Phylogenetic and expression analysis of *CENH3* and *APOLLO* genes in sexual and apomictic *Boechnera* species

Evgeny Bakin¹, Fatih Sezer², Ashhan Özbilen³, Irem Kilic³, Mike Rayko⁴, Kemal Melih Taskin², Vladimir Brukhin^{5,6, *}

¹ Bioinformatics Institute, Kantemirovskaya str. 2A, 197342 St. Petersburg, Russia (E.B.)

² Department of Molecular Biology and Genetics, Çanakkale Onsekiz Mart University, Terzioğlu Kampusu, 17100 Çanakkale, Turkey; fatihsezer@comu.edu.tr (F.S.), kmtaskin@comu.edu.tr (K.M.T.)

³ Department of Biology, Çanakkale Onsekiz Mart University, Terzioğlu Kampusu, 17100 Çanakkale, Turkey; aslihanozbilen@hotmail.com (A. Ö.), iremkilic72@gmail.com (I.K.)

⁴ Laboratory for Algorithmic Biology, Saint-Petersburg State University, Vasilievsky Island, 199004 St. Petersburg, Russia; mike.rayko@gmail.com (M. R.)

⁵ Plant Genomics Lab, ChemBio Cluster, ITMO University, Lomonosova 9, 191002, Saint-Petersburg, Russia; vbrukhin@gmail.com, brukhin@scant-itmo.ru (V.B.)

⁶ Department of Plant Embryology and Reproductive Biology, Komarov Botanical Institute RAS, Prof. Popov 2, 197376 St. Petersburg, Russia

* Correspondence: vbrukhin@gmail.com

Abstract: Apomictic plants (reproducing via asexual seeds), unlike sexual individuals, avoid meiosis and egg cell fertilization. Consequently, apomixis is very important for fixing maternal genotypes in the next plant generations. Despite the progress in the study of apomixis, molecular and genetic regulation of the latter remains poorly understood. So far *APOLLO* (Aspartate Glutamate Aspartate Aspartate histidine exonuclease) is the only described gene associated with apomixis in *Boechnera* species. The centromere-specific histone H3 variant encoded by *CENH3* gene is essential for cell division. Mutations in *CENH3* disrupt chromosome segregation during mitosis and meiosis since the attachment of spindle microtubules to a mutated form of the *CENH3* histone fails. This paper presents *in silico* characteristic of *APOLLO* and *CENH3* genes, which may affect apomixis. Also, in this research we characterize the structure of *CENH3*, study expression levels of *CENH3* and *APOLLO* in gynoeceum/silques of the natural diploid apomictic and sexual *Boechnera* species at the stages of before and after fertilization. At the premeiotic stage, the expression level of *CENH3* in the gynoeceum of apomicts was two times lower than that of the sexual *Boechnera*, it decreased in both species by the time of meiosis and increased after fertilization. By 1 DAP *CENH3* expression started dropping in sexual *B. stricta* silques and kept increasing in apomictic *B. divaricarpa* ones. That might indicate to a role of *CENH3* in apomictic development in *Boechnera* species. The expression levels of *APOLLO* also sharply decreased by the time of meiosis in gynoeceum of both species; however, by 3 DAP, the level of *APOLLO* expression in silques of apomicts was almost 1.5 times higher than that of the sexuals. While *CENH3* was a single copy gene in all *Boechnera* species, the *APOLLO* gene have several polymorphic alleles associated with sexual and apomictic reproduction in the *Boechnera* genera. We also discuss polymorphism and phylogeny of the *APOLLO* and *CENH3* genes.

Keywords: *Boechnera*; Brassicaceae; *CENH3*; *APOLLO*; apomixis associated genes; cell division; apomeiosis

1. Introduction

Sexual reproduction is the main mode of reproduction of flowering plants. The major features of sexuality are meiosis and fertilization, the latter occurs through the fusion of haploid female and male gametes, giving rise to the formation of the genetically variable progeny. Genetic mutations and meiotic recombination provide permanent genetic changes, which are the source for evolution and adaptation of the population, as well as the basis for selection in agriculture. Despite the fact that sexual reproduction is energetically very expensive, the advantage of sex is that it forms a combination of useful mutations that are absent in asexual organisms [1,2]. Thus, the benefit of sexuality is that it helps to get rid of harmful mutations (Hill-Robertson effect) [3,4] and produces useful traits and variability. However, apart from the high energy costs, the disadvantage of sexual reproduction is the segregation of beneficial traits in subsequent generations, so that the offspring can lose useful combinations of their parental genes [5,6].

One of the mechanisms that can produce “clones” of mother plants is apomixis, which is a mode of asexual reproduction through seeds that has been identified in over four hundred of plant species [7,8]. In apomicts, meiosis and fertilization are modified or completely absent. Consequently, the embryo is formed without prior meiosis (by apomeiosis) and fertilization (i.e. by parthenogenesis), while endosperm development occurs either autonomously, i.e. without fertilization, or pseudogamously (by fertilization of the central cell) [6,8–13]. During double fertilization, intrinsic to all Angiosperms, pollen tube transports two sperm cells (male gametes) to the embryo sac, one of which fertilizes the egg cell (female gamete) and the second one fertilizes the central cell, giving rise to a $2n$ embryo and a $3n$ endosperm, respectively. In sexuals departure from the 2 maternal:1paternal (2m:1p) genome ratio in endosperm nuclei results in seed abortion. While in pseudogamous apomicts, endosperm ploidy varies according to the ploidy of the sperm and central cell. Deviations from 2m:1p genome ratio sometimes occur, demonstrating that the apomictic system is more resilient compare to sexuals [14].

Apomictic plants, forming genetically identical offspring, dramatically influence the structure of the population, thereby being excellent models for studying the mechanisms of the onset of meiosis and its replacement by apomeiosis, the formation of a seed in the absence of fertilization or under pseudogamy; as well as the important ecological role of apomixis in the origin of polyploids and speciation [15,16].

However, so far, little is known about the molecular background of apomixis and the genes associated with its triggering [17].

In this paper, we characterized the protein structure, evolution, and expression patterns of the *APOLLO* (for APomixis-Linked LOcus; encodes Aspartate Glutamate Aspartate Aspartate histidine exonuclease) gene, one of the important genes associated with apomixis in *Boechea* genus. It was shown that *APOLLO* have several polymorphic sex- and apo-alleles [18,19]. This gene is down-regulated in sexual ovules at the meiosis stage and is up-regulated in apomeiotic ovules at the same stage of development in *Boechea* plants. It was reported that genomes of apomictic plants are always heterozygous carrying at least one of the apo-alleles, while sexual genotypes were always homozygous for sex-alleles [18,19]. Evolutionary *APOLLO* gene

analysis presented in this study shows that sexual and apomictic species of *Boechera* are clustered in different clades in the phylogenetic tree based on the multiple protein sequence alignment. This could be partly because the *APOLLO* apo-alleles present in the genomes of apomictic species could acquire a new function [19]. Expression analysis of the *APOLLO* presented here showed that the expression levels of the gene dramatically decreased before meiosis in gynoecia of both apomictic and sexual species, however after meiosis and fertilization, the expression of the *APOLLO* was upregulated in apomictic siliques compared to sexual ones.

We also characterize *CENH3* gene, which is essential for cell division [20,21] and might affect the apomictic events. *CENH3* encodes the centromere-specific histone H3 variant. The accumulation of *CENH3* provides an assembly site for a protein complex called kinetochore. The main function of the kinetochore is to bind chromosomes to spindle fibers during chromosome segregation in meiosis and mitosis [22]. *CENH3* harbor two domains, a DNA-binding histone fold domain (HFD) and N-terminal tail domain. HFD is structurally similar to the same domain in H3 and is highly conserved across higher eukaryotes. However, the N-terminal tail of the *CENH3* protein is highly variable even among closely related species and has meiosis-specific function [23]. Mutations at the N-terminal tail of *CENH3* disturb chromosome segregation in meiosis and often lead to sterility [20,24]. Null mutations in *CENH3* also cause chromosome elimination and were proposed as a tool to produce haploid plants [25–28].

In *Arabidopsis thaliana*, the *CENH3* mutants act like a haploid inducer parent when cross with a wild type plant [25,27] and the chromosomes in these mutants were lost in the developing embryo. The resulting plant contains a haploid set of the wild-type parent chromosomes, which can be induced to double its number. We investigated the structure and the expression profiles of *CENH3* in apomictic and sexual *Boechera* species (Brassicaceae) to further understand its role in meiotic chromosome segregation. Plants from the *Boechera* genus are attractive models for research because both sexual and apomictic accessions are present within this genus. Moreover, plants from the *Boechera* genus are close relatives of *Arabidopsis thaliana*, which is very well studied in terms of molecular genetics and functional gene annotation [29–31]. In apomictic *Boechera* species, the *Taraxacum* type diplospory with pseudogamous endosperm development that requires fertilization of the central cell is reported [31–34]. A special attention deserves apomictic accessions of *Boechera divaricarpa*, which is known as an interspecific hybrid between sexual species *B. stricta* and *B. retrofracta* or a closely related species [35]. Although diploid apomixis is an extremely rare condition in plants [36], both diploid and triploid apomictic *B. divaricarpa* lineages have been reported [31,37].

Our study revealed *CENH3* expression levels in the gynoecia/siliques during sexual and apomictic development. During meiosis and pollination *CENH3* expression demonstrated decreasing levels in both sexual and apomictic gynoecia and in the first day after pollination showed an increase in expression but in the gynoecia of sexual plants expression was almost two-fold higher than that in the gynoecia of apomicts. These results may indicate to a possible role that *CENH3* might play in apomictic development in *Boechera* species.

2. Materials and Methods

2.1 Plant material

The diploid sexual *B. stricta* (ES6, ID 500206; DG; DQ013050) and apomictic *B. divaricarpa* (ES9, 500209; BS) seeds were kindly obtained from Dr. Eric Schranz (Wageningen University, Netherlands). *Boechea* plants were grown as described in [37,38]. Seeds of ES6 and ES9 accessions were germinated on moist filter paper after vernalization at 4°C for 3 weeks in the dark to break seed dormancy. Then petri dishes containing the seeds were transferred into a growth chamber under the condition of 16 h light: 8 h dark (21°C). Germinated seeds were transferred to a peat: perlite mix (1:4) for 4 weeks. Then the plants were grown at 10° C for 6 weeks. Flowering of the plants started in 6 weeks after the vernalization.

Emasculation was performed on unopened flower buds by removing anthers with fine forceps. Hand pollination of stigmas was carried out with pollen from the same plant. Gynoecia/ siliques were collected one day before pollination (before meiosis according to gynoecium size, see [39]), right after pollination (after meiosis), on the 1st, and 3rd days after pollination. Anthers we collected before and after meiosis (see [39]). The collected samples were immediately frozen in liquid nitrogen and stored at -80 °C until use.

2.2 APOLLO and CENH3 genes retrieval and pre-processing

Sequences of the *APOLLO* (Aspartate Glutamate Aspartate Aspartate histidine exonuclease) and *CENH3* (histone H3-like centromeric protein) genes in various *Boechea* species were obtained from the sources listed in table 1 (also see in [31]). For *APOLLO* phylogenetic analysis, additionally 10 accessions from [18] listed in table 2 (5 apo-alleles, 5 sex-alleles) were used.

Table 1. Genome sequencing data for *Boechea* genus

Species	Reproduction mode	Raw data NCBI accession	Genome assembly	Genome annotation	Reference
<i>B. stricta</i>	Sexual	SRR396760 SRR396762 SRR396756	Yes	Yes	[40]
<i>B. retrofracta</i>	Sexual	SRR3929707	Yes	Yes	[19]
<i>B. puberula</i>	Sexual	ERX2578777 ERX2578776	Yes	No	[41]
<i>B. spatifolia</i> (Rosita3)	Sexual	SRR5116719	No	No	[42]
<i>B. spatifolia</i> (Tiesiding2)	Sexual	SRR5116723	No	No	[42]
<i>B. spatifolia</i> (Chicago2)	Sexual	SRR5116732	No	No	[42]
<i>B. spatifolia</i>	Sexual	SRR5116730	No	No	[42]

(Royal2)

<i>B. arcuata</i>	Sexual	SRR6448790	No	No	n/a
<i>B. divaricarpa</i>	Apomictic	SRR3500627 SRR3500628	No	No	n/a
<i>B. perennas</i>	Apomictic	SRR6448882	No	No	n/a

Table 2. *Boechea* accessions used for investigation of *APOLLO* gene (from [18])

GenBank accession	Sample ID	Allele type
KF705583.1	369S2_S3	Sex-allele
KF705582.1	376S2_S5	Sex-allele
KF705581.1	355S2_S3	Sex-allele
KF705569.1	329S2_S1	Sex-allele
KF705572.1	385S2_S11	Sex-allele
KF705596.1	43A3_A3	Apo-allele
KF705598.1	1A2_A6	Apo-allele
KF705600.1	11A2_A1	Apo-allele
KF705599.1	11A2_A3	Apo-allele
KF705588.1	33A2_A5	Apo-allele

For the species with available assembled and annotated genome a BLAST search over extracted proteins datasets was performed using *A. thaliana* peptide sequences as a query. For the other species (all *B. spatifolia* accessions from research [42] as well as *B. arcuata* and *B. divaricarpa* accessions) a local assembly using SRAssembler v1.0 was performed [43]. Consistency of the loci assembly results were checked with another approach, in which a whole-genome draft assembly was performed using Platanus software v1.2.1 [44] with a consequent genes search via Exonerate v2.2 run in protein2genome mode [45]. All the mentioned tools were run with default parameters.

The resulting parameters of the obtained assemblies are listed in Table 1S (Supplement). As one can see these assemblies have rather modest characteristics (N50 of about a few kilobases) as predicted earlier in [31]. Insufficient genome coverage in the datasets (30x in average), short reads length and high complexity of the genomes resulted in the fact, that most of the *B. spatifolia* accessions reported in [42] were excluded from the further analysis. In the excluded datasets, a significant divergence of the obtained sequences between two assembly methods and abnormally low their similarities with homologs in the other *Brassicaceae* species were detected. Thus, only Rosita3 accession was accepted for *CENH3* analysis. For *APOLLO* analysis Rosita3, Tiesiding2, Chicago2 and Royal2 accessions were accepted. In *B. puberula* and *B. perrenas* assemblies *APOLLO* gene was incomplete and thus these accessions were also excluded from this gene analysis. It is worth noting that a short read length didn't allow us to perform unambiguous diploid assemblies and distinguish different copies and alleles of the

considered genes, potentially presenting in the genomes. Thus, all the downstream analysis was performed with consensus variants of the assembled genes. Sequences for non-*Boechea* species were retrieved directly from Phytozome v. 12.1 [46] and Uniprot [47] databases.

2.3 General characteristics of *CENH3* gene

CENH3 gene length and exon structure were directly extracted from both SRAssembler and Exonerate output. The molecular weight and isoelectric point of the proteins were calculated with ExPASy's ProtParam server [48]. Subcellular localizations of the *Boechea* proteins were predicted via CELLO v.2.5: subcellular localization predictor [49]. Conserved regions were investigated with CDD (Conserved Domain Database) [50]. The promoter regions were analyzed for regulatory motifs using PLANTCARE database [51].

2.4 Evolutionary analysis of the *CENH3* and *APOLLO* genes

Multiple alignment of the obtained amino acid sequences was performed with MUSCLE algorithm [52] and then visualized with Jalview v2 [53]. The evolutionary history was inferred by the Maximum Likelihood method and w/freq. model [54]. As a result, the tree with the highest log likelihood was reconstructed. The percentage of trees in which the associated taxa clustered together is shown next to the branches (1000 iterations were performed). The tree was drawn to scale, with branch lengths measured in the number of substitutions per site. Evolutionary analyses were conducted in MEGA X [55]. Heatmaps visualization was performed by R package Pheatmap [56].

2.5 *CENH3* and *APOLLO* genes expression studies

Total RNAs were isolated from gynoecium and anthers before meiosis and siliques after meiosis 1-3 days after pollination (DAP) according [39] with Purelink RNA mini kit (Invitrogen – 12183018A) and TRIzol reagent (Invitrogen, 15596018). RNA quality was checked on 1% agarose gel electrophoresis, the concentration of RNA was measured with a Qubit 2.0 Fluorometer (Life Technologies, Carlsbad, CA, USA). Before the cDNA synthesis, 1 µg of total RNA was treated with Rnase-free Dnase I (Thermo Scientific, EN0521) for 30 minutes at 37 °C. Reverse transcription reactions were performed using 1 µg of total RNA and High-Capacity cDNA Reverse Transcription Kit (4368814; Applied Biosystems, Foster City, CA, USA) according to the manufacturer's instructions. Expression of the *CENH3* and *APOLLO* genes were analyzed with the 7500 real-time PCR system (Applied Biosystems) according to the manufacturer's instructions. The efficiency of the primers used in the study was evaluated with the standard curve experiments, only the primers with 90%-100% efficiency were used. UBQ (Polyubiquitin) gene was used as the endogenous control. No template (water instead of the template) and – RT (water instead of reverse transcriptase enzyme in cDNA synthesis) controls showed negative amplification. Primers sequences were BOECHUBQ_F: 5'-GGCTAAGATCCAGGACAAGGAAGGTAT-3', BOECHUBQ_R: 5'-CTGGATGTTATAGTCAGCCAAAGTGCG-3' for UBQ amplification; APOLLO_F: 5'-CGGAGTTCTCTCTGCACCTAC 3' and APOLLO_R: 5'-TTCGTCCGTGGAGAATGTGCG-3' for APOLLO analysis; CenH3_F 5'-CAACTCCTACAACCTCACCAGCTACTG-3' and

Cenh3_R 5'-TTGTGAACCTTGTGGCCTAGCATATC-3' for CENH3 analysis. All reactions were performed in 3 biological and 3 technical replicates. Reactions included 10 µl SYBR Green PCR Master Mix (Applied Biosystems – 4367659), 0,6 µl primers (10 µM stock), 1 µl cDNA template (50 ng/µl), 7,8 nuclease-free water. The PCR cycling conditions were as follows: pre-denaturation at 95 °C for 10 min, denaturation at 95 °C for 15 s, annealing at 60 °C for 30 s, and elongation at 72 °C for 60 s for 40 cycles. The data were analyzed with expression suite software (Applied Biosystems) with the $\Delta\Delta C_t$ method.

3. Results

3.1. Characteristics of CENH3 gene and protein isoforms

General characteristics of the *CENH3* gene and CENH3 protein in seven studied *Boechera* species is given in table 3. In all these species CENH3 protein isoforms showed high similarity to each other with similarity index not less than 97% (at both the nucleotide and protein levels). All *Boechera* CENH3 protein sequences contained a Histone H3/CENP-A (IPR000164) conserved domain, were 177 a.a. long and predicted to be localized in the nucleus. CENH3 protein alignment results are represented in Fig. 1.

Table 3. General characteristics of CENH3 gene and CENH3 protein in studied seven *Boechera* species.

Parameter	Value
Gene length	2231 - 2298 b.p.
Number of exons	10
Protein length	177 a.a.
Molecular weight	19616.82 ± 60.4
Theoretical pI	11.25 ± 0.10
Subcellular Localization	Nuclear
Conserved Domains	Histone H3/CENP-A (from 57 a.a. to 172 a.a.)

Figure 1. Multiple alignment of CENH3 proteins in considered *Boechera* sp. Multiple alignment was performed using MUSCLE software and visualized in Jalview.

Among the seven investigated *Boechera* species 10 polymorphic amino acid positions in CENH3 protein were detected (listed in Table 4). *B. stricta* and *B. divaricarpa*, which were further subjected to the comparative expression analysis of *CENH3* gene, had only two polymorphic amino acids in N-tail positions 7 and 16, which are located out of the Histone H3 conserved domain. Variability within the N-terminal tail might affect the chromosome segregation in meiosis[20,24] and lead to apomeiosis.

Table 4. Polymorphic positions in CENH3 protein

Position	<i>B. retrofracta</i>	<i>B. stricta</i>	<i>B. divaricarpa</i>	<i>B. puberula</i>	<i>B. arcuata</i>	<i>B. perennas</i>	<i>B. spatifolia</i> (Rosita3)
7	L	L	F	F	F	F	F
16	Q	W	R	R	W	R	R
47	T	P	P	T	P	P	P
54	A	A	A	S	A	A	A
55	T	T	T	T	T	T	R
56	G	G	G	G	G	G	F
67	A	A	A	A	P	A	A
91	Y	H	H	H	H	H	H
96	I	I	I	I	V	I	I
157	V	I	I	I	I	I	I

CENH3 similarity heatmap matrix and reconstructed phylogenetic tree are shown in Fig. 2, A and B respectively. The heatmap demonstrates that CENH3s in all *Boecheira* species are almost identical, which is consistent with the protein alignment analysis (Fig. 1) and are very similar ($\geq 80\%$) to *Arabidopsis* and most of the other representatives of Brassicaceae studied. *CENH3* gene is represented by only one copy per genome of each investigated species. The evolutionary history was estimated with the help of the highest log likelihood. The reconstructed phylogenetic tree (Fig. 2B) was rooted with *Capsella spp.* as an out-group. All *Boecheira* species were grouped in one clade, which was separated from *Arabidopsis spp.*, *Eutrema spp.*, *Cardamine sp.*, suggesting that unlike APOLLO (see paragraph 3.5) CENH3 protein sequences do not reflect the mode of propagation within the *Boecheira spp.*

Figure 2. (A) Similarity matrix and (B) Phylogenetic tree of CENH3 (histone H3-like centromeric protein) in seven species of interest. Sequences of *Capsella* species were used as an outgroup. All *Boecheira* species are grouped in one clade. Numbers near nodes represent corresponding bootstrap support. Phylogenetic tree was reconstructed using Maximum Likelihood method.

3.2 *CENH3* promoter characteristic

Since it was found that genomes of all studied species contain only a single copy of the *CENH3* gene, we decided to investigate and compare the structure of the *CENH3* gene promoters in two species, sexual *B. stricta* and apomictic *B. divaricarpa*. 1000 bp upstream regions from the first codon (ATG) of *BsCENH3* and *BdCENH3* genes were analyzed. Several regulatory elements and core promoter regions like TATA and CAAT-box within both *CENH3* promoters were found. Among the regulatory motifs, MYB transcription factors recognition and binding sites were common for both *CENH3* promoters. MYB transcription factors are encoded by several genes that control plant development, differentiation, stress resistance, and defense [57]. The presence of several MYB binding motifs might indicate to a role of MYB transcription factors in *CENH3* regulation. E2F transcription factor binding sites at -163 positions of *BsCENH3* and *BdCENH3* were found as well. At -115 site *BdCENH3* and *BsCENH3* had a single polymorphic site (GCGGGAAA/ GCGGGAAG). Promoter regions also included several light response motifs like Box 4, GT1-motif and TCT-motif. Hormone response elements like gibberellin responsive element (GARE-motif) and auxin-responsive elements (TGA-element) were detected in the *CENH3* promoters of both species. A cis-acting regulatory element involved in the MeJA-responsiveness (CGTCA-motif), cis-acting element involved in low-temperature responsiveness (LTR), cis-acting regulatory element involved in circadian control, cis-acting regulatory element related to meristem expression (CAT-box) and cis-regulatory element involved in endosperm expression (GCN4_motif) were also found in the *CENH3* promoter (Fig. 3). Our research of promoter regions of *BsCENH3* and *BdCENH3* genes did not reveal significant differences (Fig. 3). Therefore, we decided to check the expression levels of *BsCENH3* and *BdCENH3* genes in the premeiotic gynoecia and young siliques just after fertilization.

Figure 3. Motifs found in 1000 bp upstream of the transcription start site of *BsCENH3* and *BdCENH3*

3.3 *CENH3* expression before and after meiosis in gynoecia/siliques and anthers

Comparative expression analysis of *CENH3* gene was performed in gynoecia and anthers from unopened flower buds (before pollination) and in siliques 1-3 DAP (Fig. 4 A, B; Table 2S in Supplement) to cover developmental stages around meiosis, gametophyte development, and seed set in diploid apomict *B. divaricarpa* and diploid sexual *B. stricta*. For investigation after the pollination, only the extending siliques were used to make sure that the pollination has occurred, and seed development has started. It was found that, before meiosis/ apomeiosis, the *CENH3* gene is intensely expressed in gynoecium of both sexual *B. stricta* and apomictic *B. divaricarpa*, two folds higher expression level was observed in *B. stricta*. *CENH3* was

significantly downregulated by the time of meiosis, while the expression level in *B. stricta* is more than three folds higher than that in *B. divaricarpa* (Fig. 4A). One day after pollination (1 DAP), the levels of *CENH3* expression in gynoecium significantly increase in both sexual and apomictic *Boechera*, exceeding the values of expression before meiosis by 1.3 times in *B. stricta* and 1.7 times in *B. divaricarpa*, while the absolute level of *CENH3* expression by 1 DAP in *B. stricta* was 1.8 times higher than that in *B. divaricarpa*. By 3 DAP, the *CENH3* expression level continues to grow in *B. divaricarpa* and decreases in *B. stricta* becoming almost the same in the sexual and apomictic *Boechera* species (Fig. 4A). We also investigated the expression of the *CENH3* gene in *B. stricta* and *B. divaricarpa* anthers before and during meiosis. In contrast to gynoecium, before meiosis, the level of the gene expression was higher in the anthers of *B. divaricarpa* than in *B. stricta*. By the end of meiosis, expression of *CENH3* was down regulated almost completely in anthers of both species (Fig. 4B). The observed difference in the level of *CENH3* gene expression between sexual and apomictic accessions could be either due to differences in regulation of expression by the transcription factors and genes responsible for their expression or due to epigenetic factors that might regulate the expression at a post-translational level.

Figure 4. Expression levels of *CENH3* and *APOLLO* and genes in diploid apomict and sexual *Boechera* plants: *B. divaricarpa* and *B. stricta* in (A; C) gynoecia (before meiosis) and carpels 1-3 days after pollination (DAP); (B; D) anthers. P- pollination. Bars indicate $\pm$ standard errors of fold change levels.

3.4 *APOLLO* expression before and after meiosis in gynoecia/siliques and anthers

Similar to the analysis of the *CENH3* gene dynamics, the expression of the *APOLLO* gene was studied before meiosis in gynoecium and anthers and after meiosis in siliques of sexual diploid *B. stricta* and apomictic diploid *B. divaricarpa* 1-3 DAP (Fig. 4 C, D; Table 2S in Supplement). In premeiotic gynoecium, the levels of *APOLLO* expression were similar in *B. stricta* and *B. divaricarpa*, almost completely dropping by the time of meiosis/apomeiosis in MMC and synchronously increasing after the pollination. However, they differed markedly after the 1 DAP, gradually deviating, and by 3 DAP the expression of *APOLLO* in apomictic *B. divaricarpa* was almost 1.5 times higher than that in sexual *B. stricta* (Fig. 4 C). The level of expression of the *APOLLO* gene in premeiotic anthers was more than two folds higher in apomictic *B. divaricarpa* compared with the anthers of *B. stricta* drastically downregulating by the time of meiosis in anthers in both species (Fig. 4 D).

Revealed interesting patterns of *APOLLO* expression were somewhat different from those previously published [18], where the expression levels of this gene in gynoecia of sexual and apomictic *Boechera* are quite similar at the premeiotic, meiotic, and early post meiotic stages, however, diverging in one day after pollination may suggest a role of *APOLLO* for the development of the embryo sac and possibly for the autonomous egg cell and the central cell development. Differing patterns of *APOLLO* expression in premeiotic anthers, with considerably higher gene expression in apomictic plants, may indicate to a role of *APOLLO* in meiosis or apomeiosis during pollen formation in apomictic plants. However, this difference observed in *APOLLO* expression in our study, may also be tissue and species specific.

3.5 Phylogenetic analysis of *APOLLO* gene

A preliminary analysis of *APOLLO* gene and protein orthologs we reported previously [19], therefore, the objective of this research was to study the phylogeny of *APOLLO* gene using additional available in open access sequenced genomes of *Boechera* species with sexual and apomictic mode of reproduction by translating *APOLLO* nucleotide sequences into proteins and applying Multiple alignment method for obtaining amino acid sequences using MUSCLE software and Maximum Likelihood method to construct a phylogenetic tree. The heatmap shown in Fig. 5A demonstrates that all compared *APOLLO* proteins of *Boechera* are rather similar ($\geq 92.5\%$) despite the presence of 5 apo-alleles and 5 sex-alleles of that gene (see table 2 in Methods). Phylogenetic tree of *APOLLO* (Fig. 5B) shows that the *Boechera* species are clustered in two separate subclades according to the mode of reproduction, i.e. branches in the tree were grouped by genes rather than by species: the

first subclade contains the homozygous for sex-alleles accessions reproducing sexually and the second contains accessions with apo-alleles that reproduce via apomixis. At the same time, all *Boechera* species represented in the phylogenetic tree are clustered in a common clade, slightly separating them from the other genera of the Brassicaceae family (similarity $\geq 82.5\%$).

Figure 5. Similarity matrix (A) and Phylogenetic tree (B) of *APOLLO* in species of interest. Numbers near nodes represent corresponding bootstrap support. Phylogenetic tree

was reconstructed using Maximum Likelihood method. The *Boecheira* species are clustered in two separate subclades according to the mode of reproduction: the first subclade contains the homozygous accessions for sex-alleles and the second contains accessions with apo-alleles.

4. Discussion

Apomixis via clonal seeds produces the offspring, that are genetically identical to the maternal plant. Understanding of the genetic components that regulate apomixis is very important for studying plant development and evolution, in addition, the introduction of apomixis in agricultural plants would allow a long-term fixation of complex genotypes, including F1 hybrids, often used in agriculture. Yet, the molecular mechanisms underlying apomixis are poorly understood. Namely, the factors inducing avoidance or modification of meiosis (apomeiosis) and parthenogenesis. To study the genes regulating meiosis and embryogenesis in comparison of sexual vs. apomictic plants we use a convenient model plants from the *Boecheira* genus, that comprise species naturally reproducing both by sexual and apomictic ways and show features of hybrid origin [31].

In this paper we studied similarity of homologs and phylogeny of *CENH3* and *APOLLO* genes that might affect some components of apomixis and also compared the expression patterns of these genes in apomictic and sexual plants of *Boecheira*.

The *CENH3* gene plays an important role in cell divisions and genome elimination when mutated. Mutations in *CENH3* of *Arabidopsis thaliana* cause disturbed meiotic chromosome segregation [20,21] that was also used to induce genome elimination in *A. thaliana* and rice. In hybrids of *cenh3* mutant lines with diploid wild type plants the *cenh3* line genome was eliminated [58]. We showed that *CENH3* is a single-copy gene and that its structure is almost identical among the seven studied *Boecheira* species irrespective of the reproduction (sexual or apomictic) mode. Polymorphic sites were mostly found at the N-tail protein regions, although *B. retrofracta* and *B. arcuata* had one site at 91 a.a. and two polymorphic sites at 67 a.a. and at 96 a.a. respectively in conservative Histone H3/CENP-A domain. Variability within the N-terminal tail might lead to apomeiosis, since it influences the chromosome segregation in meiosis [20,24], although this assumption still needs to be tested on mutant lines with the replacement of the corresponding polymorphic sites. Still the similarity index between all studied *Boecheira* *CENH3* was $\geq 97\%$ (both at nucleotide and protein levels). Also, in the *CENH3* phylogenetic tree, performed by multiple alignment, all studied *Boecheira* species are clustered into the same clade, although being very close to other Brassicaceae species from *Arabidopsis*, *Eutrema*, *Cardamine* genera. *CENH3* expression profile analysis showed that before meiosis expression levels of *BsCENH3* in gynoecium of sexual *B. stricta* was more than twice as high compared with *BdCENH3* in gynoecium of apomictic *B. divaricarpa*. By the meiosis time expression levels of both these genes dramatically dropped in gynoecia. After pollination, the expression of both of these genes significantly increases, with *BsCENH3* expression growing faster than *BdCENH3*; however, one day after pollination, the level of *BsCENH3* expression begins to decline, while *BdCENH3* expression continues to increase. The lower expression levels before and after pollination in *B. divaricarpa* (ES9) could indicate to a feasible *CENH3* role in apomeiosis and initiation of parthenogenesis.

In sexually reproduced plants, two sperms enter the embryo sac, while one of them fuses with the central cell nucleus that further forms endosperm, the second fertilizes the egg, in contrast, during pseudogamous apomixis, the embryo develops without fertilization, which is the cause of the "spare sperm problem" [59,60]. In apomicts fertilization of the central cell with haploid sperm generally leads to a 4m:1p genomes ratio in endosperm cells, which causes seed abortion [61]. Apomictic species can tolerate such deviations in endosperm via changing their imprinting systems, however, preventing the 'spare sperm' from fusing with the central cell nucleus, might also be important [62]. The fusion of reduced or unreduced 'spare sperm' to the central cell can potentially affect the parental genomic ratio. Therefore, it is not known how the central cell in apomicts can avoid fertilization by enhancing a polyspermy barrier for that 'spare sperm problem' [60]. Thus, different expression behavior of *CENH3* before and after pollination in *B. divaricarpa* vs *B. stricta* gynoecium/silique might suggest its involvement in apomictic development such as elimination of the chromosomes from male gamete during endosperm development. However, this assumption requires the further proof.

Promoter analysis revealed the presence of several MYB transcription factor binding and recognition motifs within the promoter regions of *BsCENH3* and *BdCENH3*. These motifs in *Boecheira CENH3* promoters might suggest a regulation of these genes by a MYB family proteins. A study on *Arabidopsis CENH3* promoter region and its regulation revealed two E2F binding region at -163 and -115 sites [59]. Our analysis of these regions in *Boecheira* species detected the motif at -163 site in *BsCENH3* and *BdCENH3*. At -115 site *BdCENH3* and *BsCENH3* had a single polymorphic site (GCGGGAAA/ GCGGGAAG). Further functional studies on the *Boecheira CENH3* gene and its epigenetic and transcription regulation could elucidate a functional difference of *CENH3* between apomictic and sexual *Boecheira*.

As to *APOLLO*, it is the only so far found gene comprising the alleles with apomixis-associated polymorphism in *Boecheira* species [18]. Thus, identification of *APOLLO* apo-alleles might be used as molecular markers to spot apomictic individuals among the *Boecheira* species. Earlier it was shown that *APOLLO* encodes the exonuclease NEN3 and suggested an evolutionary scenario where after series of duplications one of the NEN3 protein copies of the *Boecheira* ancestors acquired an alter function leading to apomictic development from the sexual state [19]. Moreover, it was demonstrated that apo-alleles are under a positive selection [19]. In the current study we retrieved 5 apo- and 5 sex-alleles from [18] to screen the genomes of ten *Boecheira* species with different reproduction modes as well as other Brassicaceae species from the *Arabidopsis*, *Eutrema*, *Cardamine*, and *Capsella* genera and to perform *APOLLO* gene phylogenetic analysis. The results showed that sexual and apomictic *Boecheira* species are clustered into the separate sub-clades, while being very similar to each other and to other Brassicaceae species. As well as for the *CENH3* gene, functional studies of the *APOLLO* alleles and their putative epigenetic and transcriptional regulation in *Boecheira* species is required in order to find out if there are functional differences between sexual and apomictic accessions. However, the latter will only be possible when good quality diploid genome assemblies of the studied species have been implemented. Investigation of *APOLLO* expression in premeiotic gynoecium and anthers and after meiosis in siliques of sexual diploid *B. stricta* and apomictic diploid *B. divaricarpa* 1-3 DAP revealed expression patterns somewhat different from those reported before [39]. Strong suppression of the *APOLLO* gene activity during meiosis and early postmeiotic stage was observed in both apomictic and sexual plant gynoecia.

However, one day after pollination, the *APOLLO* level increased faster in siliques of apomictics as compared to sexual plants, which might affect the development of embryo sacs and parthenogenesis in apomictic plants. A higher level of *APOLLO* expression in premeiotic anthers in apomictic *B. divaricarpa* may indicate to a role of *APOLLO* in meiosis or apomeiosis during pollen formation in apomictic plants since the male gametophyte development in apomictic *B. divaricarpa* were reported to produce both reduced and nonreduced gametes [38,39]. Therefore, we infer that *APOLLO* may have species specific function for the regulation of meiosis/apomeiosis in *Boechea* and have diverse functions in pseudogamous apomicts compared to sexual relatives.

In conclusion, a detailed knowledge of the structure, phylogeny of genes related to apomixis and the dynamics of their expression can presumably help to better understand the nature and regulation of apomixis vs sexual reproduction and facilitate further study of the evolutionary, ecological, and population role of apomixis. However, for more accurate studies on the phylogeny and evolution of the *Boechea* species, it is necessary to have a good quality diploid level whole genome assembly for these species, which we are now actively working on.

Author Contributions: V.B., K.M.T. designed the study. F.S., I.K. and A.Ö. grown plants, extracted RNA, *CENH3* expression and promoter analysis. E.B., M.R. bioinformatic work, gene alignment, annotation and phylogenetic analysis. V.B., K.M.T., E.B. wrote the manuscript with revision by all other authors. All authors read and approved the final manuscript.

Funding:

The reported study was supported by the Russian Foundation for Basic Research (RFBR) research grant No. 20-54-46002 CT_a to V.B.; TÜBİTAK research grant 119N384 and research grant No. FHD-2016-1038 from Çanakkale Onsekiz Mart University BAP to K.M.T.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Muller, H.J. Some Genetic Aspects of Sex. *The American Naturalist* **1932**, *66*, 118–138, doi:10.1086/280418.
2. Fisher, R.A. *The Genetical Theory of Natural Selection: By R.A. Fisher ; Edited with a Foreword and Notes by J.H. Bennett*; A complete variorum ed.; Oxford University Press: Oxford, 1999; ISBN 978-0-19-850440-5.
3. Hill, W.G.; Robertson, A. The Effect of Linkage on Limits to Artificial Selection. *Genet. Res.* **1966**, *8*, 269–294.
4. Felsenstein, J. The Evolutionary Advantage of Recombination. *Genetics* **1974**, *78*, 737–756.
5. Spillane, C.; Curtis, M.D.; Grossniklaus, U. Apomixis Technology Development—Virgin Births in Farmers' Fields? *Nat Biotechnol* **2004**, *22*, 687–691, doi:10.1038/nbt976.
6. Brukhin, V. Molecular and Genetic Regulation of Apomixis. *Russ J Genet* **2017**, *53*, 943–964, doi:10.1134/S1022795417090046.
7. Carman, J.G. Asynchronous Expression of Duplicate Genes in Angiosperms May Cause Apomixis, Bispory, Tetraspory, and Polyembryony. *Biological Journal of the Linnean Society* **1997**, *61*, 51–94, doi:10.1111/j.1095-8312.1997.tb01778.x.

8. Johri, B.M. *Embryology of Angiosperms*; Springer Berlin Heidelberg: Berlin, Heidelberg, 1984; ISBN 978-3-642-69302-1.
9. Asker, S.E.; Jerling, L. *Apomixis in Plants*; CRC Press: Boca Raton, 1992; ISBN 978-0-8493-4545-6.
10. *The Flowering of Apomixis: From Mechanisms to Genetic Engineering*; Savidan, Y., Carman, J.G., Dresselhaus, T., International Maize and Wheat Improvement Center, Eds.; International Maize and Wheat Improvement Center; European Union; Institut de recherche pour le développement: México, D.F.: [Luxembourg?]: [Paris], 2001; ISBN 978-970-648-074-3.
11. Bicknell, R.A. Understanding Apomixis: Recent Advances and Remaining Conundrums. *THE PLANT CELL ONLINE* **2004**, *16*, S228–S245, doi:10.1105/tpc.017921.
12. *Lost Sex: The Evolutionary Biology of Parthenogenesis*; Schön, I., Martens, K., Dijk, P. van, Eds.; Springer: Dordrecht; New York, 2009; ISBN 978-90-481-2769-6.
13. Kotani, Y.; Henderson, S.T.; Suzuki, G.; Johnson, S.D.; Okada, T.; Siddons, H.; Mukai, Y.; Koltunow, A.M.G. The *LOSS OF APOMEIOSIS (LOA)* Locus in *Hieracium Praealtum* Can Function Independently of the Associated Large-Scale Repetitive Chromosomal Structure. *New Phytol* **2014**, *201*, 973–981, doi:10.1111/nph.12574.
14. Aliyu, O.M.; Schranz, M.E.; Sharbel, T.F. Quantitative Variation for Apomictic Reproduction in the Genus *Boechera* (Brassicaceae). *American Journal of Botany* **2010**, *97*, 1719–1731, doi:10.3732/ajb.1000188.
15. Hojsgaard, D.; Klatt, S.; Baier, R.; Carman, J.G.; Hörandl, E. Taxonomy and Biogeography of Apomixis in Angiosperms and Associated Biodiversity Characteristics. *Critical Reviews in Plant Sciences* **2014**, *33*, 414–427, doi:10.1080/07352689.2014.898488.
16. Hojsgaard, D. Transient Activation of Apomixis in Sexual Neotriploids May Retain Genomically Altered States and Enhance Polyploid Establishment. *Front. Plant Sci.* **2018**, *9*, 230, doi:10.3389/fpls.2018.00230.
17. Brukhin, V.; Baskar, R. A Brief Note on Genes That Trigger Components of Apomixis. *J Biosci* **2019**, *44*, 45, doi:10.1007/s12038-019-9850-1.
18. Corral, J.M.; Vogel, H.; Aliyu, O.M.; Hensel, G.; Thiel, T.; Kumlehn, J.; Sharbel, T.F. A Conserved Apomixis-Specific Polymorphism Is Correlated with Exclusive Exonuclease Expression in Premeiotic Ovules of Apomictic *Boechera* Species. *PLANT PHYSIOLOGY* **2013**, *163*, 1660–1672, doi:10.1104/pp.113.222430.
19. Kliver, S.; Rayko, M.; Komissarov, A.; Bakin, E.; Zhernakova, D.; Prasad, K.; Rushworth, C.; Baskar, R.; Smetanin, D.; Schmutz, J.; et al. Assembly of the *Boechera Retrofracta* Genome and Evolutionary Analysis of Apomixis-Associated Genes. *Genes* **2018**, *9*, 185, doi:10.3390/genes9040185.
20. Lermontova, I.; Koroleva, O.; Rutten, T.; Fuchs, J.; Schubert, V.; Moraes, I.; Koszegi, D.; Schubert, I. Knockdown of CENH3 in Arabidopsis Reduces Mitotic Divisions and Causes Sterility by Disturbed Meiotic Chromosome Segregation: *Consequences of AtCENH3 Depletion*. *The Plant Journal* **2011**, *68*, 40–50, doi:10.1111/j.1365-313X.2011.04664.x.
21. Lermontova, I.; Sandmann, M.; Mascher, M.; Schmit, A.-C.; Chabouté, M.-E. Centromeric Chromatin and Its Dynamics in Plants. *Plant J* **2015**, *83*, 4–17, doi:10.1111/tpj.12875.
22. Talbert, P.B.; Henikoff, S. Histone Variants — Ancient Wrap Artists of the Epigenome. *Nat Rev Mol Cell Biol* **2010**, *11*, 264–275, doi:10.1038/nrm2861.
23. Maheshwari, S.; Tan, E.H.; West, A.; Franklin, F.C.H.; Comai, L.; Chan, S.W.L. Naturally Occurring Differences in CENH3 Affect Chromosome Segregation in Zygotic Mitosis of Hybrids. *PLoS Genet* **2015**, *11*, e1004970, doi:10.1371/journal.pgen.1004970.
24. Ravi, M.; Shibata, F.; Ramahi, J.S.; Nagaki, K.; Chen, C.; Murata, M.; Chan, S.W.L. Meiosis-Specific Loading of the Centromere-Specific Histone CENH3 in Arabidopsis Thaliana. *PLoS Genet* **2011**, *7*, e1002121, doi:10.1371/journal.pgen.1002121.

25. Ravi, M.; Chan, S.W.L. Haploid Plants Produced by Centromere-Mediated Genome Elimination. *Nature* **2010**, *464*, 615–618, doi:10.1038/nature08842.
26. Marimuthu, M.P.A.; Jolivet, S.; Ravi, M.; Pereira, L.; Davda, J.N.; Cromer, L.; Wang, L.; Nogue, F.; Chan, S.W.L.; Siddiqi, I.; et al. Synthetic Clonal Reproduction Through Seeds. *Science* **2011**, *331*, 876–876, doi:10.1126/science.1199682.
27. Karimi-Ashtiyani, R.; Ishii, T.; Niessen, M.; Stein, N.; Heckmann, S.; Gurushidze, M.; Banaei-Moghaddam, A.M.; Fuchs, J.; Schubert, V.; Koch, K.; et al. Point Mutation Impairs Centromeric CENH3 Loading and Induces Haploid Plants. *Proc Natl Acad Sci USA* **2015**, *112*, 11211–11216, doi:10.1073/pnas.1504333112.
28. Evtushenko, E.V.; Lipikhina, Y.A.; Stepanov, P.I.; Vershinin, A.V. Cytogenetic and Molecular Characteristics of Rye Genome in Octoploid Triticale ($\times$ Triticosecale Wittmack). *CCG* **2019**, *13*, 423–434, doi:10.3897/CompCytogen.v13i4.39576.
29. Koch, M.; Haubold, B.; Mitchell-Olds, T. Molecular Systematics of the Brassicaceae: Evidence from Coding Plastidic MatK and Nuclear Chs Sequences. *Am. J. Bot.* **2001**, *88*, 534–544.
30. Al-Shehbaz, I.A. Transfer of Most North American Species of *Arabis* to *Boechera* (Brassicaceae). *Novon* **2003**, *13*, 381, doi:10.2307/3393366.
31. Brukhin, V.; Osadtchiy, J.V.; Florez-Rueda, A.M.; Smetanin, D.; Bakin, E.; Nobre, M.S.; Grossniklaus, U. The *Boechera* Genus as a Resource for Apomixis Research. *Front. Plant Sci.* **2019**, *10*, 392, doi:10.3389/fpls.2019.00392.
32. Böcher T. Experimental Taxonomical Studies in the *Arabis Holboellii* Complex. *Svensk Botanisk Tidskrift* **1954**, *48*, 31–44.
33. Naumova, T.N.; van der Laak, J.; Osadtchiy, J.; Matzk, F.; Kravtchenko, A.; Bergervoet, J.; Ramulu, K.S.; Boutilier, K. Reproductive Development in Apomictic Populations of *Arabis Holboellii* (Brassicaceae). *Sex Plant Reprod* **2001**, *14*, 195–200, doi:10.1007/s00497-001-0118-0.
34. Taskin, K.M.; Turgut, K.; Scott, R.J. Apomictic Development in *Arabis Gunnisoniana*. *Israel Journal of Plant Sciences* **2004**, *52*, 155–160, doi:10.1560/L3DE-FMVY-1XCQ-QRY5.
35. Koch, M.A. 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* **2003**, *20*, 338–350, doi:10.1093/molbev/msg046.
36. Voigt-Zielinski, M.-L.; Piwczyński, M.; Sharbel, T.F. Differential Effects of Polyploidy and Diploidy on Fitness of Apomictic *Boechera*. *Sex Plant Reprod* **2012**, *25*, 97–109, doi:10.1007/s00497-012-0181-8.
37. Schranz, M.E.; Dobes, C.; Koch, M.A.; Mitchell-Olds, T. Sexual Reproduction, Hybridization, Apomixis, and Polyploidization in the Genus *Boechera* (Brassicaceae). *American Journal of Botany* **2005**, *92*, 1797–1810, doi:10.3732/ajb.92.11.1797.
38. Schranz, M.E.; Kantama, L.; de Jong, H.; Mitchell-Olds, T. Asexual Reproduction in a Close Relative of *Arabidopsis*: A Genetic Investigation of Apomixis in *Boechera* (Brassicaceae). *New Phytol* **2006**, *171*, 425–438, doi:10.1111/j.1469-8137.2006.01765.x.
39. Taskin, K.M.; Turgut, K.; Scott, R.J. Apomeiotic Pollen Mother Cell Development in the Apomictic *Boechera* Species. *Biologia plant.* **2009**, *53*, 468–474, doi:10.1007/s10535-009-0087-y.
40. Lee, C.-R.; Wang, B.; Mojica, J.P.; Mandáková, T.; Prasad, K.V.S.K.; Goicoechea, J.L.; Perera, N.; Hellsten, U.; Hundley, H.N.; Johnson, J.; et al. Young Inversion with Multiple Linked QTLs under Selection in a Hybrid Zone. *Nat Ecol Evol* **2017**, *1*, 0119, doi:10.1038/s41559-017-0119.

41. Kiefer, C.; Willing, E.-M.; Jiao, W.-B.; Sun, H.; Piednoël, M.; Hümann, U.; Hartwig, B.; Koch, M.A.; Schneeberger, K. Interspecies Association Mapping Links Reduced CG to TG Substitution Rates to the Loss of Gene-Body Methylation. *Nat. Plants* **2019**, *5*, 846–855, doi:10.1038/s41477-019-0486-9.
42. Lovell, J.T.; Aliyu, O.M.; Mau, M.; Schranz, M.E.; Koch, M.; Kiefer, C.; Song, B.-H.; Mitchell-Olds, T.; Sharbel, T.F. On the Origin and Evolution of Apomixis in *Boechera*. *Plant Reprod* **2013**, *26*, 309–315, doi:10.1007/s00497-013-0218-7.
43. McCarthy, T.W.; Chou, H.; Brendel, V.P. SRAssembler: Selective Recursive Local Assembly of Homologous Genomic Regions. *BMC Bioinformatics* **2019**, *20*, 371, doi:10.1186/s12859-019-2949-4.
44. Kajitani, R.; Toshimoto, K.; Noguchi, H.; Toyoda, A.; Ogura, Y.; Okuno, M.; Yabana, M.; Harada, M.; Nagayasu, E.; Maruyama, H.; et al. Efficient de Novo Assembly of Highly Heterozygous Genomes from Whole-Genome Shotgun Short Reads. *Genome Res.* **2014**, *24*, 1384–1395, doi:10.1101/gr.170720.113.
45. Slater, G.; Birney, E. Automated Generation of Heuristics for Biological Sequence Comparison. *BMC Bioinformatics* **2005**, *6*, 31, doi:10.1186/1471-2105-6-31.
46. Goodstein, D.M.; Shu, S.; Howson, R.; Neupane, R.; Hayes, R.D.; Fazo, J.; Mitros, T.; Dirks, W.; Hellsten, U.; Putnam, N.; et al. Phytozome: A Comparative Platform for Green Plant Genomics. *Nucleic Acids Research* **2012**, *40*, D1178–D1186, doi:10.1093/nar/gkr944.
47. The UniProt Consortium UniProt: A Worldwide Hub of Protein Knowledge. *Nucleic Acids Research* **2019**, *47*, D506–D515, doi:10.1093/nar/gky1049.
48. Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.; Wilkins, M.R.; Appel, R.D.; Bairoch, A. Protein Identification and Analysis Tools on the ExPASy Server. In *The Proteomics Protocols Handbook*; Walker, J.M., Ed.; Humana Press: Totowa, NJ, 2005; pp. 571–607 ISBN 978-1-58829-343-5.
49. Yu, C.-S.; Chen, Y.-C.; Lu, C.-H.; Hwang, J.-K. Prediction of Protein Subcellular Localization. *Proteins* **2006**, *64*, 643–651, doi:10.1002/prot.21018.
50. Marchler-Bauer, A.; Derbyshire, M.K.; Gonzales, N.R.; Lu, S.; Chitsaz, F.; Geer, L.Y.; Geer, R.C.; He, J.; Gwadz, M.; Hurwitz, D.I.; et al. CDD: NCBI's Conserved Domain Database. *Nucleic Acids Research* **2015**, *43*, D222–D226, doi:10.1093/nar/gku1221.
51. Lescot, M. PlantCARE, a Database of Plant Cis-Acting Regulatory Elements and a Portal to Tools for in Silico Analysis of Promoter Sequences. *Nucleic Acids Research* **2002**, *30*, 325–327, doi:10.1093/nar/30.1.325.
52. Edgar, R.C. MUSCLE: Multiple Sequence Alignment with High Accuracy and High Throughput. *Nucleic Acids Research* **2004**, *32*, 1792–1797, doi:10.1093/nar/gkh340.
53. Waterhouse, A.M.; Procter, J.B.; Martin, D.M.A.; Clamp, M.; Barton, G.J. Jalview Version 2--a Multiple Sequence Alignment Editor and Analysis Workbench. *Bioinformatics* **2009**, *25*, 1189–1191, doi:10.1093/bioinformatics/btp033.
54. Jones, D.T.; Taylor, W.R.; Thornton, J.M. The Rapid Generation of Mutation Data Matrices from Protein Sequences. *Bioinformatics* **1992**, *8*, 275–282, doi:10.1093/bioinformatics/8.3.275.
55. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular Biology and Evolution* **2018**, *35*, 1547–1549, doi:10.1093/molbev/msy096.
56. Kolde, R. Pheatmap: Pretty Heatmaps. R Package Version 1.0.12 Available online: <https://CRAN.R-project.org/package=pheatmap>.
57. Ambawat, S.; Sharma, P.; Yadav, N.R.; Yadav, R.C. MYB Transcription Factor Genes as Regulators for Plant Responses: An Overview. *Physiol Mol Biol Plants* **2013**, *19*, 307–321, doi:10.1007/s12298-013-0179-1.

58. Mieulet, D.; Jolivet, S.; Rivard, M.; Cromer, L.; Vernet, A.; Mayonove, P.; Pereira, L.; Droc, G.; Courtois, B.; Guiderdoni, E.; et al. Turning Rice Meiosis into Mitosis. *Cell Res* **2016**, *26*, 1242–1254, doi:10.1038/cr.2016.117.
59. Spielman, M.; Vinkenoog, R.; Scott, R.J. Genetic Mechanisms of Apomixis. *Philos Trans R Soc Lond B Biol Sci* **2003**, *358*, 1095–1103, doi:10.1098/rstb.2003.1298.
60. Scott, R.J. Polyspermy in Apomictic Crataegus: Yes and No. *New Phytol* **2007**, *173*, 227–229, doi:10.1111/j.1469-8137.2006.01958.x.
61. Grossniklaus, U.; Spillane, C.; Page, D.R.; Köhler, C. Genomic Imprinting and Seed Development: Endosperm Formation with and without Sex. *Curr Opin Plant Biol* **2001**, *4*, 21–27, doi:10.1016/s1369-5266(00)00130-8.
62. Scott, R.J.; Armstrong, S.J.; Doughty, J.; Spielman, M. Double Fertilization in Arabidopsis Thaliana Involves a Polyspermy Block on the Egg but Not the Central Cell. *Mol Plant* **2008**, *1*, 611–619, doi:10.1093/mp/ssn016.