (Dy); (C) одноклеточный диплоспорический зародышевый мешок (DES) с остатками «мегаспоры» (“MS”); (D) каллоза в клеточной стенке диплоспорической диады (Naumova et al., 2001; Osadtchiy et al., 2017).

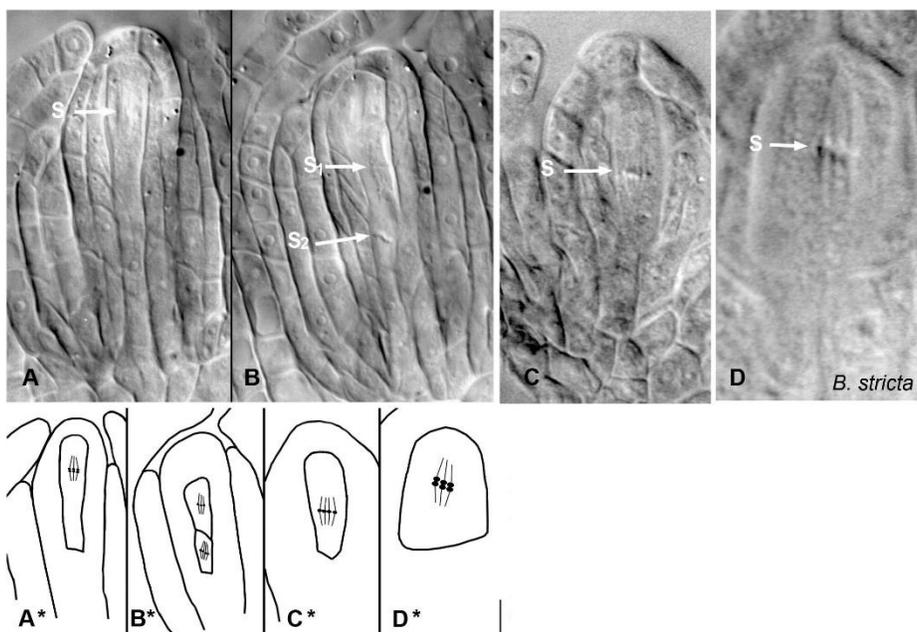


Рисунок 3. - Мейоз и апомейоз в зародышевых мешках *B. holboellii s. l.* (A-C, A*-C*) и *B. stricta* (D, D*). (A, A*) мейотическое веретено метафазы I деления мейоза (Osadtchiy et al., 2017).

В апомиктных линиях рода *Boechera* диплоспорический апомейоз по *Taraxacum*-типу (Crane, 2001) распространен наиболее широко (Bocher, 1951; de Arias, 2015; Naumova et al., 2001; Schmidt et al., 2014). При мейотической диплоспории зародышевый мешок происходит от материнской клетки мегаспор, подвергающейся мейозу с отклонениями в сегрегации хромосом, что приводит к образованию диады нередуцированных мегаспор. Характерной особенностью материнской клетки мегаспор является отсутствие отложения каллозы вокруг клетки (Nogler, 1984; Rodkiewicz, 1970), но в случае вышеупомянутой

диады каллоза была найдена между двумя ее клетками. Наиболее часто наблюдается формирование диад (Roy, 1995; Schmidt et al., 2014), хотя редкие триады и даже тетрады также встречаются (Schmidt et al., 2014). Халазальная клетка диады проходит три раунда митоза, в результате чего образуется нередуцированный восьмиядерный зародышевый мешок, морфологически соответствующий мешку *Polygonum*-типа (Crane, 2001; Rojek et al., 2018).

Ранее считалось, что апоспорический апомейоз редко встречается среди представителей *Boechera spp.*, но недавно его наличие было описано сразу у нескольких видов. Среди описанных случаев подавляющее большинство апоспорических зародышевых мешков развивалось по *Hieracium*-типу (Crane, 2001). Материнская клетка мегаспор в этом случае может дегенерировать или же подвергаться мейозу, что наблюдалось в редких случаях у *B. Microphylla* (de Arias, 2015). У гибридов *B. stricta* и *B. retrofracta* была выявлена диплоспория по *Antennaria*-типу (Crane, 2001), но ее частота оказалась крайне мала. При анализе семян методом проточной цитометрии было также показано, что процент зрелых половых семян (относительно апомиктических) у всех исследованных апомиктов значительно ниже, чем процент морфологически нормальных мейотических тетрад (относительно проявлений апомейоза). Из этого следует, что в большинстве случаев мейоз протекает с отклонениями от нормы, приводящими к формированию нежизнеспособных семян, а сам апомиксис сглаживает это явление, защищая растение от бесплодия (de Arias, 2015).

У большинства апомиктов, развитие пыльцы не подвержено каким-либо отклонениям, однако среди представителей рода *Boechera* апомейоз был также описан и для микроспорогенеза (Bocher, 1951; Taskin et al., 2009). В случае триплоидных апомиктов мейоз I нарушается, поскольку хромосомы не могут правильно образовать пары в пахитене. Хромосомы мигрируют к противоположным полюсам пыльцевой материнской клетки и деконденсируются, затем следует цитокинез. После цитокинеза диада, в отличие от мейотической тетрады, окружена стенкой каллозы. В метафазе I деления пыльцевой материнской клетки *B. holboellii* обнаружена конъюгация хромосом, приводящая к образованию би- и тривалентов, в то время как апомейоз у *B. gunnisoniana* протекает без какой-либо конъюгации. Исследования микроспорогенеза апомиктических триплоидов этих видов показали, что большая часть их пыльцевых зерен нередуцирована и образована апомейотическими диадами (98% для *B. holboellii* и 90% для *B. gunnisoniana*), тогда как оставшая часть пыльцы образовалась за счет мейоза (частично аномального), в ходе

которого формируются тетрады или триады микроспор (Taskin et al., 2009), а в случае триплоидов *B. holboellii* могут образовываться даже монады (Bocher, 1951). Диплоидные апомикты обладают большей вариативностью апомейоза. У различных изолятов *B. holboellii* конъюгация и кроссинговер в норме происходят в пахитене, что в одних случаях приводит к образованию редуцированной пыльцы, а в других — к формированию тривалентов и даже квадрилвалентов. Некоторые представители вида формируют в основном диплоидную пыльцу, тогда как иные демонстрируют наличие не только диплоидной, но и гаплоидной пыльцы (Kantama et al., 2007; Kirioukhova et al., 2018; Taskin et al., 2009). Исследователями было отмечено, как диплоидные, так и триплоидные апомикты способны давать редуцированную и нередуцированную пыльцу в различных пропорциях (Aliyu et al., 2010; Bocher, 1951; Voigt-Zielinski et al., 2012).

При половом размножении у представителей рода *Boechera* происходит двойное оплодотворение, тогда как при апомиксисе имеет место только оплодотворение полярных ядер. В редких случаях также удается обнаружить полностью автономный апомиксис (Matzk, Meister, & Schubert, 2000; Naumova et al., 2001), присущий в большей степени триплоидам (встречается с частотой до 15%), чем диплоидам (частота не более 1,33%) (Aliyu et al., 2010). Еще одним редким событием является формирование зародышей с удвоенным набором хромосом в результате слияния нередуцированных женских и мужских половых клеток (Naumova et al., 2001). Наиболее важным выводом, сделанным на основании анализа семян при помощи проточной цитометрии в соответствующих исследованиях, стало заключение о том, что все зрелые семена, формируемые апомиктами, были получены из нередуцированных женских и мужских гамет (Naumova et al., 2001).

Плоидность эндосперма апомиктов варьирует в зависимости от плоидности спермиев, но наиболее распространенным является соотношение в 2 материнских генома к 1 отцовскому. Исключения, хотя и с куда меньшей частотой, наблюдаются, что указывает на определенную гибкость системы или же на возможные ошибки в ней (Aliyu et al., 2010).

Анализ большого количества семян шестнадцати видов рода *Boechera* при помощи проточной цитометрии показал, что для диплоидов свойственен широкий спектр репродуктивных режимов: от исключительно полового до близкого к облигатному апомиксиса. Оценка частоты апомейоза, полового и партеногенетического формирования зародышей, а также полового, псевдогамного и автономного развития эндосперма показала, что все исследованные факультативные апомикты с одним и тем же геномом обладали либо очень низким (1-3%), либо очень высоким (87-99%) процентом апомейоза, в

то время как индивиды с промежуточными частотами не наблюдались. Кроме того, всем триплоидам был присущ облигатный апомейоз. Для каждого отдельно взятого генотипа корреляция между апомейозом и партеногенезом, сочетающимся с псевдогамным или автономным развитием эндосперма, продемонстрировала, что частоты последнего никогда не превышали частоту апомейоза (Aliyu et al., 2010). Такая закономерность может указывать на тесную взаимосвязь их генетического контроля, а также на ключевую роль, которая отведена апомейозу на последующих этапах развития апомикта.

1.5 Популяционно-генетические исследования апомиксиса в роде *Boechea*

Иногда апомиктические и половые линии рода *Boechea* могут пересекаться (Alexander et al., 2013; Schranz et al., 2005). Очевидно, что различающиеся по способу размножения линии должны эволюционировать под действием разных сил, а значит — в разных направлениях. Рекомбинация у половых линий увеличивает вероятность устранения вредных мутаций (Felsenstein & Yokoyama, 1976; Hill & Robertson, 1966), тогда как ее отсутствие у апомиктов напротив, не позволяет восстанавливать жизненно-важные или адаптивные аллели при их нарушении в ходе мутагенеза. Следовательно, в апомиктических линиях можно ожидать накопления вредных аллелей — феномен, известный как храповик Мюллера (Charlesworth & Charlesworth, 1997; Muller, 1964). Недавние исследования, направленные на сравнение апомиктов и половых линий *Boechea spp.*, подтвердили эти ожидания (Lovell et al., 2013, 2014, 2017). Lovell и соавторы (2017), используя апомиктические и половые популяции *B. spatifolia*, исследовали закономерности изменения нуклеотидов при обоих способах размножения с помощью полногеномного секвенирования. В апомиктических линиях ими были обнаружены повышенное разнообразие последовательностей ДНК, гетерозиготность, а также повышенная частота накопления мутаций. Схожим образом, в более широком обзоре 37 естественных популяций четырех видов *Boechea spp.* (*B. stricta*, *B. retrofracta*, *B. polyantha*, *B. pendulocarpa*), при помощи микросателлитных маркеров, была продемонстрирована аналогичная тенденция: вне зависимости от пloidности, уровень гетерозиготности у апомиктических линий был выше, чем у половых (Lovell et al., 2013).

Эволюция апомиктов происходит благодаря дрейфу генов и естественному отбору (D. Charlesworth & Wright, 2001; Glemin, Bazin, & Charlesworth, 2006). Следствием более низкой эффективности отбора, ожидаемой в апомиктических линиях, должен быть растущий риск исчезновения апомиктов ввиду накопления вредных аллелей и неспособности адаптироваться к изменениям окружающей среды (Muller, 1964). Не смотря на эти предположения, род *Boechera* достаточно разнообразен и включает в себя несколько стабильных апомиктических линий (Alexander et al., 2013). Наиболее вероятным объяснением выживания последних является внутри- и межвидовой поток генов внутри рода (Beck et al., 2012; Bocher, 1951; C. H. Dobes, Mitchell-Olds, & Koch, 2004; Lovell et al., 2013; Schilling, Gompert, Li, Windham, & Wolf, 2018; Schranz et al., 2005; T F Sharbel & Mitchell-Olds, 2001). Основной поток генов происходит по направлению от половых особей к апомиктическим, в то время как апомикты, благодаря способности формировать редуцированную пыльцу, могут опылять половые линии, передавая им доминантный признак (признаки), инициирующий апомиксис. Взаимный поток генов между линиями обоих типов может позволить провести интрогрессию, как было показано Van Dijk и соавт. (2009) на *Taraxacum spp.* Более того, апомиксис у *Boechera spp.* является факультативным: одни и те же особи способны производить как половое, так и апомиктическое потомство (Aliyu et al., 2010; Schranz et al., 2005). Таких случаев вероятного полового размножения у апомиктов может быть достаточно для очищения генома от вредных мутаций и восстановления приспособленности данных линий к меняющимся условиям среды (Schön et al., 2009).

Различия в силе естественного отбора, действующего на апомиктические и половые популяции *B. spatifolia*, были описаны Lovell и соавторами (2014). Они проанализировали градиенты отбора, посредством сопоставления генотипических признаков с мерой относительной приспособленности, и обнаружили, что сила адаптивной эволюции значительно снижена у апомиктов, относительно половых линий. Апомиктические линии обладали меньшим разнообразием количественных и молекулярно-генетических признаков между разными популяциями, нежели половые. Кроме того, различия между популяциями апомиктов не коррелировали с изменениями окружающей среды, тогда как геномная структура и количественные черты половых линий демонстрировали явную зависимость от широты, климатических условий, высоты, на которой обитала популяция. Простой эксперимент по выращиванию растений показал, что у популяций на высотных участках проводился жесткий отбор по времени цветения; данное наблюдение согласуется с

лабораторными и полевыми исследованиями рекомбинантных имбрендных линий вида *B. stricta*, для которых был показан направленный характер этого отбора (Anderson, Lee, & Mitchell-Olds, 2011).

Ряд исследователей оценил генетическую динамику и эффективность естественного отбора в популяциях *Boecheira spp.* Более ранние работы по популяционной динамике половых видов *B. fecunda* (Song & Mitchell-Olds, 2007) и *B. stricta* (Song et al., 2009), в которых были использованы данные о последовательностях нескольких ядерных локусов и микросателлитов, выявили близкие уровни полиморфизма и популяционной дифференциации в обоих видах, не смотря на значительную разницу между широко распространенной *B. stricta* и исчезающей *B. fecunda*, обладающей меньшим ареалом. Схожим образом, исследования, в которых сравнивались распространенные виды *B. stricta* и *B. latifolia* с редкими видами *B. crandallii* и *B. vivariensis*, не продемонстрировали значительной корреляции между размером ареала вида и генетическим разнообразием внутри его популяций (Lovell & McKay, 2015). Тем не менее, наиболее широко распространенным видам были свойственны высокая фенотипическая пластичность и большая дисперсия значений количественных признаков, тогда как редкие виды были подвержены более жесткому отбору по фенотипическим параметрам, о чем свидетельствовало соотношение таких часто применяемых в популяционной генетике показателей, как индекс фиксации (F_{st}) и индекс фиксации количественных признаков (Q_{st}). В данном случае соотношение $Q_{st} : F_{st}$ было выше, что и позволило сделать соответствующий вывод (Lovell & McKay, 2015).

Гибридизация широко распространена в роде *Boecheira* – было зафиксировано ее многократное и независимое появление в пределах рода (Bocher, 1951; C. N. Dobes et al., 2004; T F Sharbel & Mitchell-Olds, 2001; Windham & Al-Shehbaz, 2007). Самые ранние молекулярные доказательства наличия гибридизации были получены при исследовании внутренних транскрибируемых спейсеров и последовательностей хлоропластной ДНК; также в ходе филогеографического анализа был обнаружен поток генов между видами, ныне известными как *B. stricta* и *B. retrofracta*. (C. Dobes et al., 2004). И хотя следует отметить, что выводы о гибридной природе особей, основанные лишь на одном локусе или геноме органелл, могут не совсем точно отражать историю клады или популяции (Maddison & Knowles, 2006). Schranz и соавторы (2005) в своих обширных экспериментах успешно продемонстрировали возможность скрещивания у нескольких представителей рода. Полученные ими результаты указывают на отсутствие внутренних барьеров

репродуктивной изоляции и, следовательно, на возможное присутствие обширного потока генов среди различных видов *Boechea* spp.

До настоящего времени микросателлитные маркеры занимали центральное место в идентификации видов и предполагаемых гибридов внутри рода *Boechea* (Li, Rushworth, Beck, & Windham, 2017). Beck и соавторы (2012) проводили исследования на представителях рода, используя набор из 13 микросателлитов. Ими были получены подтверждения формирования гибридов *B. fendleri* x *B. stricta* и *B. retrofracta* x *B. stricta*. Работая с аналогичными методиками, Lovell и соавторы (2013) исследовали 231 особь из 37 естественных популяций четырех видов *Boechea* (*B. stricta*, *B. retrofracta*, *B. polyantha* и *B. pendulocarpa*). Они заключили, что все триплоидные особи были апомиктичными гибридами, в то время как этого нельзя было сказать о диплоидных апомиктах, которые в большей степени соответствовали полноценным видам, а не отдельным гибридным особям. На основании всех этих результатов был сделан вывод, что гибридизация является косвенным коррелятом апомиксиса в роду *Boechea*.

С появлением технологий секвенирования нового поколения идентификация гибридов стала намного более точной. Используя полногеномное секвенирование, Lovell и соавторы (2017) проверяли гипотезу о гибридогенном происхождении апомиктичных популяций *B. spatifolia*. Гипотеза была подтверждена, путем анализа бутстрапа на 22000 деревьев гаплотипов, при построении которых использовались данные всего генома. В другом исследовании, в основе которого лежал принцип генотипирования через секвенирование, Schilling и соавторы (2018) оценивали вариативность генома у 79 особей, принадлежащих восьми видам *Boechea*. Анализ генетических примесей различных гипотетических предков позволил с высокой точностью идентифицировать гибридные особи. Это исследование предоставило доказательства недавнего и древнего смешения генофондов видов.

1.6 Наследование и генетические аспекты апомиксиса в роду *Boechea*

Ключевая работа Nogler 1970х-90х (Nogler, 1984) годов показала, что апомиксис контролируется генетически. Последующие эксперименты по скрещиванию апомиктичных растений как доноров пыльцы, с половыми растениями продемонстрировали наследование

апомиксиса как доминантного признака (Grimanelli et al., 2001). Ранее считалось, что апомиксис наследуется как одиночный доминантный локус, например, у *Ranunculus auricomus* и *Panicum maximum* – в их случае апомейоз и партеногенез наблюдались совместно (Nogler, 1984). Однако более поздние работы показали, что отдельные компоненты апомиктического развития, такие как апомейоз, партеногенез, автономное образование эндосперма, у большинства апомиктов контролируются разными локусами. Также было обнаружено, что в областях генома, связанных с проявлением апомиксиса, наблюдается подавленная рекомбинация (Barcaccia & Albertini, 2013; R. A. Bicknell & Koltunow, 2004; Brukhin, 2017; Grimanelli et al., 2001; Hand & Koltunow, 2014; van Dijk & Vijverberg, 2005). Как упоминалось выше, апомиктическое развитие часто ассоциировано с гибридизацией и возникающей полиплоидией (Koltunow & Grossniklaus, 2003). Увеличенная геномная нагрузка может быть причиной нарушений регуляции генов, отвечающих за половое размножение (Barcaccia & Albertini, 2013; Grimanelli et al., 2001; C Spillane et al., 2001), поскольку вновь сформированный полиплоидный гибрид сталкивается с их асинхронной экспрессией (Carman, 1997; Grimanelli et al., 2001; Schön et al., 2009). Последние данные указывают на то, что апомиксис связан с повышенным разнообразием (Hojsgaard et al., 2014), что позволяет предположить его ключевую роль в создании новых полиплоидов (Hojsgaard, 2018) и диверсификации покрытосеменных растений.

Одной из уникальных характеристик рода *Boechea* является появление апомиксиса даже у диплоидных растений. Диплоидные апомикты *Boechea spp.* являются высоко гетерозиготными гибридами (Beck et al., 2012), и недавние цитогенетические и популяционные исследования половых и апомиктных видов данного рода показали, что эти диплоидные геномы могут иметь достаточно сложную структуру. Модель, в основе которой лежит анализ генетических маркеров и плоидности, предполагает, что, во-первых, генетические факторы апомейоза возникли независимо. Путем самоопыления редуцированной пылью особь, которой присущ апомейоз исключительно при генерации яйцеклетки, могла бы стабильно формировать семена с диплоидным эмбрионом и эндоспермом, обладающим 5-ю наборами хромосом. Другие растения, у которых апомейоз происходит только при формировании спермиев, могли бы продуцировать семена с различной плоидностью. Редуцированная пыльца, полученная от апомейотика по женской линии половых клеток, позволила бы скрещиваться с половыми линиями, тем самым распространяя фенотип. В конечном итоге происходит скрещивание растений, обладающих

апомейозом при формировании яйцеклеток, и растений, апомейоз у которых происходит при синтезе спермиев, вследствие чего получаются семена с диплоидным эмбрионом и гексаплоидным эндоспермом, которые вырастут в стабильные диплоиды, способные к производству нередуцированных женских и мужских гамет. Пыльца этих апомиктов также может в дальнейшем оплодотворять половые диплоиды, что приводит к формированию апомиктических триплоидов (Lovell et al., 2013).

Окрашивание метафазных хромосом методами геномной гибридизации *in situ* продемонстрировало, что все исследованные апомиктные линии обладали признаками гибридогенного происхождения. Было обнаружено, что все они являются аллоплоидами с различным количеством хромосом, унаследованных от *B. holboellii s. l.* или от *B. stricta*. Структура их хромосом была в значительной степени затронута последствиями гибридизации, что привело к анеуплоидии и замене гомологичных хромосом (Kantama et al., 2007). Таким образом, эти апомиктные линии не являются однозначными диплоидами, скорее они имеют полигаплоидное происхождение. Следует отметить, что эти цитогенетические данные не исключают до конца возможного существования истинных диплоидов среди апомиктов рода *Boechea*. Предполагается, что обнаруженные у диплоидных апомиктов хромосомы Het, Het' и Del (представлены в основном гетерохроматином) отвечают за наследование признаков, связанных с апомиксисом (Kantama et al., 2007). Согласно Kantama и соавт., все исследованные диплоидные апомикты обладали по меньшей мере четырьмя хромосомами *B. stricta*, включая Het и Del, и данной рабочей группой был сделан вывод, что комбинация этих хромосом может иметь важное значение для проявления апомиксиса.

Недавние исследования показали, что хромосома Het является измененным гомологом первой хромосомы *B. stricta*, в которой произошло накопление перичентромерного гетерохроматина, в то время как пара Het' и Del – продукт разрыва Het с последующей перичентрической инверсией в хромосоме Het' (Mandakova et al., 2015). Согласно более ранним данным, в некоторых линиях хромосома Del могла возникнуть в результате транслокации, сливающей проксимальный сегмент хромосомы *B. stricta* с дистальным сегментом хромосомы *B. holboellii s. l.* (Kantama et al., 2007). Однако гибридизация половых и апомиктических растений *Boechea* не давала апомиктического потомства, несмотря на унаследованную хромосому Het (Schranz, Kantama, de Jong, & Mitchell-Olds, 2006). При скрещивании половых диплоидов *B. stricta* с апомиктическими аллодиплоидами *B. divaricarpa*, несущими хромосому Het, потомство F1 было триплоидным и имело

низкую фертильность, но не обладало апомиксисом, даже при наличии вышеупомянутой хромосомы *Net*. В следующем поколении (F2) плоидность и число хромосом варьировали и наблюдалась низкая фертильность. Некоторые особи поколения F3, по-видимому, сохранили высокую плоидность своих родителей и продемонстрировали более высокую плодовитость, чем в F1 и F2, но все еще не достигавшую фертильности использованных первоначально растений. Апомиктического потомства так и не было произведено, а значит, что генетический контроль апомиксиса в роде *Boechea* не ограничивается наследованием абберантных хромосом.

Хромосомные регионы с подавленной рекомбинацией вокруг генов, связанных с апомиксисом, часто находящиеся в гемизиготном состоянии и обогащенные повторами и транспозонами, были найдены во многих филогенетически удаленных апомиктах, как двудольных, так и однодольных (Grimanelli et al., 2001; Kotani et al., 2014; Okada et al., 2011; Ozias-Akins, Akiyama, & Hanna, 2003; Schön et al., 2009). Предполагается, что такая рекомбинационно-инертная область может содержать несколько связанных генов с различными функциями, синергия которых способна привести к апомиктичному развитию. Однако у многих апомиктов локусы, контролирующие различные компоненты апомиксиса, находятся в различных участках генома. В случае *Boechea spp.* наиболее вероятными кандидатами на перенос таких локусов являются абберантные хромосомы *Net*, *Net'* и *Del*, речь о которых уже велась выше.

Принимая во внимание гибридогенную природу апомиктов *Boechea* в качестве механизма, инициирующего формирование и дальнейшую эволюцию таких рекомбинантно-инертных блоков, содержащих связанные с апомиксисом гены, можно сделать вывод, что, вероятно, именно гибридизация видов с неполной хромосомной гомологией способна привести к образованию гемизиготных областей с подавленной рекомбинацией, которые в дальнейшем получают свое развитие (Sharbel et al., 2010).

Несколько теорий описывают возможные механизмы, контролирующие апомиксис. Некоторые из них присваивают генетическим мутациям роль главного индуктора, согласно таким гипотезам, мутация в одном из первых генов регуляторного каскада (или перед ним) может привести к апомейозу, партеногенезу и/или развитию автономного эндосперма (Anna M Koltunow & Grossniklaus, 2003). У *Arabidopsis* были идентифицированы различные мутанты, которым свойственен апомейоз (Schmidt, Schmid, & Grossniklaus, 2015). Интересно, что их мутации, по-видимому, разделены между регуляторами клеточного цикла/основными генами мейоза и генами, задействованными в регуляторных

путях малых РНК (d'Erfurth et al., 2010, 2009; Ravi, Marimuthu, & Siddiqi, 2008; Zhao et al., 2017).

В то время как мутанты рода *Boechea* еще не изучены, были идентифицированы два локуса, коррелирующие с женским и мужским апомейозом. Первый — APOmixis Linked Locus (APOLLO) – кодирует экзонуклеазу Asp-Glu-Asp-Asp-His, его экспрессия подавляется в материнских клетках мегаспор, когда они вступают в мейоз, и повышается в апомейотических (Corral et al., 2013). APOLLO демонстрирует двуаллельное наследование с «апоаллелями» и «половыми» аллелями, эти варианты гена отличаются друг от друга по 5'- нетранслируемой области гена экзонуклеазы, в которой наблюдается 20-ти нуклеотидный полиморфизм. Все протестированные апомиктичные растения рода *Boechea* были гетерозиготными по аллелям APOLLO, они обладали по меньшей мере одним «апоаллелем» и одним «половым» аллелем, в то время как все половые особи гомозиготны и обладают лишь «половыми» вариантами гена (Corral et al., 2013). Мужской аналог APOLLO – локус под названием Unreduced Pollen GRAin Development2 (UPGRADE2), который экспрессируется исключительно в пыльцевых материнских клетках апомиктных видов. Он кодирует химерную длинную некодирующую РНК, обладающую склонностью к образованию вторичных структур. UPGRADE2 возник в результате дупликации UPGRADE1 с последующей вставкой функционального гена и дальнейшим формированием экзона в этой области, что сделало этот ген транскрипционно активным (Mau et al., 2013). Существует высокая корреляция между наличием двух вышеперечисленных локусов, связанных с апомиксисом, и реализацией растением апомиктического способа размножения (у 98,4% апомиктов в выборке из 275 особей, принадлежащих 22 видам рода *Boechea*, был найден «апоаллель» APOLLO, у 96% - UPGRADE2), хотя для половых растений также было показано наличие «апоаллеля» APOLLO в 2,27% случаев и присутствие UPGRADE2 в 34,48% случаев. Таким образом, можно заключить, что APOLLO является наиболее подходящим диагностическим индикатором апомиксиса у разных видов и линий рода *Boechea* (Mau et al., 2015), но, несмотря на это, его функция в размножении до сих пор остается невыясненной. Независимость APOLLO и UPGRADE2 согласуется с популяционно-генетическими исследованиями, которые показали, что мужской и женский апомейоз наследуются отдельно друг от друга, хотя и наблюдается их корреляция друг с другом на уровне популяции (Lovell et al., 2013).

Kliver и соавт. (2018) обнаружили две дополнительные, более удаленные, копии APOLLO, наличие которых может указывать на произошедшие в прошлом дубликации. Изучение «апоаллелей» и «половых» аллелей гена APOLLO показывает, что они возникли после отделения рода *Boechea* и образуют две отдельные клады. Учитывая, что *B. retrofracta* и *B. stricta* являются половыми видами, не вызывает удивления наличие у них «половых» аллелей APOLLO. Авторы предлагают эволюционный сценарий, в котором после трипликации, вероятно состоявшейся после обособления рода внутри *Brassicaceae*, одна из копий APOLLO могла приобрести новую функцию у общего предка *Boechea spp.*, что привело к отделению апомиктных линий. Соотношение замен по синонимичным (Ks) и несинонимичным сайтам (Ka) аллелей APOLLO показывает, что ветвь, ведущая к «апоаллелям» находится под положительным отбором ($Ka/Ks = 1.4646$), что типично для паралогов, которые приобрели новую функцию.

Предполагается, что эпигенетические изменения в регуляции генов также лежат в основе апомиксиса. Было продемонстрировано, что межвидовая гибридизация *A. thaliana* с сопутствующей полиплоидизацией приводит к кардинальным изменениям в экспрессии генов (H. S. Lee & Chen, 2001), что позволяет соединить в ней вместе как гипотезу индукции апомиксиса в ходе гибридизации, так и гипотезу индукции за счет генетических мутаций, в которой эпиаллели, а не мутантные аллели, могут играть ключевую роль в нарушении регуляции репродуктивных генов во времени и пространстве (Grimanelli et al., 2001; C Spillane et al., 2001). Малые молекулы РНК участвуют в эпигенетической смене программы развития во время гаметогенеза и в регуляции событий, следующих за оплодотворением (Martinez & Kohler, 2017); среди генов, на которые они воздействуют — AGO9, ген, взаимодействующий с 24-х нуклеотидными малыми РНК, выделенными из транспозируемых элементов яйцеклеток. Было показано, что мутанты по этому гену образуют несколько клеток, способных митотически формировать несколько зародышевых мешков в одном семязачатке. К сожалению, пока неясно, является ли апомейотический фенотип мутантов AGO9 следствием отсутствия сайленсинга транспозируемых элементов или же итогом воздействия других малых РНК (Rodriguez-Leal & Vielle-Calzada, 2012). Профайлинг экспрессии малых РНК у *Boechea spp.* выявил консервативные специфичные для данного рода и похожие на микросателлиты малые РНК, многие из которых имеют потенциальные сайты связывания в экзонных областях, причем большинство их мишеней — регуляторные факторы. Предполагается, что количественное изменение в способности микросателлитоподобных РНК связываться с их мишенями происходит в результате

полиморфизмов в длине микросателлитных участков либо предшественников таковых РНК, либо генов самих мишеней, и таким образом это может приводить к значительным сдвигам в экспрессии генов между половыми и апомиктическими линиями (Amiteye et al., 2013).

Хотя гены, осуществляющие контроль отдельных компонентов апомиксиса у *Boechnera* spp. все еще не полностью установлены, данные о наследовании апомиксиса и локусов, связанных с ним, уже сейчас предоставляют большое поле для деятельности.

1.7 Геномные ресурсы для рода *Boechnera* и проблемы геномного анализа

Появление технологий секвенирования нового поколения наряду с эволюцией биоинформатических инструментов открыло новые перспективы в изучении апомиксиса. Использование полногеномного секвенирования позволило искать соответствующие локусы, а также сравнивать и генотипировать виды и линии рода *Boechnera*, проводить их филогенетический анализ.

В настоящее время собраны и опубликованы геномы лишь двух видов *Boechnera* — *B. stricta* и *B. retrofracta*. Оба вида являются самоопыляемыми, диплоидными, половыми и обладают высокогомозиготными геномами, с которыми легко работать при сборке. Примечательно, что повторы в геноме *B. retrofracta* занимают почти 40% места, и половина из них представлена длинными терминальными повторами (Kliver et al., 2018). В случае *B. stricta* повторы занимают лишь 20% генома (C.-R. Lee et al., 2017). Разница в количестве этих последовательностей коррелирует с разницей в размерах геномов данных видов. Локусы апомиксиса некоторых апомиктических видов ассоциированы с гетерохроматином и/или с повторами, из числа наиболее широко представленных (Hand & Koltunow, 2014). Хромосомы, несущие локусы LOSS-OF-APOMEIOSIS (LOA) у *Hieracium praelatum* и APOSPORY-SPECIFIC GENOMIC REGION (ASGR) у *Pennisetum squamulatum* характеризуются большим количеством повторяющихся последовательностей и транспозон-богатых регионов (Okada et al., 2011). Область генома *Paspalum simplex*, содержащая связанные с апомиксисом локусы, подверглась значительным перестройкам из-за обилия мобильных элементов (Calderini et al., 2006). Все эти сходства в составе геномов апомиктов привели к формированию гипотезы о том, что такие богатые повторами

и мобильными элементами области могут служить хранилищем для факторов, приостанавливающих половое развитие и запускающих апомиктическое (Koltunow & Grossniklaus, 2003). В соответствии с этой идеей, было обнаружено, что некоторые апомикты *Boecheira* имеют хромосомы, представленные по большей части гетерохроматином, а отдельные семейства транспозонов получили широкое распространение внутри апомиктических линий данного рода растений. Однако ввиду обширной гибридизации среди *Boecheira spp.* обилие повторов в геноме этих растений может отражать не склонность к определенному типу размножения, а скорее их филогенетическую историю.

Последняя аннотация геномов *B. stricta* и *B. retrofracta* охватывает около 27000 генов у обоих видов. Присутствие несколько большего числа предсказанных для *B. stricta* транскриптов может быть объяснено отсутствием данных по экспрессии генов в тканях *B. retrofracta*, что сказалось на полноте аннотации в целом. Отчасти такая гипотеза подтверждается анализом BUSCO (Universal Single-Copy Orthologs — программа для поиска однокопийных ортологов эукариот).

Диплоидная сборка геномов апомиктов из рода *Boecheira* достаточно трудоемка, ввиду высокого уровня гетерозиготности (**Рисунок 4**), что является следствием их гибридогенного происхождения — произошло объединение относительно разнородного генетического материала (Beck et al., 2012). Например, уровень гетерозиготности генома *B. divaricarpa* составляет около 2.5% по оценке программы GenomeScope (Vurture et al., 2017). Из-за проблем, описанных выше, de novo сборка генома такого растения может быть сильно фрагментирована, в таком случае будет тяжело провести генетический анализ, претендующий на высокую точность. Кроме того, при сборке близких к идентичным последовательностей белок-кодирующих генов могут быть сформированы мозаичные скэффолды, не соответствующие ни одному из анализируемых вариантов. Высокий уровень фрагментации и подобные ошибки сборки могут негативно сказываться на правильности выводов об эволюции локусов, связанных с апомиксисом, и молекулярных механизмов, контролирующих его (Claros et al., 2012).

Ключевым вызовом является сложность сборки коротких прочтений в протяженные последовательности (контиги), которые в дальнейшем могут быть собраны в скэффолды, близкие по размерам к хромосомам. Еще одна непростая задача — гаплотипирование, процесс, в ходе которого каждой хромосоме из пары гомеологичных присваивают тот или иной вариант гена, локус, связанный с определенной функцией, или какой-либо другой

участок последовательности ДНК, происхождение которого играет важную роль в исследовании (Korbel & Lee, 2013).

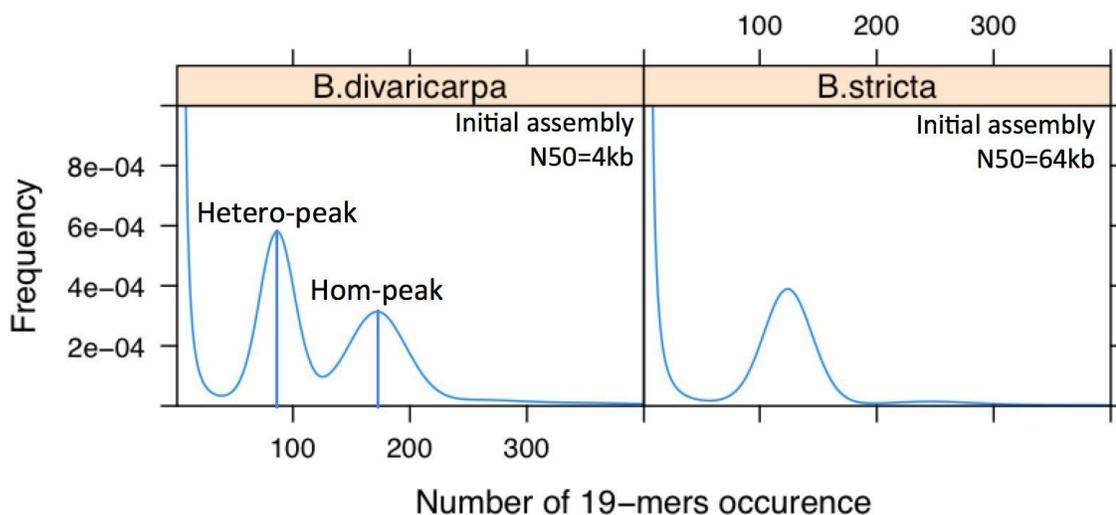


Рисунок 4. - График распределения покрытия различных 19-меров в геномах *B. stricta* и *B. divaricarpa*. По оси абсцисс отложено покрытие 19-меров (сколько раз каждый из них встретился в данных), по оси ординат — количество соответствующих K-меров. Распределение 19-меров в геноме *B. stricta*, имеющее только один пик, указывает на высокую гомозиготность генома, в то время как спектр 19-меров генома *B. divaricarpa* содержит два пика. По количеству вариантов 19-меров пики различаются приблизительно в два раза, поэтому предполагается, что первый пик, обладающий меньшей вариативностью K-меров и большей частотой их встречаемости, относится к гетерозиготной части генома, а второй — к гомозиготной. Показатели N50 основываются на сборке парных прочтений Illumina (покрытие x100) программой Platanus (Brukhin et al., 2019).

Совсем недавно были разработаны подходы, решающие проблему сборки высоко гетерозиготного генома. Особенное внимание исследователей привлекла технология секвенирования Pacific Bioscience, позволяющая получать длинные прочтения, а также алгоритмы, ориентированные на работу с ними — FALCON / FALCON-Unzip; с их помощью был собран гетерозиготный геном F1 гибрида *A. thaliana* и широко культивируемого *Vitis vinifera cv. Cabernet Sauvignon* (Chin et al., 2016). Дальнейшее развитие вышеописанных алгоритмов привело к появлению FALCON-Phase, нового метода, который собирает контиги, дополнительно используя короткие прочтения секвенатора Hi-C (технология данного секвенатора позволяет реконструировать карту пространственных взаимодействий хромосом в ядре клетки без специфических взаимодействий, не используя иммунопреципитацию). Такой комплексный подход позволяет сохранить высокую точность считывания нуклеотидов, не теряя при этом информации о том, какой именно хромосоме принадлежат конкретные гены и локусы

(Kronenberg et al., 2018). Технология секвенирования Linked-Read (“10x Genomics”) недавно была успешно применена для *de novo* сборки генома гетерозиготного гибридного диплоида *Capsicum annuum* (Hulse-Kemp et al., 2018).

Некоторое время назад были разработаны методы, использующие информацию о гаплотипе для сборки полиплоидных геномов с определением принадлежности регионов к конкретным хромосомам в случае, если известны родительские виды (Akama, Shimizu-Inatsugi, Shimizu, & Sese, 2014; Kyriakidou, Tai, Anglin, Ellis, & Stromvik, 2018). Тем не менее, видообразование в роду *Boechea* имеет очень запутанную историю — половые диплоиды породили множество апомиктических видов в ходе гибридизации, при которой формировались полиплоиды, аллоплоиды и анеуплоиды. Такое сложное видообразование привело к беспрецедентному разнообразию геномов апомиктических видов в этом роде, подкрепленному накоплением мутаций и повышенной активностью транспозонов у апомиктов (Ferreira de Carvalho, de Jager, van Gurp, Wagemaker, & Verhoeven, 2016; Lovell et al., 2017), из-за чего зачастую не удается установить родительские виды полиплоидов *Boechea spp.* Геномный анализ, основанный на референсном гаплоидном геноме, может не отражать реального положения вещей, особенно для апомиктов с высоко гетерозиготными диплоидными или полиплоидными геномами. В такой ситуации многие локусы могут полностью отсутствовать в референсном геноме, поскольку даже в идеальном случае он будет являться консенсусным.

Особенно выделяется «безреференсный» подход *de novo*, т. е. подход, при котором последовательности сравниваются без выравнивания друг на друга, очевидна его ценность в анализе геномов организмов, для которых не получен референс (Zielezinski, Vinga, Almeida, & Karlowski, 2017). Одним из наиболее популярных методов данного класса является метод сравнения частот К-меров или коротких последовательностей («слов»), но, к сожалению, его использование может быть затруднено при недостаточной глубине покрытия секвенирования или при отклонении в концентрации ДНК, содержащейся в подготовленных для анализа библиотеках. Наиболее подходящие данные для данного подхода позволяет получить технология секвенирования парных концов Illumina при работе с библиотеками, подготовленными методом случайного праймирования.

Современные разработки в области секвенирования геномов и их сборки предоставляют множество новых возможностей для исследования высоко гетерозиготных организмов, которыми и являются апомиктические гибриды из рода *Boechea*.

Глава 2. Материалы и методы исследования

2.1 Материал

В работе были использованы данные секвенирования геномной и хлоропластной ДНК двух гибридных апомиктных линий *Boechera divaricarpa*, предоставленные Центром геномной биоинформатики им. Ф. Г. Добржанского (Санкт-Петербургский Государственный Университет).

Материалом для геномного анализа послужили растения линий *Boechera divaricarpa ES517* и *Boechera divaricarpa M4B*, выращенные в помещениях Ботанического института им. В. Л. Комарова РАН. Образцы тканей были просеквенированы на платформах Illumina MiSeq и PacBio RS II, предоставивших короткие парные и длинные непарные прочтения соответственно (покрытие в обоих случаях 100x).

Все расчеты проводились с использованием вычислительных мощностей Центра геномной биоинформатики им. Ф. Г. Добржанского.

2.2 Методы работы с данными

2.2.1 Предобработка данных

Оценка качества «сырых» прочтений Illumina проводилась при помощи программы FastQC v0.11.7 (Andrews & others, 2016), после чего была проведена их коррекция

программой Trimmomatic (Bolger, Lohse, & Usadel, 2014), отрезались первые 20 нуклеотидов каждого рида.

Сбор статистики по ридам PacBio, с их последующей фильтрацией и предварительной сборкой, производился в ходе исполнения пайплайна, представленного в наборе инструментов SMRT-link. Данный пайплайн содержал следующие этапы работы:

Расчет ориентировочной длины генома, собираемого из представленных ридов

Фильтрация прочтений по качеству

Расчет средней длины ридов и сбор статистики N50 по ним

Расчет ожидаемых длин контигов, которым принадлежат случайно выбранные нуклеотиды

Подсчет общего числа оснований, содержащихся в отфильтрованных прочтениях

Расчет ожидаемого покрытия нуклеотидных оснований собираемого генома основаниями ридов

Оценка порогового значения длин ридов, фильтрация прочтений, обладающих меньшей длиной

Повторение действий с п.3 по п.6 для вновь отобранных ридов

Предварительная сборка исходных отфильтрованных прочтений в более длинные риды

Повторение действий с п.3 по п.6 для собранных ридов

Расчет покрытия предварительно собранных ридов прочтениями, полученными в п.7

Коррекция ошибок прочтений отобранных в п.7, базируемая на информации из пунктов 10-11, деление на более короткие последовательности тех ридов, в которых содержатся участки с «шумом», удаление таких участков.

Для оценки уровня гетерозиготности геномов исследуемых линий, в программе KrATeR v0.40 были построены спектры K-меров прочтений Illumina.

2.2.2 Примененные стратегии сборки генома

В ходе работы были протестированы два варианта сборки генома, после чего проводилось сравнение их качества со сборками, предоставленными коллегами из Центра геномной биоинформатики им. Ф. Г. Добржанского (**Рисунок 5**).

Реализованные в ходе работы стратегии:

1) Гибридная сборка генома de novo при помощи программы MaSuRCA (Maryland Super Read Cabog Assembler) v3.2.8 (Zimin et al., 2013) со стандартными параметрами и средним значением/стандартным отклонением размера вставки данных Illumina 500/100 (значения рекомендуются для безреференсной сборки). Сборка осуществляется одновременно из коротких парных и длинных непарных прочтений, при этом более качественные короткие риды Illumina картируются на длинные риды PacBio.

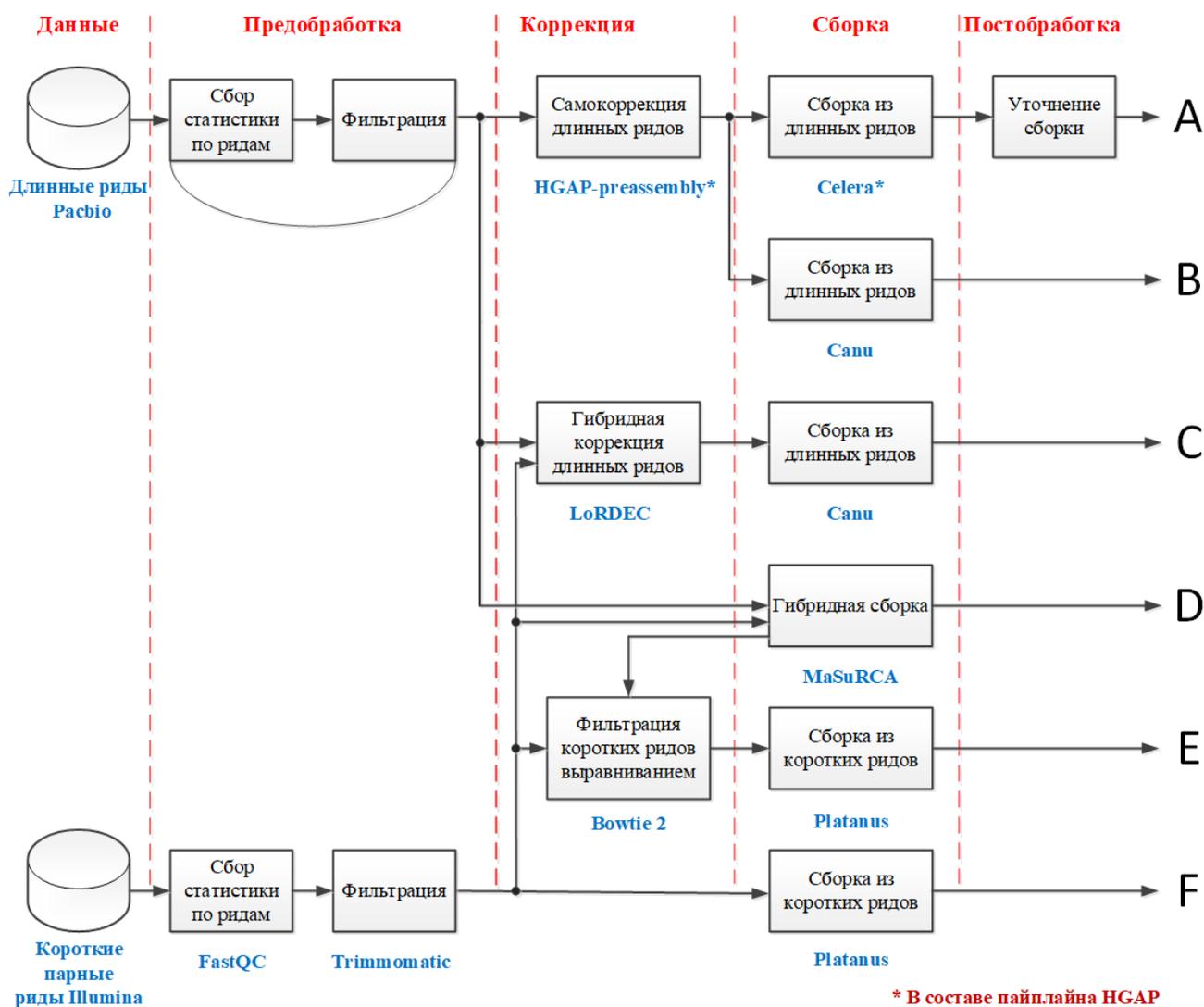


Рисунок 5. – Стратегии сборки, по которым проводилось сравнение. В ходе данной работы были реализованы сборки **D, E и F**. **A** – Сборка из длинных ридов пайплайном HGAP; **B** – Сборка из длинных ридов программой Canu; **C** – Сборка из длинных ридов, откорректированных короткими; **D** – Гибридная сборка из коротких и длинных ридов; **E** – Сборка из коротких ридов, отфильтрованных выравниванием на гибридную сборку; **F** – Сборка из исходных коротких ридов.

2) Сборка из исходных коротких ридов (после применения к ним Trimmomatic) при помощи программного обеспечения Platanus v1.2.4 (сборщик основан на построении графов де-Брюина) (Kajitani et al., 2014), производилась в три этапа:

- Сборка контигов. Строится граф де-Брюина, в узлах которого находятся K-меры, а грани представлены их перекрытиями, после чего из графа удаляются «пузыри» - разветвления топологии, возникшие в связи с различиями в парных хромосомах высокогетерозиготного организма (информация о них используется в последующих стадиях сборки).
- Сборка контигов в более протяженные скэффолды с использованием данных о парности прочтений и дальнейшее удаление пузырей, возникших на этом этапе, после чего происходит их обработка, и выбор наиболее подходящих вариантов топологии для каждого из таких участков.
- Закрытие разрывов, образовавшихся при редактировании и объединении контигов, полученных на предыдущих этапах.

Качество сборок оценивалось при помощи программного обеспечения QUAST-5.0.0, с использованием пакета BUSCO (Benchmarking Universal Single-Copy Orthologs, поиск универсальных однокопийных ортологов эукариот). Также готовые сборки были проверены на контаминации выравниванием алгоритмом BLAST на базу NCBI Nucleotide.

2.2.3 Удаление контаминаций из данных Illumina

Для удаления последовательностей, принадлежащих другим организмам, из библиотек ридов, короткие прочтения Illumina были выровнены на гибридную сборку MaSuRCA программой Bowtie 2 v2.3.4.3 (Langmead & Salzberg, 2012) с набором настроек “very-sensitive”. Отобранные риды были использованы для повторной сборки программой Platanus v1.2.4, после чего полученные контиги проверялись на контаминации при помощи выравнивания алгоритмом BLAST на базу данных NCBI Nucleotide. Помимо этого, по прочтениям, отфильтрованным от контаминирующих последовательностей, была построена новая версия спектра K-меров программой KrATER v0.40.

2.2.4 Подготовка к сборке хлоропластной ДНК

Фильтрация ридов, принадлежащих хлоропластной ДНК линии *M4B*, производилась методом выравнивания исходных библиотек на референсную последовательность хлоропластной ДНК линии *ES517*, собранную ранее сотрудниками Центра геномной биоинформатики Ф.Г. Добржанского. Прочтения PacBio выравнивались программой *bwa mem* (Burrows–Wheeler Aligner, были применены стандартные параметры), после чего выровненные последовательности были отобраны и сохранены в отдельный файл с использованием пакета *samtools*. Прочтения Illumina были выровнены программой *Bowtie 2 v2.3.5* с набором настроек «very-sensitive», после чего с использованием скрипта *hist* из пакета прикладных программ KAT (Mapleson, Garcia Accinelli, Kettleborough, Wright, & Clavijo, 2016) строились спектры K-меров отобранных ридов. K-меры с наиболее высоким покрытием (широко представленные в данных) были изъяты программой *filter kmer* из того же пакета, и далее по ним проходила финальная фильтрация коротких прочтений Illumina (отбирались риды, содержащие соответствующие K-меры).

2.2.5 Сборка и аннотация хлоропластной ДНК

Сборка хлоропластной ДНК осуществлялась программой *MaSuRCA v3.3.1* со стандартными настройками из отобранных длинных и коротких прочтений. Полученные контиги выравнивались алгоритмом *BLAST* на базу данных *NCBI Nucleotide*, с целью подтвердить их принадлежность хлоропластной ДНК. Для аннотации был выбран контиг с наибольшей длиной.

Аннотация кольцевой ДНК хлоропластов производилась веб-инструментом *GeSeq* (<https://chlorobox.mpimp-golm.mpg.de/geseq.html>) со следующими параметрами:

1. Общие опции
 - Создание multi-GenBank файла
 - Создание multi-GFF3 файла
 - Создание multi-FASTA файла
 - Привести выравнивание по кодонам
2. BLAT search

- Аннотировать инвертированные повторы пластид
- Настройки поиска — стандартные
- 3. HMMER profile search
 - Использовать набор настроек для работы с хлоропластами Высших растений
- 4. ARAGORN v1.2.38
 - Генетический код — Бактериальный/хлоропластов растений
 - Максимальная длина интрона — 3000 пар оснований
- 5. tRNAscan-SE v2.0.3
 - Источник последовательности — тРНК органелл
 - Режим поиска — стандартный
 - Генетический код — универсальный
 - Пороговое значение для записи найденных тРНК — 15
 - Отключить проверку на псевдогены
 - Вывести детальный отчет о предсказанных генах
 - Отображать происхождение первых хитов
- 6. BLAT Reference Sequences
 - Использовать MPI-MP референсы хлоропластов.

2.2.6 Методы филогенетического анализа с использованием хлоропластной ДНК

Филогенетический анализ проводился по двум молекулярным маркерам хлоропластной ДНК — межгенному спейсеру, находящемуся между генами *trnF* и *trnL*, и интрону *trnL*. Помимо рассматриваемых линий *M4B* и *ES517*, в анализе были использованы данные некоторых других представителей рода *Boechea* (последовательности генетического материала их хлоропластов имеют следующие ID в базе NCBI Nucleotide: DQ013045.1, DQ013046.1, DQ013047.1, DQ013048.1, DQ013049.1, DQ013050.1, DQ013051.1, DQ013052.1, DQ013053.1, DQ013054.1, DQ013055.1, DQ013056.1, DQ013057.1,

DQ013058.1, DQ013059.1, DQ013060.1), а также видов *Halimolobos parryii* (AF307539.1) и *Polycstenium fremontii* (AF183043.1). Хлоропластная ДНК вида *Boecheera retrofracta* также была предоставлена коллегами из Центра геномной биоинформатики им. Ф. Г. Добржанского. Все последовательности вышеупомянутых маркеров были выровнены алгоритмом множественного выравнивания MUSCLE, в качестве подхода к кластеризации был выбран метод присоединения соседей, штраф за открытие разрыва - «-400», штраф за продолжение разрыва - «0». Филогенетический анализ полученного выравнивания проводился в программе MrBayes v3.2.6 (Ronquist & Huelsenbeck, 2003) с использованием следующих настроек: 6 Марковских цепей, 6 миллионов генераций, изъятие дерева каждые 100 генераций, модель эволюции — GTR+I+G.

Полученное дерево было визуализировано в RStudio с подключенным модулем Ape.

2.2.7 Методы филогенетического анализа с использованием последовательности гена APOLLO

Для проведения филогенетического анализа по гену APOLLO из базы данных NCBI Nucleotide были загружены все последовательности данного гена, опубликованные Cogral и соавторами (ID's в базе NCBI Nucleotide: KF705570.1, KF705577.1, KF705588.1, KF705590.1, KF705592.1, KF705600.1, KF705601.1, KF705602.1). При помощи алгоритма BLAST данные последовательности были выровнены на геном *Boecheera stricta* (загружен с интернет ресурса Phytozome, <https://phytozome.jgi.doe.gov>), сборку *M4B*, а также на предоставленные коллегами из Центра геномной биоинформатики им. Ф. Г. Добржанского сборки *Boecheera divaricarpa ES517* и *Boecheera retrofracta*. По итогам работы алгоритма скриптом, написанном на языке Python, из приведенных геномных данных извлекались участки, соответствующие локусу APOLLO. Для построения более подробного филогенетического дерева из базы NCBI Nucleotide были также загружены последовательности гена, ортологичного данному, принадлежащие ряду близкородственных видов (их ID's в базе NCBI Nucleotide: NM_001334644.1, XM_006390370.2, XM_021035177.1, XM_023787194.1)

Множественное выравнивание последовательностей производилось алгоритмом MUSCLE, после чего они были подвергнуты фильтрации программой Gblocks v0.91b с использованием параметра «Использовать менее строгую фильтрацию фланкирующих позиций». Итоговое выравнивание анализировалось в программе MrBayes v3.2.6 со

следующим набором настроек: 6 Марковских цепей, 6 миллионов генераций, дерево изымалось каждые 100 генераций, модель эволюции — GTR+I+G.

Полученный вывод программы был визуализирован пакетом `ape` в Rstudio.

Глава 3. Результаты и обсуждение

3.1 Анализ и коррекция исходных прочтений

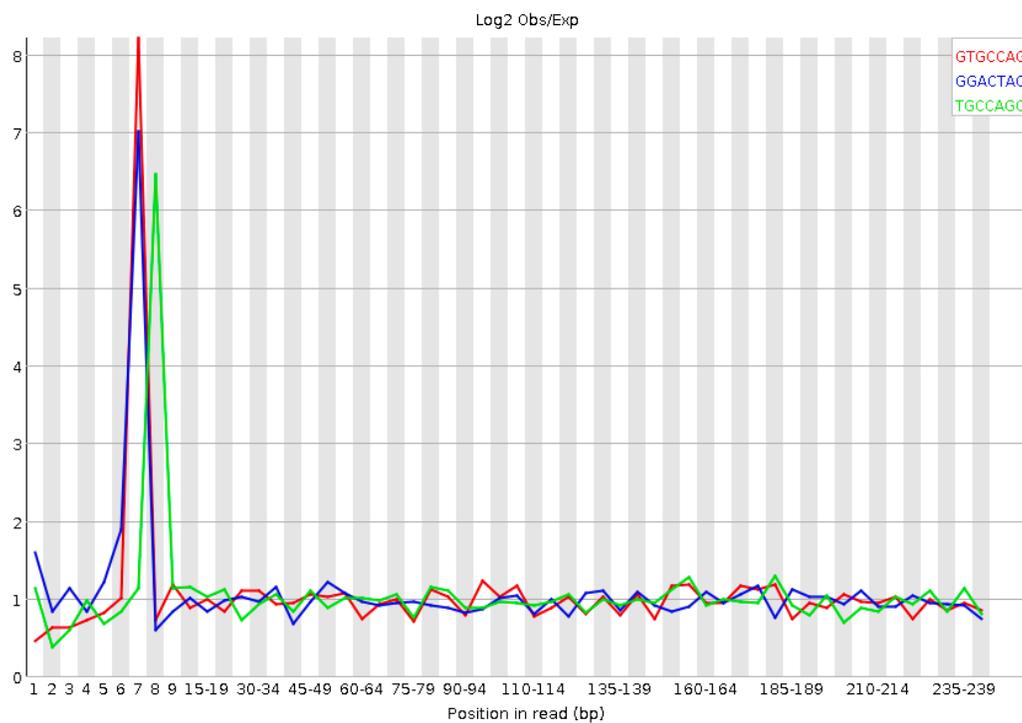
В ходе оценки качественных характеристик ридов Illumina программой FastQC v0.11.7 была выявлена проблема перепредставленности ряда К-меров в данных, принадлежащих *Boecheira divaricarpa* линии M4B (Рисунок 6). Все излишне часто встречающиеся К-меры наблюдались в области 20-ти первых нуклеотидов каждого прочтения, из чего был сделан вывод о том, что из полученных прочтений не были удалены адаптерны последовательности. Данные нуклеотиды обрезались программой Trimmomatic v0.32, после чего был проведен повторный анализ FastQC, не показавший проблем.

В отчете, сформированном высоко автоматизированной утилитой обработки ридов PacBio SMRT-link не содержалось сообщений об ошибках коррекции.

Анализ спектров К-меров данных Illumina, визуализированных в программе KrATER v0.40, показал, что геномы линий ES517 и M4B обладают высоким уровнем гетерозиготности (Рисунок 7). На это указывают два отчетливых, сопоставимых по размеру пика, наблюдаемые в обоих спектрах, но в случае линии ES517 К-меры, формирующие эти пики, обладают достаточно высоким покрытием, тогда как в спектре линии M4B они частично маскированы большим количеством низко покрытых последовательностей, которые можно объяснить наличием в библиотеках секвенирования

следовых количеств длинных молекул ДНК, принадлежащих другим организмам (контаминации с чужеродным генетическим материалом).

а) **Kmer Content**



б) **Kmer Content**

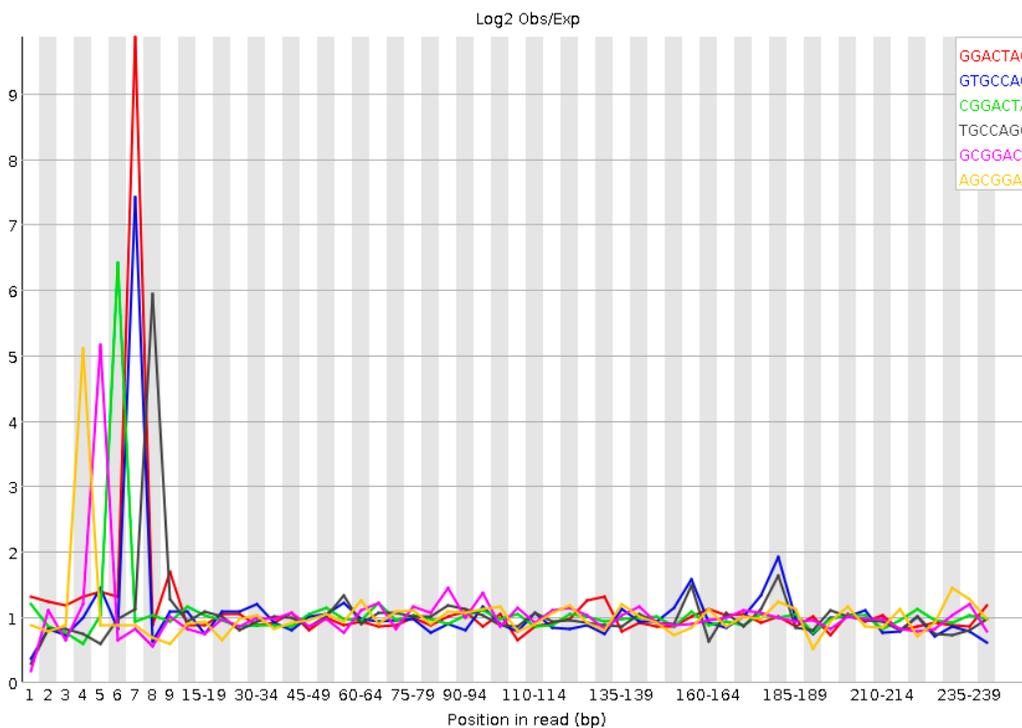


Рисунок 6. - Графики содержания К-меров в библиотеках прямых (а) и обратных (б) прочтений *Shuttle* линии *M4B*. Из графиков следует, что некоторые К-меры встречаются в данных намного чаще остальных.

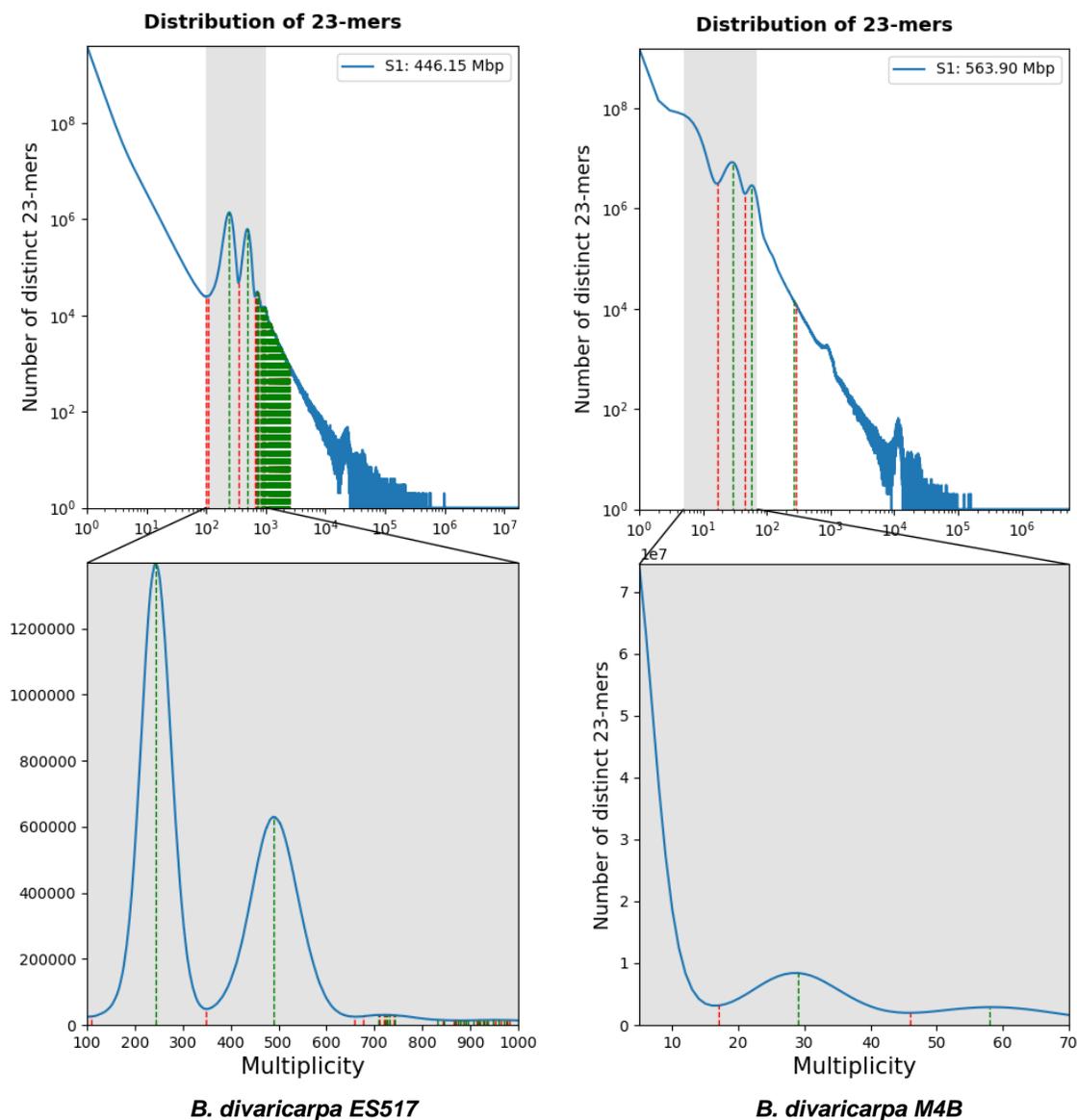


Рисунок 7. – Спектры К-меров ридов *Shumina*, принадлежащие геному *Bescheria divaricarpa* M4B. Оба графика демонстрируют по 2 пика, один из которых приходится на зону слабо покрытых ридов (гетерозиготная часть генома), а другой – на зону, представляющую К-меры с высоким покрытием ридов (гомозиготная часть генома).

Помимо того, стоит обратить внимание на соотношение высоты пиков каждого спектра: в обоих случаях первый пик выше второго, что, также свидетельствует о высокой гетерозиготности исследуемого генома. Полученные данные были учтены в дальнейшей работе с «сырыми» ридами — предпочтение отдавалось подходам, решающим проблемы сборки высоко гетерозиготных геномов.

3.2 Выбор наилучшей стратегии сборки высоко гетерозиготного генома *Boechera divaricarpa* M4B

Данные, полученные в ходе анализа библиотек прочтений, были учтены при выборе подходов к сборке геномов — предпочтение отдавалось методам, решающим проблемы генетического анализа высоко гетерозиготных организмов. В рамках данной работы было решено протестировать две программы для сборки генома, удовлетворяющие этому условию — MaSuRCA v3.2.8 и Platanus v1.2.4, и провести сравнительный анализ итогов их работы со сборками, предоставленными сотрудниками Центра геномной биоинформатики им. Ф. Г. Добржанского.

3.2.1 De novo сборка генома *Boechera divaricarpa* M4B программой MaSuRCA

При помощи программы MaSuRCA v3.2.8 из коротких и длинных ридов была произведена гибридная сборка генома *Boechera divaricarpa* линии M4B.

Короткие риды Illumina обладают высоким качеством, в то время как риды PacBio во много раз превосходят их по длине. Алгоритм, лежащий в основе сборщика MaSuRCA, позволяет объединить преимущества обоих типов данных — длинные риды PacBio являются референсом для картирования на них коротких ридов Illumina, в следствие чего появляется возможность избежать образования “пузырей” в графах де-Брюина при формировании контигов, не жертвуя при этом точностью секвенирования. Стоит отметить, что, не смотря на использование длинных прочтений, сборка, полученная данным методом является консенсусной.

В ходе анализа данной сборки генома линии M4B программой QUAST-5.0.0 были собраны следующие метрики:

- Число контигов — 9065
- Общая длина (в парах оснований, учтены только контиги длиннее 500 п.н.) - 392375948
- Наибольший контиг (в парах оснований) — 4529469

- N50 (в парах оснований, данный показатель отражает длину контига, который, вместе со всеми контигами большей длины, присутствующими в сборке, покрывает 50% генома или более) — 396913
- Метрика Complete BUSCO (процент найденных полных ортологов эукариот) — 90.76
- Метрика Partial BUSCO (процент частично собранных ортологов эукариот) — 2.31

3.2.2 De novo сборка генома *Boechera divaricarpa M4B* программой Platanus

Сборщик Platanus v1.2.4 также был выбран для работы, поскольку он разработан специально для de novo сборки геномов высоко гетерозиготных диплоидных организмов из коротких прочтений. Его алгоритм включает несколько стадий, позволяющих эффективно устранять «пузыри» из построенных графов де-Брюина, а также стадию, в которой используется информация о парности прочтений, благодаря чему сокращается вероятность неправильного объединения контигов в скэффолды.

Анализ программой QUAST-5.0.0 предоставил следующие сведения о сборке *B. divaricarpa M4B*, произведенной Platanus:

- Число контигов — 105470
- Общая длина (в парах оснований, учтены только контиги длиннее 500 п.н.) - 266829223
- Наибольший контиг (в парах оснований) — 132076
- N50 (в парах оснований) — 4610
- Метрика Complete BUSCO (процент найденных полных ортологов эукариот) — 71.62
- Метрика Partial BUSCO (процент частично собранных ортологов эукариот) — 13.86

3.2.3 Сравнительная характеристика различных стратегий сборки высоко гетерозиготного генома *Boechera divaricarpa M4B*

Статистические показатели двух вышеупомянутых сборок генома, полученные с помощью программы QUAST-5.0.0, сравнивались между собой, а также с показателями сборок, предоставленных коллегами из Центра геномной биоинформатики им. Ф. Г.

Добржанского (**Таблица 1**). Как следует из приведенной таблицы, наилучший результат по метрике N50 показала сборка программы MaSuRCA v3.2.8, в то же время она заняла второе место по суммарной длине получившихся контигов и незначительно отстала от лидера по проценту найденных BUSCO ортологов. Наивысшие показатели суммарной длины сборки и найденных однокопийных ортологов эукариот принадлежат сборке генома, проведенной с помощью Canu. К сожалению, значение N50 и величина самого длинного контига данной сборки значительно меньше, чем таковые у сборки MaSuRCA, что свидетельствует о большей дискретности контигов сборки Canu. В исследованиях молекулярных механизмов регуляции апомиксиса у высоко гетерозиготных гибридов принадлежность конкретных аллелей генов геному того или иного родителя может играть ключевую роль, поэтому для анализа следует использовать контиги, наиболее близкие к реконструкциям полных хромосом. В связи с вышесказанным, сборка, осуществленная программой MaSuRCA, была определена как самая эффективная, поскольку ее статистические показатели говорят о высоком покрытии собранного генома контигами, превышающими по размерам контиги всех прочих сборок.

Выравнивание данной сборки алгоритмом BLAST на базу данных NCBI Nucleotide не выявило явных контаминаций с генетическим материалом других организмов, что также свидетельствует в пользу ее высокого качества.

3.2.4 Очистка исходных коротких прочтений генома *Boechera divaricarpa* M4B выравниванием на наиболее успешную сборку

Исходные короткие прочтения Illumina генома линии M4B, необходимые для дальнейшей работы, были отфильтрованы от контаминирующих чужеродных последовательностей путем выравнивания программой Bowtie 2 v2.3.4.4 на референсную последовательность, в роли которой выступала сборка генома, полученная в результате исполнения программы MaSuRCA v3.2.8. Из отобранных данных была произведена сборка генома программой Platanus v1.2.4, затем проводился анализ полученной сборки программами QUAST-5.0.0 и BLAST. Отчет программы QUAST содержал следующую информацию:

- Число контигов — 97679
- Общая длина (в парах оснований, учтены только контиги длинее 500 п.н.) - 251482452
- Наибольший контиг (в парах оснований) — 132084

- N50 (в парах оснований) — 4621
- Метрика Complete BUSCO (процент найденных полных ортологов эукариот) — 66.67
- Метрика Partial BUSCO (процент частично собранных ортологов эукариот) — 16.67

Как можно заметить, по сравнению с первой сборкой *Platanus* уменьшились значения таких показателей, как число контигов, суммарная длина всех собранных последовательностей, процент полностью собранных однокопийных ортологов эукариот, в то время как N50 и длина наибольшего контига получили лишь незначительный прирост. Два последних показателя свидетельствуют о почти не изменившемся количестве длинных контигов, а также об их размерах, сохранивших прежние пределы. Из этого можно сделать вывод о том, что падение качества, вероятно, вызвано исключением из сборки прочтений генетического материала других организмов, поскольку они, будучи низко покрытыми, собираются преимущественно в небольшие контиги, не влияющие на значение N50 и величину самой длинной собранной последовательности.

Выравнивание полученных контигов алгоритмом BLAST не показало значимых контаминаций, из чего можно сделать вывод об успешном изъятии посторонних данных.

Чтобы узнать, увеличилось ли после фильтрации покрытие ридов, принадлежащих геному *Boechera divaricarpa* линии *M4B*, по новым данным был построен спектр К-меров в программе KrATeR v0.40 (**Рисунок 8**). На графике наблюдается заметный прирост покрытия К-мерами ридов, относящихся к геномной ДНК (два пика в левой части графика). Такое демаскирование также свидетельствует об эффективном удалении чужеродных прочтений.

Исходные данные	Основные программы	Число контигов	Длина наибольшего контига (в парах нуклеотидов)	N50 (в парах нуклеотидов)	Суммарная длина (в парах нуклеотидов)	Complete BUSCO (%)	Partial BUSCO (%)
Исходные короткие риды	Platanus	105470	132076	4610	266829223	71,62	13,86
Исходные короткие и длинные риды (гибридная сборка)	MaSuRCA	9065	4529469	396913	392375948	90,76	2,31
Исходные длинные риды	SMRT-link, Canu	4106	2080822	225982	410681519	91,42	0,99
Исходные длинные риды	SMRT-link, HGAP	2223	2203066	327094	325620120	88,12	3,63
Откорректированные длинные риды	LoRDEC, Canu	20883	181615	6410	111378385	79,21	4,29

Таблица 1. Статистические показатели сборок генома *Boechera divaricarpa* M4B

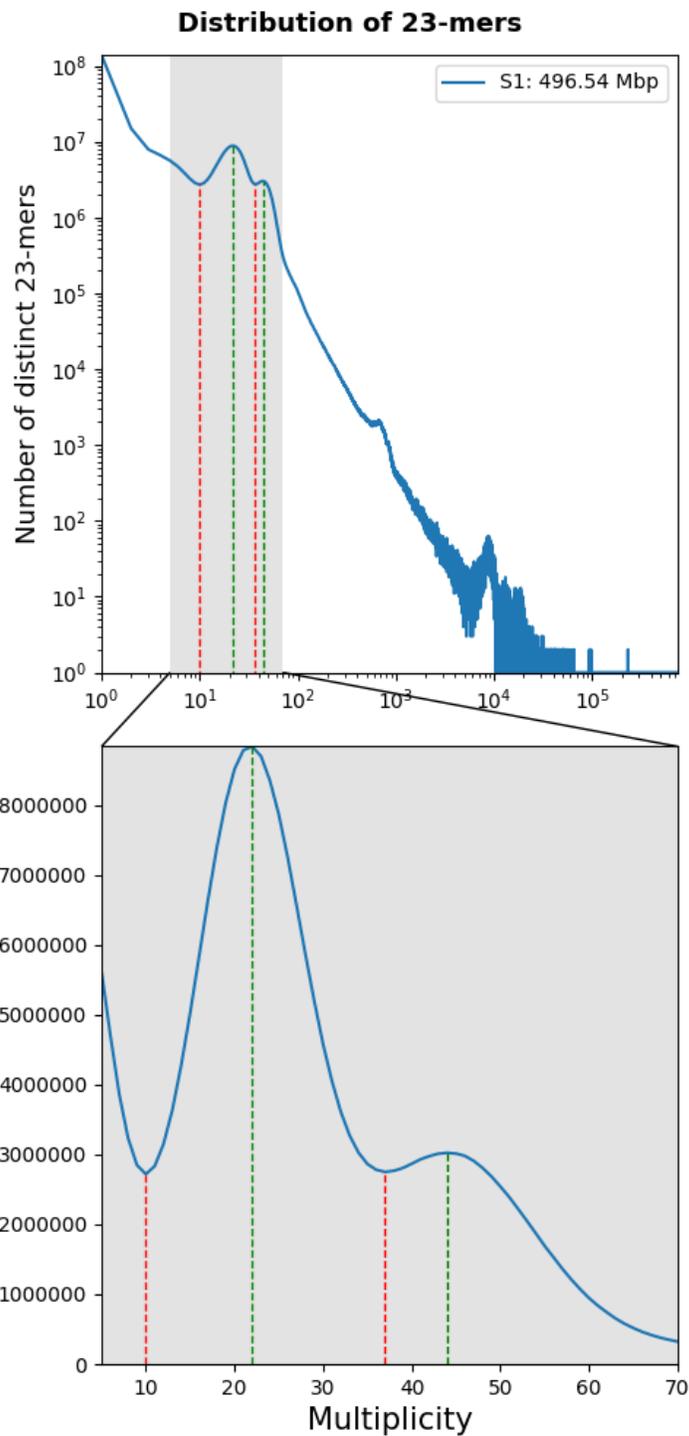


Рисунок 8. – Спектр К-меров прочтений Illumina, отфильтрованных выравниванием на референсную последовательность - сборку геномной ДНК *Boechera divaricarpa M4B*, произведенную программой MaSuRCA. На графике наблюдается возросшее покрытие (по сравнению с рис.7 *B. divaricarpa M4B*) на участках, относящихся к геномной ДНК.

3.3 Анализ хлоропластной ДНК *Boechera divaricarpa* линий M4B и ES517

3.3.1 Подготовка прочтений для сборки хлоропластной ДНК

Для сборки хлоропластной ДНК была выбрана программа MaSuRCA, поскольку она показала наиболее высокие результаты по сравнению с другими сборщиками. Ввиду этого понадобилось отделить как короткие, так и длинные прочтения генетического материала пластид от ридов, принадлежащих ядерной ДНК.

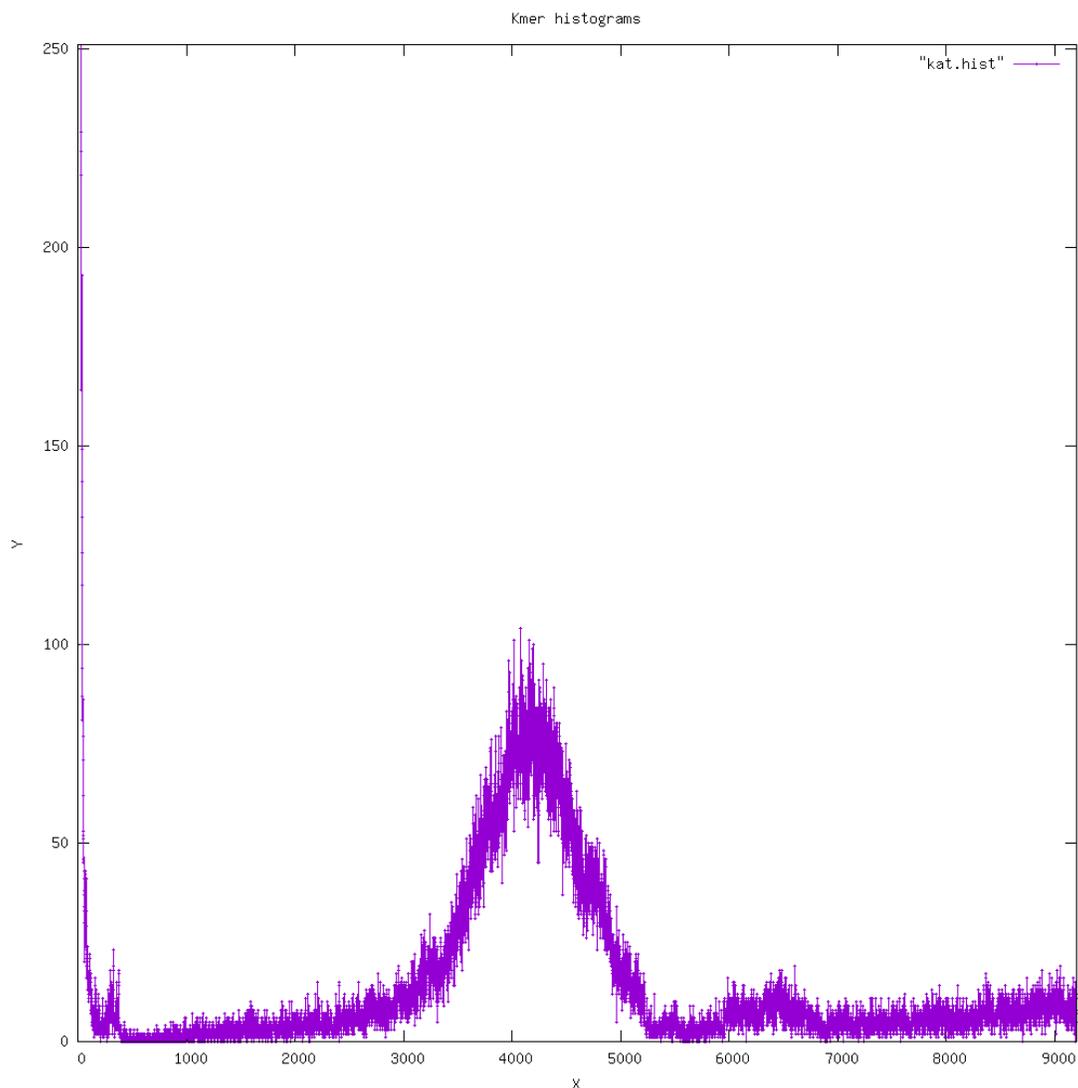


Рисунок 9. – Спектр К-меров для отфильтрованных прямых ридов Illumina.

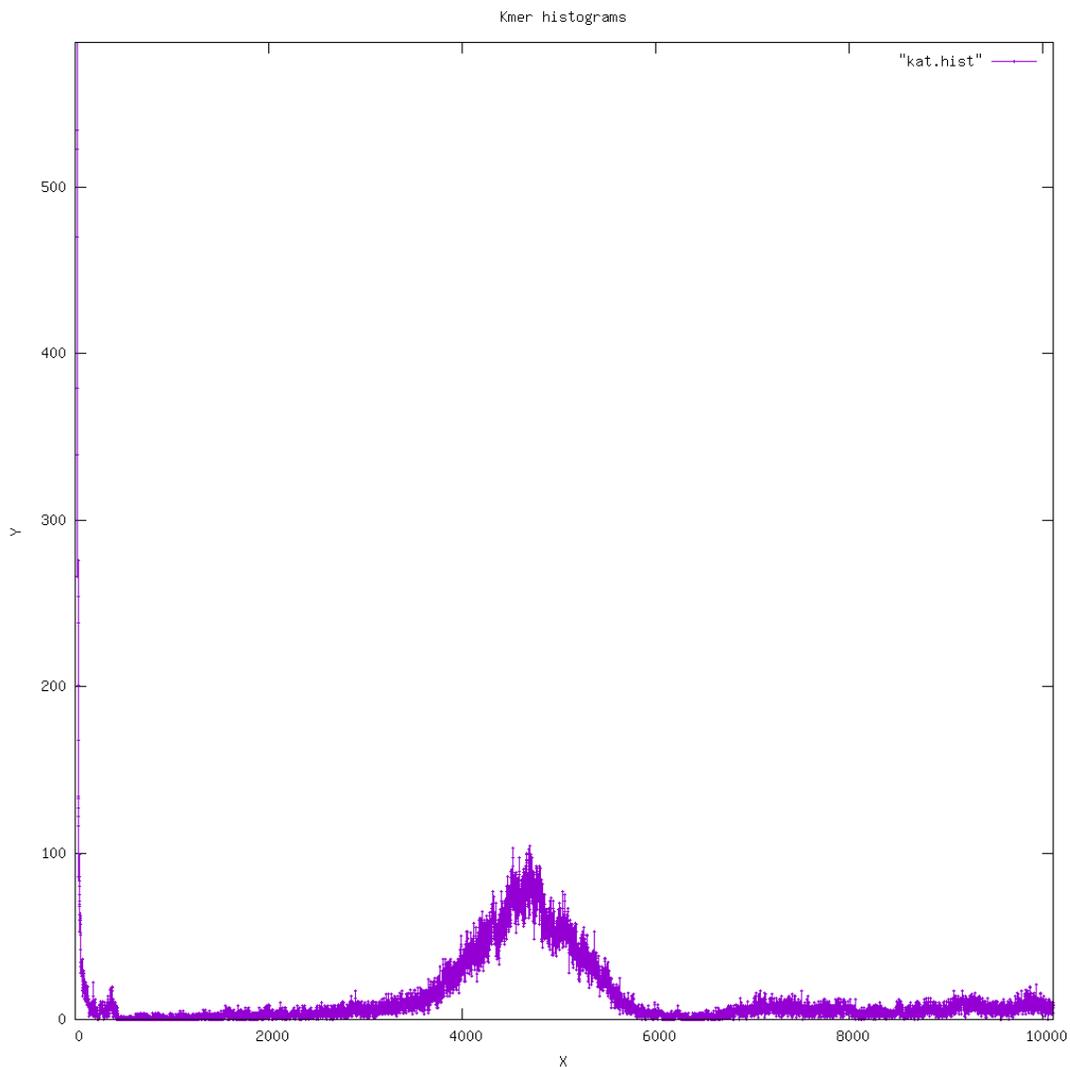


Рисунок 10. – Спектр К-меров для отфильтрованных обратных ридов *Shumina*.

Референсной последовательностью для отбора необходимого материала послужила сборка хлоропластной ДНК растения *Boechera divaricarpa* линии *ES517*, которая была предоставлена сотрудниками Центра геномной биоинформатики им. Ф.Г. Добржанского. Риды *Shumina* выравнивались на данную последовательность программой Bowtie 2, и далее с помощью программ, входящих в пакет KAT, по отфильтрованным прочтениям были построены спектры К-меров (**Рисунки 9, 10**). После анализа полученных графиков были отобраны К-меры с покрытием от 3500 до 6000 для библиотеки прямых прочтений, и с покрытием от 3000 до 5500 для библиотеки обратных, поскольку в этих границах лежали пики их покрытия. В дальнейшем из библиотек прочтений были изъяты риды, которым принадлежали отобранные К-меры.

Длинные прочтения PacBio также фильтровались выравниванием на референсную последовательность хлоропластной ДНК линии *ES517*.

3.3.2 Реконструкция хлоропластной ДНК и поиск молекулярных маркеров для филогенетического анализа

В ходе сборки ДНК из вновь отфильтрованных ридов было получено 2 контига, длиной 139622 и 95646 пар нуклеотидов, после чего путем выравнивания обоих контигов на базу данных NCBI Nucleotide была установлена их принадлежность хлоропластной ДНК. Для аннотации был выбран контиг длиной 139622 п. н., поскольку его размер в большей степени соответствовал длине хлоропластной ДНК, свойственной близкородственным организмам (~154000 п. н.).

Аннотация выбранной последовательности, выполненная с помощью инструмента GeSeq, дала следующие результаты (**Рисунок 11.**):

Алгоритм ARAGORN обнаружил 36 генов тРНК хлоропластов

Алгоритм tRNAscan-SE обнаружил 36 генов тРНК хлоропластов

Алгоритм hmmer обнаружил 111 генов хлоропластной ДНК

Для филогенетического анализа были выбраны участки контига, принадлежащие межгенному спейсеру *trnF-trnL* и интрону гена *trnL* (координаты соответствующих участков были изъяты из файла, сформированного в ходе аннотации).

3.3.3 Определение донора хлоропластов для растений *Boechera divaricarpa* линий *M4B* и *ES517*

Гибридогенная природа *Boechera divaricarpa* подразумевает, что генетический материал хлоропластов растений этого вида наследуется от одного из видов-родителей (*B. stricta* или *B. retrofracta*). Чтобы узнать, каким половым видам принадлежит хлоропластная ДНК растений *B. divaricarpa M4B* и *B. divaricarpa ES517*, был проведен филогенетический анализ по последовательностям спейсера *trnF-trnL* и интрона гена *trnL* (**Рисунок 12.**).

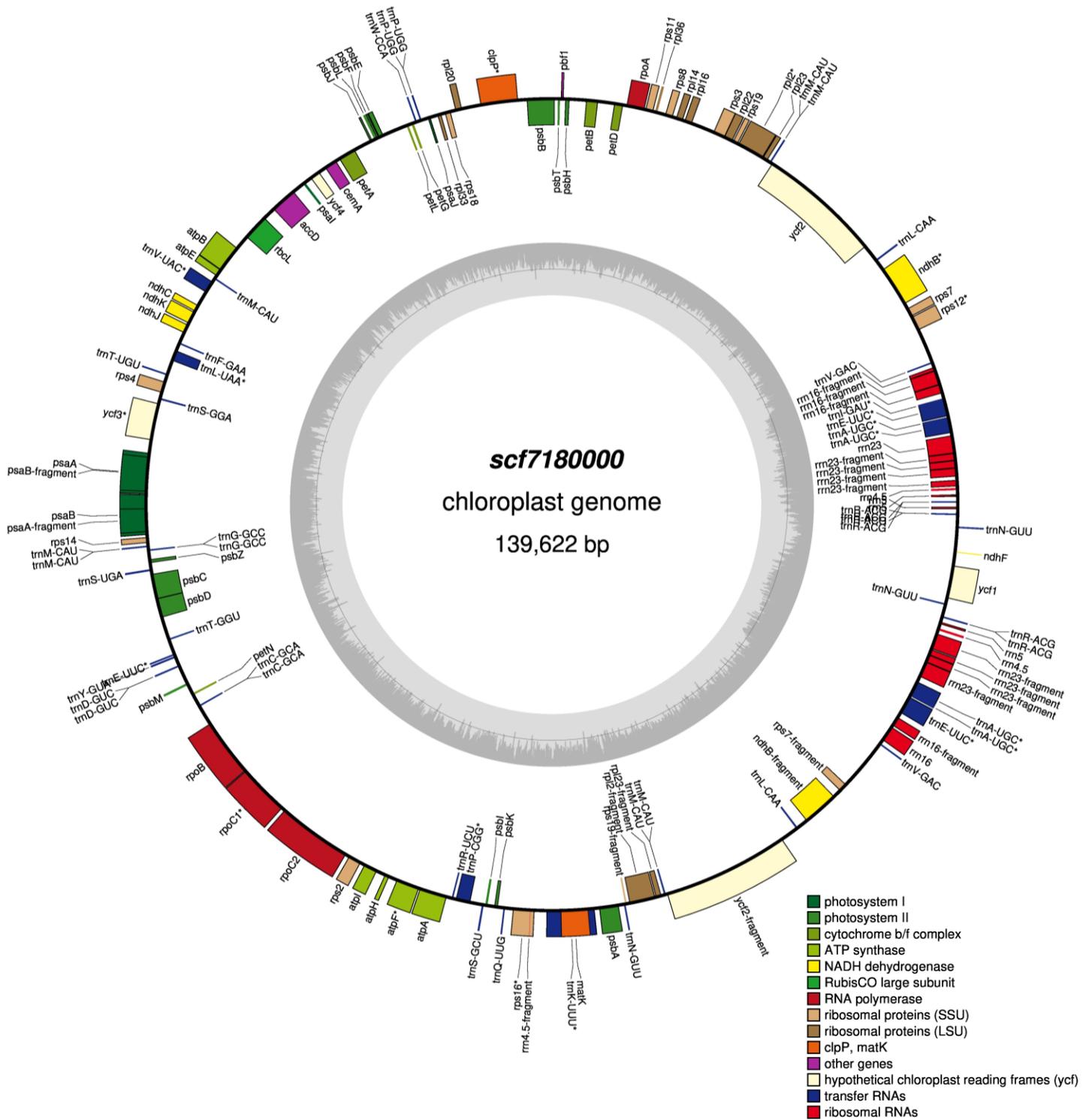


Рисунок 11. – Визуализация аннотации хлоропластной ДНК *Boechera divaricarpa* M4B, собранной программой MaSuRCA. Для близкородственного филогенетического анализа наибольший интерес представляет регион, содержащий гены *trnF* и *trnL*.

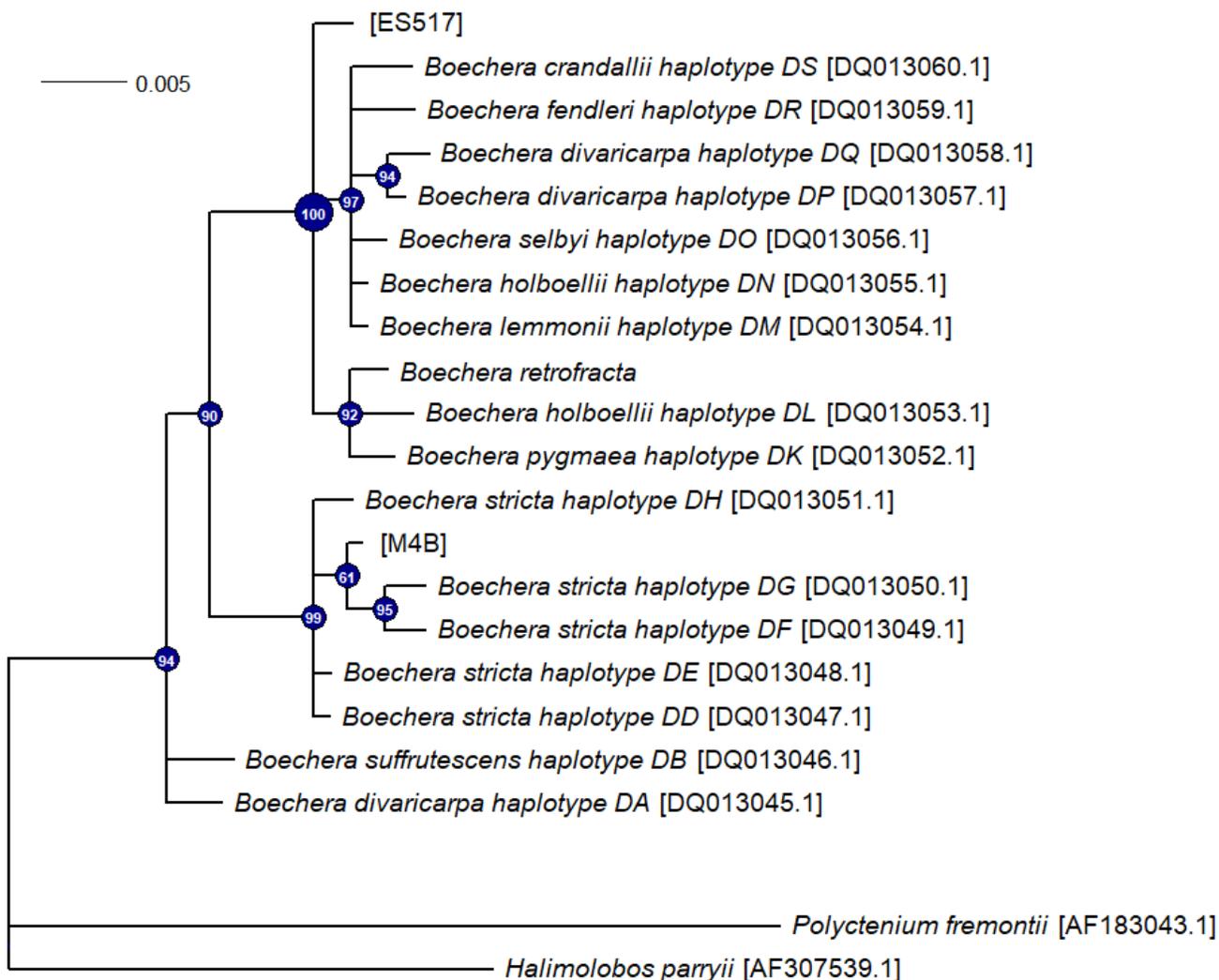


Рисунок 12. – Филогенетическое дерево, построенное на базе Байесовского анализа последовательностей спейсера trnF-trnL и интрона trnL. Масштабная шкала отражает количество замен, приходящихся на один сайт, значения в узлах – Байесовские вероятности.

Из топологии приведенного дерева можно сделать вывод, что генетический материал хлоропластов линии *M4B* был унаследован от *B. stricta*, поскольку эта линия попала в одну кладу с представителями данного вида.

Построенное дерево свидетельствует о большем родстве последовательностей генетических маркеров линии *ES517* с таковыми вида *Boechera retrofracta*, чем с последовательностями, принадлежащими *B. stricta*, что служит основанием для гипотезы о наследовании пластидной ДНК от первого из указанных видов. Дальнейший

биоинформатический анализ гаплотипов иных представителей рода *Boechea* может привести к лучшей кластеризации *B. divaricarpa ES517* и *B. retrofracta*.

3.4 Анализ ядерной ДНК *Boechea divaricarpa* линий *ES517* и *M4B*

В ходе выравнивания различных вариантов последовательностей гена APOLLO на сборки *B. divaricarpa ES517* и *B. divaricarpa M4B*, были отобраны контиги, на которые последовательности выравнивались полностью и с высокой идентичностью (>97%). В случае линии *ES517* был найден лишь один такой участок, в то время как в сборке генома линии *M4B* помимо одной цельной последовательности был обнаружен ряд частично покрывающих последовательность искомого гена. Участки контигов, на которых наиболее успешно прошло выравнивание, были изъяты для дальнейшего анализа. Таким же образом были отобраны филогенетические маркеры из референсной последовательности *B. stricta* и сборки *B. retrofracta*, проведенной сотрудниками Центра геномной биоинформатики им. Ф. Г. Добржанского, после чего все полученные последовательности гена APOLLO и загруженные дополнительно последовательности генов ортологов были подвергнуты множественному выравниванию, очистке и последующему анализу в программе MrBayes (Рисунок 13.).

Важно отметить, что последовательности *Boechea sp.* 1A2 S7 (KF705602.1), *Boechea sp.* 33A2 S6 (KF705570.1), *Boechea sp.* 11A2 S8 (KF705601.1) и *Boechea sp.* 111A2 S13 (KF705577.1) принадлежат половым аллелям гена APOLLO, тогда как последовательности *Boechea sp.* 1A2 A3 (KF705590.1), *Boechea sp.* 33A2 A5 (KF705588.1), *Boechea sp.* 11A2 A1 (KF705600.1) и *Boechea sp.* 111A2 A8 (KF705592.1) представляют собой апомиктические аллели этого гена (их экспрессия ощутимо коррелирует с проявлением апомиксиса). Полученное филогенетическое дерево демонстрирует значительную кластеризацию гена линии *ES517* с апомиктическими аллелями *Boechea sp.*, что, вероятно, свидетельствует о ее гомозиготности по данному признаку, поскольку, как упоминалось выше, сборка MaSuRCA является консенсусной, а значит последовательность гена APOLLO в ней является усредненной версией всех покрывающих ее ридов. В то же время оба родительских вида данной линии, *B. retrofracta* в большей степени, и *B. stricta* в меньшей, кластеризовались с половыми аллелями гена. Отдельно стоит отметить положение гена линии *M4B*,

тяготеющего к половым аллелям APOLLO. Исходя из тех же соображений о консенсусе сборки и сведений о присутствии в ней наряду с проанализированной последовательностью участков, частично покрывающих референсный аллель, такая кластеризация может свидетельствовать о гетерозиготности данной линии по аллелям гена APOLLO. Тем не менее, возможное присутствие дополнительных копий гена могло повлиять на результаты консенсуса, и поэтому для подтверждения гипотез о составе аллелей APOLLO в геномах данных линий требуются дополнительные экспериментальные данные, способные доказать однокопийность гена.

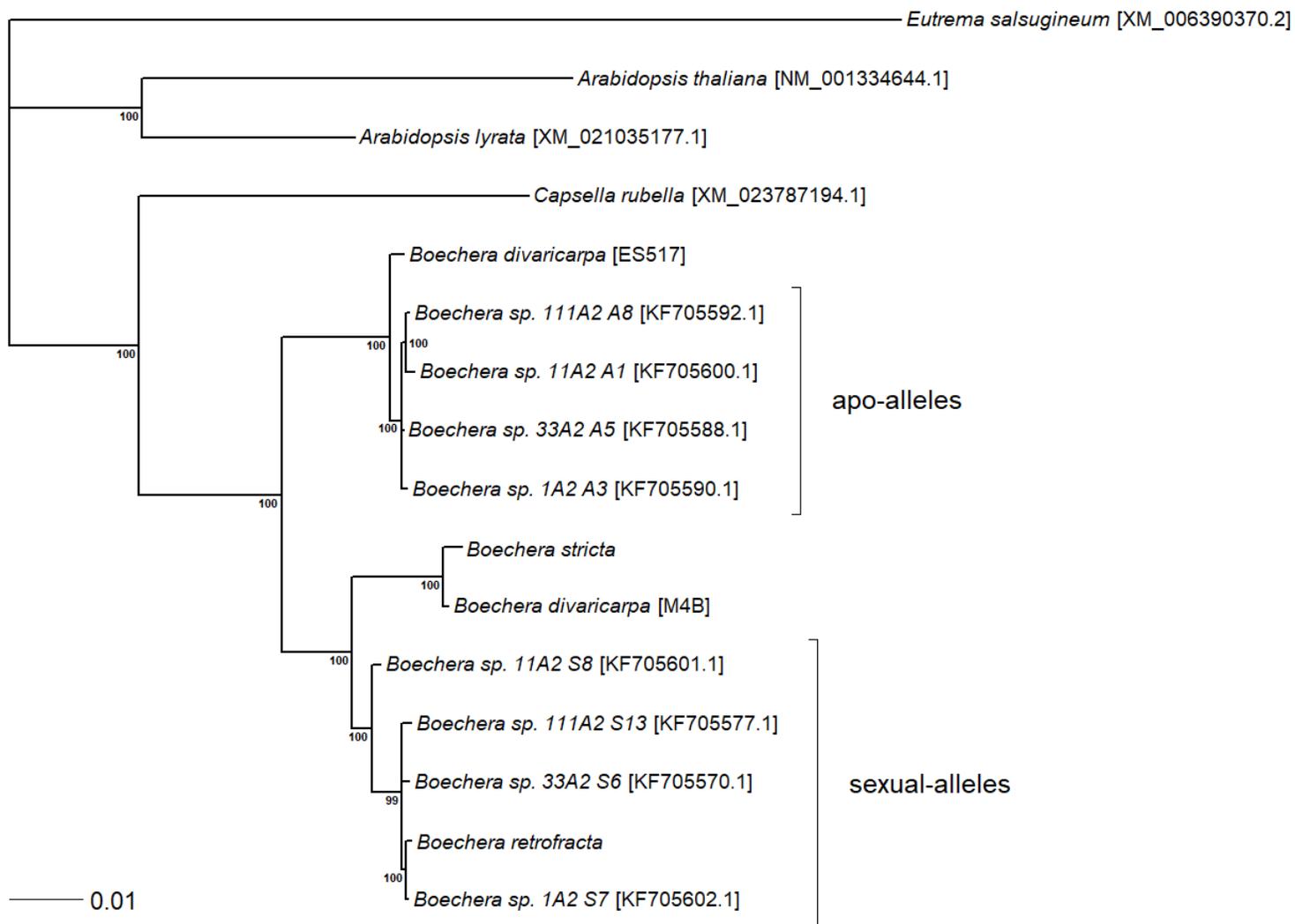


Рисунок 13. - Филогенетическое дерево, построенное на базе Байесовского анализа последовательностей гена APOLLO и его ортологов. Масштабная шкала отражает количество замен, приходящихся на один сайт, значения в узлах – Байесовские вероятности.

Заключение

В работе были реализованы различные стратегии сборки высоко гетерозиготных геномов, с дальнейшей статистической обработкой их результатов. Полученные сборки сравнивались между собой и со сборками, предоставленными сотрудниками Центра геномной биоинформатики им. Ф. Г. Добржанского, в ходе чего была определена наиболее эффективная методика для работы с сырыми прочтениями геномов высоко гетерозиготных организмов – гибридная сборка из длинных и коротких прочтений программой MaSuRCA.

С использованием отработанного метода гибридной сборки ДНК, был реконструирован генетический материал хлоропластов *Boechea divaricarpa M4B* и проведен его филогенетический анализ с использованием Байесовской статистики. В ходе анализа были получены уникальные данные о наследовании хлоропластной ДНК линии *M4B* от вида *Boechea stricta*, а также продемонстрировано, что линия *ES517* унаследовала генетический материал пластид не от представителей этого вида. Чтобы уточнить филогению хлоропластной ДНК линии *ES517* требуется рассмотреть большее количество изолятов различных видов, и, в частности, изолятов вида *B. retrofracta*.

В собранном геноме линии *M4B* была найдена последовательность гена APOLLO, после чего был проведен филогенетический анализ, включающий последовательности обеих рассматриваемых линий *B. divaricarpa*. В ходе анализа были сделаны выводы о вероятной гомозиготности линии *ES517* по апомиктичному аллелю гена, и гетерозиготности линии *M4B*, тем не менее, для точного определения состава аллелей необходим анализ копийности гена APOLLO в геномах рассмотренных линий.

Выводы

1. Определена наиболее эффективная стратегия сборки геномов высоко гетерозиготных организмов – гибридная сборка программой MaSuRCA (v3.2.8).
2. Установлено, что *Boechera divaricarpa* M4B унаследовала хлоропластную ДНК от *B. Stricta*, в то время как линия ES517 вероятно наследует генетический материал хлоропластов от вида *B. retrofracta*. Для уточнения данных о наследовании линии ES517 следует провести анализ большего количества изолятов вида *B. retrofracta*., а также изолятов других близких видов.
3. Консенсусная последовательность гена APOLLO линии ES517 показала высокую степень кластеризации с апомиктичными аллелями, что с высокой вероятностью указывает на гомозиготность данной линии по этому аллелю. Линия M4B заняла промежуточное положение между половыми и апомиктичными аллелями на дереве, полученном в ходе филогенетического анализа последовательности гена APOLLO, что свидетельствует о возможном наличии обоих аллелей в геноме этой линии. Для уточнения полученных данных следует удостовериться в однокопийности изучаемого гена у представленных линий *Boechera divaricarpa*.

Список литературы

1. Akama, S., Shimizu-Inatsugi, R., Shimizu, K. K., & Sese, J. (2014). Genome-wide quantification of homeolog expression ratio revealed nonstochastic gene regulation in synthetic allopolyploid *Arabidopsis*. *Nucleic Acids Research*. <https://doi.org/10.1093/nar/gkt1376>
2. Al-Shehbaz, I. A. (2003). Transfer of most North American species of *Arabis* to *Boechera* (Brassicaceae). *Novon*, 381–391.
3. Al-Shehbaz, I. A. (2005). Nomenclatural notes on Eurasian *Arabis* (Brassicaceae). *Novon*, 519–524.
4. Alexander, P. J., Windham, M. D., Beck, J. B., Al-Shehbaz, I. A., Allphin, L., & Bailey, C. D. (2013). Molecular phylogenetics and taxonomy of the genus *Boechera* and related genera (Brassicaceae: Boechereae). *Systematic Botany*, 38(1), 192–209.
5. Aliyu, O. M., Schranz, M. E., & Sharbel, T. F. (2010). Quantitative variation for apomictic reproduction in the genus *Boechera* (Brassicaceae). *American Journal of Botany*, 97(10), 1719–1731. <https://doi.org/10.3732/ajb.1000188>
6. Amiteye, S., Corral, J. M., Vogel, H., Kuhlmann, M., Mette, M. F., & Sharbel, T. F. (2013). Novel microRNAs and microsatellite-like small RNAs in sexual and apomictic *Boechera* species. *MicroRNA (Sharjah, United Arab Emirates)*, 2(1), 45–62.
7. Anderson, J. T., Lee, C.-R., & Mitchell-Olds, T. (2011). Life-history QTLs and natural selection on flowering time in *Boechera stricta*, a perennial relative of *Arabidopsis*. *Evolution; International Journal of Organic Evolution*, 65(3), 771–787. <https://doi.org/10.1111/j.1558-5646.2010.01175.x>
8. Andrews, S., & others. (2016). *FastQC: a quality control tool for high throughput*

sequence data. 2010.

9. Asker, S., & Jerling, L. (1992). *Apomixis in plants*. CRC press.
10. Barcaccia, G., & Albertini, E. (2013). Apomixis in plant reproduction: a novel perspective on an old dilemma. *Plant Reproduction*, 26(3), 159–179. <https://doi.org/10.1007/s00497-013-0222-y>
11. Beck, J. B., Alexander, P. J., Allphin, L., Al-Shehbaz, I. A., Rushworth, C., Bailey, C. D., & Windham, M. D. (2012). Does hybridization drive the transition to asexuality in diploid *Boechera*? *Evolution; International Journal of Organic Evolution*, 66(4), 985–995. <https://doi.org/10.1111/j.1558-5646.2011.01507.x>
12. Bicknell, R. A., & Koltunow, A. M. (2004). Understanding apomixis: recent advances and remaining conundrums. *The Plant Cell*, 16 Suppl, S228-45. <https://doi.org/10.1105/tpc.017921>
13. Bicknell, R., Catanach, A., Hand, M., & Koltunow, A. (2016). Seeds of doubt: Mendel’s choice of *Hieracium* to study inheritance, a case of right plant, wrong trait. *TAG. Theoretical and Applied Genetics. Theoretische Und Angewandte Genetik*, 129(12), 2253–2266. <https://doi.org/10.1007/s00122-016-2788-x>
14. Bocher, T. W. (1951). *Cytological and embryological studies in the amphi-apomictic Arabis holboellii complex*.
15. BOCHER, T. W. (1947). Cytological studies of *Arabis Holboellii* Hornem. *Hereditas*, 33(4), 573.
16. Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114–2120.
17. Brukhin, V. (2017). Молекулярно-генетическая регуляция апомиксиса, “Генетика.” *Генетика*, (9), 1001–1024. <https://doi.org/10.7868/S0016675817090041>
18. Brukhin, V., Osadtchiy, J. V., Florez-Rueda, A. M., Smetanin, D., Bakin, E., Nobre, M. S., & Grossniklaus, U. (2019). The *Boechera* Genus as a Resource for Apomixis Research. *Frontiers in Plant Science*, 10. <https://doi.org/10.3389/fpls.2019.00392>
19. Calderini, O., Chang, S. B., de Jong, H., Busti, A., Paolucci, F., Arcioni, S., ... Pupilli, F. (2006). Molecular cytogenetics and DNA sequence analysis of an apomixis-linked BAC in *Paspalum simplex* reveal a non pericentromere location and partial microcolinearity with rice. *TAG. Theoretical and Applied Genetics. Theoretische Und Angewandte Genetik*, 112(6), 1179–1191. <https://doi.org/10.1007/s00122-006-0220-7>

20. Carman, J. G. (1997). Asynchronous expression of duplicate genes in angiosperms may cause apomixis, bispory, tetraspory, and polyembryony. *Biological Journal of the Linnean Society*, *61*(1), 51–94.
21. Charlesworth, B., & Charlesworth, D. (1997). Rapid fixation of deleterious alleles can be caused by Muller’s ratchet. *Genetical Research*, *70*(1), 63–73.
22. Charlesworth, D., & Wright, S. I. (2001). Breeding systems and genome evolution. *Current Opinion in Genetics & Development*, *11*(6), 685–690.
23. Chin, C.-S., Peluso, P., Sedlazeck, F. J., Nattestad, M., Concepcion, G. T., Clum, A., ... Schatz, M. C. (2016). Phased diploid genome assembly with single-molecule real-time sequencing. *Nature Methods*, *13*(12), 1050–1054. <https://doi.org/10.1038/nmeth.4035>
24. Claros, M. G., Bautista, R., Guerrero-Fernandez, D., Benzerki, H., Seoane, P., & Fernandez-Pozo, N. (2012). Why assembling plant genome sequences is so challenging. *Biology*, *1*(2), 439–459. <https://doi.org/10.3390/biology1020439>
25. Conner, J. A., & Ozias-Akins, P. (2017). Apomixis: Engineering the Ability to Harness Hybrid Vigor in Crop Plants. *Methods in Molecular Biology (Clifton, N.J.)*, *1669*, 17–34. https://doi.org/10.1007/978-1-4939-7286-9_2
26. Corral, J. M., Vogel, H., Aliyu, O. M., Hensel, G., Thiel, T., Kumlehn, J., & Sharbel, T. F. (2013). A conserved apomixis-specific polymorphism is correlated with exclusive exonuclease expression in premeiotic ovules of apomictic boechera species. *Plant Physiology*, *163*(4), 1660–1672. <https://doi.org/10.1104/pp.113.222430>
27. Crane, C. F. (2001). *Classification of apomictic mechanisms*.
28. d’Erfurth, I., Cromer, L., Jolivet, S., Girard, C., Horlow, C., Sun, Y., ... Mercier, R. (2010). The cyclin-A CYCA1;2/TAM is required for the meiosis I to meiosis II transition and cooperates with OSD1 for the prophase to first meiotic division transition. *PLoS Genetics*, *6*(6), e1000989. <https://doi.org/10.1371/journal.pgen.1000989>
29. d’Erfurth, I., Jolivet, S., Froger, N., Catrice, O., Novatchkova, M., & Mercier, R. (2009). Turning meiosis into mitosis. *PLoS Biology*, *7*(6), e1000124. <https://doi.org/10.1371/journal.pbio.1000124>
30. de Arias, M. (2015). *Effects of plant stress on facultative apomixis in Boechera (Brassicaceae)*.
31. Dobes, C. H., Mitchell-Olds, T., & Koch, M. A. (2004). Extensive chloroplast haplotype

- variation indicates Pleistocene hybridization and radiation of North American *Arabis drummondii*, *A. x divaricarpa*, and *A. holboellii* (Brassicaceae). *Molecular Ecology*, 13(2), 349–370.
32. Dobes, C., Mitchell-Olds, T., & Koch, M. A. (2004). Intraspecific diversification in North American *Boechera stricta* (= *Arabis drummondii*), *Boechera xdivaricarpa*, and *Boechera holboellii* (Brassicaceae) inferred from nuclear and chloroplast molecular markers--an integrative approach. *American Journal of Botany*, 91(12), 2087–2101. <https://doi.org/10.3732/ajb.91.12.2087>
 33. Dresselhaus, T., Sprunck, S., & Wessel, G. M. (2016). Fertilization Mechanisms in Flowering Plants. *Current Biology: CB*, 26(3), R125-39. <https://doi.org/10.1016/j.cub.2015.12.032>
 34. Felsenstein, J., & Yokoyama, S. (1976). The evolutionary advantage of recombination. II. Individual selection for recombination. *Genetics*, 83(4), 845–859.
 35. Ferreira de Carvalho, J., de Jager, V., van Gorp, T. P., Wagemaker, N. C. A. M., & Verhoeven, K. J. F. (2016). Recent and dynamic transposable elements contribute to genomic divergence under asexuality. *BMC Genomics*, 17(1), 884. <https://doi.org/10.1186/s12864-016-3234-9>
 36. Glemin, S., Bazin, E., & Charlesworth, D. (2006). Impact of mating systems on patterns of sequence polymorphism in flowering plants. *Proceedings. Biological Sciences*, 273(1604), 3011–3019. <https://doi.org/10.1098/rspb.2006.3657>
 37. Grimanelli, D., Leblanc, O., Perotti, E., & Grossniklaus, U. (2001). Developmental genetics of gametophytic apomixis. *Trends in Genetics: TIG*, 17(10), 597–604.
 38. Grossniklaus, U., Nogler, G. A., & van Dijk, P. J. (2001). How to avoid sex: the genetic control of gametophytic apomixis. *The Plant Cell*, 13(7), 1491–1498. <https://doi.org/10.1105/tpc.13.7.1491>
 39. Hand, M. L., & Koltunow, A. M. G. (2014). The genetic control of apomixis: asexual seed formation. *Genetics*, 197(2), 441–450. <https://doi.org/10.1534/genetics.114.163105>
 40. Hill, W. G., & Robertson, A. (1966). The effect of linkage on limits to artificial selection. *Genetical Research*, 8(3), 269–294.
 41. Hojsgaard, D. (2018). Transient Activation of Apomixis in Sexual Neotriploids May Retain Genomically Altered States and Enhance Polyploid Establishment. *Frontiers in Plant Science*, 9, 230. <https://doi.org/10.3389/fpls.2018.00230>

42. Hojsgaard, D., Klatt, S., Baier, R., Carman, J. G., & Horandl, E. (2014). Taxonomy and Biogeography of Apomixis in Angiosperms and Associated Biodiversity Characteristics. *Critical Reviews in Plant Sciences*, 33(5), 414–427. <https://doi.org/10.1080/07352689.2014.898488>
43. Huang, C.-H., Sun, R., Hu, Y., Zeng, L., Zhang, N., Cai, L., ... Ma, H. (2016). Resolution of Brassicaceae Phylogeny Using Nuclear Genes Uncovers Nested Radiations and Supports Convergent Morphological Evolution. *Molecular Biology and Evolution*, 33(2), 394–412. <https://doi.org/10.1093/molbev/msv226>
44. Hulse-Kemp, A. M., Maheshwari, S., Stoffel, K., Hill, T. A., Jaffe, D., Williams, S. R., ... Van Deynze, A. (2018). Reference quality assembly of the 3.5-Gb genome of *Capsicum annuum* from a single linked-read library. *Horticulture Research*, 5, 4. <https://doi.org/10.1038/s41438-017-0011-0>
45. Jefferson, R. A. (1994). Apomixis: A social revolution for agriculture. *Biotechnology and Development Monitor*, 19, 14–16.
46. Kajitani, R., Toshimoto, K., Noguchi, H., Toyoda, A., Ogura, Y., Okuno, M., ... others. (2014). Efficient de novo assembly of highly heterozygous genomes from whole-genome shotgun short reads. *Genome Research*, 24(8), 1384–1395.
47. Kantama, L., Sharbel, T. F., Schranz, M. E., Mitchell-Olds, T., de Vries, S., & de Jong, H. (2007). Diploid apomicts of the *Boechera holboellii* complex display large-scale chromosome substitutions and aberrant chromosomes. *Proceedings of the National Academy of Sciences of the United States of America*, 104(35), 14026–14031. <https://doi.org/10.1073/pnas.0706647104>
48. Kiefer, C., Dobeš, C., & Koch, M. A. (2009). *Boechera* or not? Phylogeny and phylogeography of eastern North American *Boechera* species (Brassicaceae). *Taxon*, 58(4), 1109–1121.
49. Kirioukhova, O., Shah, J. N., Larsen, D. S., Tayyab, M., Mueller, N. E., Govind, G., ... Johnston, A. J. (2018). Aberrant imprinting may underlie evolution of parthenogenesis. *Scientific Reports*, 8(1), 10626. <https://doi.org/10.1038/s41598-018-27863-7>
50. Kliver, S., Rayko, M., Komissarov, A., Bakin, E., Zhernakova, D., Prasad, K., ... Brukhin, V. (2018). Assembly of the *Boechera retrofracta* Genome and Evolutionary Analysis of Apomixis-Associated Genes. *Genes*, 9(4). <https://doi.org/10.3390/genes9040185>

51. Koch, M. A., Dobes, C., & Mitchell-Olds, T. (2003). Multiple hybrid formation in natural populations: concerted evolution of the internal transcribed spacer of nuclear ribosomal DNA (ITS) in North American *Arabis divaricarpa* (Brassicaceae). *Molecular Biology and Evolution*, 20(3), 338–350. <https://doi.org/10.1093/molbev/msg046>
52. Koltunow, A M. (1993). Apomixis: Embryo Sacs and Embryos Formed without Meiosis or Fertilization in Ovules. *The Plant Cell*, 5(10), 1425–1437. <https://doi.org/10.1105/tpc.5.10.1425>
53. Koltunow, Anna M, & Grossniklaus, U. (2003). Apomixis: a developmental perspective. *Annual Review of Plant Biology*, 54, 547–574. <https://doi.org/10.1146/annurev.arplant.54.110901.160842>
54. Korb, J. O., & Lee, C. (2013, December). Genome assembly and haplotyping with Hi-C. *Nature Biotechnology*, Vol. 31, pp. 1099–1101. <https://doi.org/10.1038/nbt.2764>
55. Kotani, Y., Henderson, S. T., Suzuki, G., Johnson, S. D., Okada, T., Siddons, H., ... Koltunow, A. M. G. (2014). The LOSS OF APOMEIOSIS (LOA) locus in *Hieracium praealtum* can function independently of the associated large-scale repetitive chromosomal structure. *The New Phytologist*, 201(3), 973–981. <https://doi.org/10.1111/nph.12574>
56. Kronenberg, Z. N., Hall, R. J., Hiendleder, S., Smith, T. P. L., Sullivan, S. T., Williams, J. L., & Kingan, S. B. (2018). FALCON-Phase: Integrating PacBio and Hi-C data for phased diploid genomes. *Biorxiv*, 327064.
57. Kyriakidou, M., Tai, H. H., Anglin, N. L., Ellis, D., & Stromvik, M. V. (2018). Current Strategies of Polyploid Plant Genome Sequence Assembly. *Frontiers in Plant Science*, 9, 1660. <https://doi.org/10.3389/fpls.2018.01660>
58. Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, 9(4), 357.
59. Lee, C.-R., Wang, B., Mojica, J. P., Mandakova, T., Prasad, K. V. S. K., Goicoechea, J. L., ... Mitchell-Olds, T. (2017). Young inversion with multiple linked QTLs under selection in a hybrid zone. *Nature Ecology & Evolution*, 1(5), 119. <https://doi.org/10.1038/s41559-017-0119>
60. Lee, H. S., & Chen, Z. J. (2001). Protein-coding genes are epigenetically regulated in *Arabidopsis* polyploids. *Proceedings of the National Academy of Sciences of the United States of America*, 98(12), 6753–6758.

<https://doi.org/10.1073/pnas.121064698>

61. Li, F.-W., Rushworth, C. A., Beck, J. B., & Windham, M. D. (2017). Boechera microsatellite website: an online portal for species identification and determination of hybrid parentage. *Database : The Journal of Biological Databases and Curation*, 2017(1). <https://doi.org/10.1093/database/baw169>
62. Love, A., & Love, D. (1975). Nomenclatural notes on arctic plants. *Botaniska Notiser*.
63. Lovell, J. T., Aliyu, O. M., Mau, M., Schranz, M. E., Koch, M., Kiefer, C., ... Sharbel, T. F. (2013). On the origin and evolution of apomixis in Boechera. *Plant Reproduction*, 26(4), 309–315. <https://doi.org/10.1007/s00497-013-0218-7>
64. Lovell, J. T., Grogan, K., Sharbel, T. F., & McKay, J. K. (2014). Mating system and environmental variation drive patterns of adaptation in Boechera spatifolia (Brassicaceae). *Molecular Ecology*, 23(18), 4486–4497. <https://doi.org/10.1111/mec.12879>
65. Lovell, J. T., & McKay, J. K. (2015). Ecological genetics of range size variation in Boechera spp. (Brassicaceae). *Ecology and Evolution*, 5(21), 4962–4975. <https://doi.org/10.1002/ece3.1746>
66. Lovell, J. T., Williamson, R. J., Wright, S. I., McKay, J. K., & Sharbel, T. F. (2017). Mutation Accumulation in an Asexual Relative of Arabidopsis. *PLoS Genetics*, 13(1), e1006550. <https://doi.org/10.1371/journal.pgen.1006550>
67. Maddison, W. P., & Knowles, L. L. (2006). Inferring phylogeny despite incomplete lineage sorting. *Systematic Biology*, 55(1), 21–30. <https://doi.org/10.1080/10635150500354928>
68. Maheshwari, P. (n.d.). *1950An introduction to the embryology of angiosperms*. New York: McGraw-Hill.
69. Mandakova, T., Schranz, M. E., Sharbel, T. F., de Jong, H., & Lysak, M. A. (2015). Karyotype evolution in apomictic Boechera and the origin of the aberrant chromosomes. *The Plant Journal : For Cell and Molecular Biology*, 82(5), 785–793. <https://doi.org/10.1111/tpj.12849>
70. Mapleson, D., Garcia Accinelli, G., Kettleborough, G., Wright, J., & Clavijo, B. J. (2016). KAT: a K-mer analysis toolkit to quality control NGS datasets and genome assemblies. *Bioinformatics*, 33(4), 574–576.
71. Martinez, G., & Kohler, C. (2017). Role of small RNAs in epigenetic reprogramming during plant sexual reproduction. *Current Opinion in Plant Biology*, 36, 22–28.

<https://doi.org/10.1016/j.pbi.2016.12.006>

72. Matzk, F., Meister, A., & Schubert, I. (2000). An efficient screen for reproductive pathways using mature seeds of monocots and dicots. *The Plant Journal : For Cell and Molecular Biology*, 21(1), 97–108.
73. Mau, M., Corral, J. M., Vogel, H., Melzer, M., Fuchs, J., Kuhlmann, M., ... Sharbel, T. F. (2013). The conserved chimeric transcript UPGRADE2 is associated with unreduced pollen formation and is exclusively found in apomictic *Boechera* species. *Plant Physiology*, 163(4), 1640–1659. <https://doi.org/10.1104/pp.113.222448>
74. Mau, M., Lovell, J. T., Corral, J. M., Kiefer, C., Koch, M. A., Aliyu, O. M., & Sharbel, T. F. (2015). Hybrid apomicts trapped in the ecological niches of their sexual ancestors. *Proceedings of the National Academy of Sciences of the United States of America*, 112(18), E2357-65. <https://doi.org/10.1073/pnas.1423447112>
75. MULLER, H. J. (1964). THE RELATION OF RECOMBINATION TO MUTATIONAL ADVANCE. *Mutation Research*, 106, 2–9.
76. Naumova, T. N., van der Laak, J., Osadtchiy, J., Matzk, F., Kravtchenko, A., Bergervoet, J., ... Boutilier, K. (2001). Reproductive development in apomictic populations of *Arabis holboellii* (Brassicaceae). *Sexual Plant Reproduction*, 14(4), 195–200. <https://doi.org/10.1007/s00497-001-0118-0>
77. Nogler, G. A. (1984). Gametophytic Apomixis. In *Embryology of Angiosperms* (pp. 475–518). https://doi.org/10.1007/978-3-642-69302-1_10
78. Okada, T., Ito, K., Johnson, S. D., Oelkers, K., Suzuki, G., Houben, A., ... Koltunow, A. M. (2011). Chromosomes carrying meiotic avoidance loci in three apomictic eudicot *Hieracium* subgenus *Pilosella* species share structural features with two monocot apomicts. *Plant Physiology*, 157(3), 1327–1341. <https://doi.org/10.1104/pp.111.181164>
79. Ortiz, J. P. A., Quarin, C. L., Pessino, S. C., Acuna, C., Martinez, E. J., Espinoza, F., ... Pupilli, F. (2013). Harnessing apomictic reproduction in grasses: what we have learned from *Paspalum*. *Annals of Botany*, 112(5), 767–787. <https://doi.org/10.1093/aob/mct152>
80. Osadtchiy, J. V, Naumova, T. N., & Brukhin, V. B. (2017). Apomixis in the genus *Boechera* (Brassicaceae). *Russ Botanical Journal*, 102, 1587–1607.
81. Ozias-Akins, P., Akiyama, Y., & Hanna, W. W. (2003). Molecular characterization of the genomic region linked with apomixis in *Pennisetum/Cenchrus*. *Functional &*

- Integrative Genomics*, 3(3), 94–104. <https://doi.org/10.1007/s10142-003-0084-8>
82. Pupilli, F., & Barcaccia, G. (2012). Cloning plants by seeds: Inheritance models and candidate genes to increase fundamental knowledge for engineering apomixis in sexual crops. *Journal of Biotechnology*, 159(4), 291–311. <https://doi.org/10.1016/j.jbiotec.2011.08.028>
83. Ravi, M., Marimuthu, M. P. A., & Siddiqi, I. (2008). Gamete formation without meiosis in Arabidopsis. *Nature*, 451(7182), 1121–1124. <https://doi.org/10.1038/nature06557>
84. Rodkiewicz, B. (1970). Callose in cell walls during megasporogenesis in angiosperms. *Planta*, 93(1), 39–47. <https://doi.org/10.1007/BF00387650>
85. Rodriguez-Leal, D., & Vielle-Calzada, J.-P. (2012). Regulation of apomixis: learning from sexual experience. *Current Opinion in Plant Biology*, 15(5), 549–555. <https://doi.org/10.1016/j.pbi.2012.09.005>
86. Rojek, J., Kapusta, M., Kozieradzka-Kiszkurno, M., Majcher, D., Gorniak, M., Sliwinska, E., ... Bohdanowicz, J. (2018). Establishing the cell biology of apomictic reproduction in diploid *Boechera stricta* (Brassicaceae). *Annals of Botany*, 122(4), 513–539. <https://doi.org/10.1093/aob/mcy114>
87. Ronquist, F., & Huelsenbeck, J. P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, 19(12), 1572–1574.
88. Roy, B. A. (1995). The breeding systems of six species of *Arabis* (Brassicaceae). *American Journal of Botany*, 82(7), 869–877.
89. Rushworth, C. A., Song, B.-H., Lee, C.-R., & Mitchell-Olds, T. (2011). *Boechera*, a model system for ecological genomics. *Molecular Ecology*, 20(23), 4843–4857. <https://doi.org/10.1111/j.1365-294X.2011.05340.x>
90. Savidan, Y., & others. (2000). Apomixis: genetics and breeding. *Plant Breeding Reviews*, 18, 13–86.
91. Schilling, M. P., Gompert, Z., Li, F.-W., Windham, M. D., & Wolf, P. G. (2018). Admixture, evolution, and variation in reproductive isolation in the *Boechera puberula* clade. *BMC Evolutionary Biology*, 18(1), 61. <https://doi.org/10.1186/s12862-018-1173-6>
92. Schmidt, A., Schmid, M. W., & Grossniklaus, U. (2015). Plant germline formation: common concepts and developmental flexibility in sexual and asexual reproduction. *Development (Cambridge, England)*, 142(2), 229–241. <https://doi.org/10.1242/dev.102103>

93. Schmidt, A., Schmid, M. W., Klostermeier, U. C., Qi, W., Guthorl, D., Sailer, C., ... Grossniklaus, U. (2014). Apomictic and sexual germline development differ with respect to cell cycle, transcriptional, hormonal and epigenetic regulation. *PLoS Genetics*, *10*(7), e1004476. <https://doi.org/10.1371/journal.pgen.1004476>
94. Schön, I., Martens, K., & van Dijk, P. (2009). Lost sex. *The Evolutionary Biology of Parthenogenesis*.
95. Schranz, M. E., Dobes, C., Koch, M. A., & Mitchell-Olds, T. (2005). Sexual reproduction, hybridization, apomixis, and polyploidization in the genus *Boechera* (Brassicaceae). *American Journal of Botany*, *92*(11), 1797–1810. <https://doi.org/10.3732/ajb.92.11.1797>
96. Schranz, M. E., Kantama, L., de Jong, H., & Mitchell-Olds, T. (2006). Asexual reproduction in a close relative of *Arabidopsis*: a genetic investigation of apomixis in *Boechera* (Brassicaceae). *The New Phytologist*, *171*(2), 425–438. <https://doi.org/10.1111/j.1469-8137.2006.01765.x>
97. Sharbel, T F, & Mitchell-Olds, T. (2001). Recurrent polyploid origins and chloroplast phylogeography in the *Arabis holboellii* complex (Brassicaceae). *Heredity*, *87*(Pt 1), 59–68. <https://doi.org/10.1046/j.1365-2540.2001.00908.x>
98. Sharbel, T F, Mitchell-Olds, T., Dobes, C., Kantama, L., & de Jong, H. (2005). Biogeographic distribution of polyploidy and B chromosomes in the apomictic *Boechera holboellii* complex. *Cytogenetic and Genome Research*, *109*(1–3), 283–292. <https://doi.org/10.1159/000082411>
99. Sharbel, Timothy F, Voigt, M.-L., Corral, J. M., Galla, G., Kumlehn, J., Klukas, C., ... Rotter, B. (2010). Apomictic and sexual ovules of *Boechera* display heterochronic global gene expression patterns. *The Plant Cell*, *22*(3), 655–671. <https://doi.org/10.1105/tpc.109.072223>
100. Song, B.-H., & Mitchell-Olds, T. (2007). High genetic diversity and population differentiation in *Boechera fecunda*, a rare relative of *Arabidopsis*. *Molecular Ecology*, *16*(19), 4079–4088. <https://doi.org/10.1111/j.1365-294X.2007.03500.x>
101. Song, B.-H., Windsor, A. J., Schmid, K. J., Ramos-Onsins, S., Schranz, M. E., Heidel, A. J., & Mitchell-Olds, T. (2009). Multilocus patterns of nucleotide diversity, population structure and linkage disequilibrium in *Boechera stricta*, a wild relative of *Arabidopsis*. *Genetics*, *181*(3), 1021–1033. <https://doi.org/10.1534/genetics.108.095364>

102. Spillane, C, Steimer, A., & Grossniklaus, U. (2001). Apomixis in agriculture: the quest for clonal seeds. *Sexual Plant Reproduction*, 14(4), 179–187. <https://doi.org/10.1007/s00497-001-0117-1>
103. Spillane, Charles, Curtis, M. D., & Grossniklaus, U. (2004). Apomixis technology development-virgin births in farmers' fields? *Nature Biotechnology*, 22(6), 687–691. <https://doi.org/10.1038/nbt976>
104. Stebbins, G. L. (1950). *Variation and evolution in plants*. Geoffrey Cumberlege.; London.
105. Talent, N. (2009). Evolution of gametophytic apomixis in flowering plants: an alternative model from Maloid Rosaceae. *Theory in Biosciences = Theorie in Den Biowissenschaften*, 128(2), 121–138. <https://doi.org/10.1007/s12064-009-0061-4>
106. Taskin, K. M., Turgut, K., & Scott, R. J. (2009). Somatic embryogenesis in apomict *Boechera holboellii*. *Acta Biologica Hungarica*, 60(3), 301–307. <https://doi.org/10.1556/ABiol.60.2009.3.7>
107. van Dijk, P. J., & Vijverberg, K. (2005). The significance of apomixis in the evolution of the angiosperms: a reappraisal. *Regnum Vegetabile*, 143, 101.
108. Voigt-Zielinski, M.-L., Piwczynski, M., & Sharbel, T. F. (2012). Differential effects of polyploidy and diploidy on fitness of apomictic *Boechera*. *Sexual Plant Reproduction*, 25(2), 97–109. <https://doi.org/10.1007/s00497-012-0181-8>
109. Vurture, G. W., Sedlazeck, F. J., Nattestad, M., Underwood, C. J., Fang, H., Gurtowski, J., & Schatz, M. C. (2017). GenomeScope: fast reference-free genome profiling from short reads. *Bioinformatics (Oxford, England)*, 33(14), 2202–2204. <https://doi.org/10.1093/bioinformatics/btx153>
110. Windham, M. D., & Al-Shehbaz, I. A. (2007). New and noteworthy species of *Boechera* (Brassicaceae) III: Additional sexual diploids and apomictic hybrids. *Harvard Papers in Botany*, 12(1), 235–258.
111. Zhao, X., Bramsiepe, J., Van Durme, M., Komaki, S., Prusicki, M. A., Maruyama, D., ... Schnittger, A. (2017). RETINOBLASTOMA RELATED1 mediates germline entry in *Arabidopsis*. *Science (New York, N.Y.)*, 356(6336). <https://doi.org/10.1126/science.aaf6532>
112. Zielezinski, A., Vinga, S., Almeida, J., & Karlowski, W. M. (2017). Alignment-free sequence comparison: benefits, applications, and tools. *Genome Biology*, 18(1), 186. <https://doi.org/10.1186/s13059-017-1319-7>

113. Zimin, A. V, Marçais, G, Puiu, D., Roberts, M., Salzberg, S. L., & Yorke, J. A. (2013).
The MaSuRCA genome assembler. *Bioinformatics*, 29(21), 2669–2677.