A Conserved Apomixis-Specific Polymorphism Is Correlated with Exclusive Exonuclease Expression in Premeiotic Ovules of Apomictic *Boechera* Species[1][W][OPEN]

José M. Corral, Heiko Vogel, Olawale M. Aliyu, Götz Hensel, Thomas Thiel, Jochen Kumlehn, and Timothy F. Sharbel*

Leibniz Institute of Plant Genetics and Crop Plant Research, D–06466 Gatersleben, Germany (J.M.C., O.M.A., G.H., T.T., J.K., T.F.S.); and Genomics Research Group, Department of Entomology, Max Planck Institute for Chemical Ecology, D–07745 Jena, Germany (H.V.)

ORCID ID: 0000-0001-7110-5756 (T.F.S.).

Apomixis (asexual seed production) is characterized by meiotically unreduced egg cell production (apomeiosis) followed by its parthenogenetic development into offspring that are genetic clones of the mother plant. Fertilization (i.e. pseudogamy) of the central cell is important for the production of a functional endosperm with a balanced 2:1 maternal:paternal genome ratio. Here, we present the *APOLLO* (for apomixis-linked locus) gene, an Aspartate Glutamate Aspartate Aspartate histidine exonuclease whose transcripts are down-regulated in sexual ovules entering meiosis while being up-regulated in apomeiotic ovules at the same stage of development in plants of the genus *Boechera*. *APOLLO* has both "apoalleles," which are characterized by a set of linked apomixis-specific polymorphisms, and "sexalleles." All apomictic *Boechera* spp. accessions proved to be heterozygous for the *APOLLO* gene (having at least one apoallele and one sexallele), while all sexual genotypes were homozygous for sexalleles. Apoalleles contained a 20-nucleotide polymorphism present in the 5′ untranslated region that contains specific transcription factor-binding sites for *ARABIDOPSIS THALIANA HOMEOBOX PROTEIN5*, *LIM1* (for *LINEAGE ABNORMAL11*, *INSULIN1*, *MECHANOSENSORY PROTEIN3*), *SORLIP1AT* (for *SEQUENCES OVERREPRESENTED IN LIGHT-INDUCED PROMOTERS IN ARABIDOPSIS THALIANA1*), *SORLIP2AT*, and *POLYA SIGNAL1*. In the same region, sexalleles contain transcription factor-binding sites for *DNA BINDING WITH ONE FINGER2*, *DNA BINDING WITH ONE FINGER3*, and *PROLAMIN BOX-BINDING FACTOR*. Our results suggest that the expression of a single deregulated allele could induce the cascade of events leading to asexual female gamete formation in an apomictic plant.

A mixture of perplexity and frustration were Gregor Mendel's feelings when he tried to verify the fundamental laws of genetic inheritance by transferring his crossing experiments from genus *Pisum* to genus *Hieracium*. Instead of obtaining a variable F2 progeny due to trait segregation, genus *Hieracium* F2s and subsequent generations were morphologically identical to the F1 mother plants. In the following years, two different types of inheritance in plants were defined: the genus *Pisum* type, where traits segregate, and the genus *Hieracium* type, where traits rarely segregate (Nogler, 2006; Koltunow et al., 2011). The latter type corresponds to the now well-defined phenomenon of apomixis, asexual seed production whereby progeny are genetic clones of the mother lineage.

Although apomixis is found in more than 400 species (40 families) of angiosperms (Carman, 1997; Ozias-Akins, 2006), it is poorly represented among crop species, and its molecular mechanisms have not been clearly understood. Introducing apomixis into sexual crops represents a potential agricultural revolution, as F1 vigor could be genetically fixed through successive seed generations. While it represents a disruptive technology, apomixis could bring significant economic and social benefits to society; for example, the production of hybrid rice (*Oryza sativa*) alone would provide an estimated benefit of $2.5 billion per annum (McMeniman and Lubulwa, 1997).

It has been demonstrated that apomixis is under genetic control (Savidan, 1982; Nogler, 1984; Barcaccia et al., 1998; Grimanelli et al., 1998; Ozias-Akins et al., 1998; Pessino et al., 1998; Morell, 1999; Noyes and Rieseberg, 2000). For the genera *Panicum*, *Ranunculus*, and *Hieracium*, apomixis is inherited as a single Mendelian trait (Savidan, 1982; Nogler, 1984; Morell, 1999), which could reflect a single master regulatory gene controlling all three apomixis elements (apomeiosis, parthenogenesis, and endosperm development) or, alternatively, a complex of several tightly linked genes residing in a recombinationally suppressed region (Grossniklaus et al., 2001). On the other hand, the hypothesis that apomixis arises through deregulation of the developmental pathway leading to sexual seed

[1] This work was supported by the Leibniz Institute of Plant Genetics and Crop Plant Research and the European Cooperation in Science and Technology action Harnessing Plant Reproduction for Crop Improvement (grant no. FA0903).

* Address correspondence to sharbel@ipk-gatersleben.de.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) is: Timothy F. Sharbel (sharbel@ipk-gatersleben.de).

[W] The online version of this article contains Web-only data.

[OPEN] Articles can be viewed online without a subscription.

www.plantphysiol.org/cgi/doi/10.1104/pp.113.222430

formation (Belshaw and Quicke, 2003; Koltunow and Grossniklaus, 2003) is supported by heterochrony during early megaspore formation in apomictic *Tripsacum* spp. (Grimanelli et al., 2003) and heterochronic global gene expression patterns in developing ovules of apomictic *Boechera* spp. (Sharbel et al., 2010).

The North American genus *Boechera* (Brassicaceae) has become an exemplary system for ecological and evolutionary studies (Lovell, 2011; Rushworth et al., 2011), one reason being the plethora of applicable molecular genetic resources adopted from its close relative *Arabidopsis thaliana*. The genus is undergoing phylogenetic and taxonomic revision (Templeton et al., 1987; Michiels and Newman, 1998; Savolainen et al., 2000; Rushworth et al., 2011), as species determination is confounded by reticulation, apomixis, polyploidy, and widespread and recurrent hybridization (Böcher, 1951, 1954; Roy and Rieseberg, 1989; Sharbel et al., 2005). The genus *Boechera* is predominantly self-compatible, and its breeding system is variable, consisting of sexual, apomictic, and facultative apomictic forms (Böcher, 1951; Naumova et al., 2001; Aliyu et al., 2010). High levels of inbreeding and homozygosity are characteristic of sexual diploid *Boechera* spp. (McQuade et al., 1994), whereas apomicts are highly heterozygous (Rushworth et al., 2011). Compounding this variability is the wide distribution of polyploidy (mostly $2n = 3x$) and aneuploidy ($2n = 2x + 1$ or $2n = 3x + 1$; Böcher, 1951; Sharbel et al., 2004). Importantly, genus *Boechera* contains the most thoroughly documented case of apomixis at the diploid level (Böcher, 1951; Schranz et al., 2006; Kantama et al., 2007; Aliyu et al., 2010), an extremely rare phenomenon considering that most asexual organisms are polyploid.

Apomictic *Boechera* spp. accessions are characterized by genus *Taraxacum*-type diplospory, whereby the megaspore mother cell (MMC) goes through meiosis I without completing the reductional phase (apomeiosis), followed by normal meiosis II, leading to a dyad of unreduced megaspore-like cells that each have the same ploidy as the mother plant. The micropylar megaspore-like cell of the diplosporous dyad degenerates, while the chalazal megaspore-like cell increases in size, undergoing vacuolization and three subsequent mitoses to form an eight-nucleate embryo sac (Böcher, 1951; Naumova et al., 2001). As with many asexual taxa, microsporogenesis is irregular in apomictic individuals, and normal reduced, nonreduced, and aneuploid pollen can be found within and between different genotypes (Böcher, 1951; Dobeš et al., 2004; Sharbel et al., 2005; Voigt et al., 2007). Analyses of meiosis in pollen cells have additionally demonstrated a range of chromosomal synapsis (univalent and multivalent) in apomictic genotypes (Böcher, 1951; Kantama et al., 2007), and pseudogamous as well as autonomous endosperm formation have been identified (Naumova et al., 2001; Voigt et al., 2007).

Several candidate apomixis genes have been found by comparing sexual and apomictic flower transcriptomes in different plant species (Vielle-Calzada et al., 1996; Chen et al., 1999; Pessino et al., 2001; Rodrigues et al., 2003; Albertini et al., 2004). Here, we describe the *APOLLO* gene (for apomixis-linked locus), which shows differential expression between apomictic and sexual premeiotic ovules with regard to a specific group of alleles that share several apomixis-specific polymorphisms.

RESULTS

Expression Profiling of Microdissected Apomictic and Sexual Ovules Reveals a Single Differentially Regulated Gene

The initial goal of this work was to perform a comparative transcriptome analysis of sexual and apomictic ovules in a large number of biological replicates in order to statistically correct for genetic- and ploidy-mediated background noise in subsequent analyses of differential gene expression. Based upon geography (California to the American Midwest), chloroplast haplotypes, nuclear markers, morphological traits, and seed screen analysis (Sharbel and Mitchell-Olds, 2001; Schranz et al., 2005; Sharbel et al., 2005; Aliyu et al., 2010; Kiefer and Koch, 2012), seven different diploid sexual genotypes and 10 different diploid or triploid apomictic genotypes (a total of 17 genotypes) were chosen for a gene expression analysis of 20 to 40 microdissected live ovules pooled from a single flower from each genotype using a custom (designed by Agilent Technologies) 105K genus *Boechera* microarray (Table I).

Analyses were performed using GeneSpring GX software (version 10; Agilent Technologies), and candidate probes significantly differentially expressed ($P \leq 0.05$) between apomictic and sexual plants were selected based on the following parameters: percentile shift 75 normalization, median as baseline, reproductive mode (apomictic or sexual) as interpretation, Student's t test unpaired as statistical analysis, and Benjamini-Hochberg multiple test corrections. This analysis led to the identification of four different spots (0123118, 0640787, 1189846, and 1650340) whose expression levels were significantly higher in all apomictic samples ($P < 0.01$ for the first three and $P < 0.05$ for the fourth). Furthermore, the four spots had the highest $-\log_{10}$-corrected P values of all array probes when all sexual samples were compared with all apomictic samples. Importantly, when the oligonucleotide sequences of these four spots were BLASTed to a genus *Boechera* complementary DNA (cDNA) database, all four matched perfectly to ESTs belonging to the same genus *Boechera* gene (*APOLLO*), whose amino acid sequence showed highly significant similarities (BLASTp) to the DEDDh (for Aspartate Glutamate Aspartate Aspartate histidine) exonucleases from *Arabidopsis lyrata lyrata* (XM_002888928; E-value 0, 88% identity, 93% positivity), *A. thaliana* (AT1G74390; E-value 0, 85% identity, 91% positivity), and *Brassica rapa pekinensis* (AC189574; E-value 0, 77% identity, 86% positivity).

Table I. *Boechera spp. genotypes*

Experiments were as follows: 1, flow cytometry seed screen; 2, microarrays; 3, gene sequencing; 4 qPCR analyses; 5, RACE; and 6, cDNA sequencing. Alleles are listed with their GenBank accession numbers in parentheses.

Genotype	Reproductive Mode	Ploidy	Collection Locality	Apoalleles	Sexalleles	Experiment
1A2	Apomixis	2	Birch Creek, MT	1A2_A3 (KF705590) 1A2_A6 (KF705598)	1A2_S7 (KF705602) 1A2_S6 (KF705598)	1, 2, 3, 4
11A2	Apomixis	2	Sliderock, Ranch Creek, Granite, MT	11A2_A1 (KF705600) 11A2_A3 (KF705599) 11A2_A5 (KF705591)	11A2_S8 (KF705601)	1, 2, 3, 4
33A2	Apomixis	2	Mule Ranch, MT	33A2_A5 (KF705588)	33A2_S6 (KF705570)	1, 2, 3, 4
111A2	Apomixis	2	Morgan Switch Back, ID	111A2_A8 (KF705592)	111A2_S13 (KF705577)	1, 2, 3, 4
81A2	Apomixis	2	Vipond Park, Beaverhead, MT	–	–	1, 2, 5, 6
168A2	Apomixis	2	Vipond Park, Beaverhead, MT	168A2_A6 (KF705589)	168A2_S4 (KF705585)	1, 2, 3, 4, 5, 6
43A3	Apomixis	3	Highwood Mountains, MT	43A3_A3 (KF705596) 43A3_A7 (KF705593)	43A3_S4 (KF705578)	1, 2, 3, 4
66A3	Apomixis	3	Highwood Mountains, MT	66A3_A3 (KF705597)	66A3_S8 (KF705584)	1, 2, 3, 4, 6
104A3	Apomixis	3	Lost Trail Meadow, ID	104A3_A4 (KF705595)	104A3_S7 (KF705580)	1, 2, 3, 4, 5
215A3	Apomixis	3	Blue Lakes Road, CA	215A3_A7 (KF705594)	215A3_S13 (KF705579) 215A3_S5 (KF705587)	1, 2, 3, 4, 5
369S2	Sex	2	Twin Saddle, ID	–	369S2_S3 (KF705583)	1, 2, 3, 4
376S2	Sex	2	Sagebrush Meadow, MT	–	376S2_S5 (KF705582)	1, 3, 4, 5
380S2	Sex	2	Buffalo Pass, CO	–	380S2_S13 (KF705586)	1, 2, 3, 4
355S2	Sex	2	Gold Creek, CO	–	355S2_S3 (KF705581)	1, 2, 3, 4
329S2	Sex	2	Big Hole Pass, MT	–	329S2_S1 (KF705569)	1, 2, 3, 4
385S2	Sex	2	Parker Meadow, ID	–	385S2_S11 (KF705572) 385S2_S3 (KF705571)	1, 2, 3, 4
344S2	Sex	2	Bandy Ranch, MT	–	344S2_S7 (KF705576)	1, 2, 3, 4, 5
390S2	Sex	2	Panther Creek, ID	–	390S2_S16 (KF705573) 390S2_S1 (KF705574)	1, 2, 3, 4

APOLLO Is Characterized by an Apomixis-Specific Allele

The *APOLLO* gene (approximately 2,538 nucleotides) was cloned and sequenced from 18 genotypes (Table I). The gene is highly polymorphic (S = 139; h = 30; Hd = 0.993; π = 0.01537; θ = 0.01445; for parameter definitions, see "Materials and Methods") and is characterized by different levels of polymorphism between sexual (S = 89; h = 18; Hd = 0.989; π = 0.00868; θ = 0.01046) and apomictic (S = 21; h = 12; Hd = 0.987; π = 0.00153; θ = 0.00273) haplotypes. Nevertheless, several linked apomixis-specific polymorphisms were identified, being present in all 10 apomictic genotypes but not in any sexual genotype, which enabled the grouping of sequences into "apoalleles" (i.e. 13 alleles in total) and "sexalleles" (i.e. 21 alleles in total; Table I). All apomictic plants were heterozygous for the apomixis-specific polymorphisms, having at least one apoallele and one sexallele, whereas all sexual plants were homozygous for sexalleles only.

Considering the sexallele consensus sequence as a reference, the apomixis-specific polymorphisms could be classified into three groups: (1) a single 20-nucleotide insertion/deletion polymorphism in the 5′ untranslated region (UTR; Fig. 1); (2) 12 nonsynonymous polymorphisms (Table II) in the coding sequence; and (3) 16 synonymous polymorphisms located in the coding sequence or intronic regions. Interestingly, the apomixis-specific 20-nucleotide insertion in the 5′ UTR is found both in *A. thaliana* and *B. rapa pekinensis* (with more than 80% similarity; Fig. 1).

Entire 5′ ends of the genus *Boechera APOLLO* transcripts for both apoalleles and sexalleles were obtained by RACE from genotypes 81A2, 168A2, 104A3, 215A3, 376S2, and 344AS2 (Table I). Two classes of transcripts were identified: (1) truncated apoalleles having a 5′ UTR consisting of eight to 58 nucleotides with no further sequence upstream of the 20-nucleotide apomixis-specific polymorphism (Fig 1B, asterisks); and (2) sexalleles having 111- to 407-nucleotide 5′ UTRs (Fig. 1).

Genus *Boechera APOLLO* mRNA varies in length from a minimum of 1,663 nucleotides (apoalleles) to a maximum of 2,062 nucleotides (sexalleles). It furthermore contains an open reading frame of 1,491 nucleotides that encodes a protein of 496 amino acids (as predicted by DNASTAR SeqBuilder version 9.1.0) and has a 3′ UTR of approximately 160 nucleotides. The *APOLLO* gene contains six exons and five introns (Fig. 1) and has a conserved exon/intron structure, as predicted for the mature mRNA of the DEDDh exonuclease from *A. thaliana* (AT1G74390).

Genotype-specific polymorphisms were also identified. For example, the sexallele of the apomictic plant 11A2 contains a duplicated block of 10 amino acids. The sexallele of the apomictic plant 66A3 has a deletion of four consecutive amino acids. The sexual plant 380S2 only showed a single sexallele, which, due to a seven-nucleotide deletion, encodes a truncated 145-amino acid polypeptide. Finally, the sexallele of the apomictic plant 43A3 codes for a truncated 38-amino acid polypeptide due to a single nucleotide deletion.

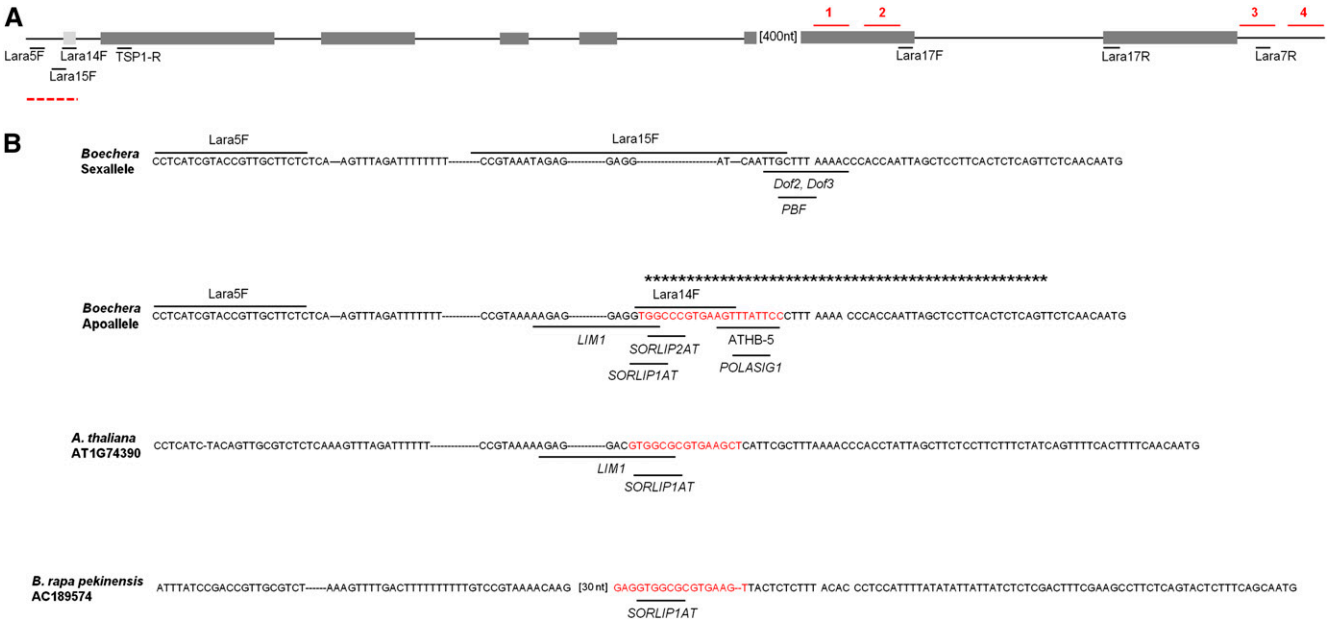

Figure 1. The *APOLLO* gene (5′→3′). A, Dark gray boxes mark coding exons, and the light gray box at the 5′ end marks the 20-nucleotide apomixis-specific polymorphism in genus *Boechera*. The positions of the four positive microarray probes are shown in red above the gene: 1, 1650340; 2, 0123118; 3, 1189846; 4, 0640787 (see "Materials and Methods"). The intron-exon structure is drawn to scale, and hence, a block of 400 nucleotides was removed from exon 5 to accommodate the figure size. Black lines over "Lara" fragments show the positions of PCR primers used in this study (see "Materials and Methods"). The red dashed line designates the variable part of the 5′ UTR, which is expanded in B. B, Fragment of the 5′ UTR that shows the apomixis-specific polymorphism (red letters) in the context of the *Boechera* spp. sexalleles and apoalleles and its orthologous region in *A. thaliana* and *B. rapa pekinensis*. Asterisks encompass the range of 5′ limits of apoallele transcripts isolated and sequenced in this work, while sexallele transcripts begin around 350 nucleotides upstream of this region. Black lines under "Lara" fragments show positions of PCR primers used in this study (see "Materials and Methods"). Black lines below each allele correspond to TBSs (in italics if they lay on the negative strand; see Supplemental Table S2).

A National Center for Biotechnology Information BLASTp of the *APOLLO* predicted amino acid sequence found similarity to related exonucleases, not only in *A. thaliana* and *Brassica* spp. but also in grape (*Vitis vinifera*), *Ricinus communis*, soybean (*Glycine max*), poplar (*Populus trichocarpa*), *Brachypodium distachyon*, rice (*Oryza sativa*), maize (*Zea mays*), sorghum (*Sorghum bicolor*), and barley (*Hordeum vulgare*; E-values from 0.0 to 1e-90, 45%–88% identity, 76%–93% positivity).

Phylogenetic analysis of the different genomic alleles, coding sequences, and polypeptides clustered the apoalleles and sexalleles into two significantly differentiated groups (bootstrap values of 93–100; Fig. 2; Supplemental Fig. S1).

Apomixis-Specific 5′ UTR Polymorphisms in *APOLLO* Correspond to Specific Transcription Factor-Binding Sites

The detection of cis-regulatory elements by PlantPAN analysis of all cloned *APOLLO* 5′ UTRs revealed that the 20-nucleotide apomixis-specific polymorphism (TGGCCCGTGAAGTTTATTCC) was characterized by a plus-strand transcription factor-binding site (TBS; agtTTATTc) for the ATHB-5 (for *ARABIDOPSIS THALIANA HOMEOBOX PROTEIN5*) transcription factor, which is absent in all sexalleles (Fig. 1; Supplemental Table S2). The same 20-nucleotide apomixis-specific polymorphism is characterized by TBSs in the minus strand for *Lim1* (for *LINEAGE ABNORMAL11, INSULIN1, MECHANOSENSORY PROTEIN3*; aagaggaGGTGG), *SORLIP1AT* (for

Table II. *Nonsynonymous apoallele-specific amino acid changes and their equivalents in A. lyrata lyrata, A. thaliana, and B. rapa pekinensis*

The sexallele 390S2_S16 was used as a reference for the amino acid positions. Codes in the left column refer to the reference (sexallele 390S2_S16) amino acid state (first letter), the relative nucleotide position of the polymorphism in allele 390S2_S16 (middle number), and derived apomixis-specific amino acid state (last letter).

Apoallele-Specific Amino Acid Change	*A. lyrata lyrata*	*A. thaliana*	*B. rapa pekinensis*
C11N	S	S	S
G119E	E	E	E
S123P	C	S	S
V177I	V	V	V
Y179H	Y	Y	H
K195T	T	T	T
F200L	F	F	F
S201F	S	S	S
T243A	I	I	D
R285T	T	T	T
G331R	G	G	G
E412K	E	E	K

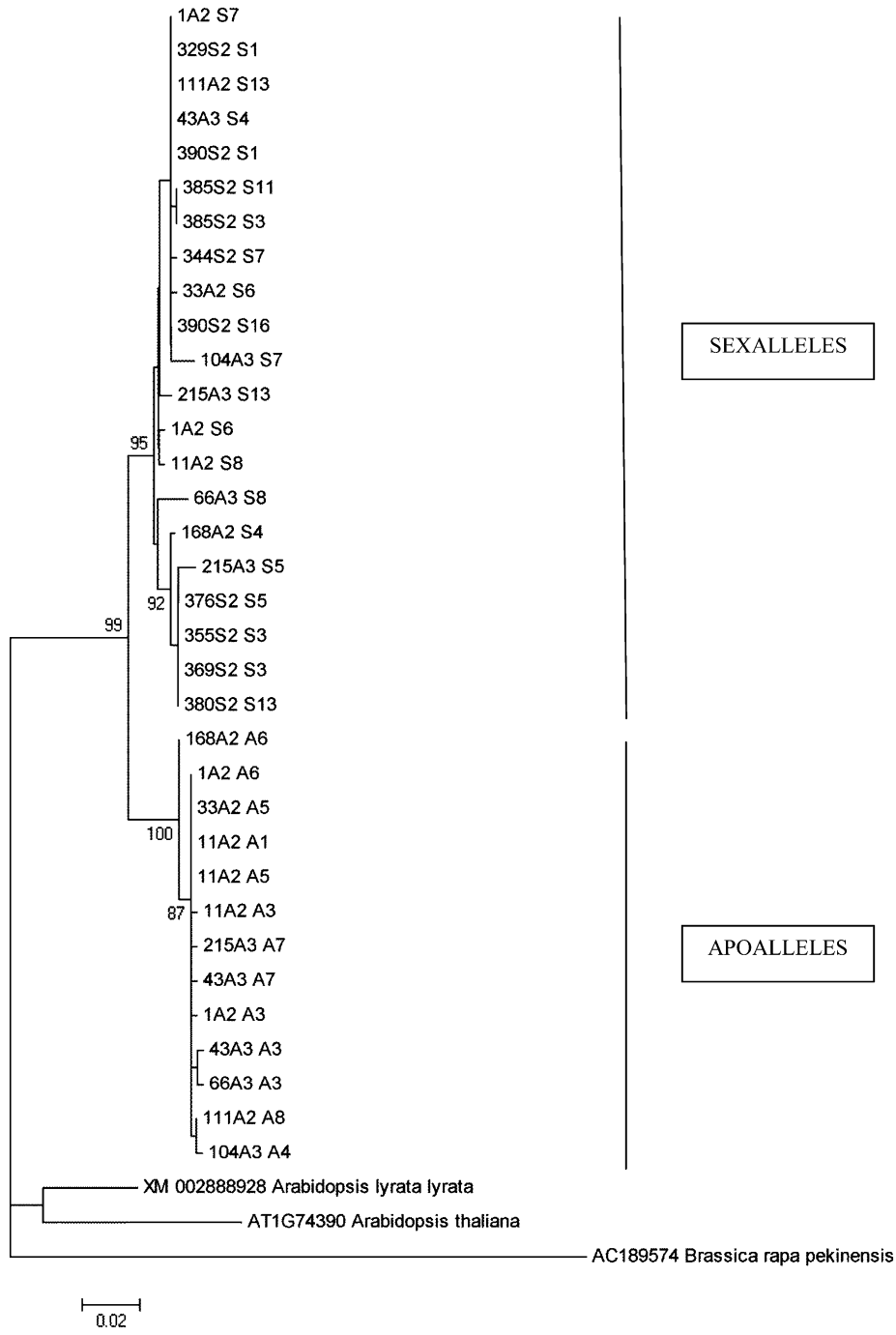

Figure 2. Tree of maximum likelihood obtained with MEGA5 based on APOLLO amino acid sequences (for allele nomenclature, see Table I). Bootstrap values are provided at nodes throughout the tree.

SEQUENCES OVERREPRESENTED IN LIGHT-INDUCED PROMOTERS IN ARABIDOPSIS THALIANA1; GTGGC), *SORLIP2AT* (GGCCC), and *POLYA SIGNAL1* (*POLASIG1*; TTTATT; Fig. 1; Supplemental Table S2). In the same region, sexalleles contained TBSs for *DNA BINDING WITH ONE FINGER3* (*Dof2*)/*Dof3* (ttGCTTTaaaa and TGCTTT) and PBF (for PROLAMIN BOX-BINDING FACTOR; GCTTT; Fig. 1). Neither tandem repeats nor CpNpG islands were detected along the transcript of any allele, but both apoalleles and sexalleles have a putative ath-miR399d microRNA target site (CGGGGCAGCCUUUUCGg) 82 bp downstream from the start codon.

APOLLO Is Expressed in Apomictic, But Not Sexual, Ovules

Sexual plants only have *APOLLO* sexalleles, whereby all apomictic plants are heterozygous for both sexalleles and apoalleles. Hence, only sexallele transcripts contribute to *APOLLO* gene expression in sexual plants, and *APOLLO* is significantly down-regulated in sexual

ovules, compared with anthers and leaves (Mann-Whitney test; $P = 0.001$; Fig. 3). In contrast, *APOLLO* was significantly up-regulated in apomictic ovules when compared with sexuals ($P < 0.001$), an indication of preferential apoallele expression (apoallele versus sexallele expression in apomictic ovules; $P = 0.021$) in apomictic ovules (Fig. 3). Only genotype 1A2 showed an exceptional presence of sexallele expression in ovules (Supplemental Table S1). In apomicts, the apoallele transcripts were overrepresented in leaves ($P = 0.002$) but not in anthers ($P = 0.08$). Comparing sexual versus apomictic plants, *APOLLO* was up-regulated in anthers ($P = 0.001$) and leaves ($P = 0.02$) of the latter (Fig. 3).

APOLLO Apoalleles Have Arisen via Genomic Fragment Duplication

Copy number variation (CNV) exists for *APOLLO* between sexual and apomictic *Boechera* spp. In diploid sexuals, two copies of *APOLLO* were found in all but genotype 369S2 (which showed only one copy), and all copies corresponded to sexalleles (Table III). However, apomictic plants showed increased variation for *APOLLO* copy number (i.e. CNV), ranging from four to six copies in diploid plants and from five to eight copies in triploid plants (Table III). All apomictic plants had at least one copy of both apoalleles and sexalleles. Kendall's correlation τ coefficients, as calculated between *APOLLO* gene and allele expression levels and genomic copy number, showed for the most part positive correlations between them; hence, increased gene expression levels appear to have been attained via *APOLLO* duplication (Supplemental Table S3).

Three different assembled contigs, BAC2b_4, BAC5_7, and BAC6, were obtained by sequencing eight positive bacterial artificial chromosomes (BACs) from the apomictic *Boechera* spp. library. Apoalleles were present in a single copy each in BAC2b_4 and BAC6, and a single sexallele was present in BAC5_7 (Fig. 4). The two BACs (BAC2b_4 and BAC6) containing apoalleles were highly similar in local collinear block presence and order, with evidence of at least two inversions in their distal ends (Fig. 4). In all three BACs, *APOLLO* (both sexalleles and apoalleles) was flanked by a *XYLOGLUCAN 6-XYLOSYLTRANSFERASE XXT5* (AT1G74380; Fig. 4). Furthermore, both apoalleles were flanked by two zinc finger protein coding genes

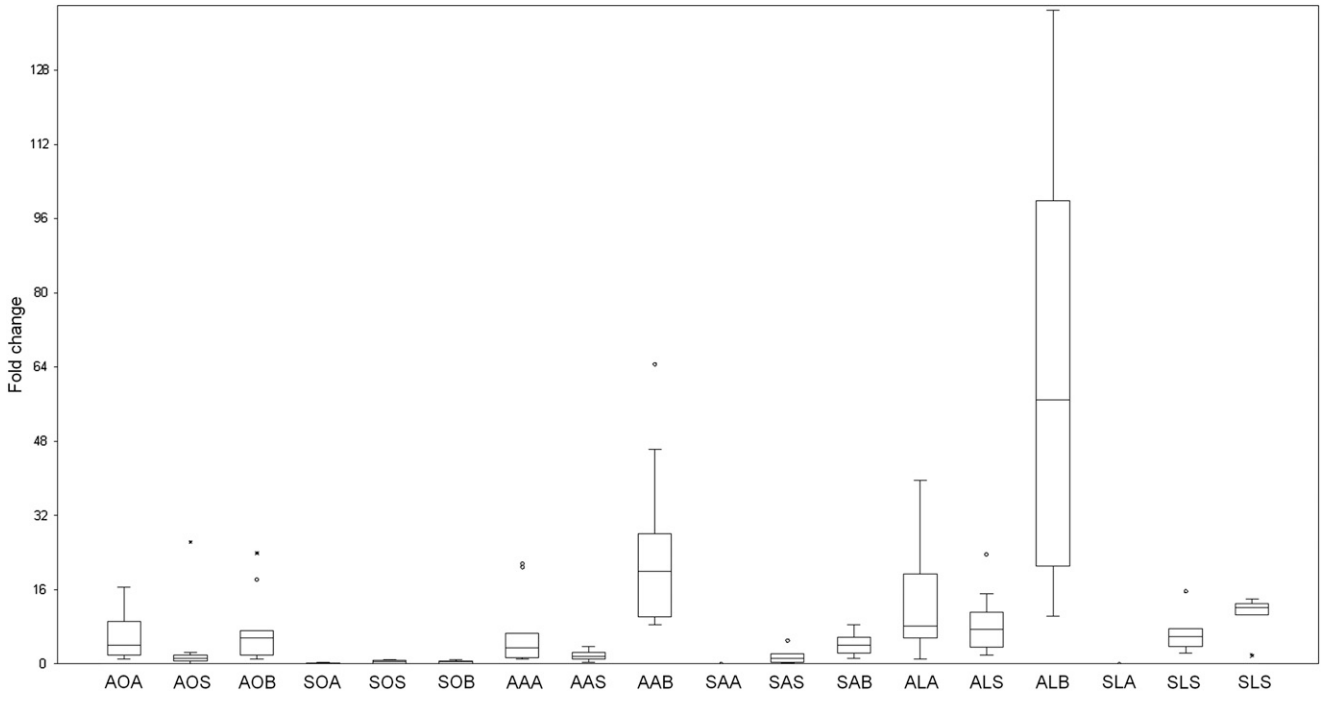

Figure 3. Quantitative reverse transcription-PCR validation of the *APOLLO* gene and sexallele and apoallele expression. Distributions are based on the average of three technical replicates from the 18 apomictic and sexual individuals corresponding to all genotypes described in Table I. Fold change was calculated using *UBQ10* expression as a reference. Different primer combinations were used for the detection of the apoallele (Lara14F and TSP1-R), sexallele (Lara15F and TSP1-R), and both alleles together (Lara17F and Lara17R; see "Materials and Methods"; Fig. 1). AOA, Apomictic ovule apoallele; AOS, apomictic ovule sexallele; AOB, apomictic ovule both alleles; SOA, sexual ovule apoallele; SOS, sexual ovule sexallele; SOB, sexual ovule both alleles; AAA, apomictic anther apoallele; AAS, apomictic anther sexallele; AAB, apomictic anther both alleles; SAA, sexual anther apoallele; SAS, sexual anther sexallele; SAB, sexual anther both alleles; ALA, apomictic leaf apoallele; ALS, apomictic leaf sexallele; ALB, apomictic leaf both alleles; SLA, sexual leaf apoallele; SLS, sexual leaf sexallele, SLB, sexual leaf both alleles. Outliers and far outliers are depicted by white circles and black squares, respectively.

Table III. *Genomic copy number of APOLLO, apoalleles, and sexalleles as determined by qPCR*

Different primer combinations were used for the detection of apoalleles (Lara14F and TSP1-R), sexalleles (Lara15F and TSP1-R), and both alleles together in *APOLLO* (Lara17F and Lara17R; see "Materials and Methods"; Fig. 3). Note that differences in primer efficiencies may explain slight incongruities when comparing apoalleles + sexalleles versus global *APOLLO* copy numbers.

Genotype	Reproduction	*APOLLO*	Apoalleles	Sexalleles	Plant Ploidy
1A2	Apomixis	6	3	3	2
11A2	Apomixis	4	3	2	2
33A2	Apomixis	6	5	1	2
111A2	Apomixis	5	3	2	2
81A2	Apomixis	4	2	1	2
168A2	Apomixis	5	2	1	2
43A3	Apomixis	5	2	3	3
66A3	Apomixis	5	2	3	3
104A3	Apomixis	5	4	1	3
215A3	Apomixis	8	6	3	3
369S2	Sex	1	0	1	2
376S2	Sex	2	0	1	2
380S2	Sex	2	0	1	2
355S2	Sex	2	0	1	2
329S2	Sex	2	0	3	2
385S2	Sex	2	0	2	2
344S2	Sex	2	0	2	2
390S2	Sex	2	0	2	2

(AT1G14580 and AT1G19260) in the same order (Fig. 4). One apoallele (BAC2b_4) and the sexallele (BAC5_7) were flanked by two genes in the same order (AT1G74370 and AT1G74360) on one side, while the sexallele and second apoallele (BAC6) were both flanked by a single gene (AT3G15310) on the opposite side (Fig. 4). The sexallele was furthermore distantly flanked by AT1G18750, the *AGAMOUS-LIKE65* gene, which is involved with pollen development (Adamczyk and Fernandez, 2009). None of the genes flanking *APOLLO* in the three BAC assemblies (Fig. 4) was characterized by a sex- or apomixis-specific CNV, as measured by an array-based comparative genomic hybridization experiment comparing 10 sexual and 10 apomictic *Boechera* spp. genotypes (Aliyu et al., 2013).

DISCUSSION

A Single Gene Is Differentially Expressed between Sexual and Apomictic Ovules in Genetically and Geographically Diverse *Boechera* Spp.

Using a custom-made genus *Boechera*-specific microarray to compare gene expression differences in live microdissected ovules at the MMC stage, between seven diploid sexual and 10 diploid and triploid apomictic genotypes (Table I), four array probes corresponding to a single apomeiosis candidate, the *APOLLO* gene, were identified. Importantly, we show that apoalleles (those that contained conserved apomixis-specific sequence polymorphisms) are exclusively expressed in apomictic ovules, while sexalleles, which are present in both sexual and apomictic *Boechera* spp., are not expressed in either sexual or apomictic ovules (with the exception of genotype 1A2; Supplemental Table S1). Considering that the diverse *Boechera* spp. genotypes used here originated from California to the

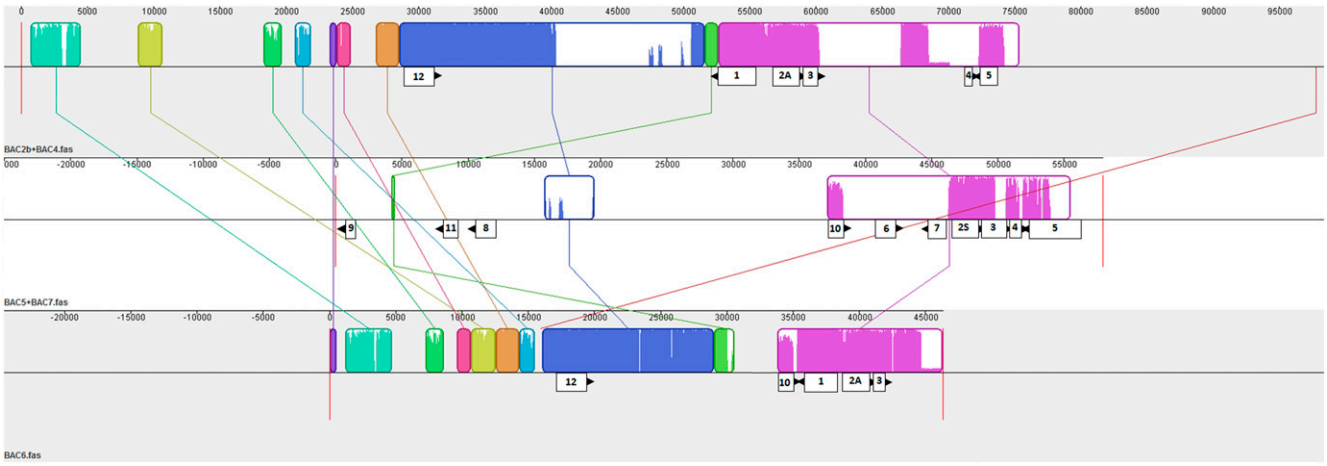

Figure 4. Local colinear sequence blocks in the three BACs containing *APOLLO*. Vertical lines inside each local colinear sequence block (multicolored block) refer to the average nucleotide similarity shared by two genomic regions. Numbers below blocks refer to BAC sequences sharing homology to the following *A. thaliana* genes: (1) AT1G14580, C2H2-like *ZINC FINGER PROTEIN* gene; (2) AT1G74390, putative *EXONUCLEASE* gene *APOLLO* apoallele (A) and sexallele (S); (3) AT1G74380, *XYLOGLUCAN 6-XYLOSYLTRANFERASE XXT5*; (4) AT1G74370, RING/U-box domain-containing protein gene; (5) AT1G74360, putative *LRR RECEPTOR-LIKE SER/THR-PROTEIN KINASE* gene; (6) AT1G74420, *FUCOSYLTRANSFERASE3*; (7) AT1G74400, pentatricopeptide repeat-containing protein gene; (8) AT1G74450, uncharacterized protein gene; (9) AT1G74470, *GERANYLGERANYL REDUCTASE* gene; (10) AT3G15310, unknown protein gene; (11) AT1G18750, *AGAMOUS-LIKE65*; (12) AT1G19260, Transposes Transcription Factor-type zinc finger protein with a HAT dimerization domain gene. GenBank accession numbers are: BAC2b_4 (KF705603), BAC5_7 (KF705604), and BAC6 (KF705605).

American Midwest (i.e. thousands of kilometers apart; Table I), the sharing of an apomixis-specific polymorphism and expression pattern in all apomicts, regardless of different genetic, ploidy, or geographic backgrounds, is highly significant.

As we have focused our analyses on *Boechera* spp. ovules entering apomeiosis/meiosis, it is unclear whether *APOLLO* expression would be exclusively present at the first step of apomixis (apomeiosis) or whether it would be also implied in the subsequent processes that lead to apomictic seed formation (parthenogenesis and endosperm development). Quantitative PCR (qPCR) experiments in ovules, anthers, and leaves have revealed correlated expression of *APOLLO* between anthers and ovules in apomictic genotypes, indicating that the apomeiosis-specific regulatory effects are shared in both reproductive tissues (Fig. 3; Supplemental Table S3).

The *APOLLO* gene described here codes for a DEDD 3′→5′ exonuclease. Enzymes of the DEDD 3′→5′ exonuclease superfamily (also named the DnaQ superfamily) are characterized by four acidic residues, three Asp (D) and one Glu (E), distributed over three separate sequence segments (Exo I, Exo II, and Exo III; Moser et al., 1997; Zuo and Deutscher, 2001). Residues from this DEDD motif constitute binding sites A and B for two divalent metal ions that are essential for the catalysis of nucleoside monophosphates at DNA or RNA termini in the 3′→5′ direction (Derbyshire et al., 1995; Brucet et al., 2008). Besides the substrate-binding sites, the presence of either Tyr or His at the active sites defines the DEDDy or DEDDh family, respectively. The overall structure of DEDDh family members is well preserved, although sequence similarities can be very low outside the Exo segments (Zuo and Deutscher, 2001). DEDDh exonucleases have been shown to be proofreading subunits of bacterial DNA polymerase III, which has an important chromosomal replicase function.

APOLLO Is Characterized by Conserved Apomixis-Specific Polymorphisms and Expression Patterns

Apomictic lineages are expected to accumulate mutations through time (i.e. Muller's ratchet; Muller, 1964), and this is reflected by structural variability in the genomes of apomicts relative to their sexual relatives (Kantama et al., 2007), the result of their hybrid nature and disturbed or absent meiotic segregation. Hence, the conservation of derived molecular genetic variation (in this case, complex apomixis-specific polymorphisms; Fig. 1) between genetically diverse asexual lineages (Table I) implies either (1) the common origin of the trait and/or (2) importance (i.e. selection) for apomictic reproduction.

qPCR validations of *APOLLO* in microdissected ovules at the MMC stage show that it is expressed in apomictic and not sexual ovules. Cloning and sequencing of *APOLLO* have furthermore demonstrated that genetically and geographically diverse apomictic *Boechera* spp. share apoalleles that are characterized by conserved polymorphisms (Fig. 1). Subsequent allele-specific qPCR finally revealed that *APOLLO* expression in apomictic ovules is characterized by the enrichment of apoallele transcripts (Fig. 3). In contrast, somatic (leaf) tissue from apomictic and sexual plants showed comparable patterns of up-regulated *APOLLO* transcripts (Fig. 3). Thus, conserved expression of *APOLLO* in the MMCs of apomictic *Boechera* spp. is tightly linked to conserved apomixis-specific polymorphisms and is suggestive of its importance for apomictic seed formation.

Quantification of copy number revealed that apomictic plants have two to six apoalleles and one to three sexalleles, whereas sexual plants have one to three sexalleles (Table III). Consistent with these data, the sequence contigs of three BACs (isolated from an apomict-derived BAC library) that were positive for *APOLLO* showed that (1) sexalleles and apoalleles do not share similar flanking genomic regions, and hence reside in either duplicated or hemizygous regions, and (2) the flanking regions of two apoalleles from the same individual are reflective of genomic fragment duplication (Fig. 4). Hence, the apoalleles of *APOLLO* are apparently subject to mutation accumulation processes that characterize the apomictic *Boechera* spp. genome (Kantama et al., 2007). This is furthermore supported by the fact that genotype-specific polymorphisms leading to the loss of amino acids or polypeptide chain truncation were also identified in three apomictic genotypes (Fig. 1; see "Results"). Interestingly, reciprocal crosses between apomictic and sexual *Boechera* spp. lineages have led to the hypothesis that at least two doses of an "apomixis factor" are required for producing viable offspring (Schranz et al., 2005), an observation that is concordant with the CNV data presented here.

The apomixis-specific amino acid mutations in the *Boechera* spp. APOLLO polypeptide do not affect the active sites (Fig. 1; Table II), and it is unclear whether these changes substantially affect its enzymatic structure. Nonetheless, as apoallele mRNAs are truncated upstream of the 20-nucleotide polymorphism while sexalleles are not (Fig. 1), translation initiation could hypothetically differ between the alleles. Alternatively, the connection between APOLLO DEDDh catalytic activity and apomeiosis could be explained by the regulatory nature of its own catalytic activity as a cleaver of transcripts, or other polynucleotides, that could be implied in meiosis. For example, the human 3′hExo DEDDh contributes to histone mRNA degradation (Mullen and Marzluff, 2008), and the *Caenorhabditis elegans* homolog, ERI-1, not only degrades small interfering RNA and functions as a negative regulator of RNA interference in neuronal cells but is also required in a complex with Dicer for the accumulation of several endogenous small RNA species (Duchaine et al., 2006). Considering these data together, and assuming that *APOLLO* must not be translated during normal sexual

MMC development (Fig. 3), we propose three scenarios to explain the reproduction-specific expression patterns of *APOLLO* sexalleles and apoalleles: CNV, derepression, and activation.

First, as the number of *APOLLO* copies is higher in apomictic plants (Table III), CNV for *APOLLO* is a potential factor that influences its expression. Assuming that the absence of *APOLLO* transcripts in sexual MMCs arises via posttranscriptional regulation (e.g. via microRNAs), then such mechanisms may not be able to efficiently degrade *APOLLO* mRNAs in developing apomictic MMCs if all duplicated copies of *APOLLO* are being transcribed. Significant positive correlations between CNV and allelic expression in sexual and apomictic tissues are evident (Supplemental Table S3), but as the correlations involve specific alleles and not CNV for *APOLLO* per se (i.e. both alleles), the data are consistent with an additional level of regulatory interaction (see below). Consistent with this idea, the apomictic genotype 1A2 shows an up-regulation of sexalleles in the premeiotic ovules analyzed by qPCR (Supplemental Table S1), an observation that could be explained by mutation accumulation (i.e. Muller's ratchet; Muller, 1964), leading to the perturbation of normal sexallele regulation during MMC formation.

Alternatively, the presence of the 20-nucleotide polymorphism in the 5′ UTR of apoalleles could lead to derepression and/or activation of *APOLLO* in apomictic MMCs. Derepression would arise via loss of the recognition site for a negatively regulating interaction element, while activation would arise if the specific 20-nucleotide polymorphism in the 5′ UTR in apoalleles generates a recognition site for a regulatory element that maintains *APOLLO* expression during MMC development. In the same light, sequence analysis for TBSs on the 5′ UTR revealed that the 20-nucleotide polymorphism is implicated, fully or partially, with specific TBSs for *ATHB-5*, *LIM1*, *SORLIP1AT*, *SORLIP2AT*, and *POLASIG1* in apoalleles. The region in sexalleles contained specific TBSs for *Dof2*, *Dof3*, and *PBF* (Supplemental Table S2). As *APOLLO* apomixis-specific amino acidic polymorphisms do not affect the active sites, the relationship of apomeiosis and the presence of the APOLLO protein in apomictic premeiotic ovules could be explained by the regulatory nature of its own exonuclease catalytic activity as a cleaver of factors implied in meiosis I.

ATHB-5 is a class I homeodomain-Leu zipper protein that positively regulates abscisic acid responsiveness during seedling establishment (Johannesson et al., 2003; Henriksson et al., 2005) and is maternally expressed in *A. thaliana* siliques (McKeown et al., 2011). Knockdown and overexpressing mutants (Abe et al., 2003; Johannesson et al., 2003) suggest that ATHB-5 interacts with other unknown factors, but in no case have unreduced gametes or similar anomalies been detected. *SORLIP1AT* and *SORLIP2AT* are overrepresented in light-induced promoters in *A. thaliana*, with the former being strand independent (Hudson and Quail, 2003). *POLASIG1*, which here was identified on the negative strand, is a highly conserved canonical nucleotide sequence (AAUAAA) across the majority of pre-mRNAs that provides a signal for the cleavage and polyadenylation specificity factor involved in the cleavage of the 3′ signaling region from a pre-mRNA (Loke et al., 2005).

Dof is a family of plant transcription factors that share a highly conserved and unique DNA-binding domain with one Cys-2/Cys-2 zinc finger motif (Umemura et al., 2004). The diversity of plant promoters that are recognized by Dof proteins suggests that these transcription factors regulate a variety of signal-responsive and/or tissue-specific gene expression in plants (Yanagisawa and Sheen, 1998; Yanagisawa and Schmidt, 1999). PBF is an endosperm-specific Dof protein whose binding activity has been detected in maize endosperm nuclei, and in combination with the Leu zipper transcription factor OPAQUE2, it is important for the regulation of *ZEIN* gene expression in endosperm (Vicente-Carbajosa et al., 1997).

To summarize, the conserved apomixis- and sex-specific sequence polymorphisms in the 5′ UTR of *APOLLO* provide binding sites for a number of known regulatory factors (Fig. 1), which could lead to its differential expression at the MMC stage.

APOLLO and the Evolutionary History of Apomixis in *Boechera* Spp.

The widespread occurrence of diploid apomixis in *Boechera* spp. has pointed to hybridization, rather than

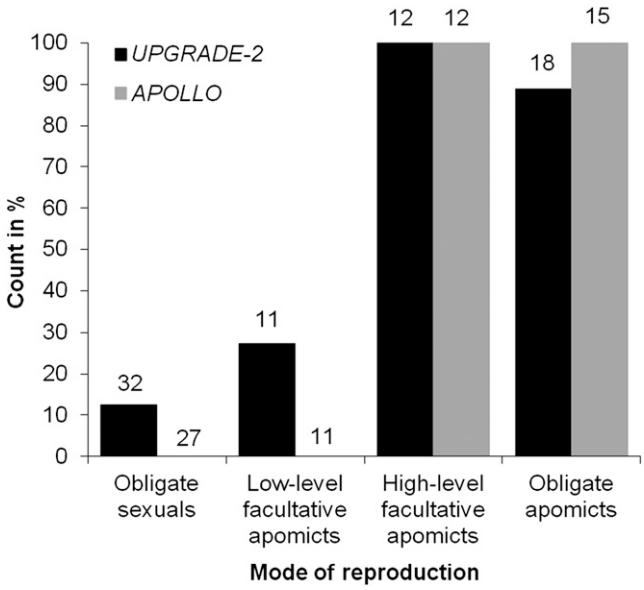

Figure 5. Correlation between *UPGRADE2* (Mau et al., 2013) and *APOLLO* with apomictic *Boechera* spp. The high correlation of *APOLLO* with 73 diploid and triploid apomicts (Aliyu et al., 2010) is also mirrored by similar results for the male meiosis marker *UPGRADE2*.

polyploidy, as the underlying phenomenon, which is correlated with the switch from sex to apomixis (Sharbel et al., 2009, 2010; Beck et al., 2012). In considering hybridization as a driving factor for the origin and spread of apomixis in *Boechera* spp., it could act globally to change the expression of genes associated with sexual seed production (Carman, 1997; Sharbel et al., 2010), or alternatively, hybridization could be a mechanism through which apomixis-specific factors could spread (Schranz et al., 2005).

The data presented here suggest that the *APOLLO* apoallele is monophyletic, having arisen in a single ancestral population, with subsequent gene flow leading to its spread into different sexual backgrounds. This is supported by the facts that (1) 28 single-nucleotide polymorphisms and a 20-nucleotide polymorphism were conserved and specific to apoalleles sequenced from genetically and geographically diverse *Boechera* spp.; and (2) the *APOLLO* sexalleles from both sexual and apomictic plants are closely related and distinct from apoalleles (Fig. 2; Supplemental Fig. S1).

Monophyly of the apoallele is furthermore strongly supported by the presence of the 20-nucleotide 5′ UTR apoallele polymorphism in *A. thaliana* and *B. rapa* (Fig. 1) and points to the deregulation of *APOLLO* as a potential inducer for apomixis, rather than the apomixis-specific polymorphisms themselves. Assuming that sexual *Boechera* spp. accessions down-regulate *APOLLO* during MMC formation by specific interactions between *APOLLO* and regulatory factors, coevolutionary balance (e.g. via binding site homology) between *APOLLO* and these factors is likely to have been attained early in the evolutionary history of genus *Boechera*, as "normal" sexual seed formation (i.e. fitness) would be under strong selection pressure. Second, as *A. thaliana* and *Brassica* spp. (both sexual taxa) share the 5′ UTR apomixis-specific polymorphism (from genus *Boechera*), it can be concluded that (1) the polymorphism is ancestral and (2) the coevolutionary balance between *APOLLO* and its regulatory factors in *A. thaliana* and *Brassica* spp. led to sexual seed formation. Thus, we propose a two-step evolutionary hypothesis in which sequence polymorphisms between *APOLLO* and its regulatory factors first diverged in the early evolution of sexual *Boechera* spp. (or a closely related taxon), followed by reintroduction of the ancestral "apomixis-specific" 5′ UTR polymorphism via an evolutionarily "early" cross with a closely related ancestor in which this polymorphism existed. Hence, hybridization between two closely related taxa could have brought together *APOLLO* and a divergent regulatory factor, which together led to the deregulation of *APOLLO* during MMC development.

In previous work involving flow cytometric analyses of over 20,000 single seeds (Aliyu et al., 2010), we have demonstrated that apomictic *Boechera* spp. are typically characterized by balanced endosperm (6× in diploids and 9× in triploids). These data suggest that both meiotically unreduced egg and pollen are under selection for stable apomictic seed production (Aliyu et al., 2010). Hence, in addition to our attempt to identify factors behind unreduced egg cell formation (this paper), we have taken a similar approach to analyze unreduced pollen formation. This second approach has led to the identification of *UPGRADE2*, a factor that is correlated with unreduced pollen formation in apomictic *Boechera* spp. (Mau et al., 2013).

In order to test the hypothesis that both factors should be required to form stable apomictic seed, the presence of both factors (*APOLLO* and *UPGRADE2*) was assessed in the same 73 genotypes analyzed by Aliyu et al. (2010). As expected, *UPGRADE2* and *APOLLO* are almost 100% correlated in both high-level facultative and obligate apomicts (Fig. 5; Mau et al., 2013). At this time, we have no information regarding the physical linkage of both factors within the genus *Boechera* genome. Nonetheless, the high correlation of two conserved apomixis-specific factors from geographically widespread apomictic genotypes is consistent with the hypothesis that both are required for balanced apomictic seed formation.

CONCLUSION

We anticipate our results to be a starting point for more intensive experiments in genus *Boechera* and other apomictic plants. For example, homologs to this gene are found in sexual crop and model plants; therefore, gene expression analysis and controlled ovule-specific expression/suppression of the appropriate alleles will be relevant not only for the development of agronomic applications but also for understanding how reproductive pathways in ovule are regulated. Validation of the apoallele by means of ectopic expression in sexual accessions will elucidate not only whether the apoallele is required but also whether it is sufficient to confer apomictic seed formation.

MATERIALS AND METHODS

Sample Selection, Ovule Staging, Microdissection, and RNA Extraction

Boechera spp. plants were grown from seedlings onward in a phytotron under controlled environmental conditions. We used the flow cytometric seed screen to analyze and select 18 *Boechera* spp. genotypes that produced 100% sexual or 100% apomictic seeds for the microarray analysis (Table I). In order to compare gene expression changes between meiosis and apomeiosis (Springer and Stupar, 2007; Sharbel et al., 2010), ovules at megasporogenesis between stages 2-II and 2-IV (MMC is differentiated and inner and outer integument are initiated; Lilly and Havey, 2001) were live microdissected from each genotype separately.

The gynoecia of sexual and apomictic *Boechera* spp. were dissected out from nonpollinated flowers at the megasporogenesis stage in 0.55 M sterile mannitol solution at a standardized time (between 8 and 9 AM) over multiple days. Microdissections were done in a sterile laminar air-flow cabinet using a stereoscopic microscope (1000 Stemi; Carl Zeiss) under 2× magnification. The gynoecium was held with forceps while a sterile scalpel was used to cut longitudinally, such that the halves of the silique along with the ovules were immediately exposed to the mannitol. Individual live ovules were subsequently collected using an inverted microscope (Axiovert 200M; Carl Zeiss) in sterile conditions, using sterile glass needles (self-made using a Narishige

PC-10 puller and bent to an angle of about 100°) to isolate ovules from placental tissue. Using a glass capillary (with an opening of 150 μm interior diameter) interfaced to an Eppendorf Cell Tram Vario, the ovules were collected in sterile Eppendorf tubes containing 100 μL of RNA-stabilizing buffer (RNA later; Sigma). Forty ovules per genotype were collected in this way, frozen directly in liquid nitrogen, and stored at −80°C. Total RNA extractions were carried out using the PicoPure RNA isolation kit (Arcturus Bioscience). RNA integrity and quantity were verified on an Agilent 2100 Bioanalyzer using RNA Pico chips (Agilent Technologies).

cDNA Library and Microarray Design

454 (FLX) technology was used to sequence normalized cDNA from flowers pooled at multiple developmental stages from three sexual and three apomictic *Boechera* spp. genotypes (Sharbel et al., 2009, 2010) as a first step in the design of high-density genus *Boechera*-specific microarrays for use in comparisons of gene expression. The 454 sequences were assembled using the CLC Genomics workbench (CLC bio version 5.5) using standard assembly parameters for long-read high-throughput sequences, after trimming of all reads using internal sequence quality scores. In doing so, 36,289 contig sequences and 154,468 nonassembled singleton sequences were obtained, and these data were provided to ImaGenes for microarray development using its preselection strategy service.

The preselection strategy service worked as follows: 14 different oligonucleotides (each 60 bp in length) per contig and eight oligonucleotides per singleton, including the "antisense" sequence of each oligonucleotide, were bioinformatically designed and spotted onto two one-million-spot test arrays. These test arrays were probed using (1) a "complex copy RNA mixture" (obtained by pooling tissues and harvesting all RNA from them) and (2) genomic DNA extracted from leaf tissue pooled from a sexual and an apomictic individual. Based upon the separate hybridization results from the copy RNA and genomic DNA samples, and after all quality tests, a final 2- × 105,000-spot array was designed. The final array thus contained multiple oligonucleotides (i.e. technical replicates) of every gene expressed during *Boechera* spp. flower development.

Microarray Hybridization and Analysis

For each of the 18 selected *Boechera* spp. genotypes, copy RNA was prepared from the microdissected ovules, labeled using the Quick-Amp One-Color labeling kit (Agilent Technologies), and hybridized to the Agilent custom *Boechera* spp. arrays (total of 18 microarrays; Table I). Analyses were performed using GeneSpring GX software (version 10), and candidate probes significantly differentially expressed ($P \leq 0.05$) between apomictic and sexual ovules were selected based on the following parameters: percentile shift 75 normalization, median as baseline, and reproductive mode (apomictic or sexual) as interpretation.

Candidate Gene Cloning

Genomic DNA from leaves was extracted using the DNeasy Plant Mini kit (Qiagen). The full-length transcript from all 18 genotypes (Table I) was amplified using proofreading polymerase (Accuprime), followed by cloning (TOPO-TA cloning kit; Invitrogen) and Sanger sequencing (16–32 clones per genotype were sequenced). Primers Lara5F (5′-CCTCATCGTACCG-TTGCTTCTCTC-3′) and Lara7R (5′-AAGTGCTCACCGACTGACCAG-GAA-3′) used in these amplifications were designed based on *Boechera* spp. cDNA library sequences that were homologous to the four microarray identified spots: 1650340, 0123118, 1189846, and 0640787. Only those sequence variants that were validated in at least two independent PCRs were selected for further analysis.

Sequence Analysis

Obtained genic sequences were edited using SeqMan Pro 9.1.1 (DNASTAR) and aligned with homologous sequences of *Arabidopsis lyrata lyrata* (XM_002888928), *A. thaliana* (AT1G74390), and *Brassica rapa pekinensis* (AC189574) using ClustalV (Higgins et al., 1992) incorporated in MegAlign 9.1.0 (DNASTAR). Phylogenetic analyses were performed with MEGA5 using the maximum likelihood method (Supplemental Fig. S1). The same pipeline was used for coding sequences (Supplemental Fig. S2) and amino acidic sequences (Fig. 2). UTRs from *A. lyrata lyrata* are unknown; therefore, its sequence was only used for the coding sequences and amino acidic sequence comparisons. All alignments are available from the authors upon request.

The following parameters were used to estimate genetic variability between alleles: number of polymorphic sites (S), number of haplotypes (h), haplotype diversity (Hd; Nei, 1987), nucleotide diversity (π; Lynch and Crease, 1990) using the correction of Jukes and Cantor (1969), and number of polymorphic nucleotide sites per nucleotide (θ; Nei 1987). These parameters were obtained with DnaSP version 5.10.01 (Librado and Rozas, 2009).

For promoter analysis of regulatory elements, PlantPAN was used (Plant Promoter Analysis Navigator; http://plantpan.mbc.nctu.edu.tw/), including all known organisms available in its database.

BAC Analysis

Pooled DNA of all tissues and genotypes was used as a template for hybridization probe generation. Two probes of different size (1.6 and 2.3 kb) were prepared by PCR amplification using two pairs of specific primers of the candidate gene genomic sequence. Both probes were labeled and used for hybridization on an apomictic *Boechera* spp. BAC library. There were eight positive hybridizations. The isolated BACs (PureLink Plasmid DNA purification kit) were named 1, 2a, 2b, 3, 4, 5, 6, and 7. Selected BACs were retested using specific primers for the candidate gene. For all BACs except BAC-3, PCR confirmed the presence of the candidate gene. The remaining seven BACs were fingerprinted by restriction enzyme digestion. The restriction patterns of BAC-1 and BAC-2a appeared to be redundant with the other BACs and were removed from downstream analysis, while BACs 2b, 4, 5, 6, and 7 were further processed for sequencing.

High-M_r BAC DNA was isolated using the Nucleobond Xtra Midi kit according to the manufacturer's instructions (Macherey-Nagel). BAC genomic DNA was sheared into two different size ranges (1–1.5 kb and 4–5 kb) with a Hydroshear device (Molecular Devices), blunted with the Quick Blunting kit (New England Biolabs), isolated from an agarose gel, column purified, and ligated into the pUC19-*Sma*I vector (Fermentas). Ligations were transformed into *Escherichia coli* ELECTROMAX DH5α-E electro-competent cells (Invitrogen). Plasmid minipreparation from bacterial colonies grown on 96-deep-well plates was performed using the 96-plasmid DNA isolation kit (Nexttec) on an Evo Freedom 150 robotic platform (Tecan). Sequencing of both termini of BAC shotgun clonal libraries was carried out on an ABI 3730 xl automatic DNA sequencer (PE Applied Biosystems).

Vector clipping, quality trimming, and BAC sequence assembly were done with the STADEN package (version 1.7.0) and the Lasergene DNASTAR SeqMan (version 9.1.0) assembler. For both assemblers, the resulting contigs were visually inspected and corrected if required. Scaffolding was performed by utilizing the combined contig information from both assemblies. If required, gap closing was performed by either direct BAC sequencing using specific sequencing primers next to a gap or by sequencing PCR products resulting from gap-spanning PCRs. BAC sequences could be assembled together for the BAC pairs 2b-4 and 5-7, whereas BAC-6 remained alone. BAC sequences were characterized by BLAST with *A. thaliana* sequences in the National Center for Biotechnology Information.

Collinear sequence block detection was done with Mauve version 2.3.1 (progressive Mauve, default parameters; Darling et al., 2004).

RACE and qPCR

Ovule and anther RNA was extracted using the Arcturus PicoPure RNA isolation kit (Applied Biosystems), and leaf RNA was extracted using the Qiagen RNeasy Plant Mini kit. Ovule, anther, and leaf cDNA was prepared using RevertAid H Minus reverse transcriptase.

The characterization of *APOLLO* transcript 5′ ends was performed using the SMARTer RACE method (Clontech) and the Advantage 2 PCR kit (Clontech). PCR fragments were cloned using the TOPO TA cloning kit for sequencing (Life Technologies) and Sanger sequenced. Complete cDNA sequences were obtained by end-to-end PCR from leaf (genotypes 81A2 and 168A2) and ovule (genotype 66A3) cDNA and were then compared with the corresponding contigs of the genus *Boechera* cDNA library.

For qPCR, the SYBR Green PCR Master Mix (Applied Biosystems) was used. qPCR amplifications were carried out in a 7900HT Fast RT-PCR System machine (Applied Biosystems) with the following temperature profile for SYBR Green assays: initial denaturation at 90°C for 10 min, followed by 40 cycles of 95°C for 15 s and 60°C for 1 min. For checking amplicon quality,

a melting curve gradient was obtained from the product at the end of the amplification. The cycle threshold, defined as the PCR cycle at which a statistically significant increase of reporter fluorescence is first detected, was used as a measure for the starting copy numbers of the target gene. The mean expression level and SD for each set of three technical replicates for each cDNA were calculated. Relative quantification and normalization of the amplified targets were performed by the comparative ΔΔCt method using a calibrator sample in reference to the expression levels of the housekeeping gene *UBQ10* (Pellino et al., 2011).

qPCR analysis of *APOLLO* cDNA from microdissected live ovules (MMC stage), anthers from the same flowers, and leaves from all genotypes used in the microarray (three technical replicates per genotype) were completed. First, a non-allele-specific qPCR analysis was performed with the same samples by using a pair of primers corresponding to conserved regions flanking an intron: Lara17F (5′-CAACTATCCCACAGCCAGACTGCA-3′) and Lara17R (5′-TGGGTTCCATTGGATTCTGAGCTTATT-3′; Fig. 1). Second, using the reverse primer TSP1-R (5′-GATAGCCCCAAACTCCAAAATCGC-3′) in combination with two different forward PCR primers, Lara14F (5′-GTGGCCCGTGAAGTTTATTCCCT-3′) and Lara15F (5′-CCGTAAATAGAGGAGGATCAATTGCTT-3′), which spanned the apomixis-specific 5′ UTR polymorphism, it was possible to measure transcript abundance for both the apoalleles and sexalleles separately (Fig. 3).

APOLLO, apoallele, and sexallele genomic copies were determined using the same combination of primers described for expression quantification but using as a template genomic DNA.

Supplemental Data

The following materials are available in the online version of this article.

Supplemental Figure S1. Tree of maximum likelihood for genomic *APOLLO* sequences.

Supplemental Table S1. Quantitative reverse transcription-PCR validation (relative mRNA levels) of the *APOLLO* gene.

Supplemental Table S2. Transcription factor binding sites.

Supplemental Table S3. Kendall's correlation tau coefficients.

ACKNOWLEDGMENTS

We thank Heidi Block, Vera Katzorke, Leane Boerner, Stephanie Mewes, Silke Fischer, Sheila Milker, and Ines Walde for technical support. We thank Ingo Schubert for constructive comments on the manuscript, Dr. Florian Mette for technical advice, and Tom Mitchell-Olds and Eric Schranz for seed material.

Received June 4, 2013; accepted October 23, 2013; published October 25, 2013.

LITERATURE CITED

Abe H, Urao T, Ito T, Seki M, Shinozaki K, Yamaguchi-Shinozaki K (2003) *Arabidopsis* AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signaling. Plant Cell **15:** 63–78

Adamczyk BJ, Fernandez DE (2009) MIKC* MADS-domain heterodimers are required for pollen maturation and tube growth in Arabidopsis. Plant Physiol **149:** 1713–1723

Albertini E, Marconi G, Barcaccia G, Raggi L, Falcinelli M (2004) Isolation of candidate genes for apomixis in Poa pratensis L. Plant Mol Biol **56:** 879–894

Aliyu OM, Schranz ME, Sharbel TF (2010) Quantitative variation for apomictic reproduction in the genus Boechera (Brassicaceae). Am J Bot **97:** 1719–1731

Aliyu OM, Seifert M, Corral JM, Fuchs J, Sharbel TF (October 29, 2013) Copy number variation in transcriptionally-active regions of sexual and apomictic *Boechera* demonstrates independently-derived apomictic lineages. Plant Cell http://dx.doi.org/10.1105/tpc.113.113860

Barcaccia G, Mazzucato A, Albertini E, Zethof J, Gerats A, Pezzotti M, Falcinelli M (1998) Inheritance of parthenogenesis in *Poa pratensis* L.: auxin test and AFLP linkage analyses support monogenic control. Theor Appl Genet **97:** 74–82

Beck JB, Alexander PJ, Allphin L, Al-Shehbaz IA, Rushworth C, Bailey CD, Windham MD (2012) Does hybridization drive the transition to asexuality in diploid Boechera? Evolution **66:** 985–995

Belshaw R, Quicke DL (2003) The cytogenetics of thelytoky in a predominantly asexual parasitoid wasp with covert sex. Genome **46:** 170–173

Böcher TW (1951) Cytological and embryological studies in the amphi-apomictic *Arabis holboellii* complex. Det Kongelige Danske Videnskabernes Selskab **6:** 1–59

Böcher TW (1954) Experimental taxonomical studies in the *Arabis holboellii* complex. Sven Bot Tidskr **48:** 31–44

Brucet M, Querol-Audi J, Bertlik K, Lloberas J, Fita I, Celada A (2008) Structural and biochemical studies of TREX1 inhibition by metals: identification of a new active histidine conserved in DEDDh exonucleases. Protein Sci **17:** 2059–2069

Carman JG (1997) Asynchronous expression of duplicate genes in angiosperms may cause apomixis, bispory, tetraspory, and polyembryony. Biol J Linn Soc Lond **61:** 51–94

Chen L, Miyazaki C, Kojima A, Saito A, Adachi T (1999) Isolation and characterization of a gene expressed during early embryo sac development in apomictic guinea grass (Panicum maximum). J Plant Physiol **154:** 55–62

Darling ACE, Mau B, Blattner FR, Perna NT (2004) Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res **14:** 1394–1403

Derbyshire V, Pinsonneault JK, Joyce CM (1995) Structure-function analysis of 3′–>5′-exonuclease of DNA polymerases. Methods Enzymol **262:** 363–385

Dobeš C, Mitchell-Olds T, Koch MA (2004) Intraspecific diversification in North American *Boechera stricta* (=*Arabis drummondii*), *Boechera* × *divaricarpa*, and *Boechera holboellii* (Brassicaceae) inferred from nuclear and chloroplast molecular markers: an integrative approach. Am J Bot **91:** 2087–2101

Duchaine TF, Wohlschlegel JA, Kennedy S, Bei Y, Conte D Jr, Pang K, Brownell DR, Harding S, Mitani S, Ruvkun G, et al (2006) Functional proteomics reveals the biochemical niche of *C. elegans* DCR-1 in multiple small-RNA-mediated pathways. Cell **12:** 343–354

Grimanelli D, García M, Kaszas E, Perotti E, Leblanc O (2003) Heterochronic expression of sexual reproductive programs during apomictic development in Tripsacum. Genetics **165:** 1521–1531

Grimanelli D, Leblanc O, Espinosa E, Perotti E, González de León D, Savidan Y (1998) Mapping diplosporous apomixis in tetraploid Tripsacum: one gene or several genes? Heredity (Edinb) **80:** 33–39

Grossniklaus U, Nogler GA, van Dijk PJ (2001) How to avoid sex: the genetic control of gametophytic apomixis. Plant Cell **13:** 1491–1498

Henriksson E, Olsson AS, Johannesson H, Johansson H, Hanson J, Engström P, Söderman E (2005) Homeodomain leucine zipper class I genes in Arabidopsis: expression patterns and phylogenetic relationships. Plant Physiol **139:** 509–518

Higgins DG, Bleasby AJ, Fuchs R (1992) CLUSTAL V: improved software for multiple sequence alignment. Comput Appl Biosci **8:** 189–191

Hudson ME, Quail PH (2003) Identification of promoter motifs involved in the network of phytochrome A-regulated gene expression by combined analysis of genomic sequence and microarray data. Plant Physiol **133:** 1605–1616

Johannesson H, Wang Y, Hanson J, Engström P (2003) The Arabidopsis thaliana homeobox gene ATHB5 is a potential regulator of abscisic acid responsiveness in developing seedlings. Plant Mol Biol **51:** 719–729

Jukes TH, Cantor CR (1969) Evolution of protein molecules. In HN Munro, ed, Mammalian Protein Metabolism. Academic Press, New York, pp 21–132

Kantama L, Sharbel TF, Schranz ME, 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. Proc Natl Acad Sci USA **104:** 14026–14031

Kiefer C, Koch MA (2012) A continental-wide perspective: the genepool of nuclear encoded ribosomal DNA and single-copy gene sequences in North American Boechera (Brassicaceae). PLoS ONE **7:** e36491

Koltunow AM, Grossniklaus U (2003) Apomixis: a developmental perspective. Annu Rev Plant Biol **54:** 547–574

Koltunow AM, Johnson SD, Okada T (2011) Apomixis in hawkweed: Mendel's experimental nemesis. J Exp Bot **62:** 1699–1707

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics **25:** 1451–1452

Lilly JW, Havey MJ (2001) Small, repetitive DNAs contribute significantly to the expanded mitochondrial genome of cucumber. Genetics **159:** 317–328

Loke JC, Stahlberg EA, Strenski DG, Haas BJ, Wood PC, Li QQ (2005) Compilation of mRNA polyadenylation signals in Arabidopsis revealed a new signal element and potential secondary structures. Plant Physiol **138:** 1457–1468

Lynch M, Crease T (1990) The analysis of population survey data on DNA sequence variation. Mol Biol Evol **7:** 377–394

Lovell JT (2011) Boechera Summit 2011. Mol Ecol **20:** 4840–4842

Mau M, Corral JM, Vogel H, Melzer M, Fuchs J, Kuhlmann M, de Storme N, Geelen D, Sharbel TF (2013) The conserved chimeric transcript UPGRADE2 is associated with unreduced pollen formation and is exclusively found in apomictic Boechera species. Plant Physiol **163:** 1640–1659

McKeown PC, Laouielle-Duprat S, Prins P, Wolff P, Schmid MW, Donoghue MT, Fort A, Duszynska D, Comte A, Lao NT, et al (2011) Identification of imprinted genes subject to parent-of-origin specific expression in Arabidopsis thaliana seeds. BMC Plant Biol **11:** 113

McMeniman S, Lubulwa G (1997) Project Development Assessment: An Economic Evaluation of the Potential Benefits of Integrating Apomixis into Hybrid Rice. Australian Centre for International Agricultural Research, Canberra, Australia

McQuade LR, Hill RJ, Francis D (1994) B-chromosome systems in the greater glider, Petauroides volans (Marsupialia: Pseudocheiridae). II. Investigation of B-chromosome DNA sequences isolated by micromanipulation and PCR. Cytogenet Cell Genet **66:** 155–161

Michiels NK, Newman LJ (1998) Sex and violence in hermaphrodites. Nature **391:** 647

Morell V (1999) Ecology returns to speciation studies. Science **284:** 2106–2108

Moser MJ, Holley WR, Chatterjee A, Mian IS (1997) The proofreading domain of Escherichia coliDNA polymerase I and other DNA and/or RNA exonuclease domains. Nucleic Acids Res **25:** 5110–5118

Mullen TE, Marzluff WF (2008) Degradation of histone mRNA requires oligouridylation followed by decapping and simultaneous degradation of the mRNA both 5′ to 3′ and 3′ to 5′. Genes Dev **22:** 50–65

Muller HJ (1964) The relation of recombination to mutational advance. Mutat Res **106:** 2–9

Naumova TN, van der Laak J, Osadtchiy J, Matzk F, Kravtchenko A, Bergervoet J, Ramulu KS, Boutilier K (2001) Reproductive development in apomictic populations of *Arabis holboellii* (Brassicaceae). Sex Plant Reprod **14:** 195–200

Nei M (1987) Molecular Evolutionary Genetics. Columbia University Press, New York

Nogler GA (1984) Genetics of apospory in apomictic *Ranunculus auricomus*. V. Conclusion. Bot Helv **94:** 411–422

Nogler GA (2006) The lesser-known Mendel: his experiments on Hieracium. Genetics **172:** 1–6

Noyes RD, Rieseberg LH (2000) Two independent loci control agamospermy (apomixis) in the triploid flowering plant Erigeron annuus. Genetics **155:** 379–390

Ozias-Akins P (2006) Apomixis: developmental characteristics and genetics. Crit Rev Plant Sci **25:** 199–214

Ozias-Akins P, Roche D, Hanna WW (1998) Tight clustering and hemizygosity of apomixis-linked molecular markers in *Pennisetum squamulatum* implies genetic control of apospory by a divergent locus that may have no allelic form in sexual genotypes. Proc Natl Acad Sci USA **95:** 5127–5132

Pellino M, Sharbel TF, Mau M, Amiteye S, Corral JM (2011) Selection of reference genes for quantitative real-time PCR expression studies of microdissected reproductive tissues in apomictic and sexual Boechera. BMC Res Notes **4:** 303

Pessino SC, Espinoza F, Martínez EJ, Ortiz JP, Valle EM, Quarín CL (2001) Isolation of cDNA clones differentially expressed in flowers of apomictic and sexual Paspalum notatum. Hereditas **134:** 35–42

Pessino SC, Evans C, Ortiz JP, Armstead I, Do Valle C, Hayward MD (1998) A genetic map of the apospory region in *Brachiaria* hybrids: identification of two markers closely associated with the trait. Hereditas **128:** 153–158

Rodrigues JC, Cabral GB, Dusi DM, de Mello LV, Rigden DJ, Carneiro VT (2003) Identification of differentially expressed cDNA sequences in ovaries of sexual and apomictic plants of Brachiaria brizantha. Plant Mol Biol **53:** 745–757

Roy BA, Rieseberg LH (1989) Evidence for apomixis in *Arabis*. J Hered **80:** 506–508

Rushworth CA, Song BH, Lee CR, Mitchell-Olds T (2011) Boechera, a model system for ecological genomics. Mol Ecol **20:** 4843–4857

Savidan YH (1975) Hérédité de l'apomixie: contribution à l'étude de l'hérédité de l'apomixie sur *Panicum maximum* Jacq. Analyse des sacs embryonnaires. Cah ORTSOM Ser Biol **10:** 91–95

Savolainen O, Langley CH, Lazzaro BP, Freville H (2000) Contrasting patterns of nucleotide polymorphism at the alcohol dehydrogenase locus in the outcrossing Arabidopsis lyrata and the selfing Arabidopsis thaliana. Mol Biol Evol **17:** 645–655

Schranz ME, Dobeš C, Koch MA, Mitchell-Olds T (2005) Sexual reproduction, hybridization, apomixis, and polyploidization in the genus Boechera (Brassicaceae). Am J Bot **92:** 1797–1810

Schranz ME, 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). New Phytol **171:** 425–438

Sharbel TF, Mitchell-Olds T (2001) Recurrent polyploid origins and chloroplast phylogeography in the *Arabis holboellii* complex (Brassicaceae). Heredity **87:** 59–68

Sharbel TF, Mitchell-Olds T, Dobes C, Kantama L, de Jong H (2005) Biogeographic distribution of polyploidy and B chromosomes in the apomictic *Boechera holboellii* complex. Cytogenet Genome Res **109:** 283–292

Sharbel TF, Voigt ML, Corral JM, Galla G, Kumlehn J, Klukas C, Schreiber F, Vogel H, Rotter B (2010) Apomictic and sexual ovules of Boechera display heterochronic global gene expression patterns. Plant Cell **22:** 655–671

Sharbel TF, Voigt ML, Corral JM, Thiel T, Varshney A, Kumlehn J, Vogel H, Rotter B (2009) Molecular signatures of apomictic and sexual ovules in the Boechera holboellii complex. Plant J **58:** 870–882

Sharbel TF, Voigt ML, Mitchell-Olds T, Kantama L, de Jong H (2004) Is the aneuploid chromosome in an apomictic *Boechera holboellii* a genuine B chromosome? Cytogenet Genome Res **106:** 173–183

Springer NM, Stupar RM (2007) Allele-specific expression patterns reveal biases and embryo-specific parent-of-origin effects in hybrid maize. Plant Cell **19:** 2391–2402

Templeton AR, Boerwinkle E, Sing CF (1987) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. I. Basic theory and an analysis of alcohol dehydrogenase activity in Drosophila. Genetics **117:** 343–351

Umemura Y, Ishiduka T, Yamamoto R, Esaka M (2004) The Dof domain, a zinc finger DNA-binding domain conserved only in higher plants, truly functions as a Cys2/Cys2 Zn finger domain. Plant J **37:** 741–749

Vicente-Carbajosa J, Moose SP, Parsons RL, Parsons RL, Schmidt RJ (1997) A maize zinc-finger protein binds the prolamin box in zein gene promoters and interacts with the basic leucine zipper transcriptional activator Opaque2. Proc Natl Acad Sci USA **94:** 7685–769

Vielle-Calzada JP, Nuccio ML, Budiman MA, Thomas TL, Burson BL, Hussey MA, Wing RA (1996) Comparative gene expression in sexual and apomictic ovaries of Pennisetum ciliare (L.) Link. Plant Mol Biol **32:** 1085–1092

Voigt ML, Melzer M, Rutten T, Mitchell-Olds T, Sharbel TF (2007) Gametogenesis in the apomictic *Boechera holboellii* complex: the male perspective. In E Hörandl, U Grossniklaus, P van Dijk, Shabel T, eds, Apomixis: Evolution, Mechanisms and Perspectives. Gantner Verlag, Ruggell, Liechtenstein, pp 235–257

Yanagisawa S, Schmidt RJ (1999) Diversity and similarity among recognition sequences of Dof transcription factors. Plant J **17:** 209–214

Yanagisawa S, Sheen J (1998) Involvement of maize Dof zinc finger proteins in tissue-specific and light-regulated gene expression. Plant Cell **10:** 75–89

Zuo Y, Deutscher MP (2001) Exoribonuclease superfamilies: structural analysis and phylogenetic distribution. Nucleic Acids Res **29:** 1017–1026