Gene Function Rather than Reproductive Mode Drives the Evolution of RNA Helicases in Sexual and Apomictic *Boechera*

Markus Kiefer[1,†], Berit H. Nauerth[1], Christopher Volkert[1], David Ibberson[2], Anna Loreth[1], and Anja Schmidt [1,*†]

[1]Department of Biodiversity and Plant Systematics, Centre for Organismal Studies (COS) Heidelberg, Heidelberg University, Heidelberg, Germany

[2]Deep Sequencing Core Facility, CellNetworks Excellence Cluster, Heidelberg University, Heidelberg, Germany

[†]These authors contributed equally to this work.

*Corresponding author: E-mail: anja.schmidt@cos.uni-heidelberg.de.

Accepted: April 11, 2020

Data deposition: Raw data have been deposited at SRA under the accession PRJNA563290. Sequence alignments have been deposited at Dryad (doi:10.5061/dryad.xksn02vc6).

Abstract

In higher plants, sexual and asexual reproductions through seeds (apomixis) have evolved as alternative strategies. Evolutionary advantages leading to coexistence of both reproductive modes are currently not well understood. It is expected that accumulation of deleterious mutations leads to a rapid elimination of apomictic lineages from populations. In this line, apomixis originated repeatedly, likely from deregulation of the sexual pathway, leading to alterations in the development of reproductive lineages (germlines) in apomicts as compared with sexual plants. This potentially involves mutations in genes controlling reproduction.

Increasing evidence suggests that RNA helicases are crucial regulators of germline development. To gain insights into the evolution of 58 members of this diverse gene family in sexual and apomictic plants, we applied target enrichment combined with next-generation sequencing to identify allelic variants from 24 accessions of the genus *Boechera*, comprising sexual, facultative, and obligate apomicts. Interestingly, allelic variants from apomicts did not show consistently increased mutation frequency. Either sequences were highly conserved in any accession, or allelic variants preferentially harbored mutations in evolutionary less conserved C- and N-terminal domains, or presented high mutation load independent of the reproductive mode. Only for a few genes allelic variants harboring deleterious mutations were only identified in apomicts. To test if high sequence conservation correlates with roles in fundamental cellular or developmental processes, we analyzed *Arabidopsis thaliana* mutant lines in *VASA-LIKE* (*VASL*), and identified pleiotropic defects during ovule and reproductive development. This indicates that also in apomicts mechanisms of selection are in place based on gene function.

Key words: apomixis, *Boechera*, evolution, reproduction, RNA helicases, sequence variation.

Introduction

Apart from sexual reproduction, asexual reproduction through seeds (apomixis) established in the course of evolution in more than 400 species (Carman 1997; Hojsgaard et al. 2014). Unlike through sexual reproduction leading to genetic diversity, clonal offspring is formed by apomixis, making apomixis exceptionally attractive for potential applications in agriculture (Spillane et al. 2004; Barcaccia and Albertini 2013; Conner and Ozias-Akins 2017). In terms of evolution however, asexual reproduction is perceived as a dead end, as deleterious mutations are predicted to agglomerate due to the absence of meiotic recombination (Muller 1964). Although mathematical models support this theory of Muller's ratchet under different assumptions (Audiffren and Pardoux 2013), empirical data on asexual taxa are inconsistent as exemplified by the unexpected age and fitness of the asexual Amazon molly contradicting the predictions (Warren et al. 2018). Nevertheless, from a study of apomixis in plants, a slightly higher genome-wide mutation rate has recently been described in sympatric apomictic as compared with sexual *Boechera spatifolii* (Lovell et al. 2017).

As model system, the genus *Boechera* increasingly gains attention (Rushworth et al. 2011; Lovell et al. 2013, 2017; Mau et al. 2015; Lee et al. 2017; Li et al. 2017; Kliver et al.

2018; Schilling et al. 2018). First, it is closely related to the sexual model species *Arabidopsis thaliana*, thus simplifying identification of homologs and comparative analyses. In *Boechera*, more than 100 taxa have been identified consistent of sexual, facultative apomictic, and obligate apomictic accessions at low ploidy levels including diploids (Aliyu et al. 2010; Windham 2010; Mau et al. 2015). Evolutionary, apomixis in *Boechera* likely recurrently derived by hybridizations, leading to heterozygosity and polyploidy typically associated with apomixis (Lovell et al. 2013). Together with copy number variations, this might be instrumental to buffer the effects of deleterious mutations (Aliyu et al. 2013; Lovell et al. 2017).

According to common hypotheses, mutation load is not only expected to be high in apomicts but in addition to differ based on ploidy (Hodač et al. 2019). However, the question remains if mutation accumulation also relates to gene function. In contrast to genome-wide studies, knowledge on the evolution of coding regions in apomicts is scarce to date. Specifically, deleterious mutations in genes essential for fundamental cellular processes such as regulation of translation initiation or for reproductive development might not be tolerated but lead to rapid extinction.

Developmental processes to form seeds are largely similar in apomicts and sexual plants, except for a few distinctive changes: Unlike in sexual plants, where the first cell of the female germline, the megaspore mother cell (MMC), undergoes meiosis, meiotic recombination and reduction is circumvented in apomicts by the corresponding apomictic initial cell (AIC). Two major types of apomixis, diplospory and apospory, are thereby distinguished (Schmidt et al. 2015). They differ by the origin and fate of the AIC: In diplospory, the AIC specifies instead of the MMC in the reproductive flower tissues. It typically undergoes a mitotic-like division instead of meiosis. In apospory, an additional sporophytic cell adjacent to the sexual MMC gets selected as AIC which completely omits meiosis. In contrast to the functional megaspore (FMS) in sexual plants, the founder cell of the gametophytic lineage remains unreduced in apomicts and typically genetically identical to the AIC. Through three rounds of mitotic divisions and cellularization, female gametophytes are formed enclosing the female gametes (egg and central cell). To give rise to formation of the embryo and its nourishing tissue, the endosperm, during sexual reproduction both female gametes get fertilized. In contrast, in apomicts, the egg cell forms the embryo parthenogenetically (without fertilization) and the central cell develops into endosperm either autonomously or dependent on fertilization.

Both in sexual plants and apomicts, germline development requires a tight regulatory system controlling cell fate specifications and other developmental decisions. An increasing number of genes involved have recently been identified (Schmidt et al. 2015; Tekleyohans et al. 2017; Zhou et al. 2017; Nakajima 2018; Nonomura 2018). The regulation of gene activity includes translational control, splicing, and degradation of previously transcribed RNAs (Hafidh et al. 2011). This requires activity of RNA helicases that are often associated to ribonucleoprotein complexes mediating storage and subcellular localization of RNAs (Arkov and Ramos 2010; Hafidh et al. 2011; Gao and Arkov 2013). Little is known about the roles of RNA helicases during plant reproduction despite their abundant expression upon germline formation, in contrast to the animal germline where the conserved roles to regulate different aspects of development are well described (Schmidt et al. 2011; Friday and Keiper 2015). Here, especially RNA helicases from the VASA clade are crucial, that is, for stem cell maintenance, cell fate decisions, cell cycle progression, and for the protection of the germline from the activity of transposable elements (Yajima and Wessel 2011). Although RNA helicases appear to be required for reproduction both in animals and plants, the VASA clade derived only during animal evolution (Gustafson and Wessel 2010).

Evolutionary, the gene family of RNA helicases is conserved throughout all kingdoms of life and has largely diversified in higher eukaryotes (Linder and Owttrim 2009). Of the six superfamilies (SFs) described, SF2 represents the largest group (Linder and Owttrim 2009). Helicase genes share a conserved core consistent of two RecA-like domains (Linder and Jankowsky 2011). Generally, they are functional in NTP-dependent unwinding of RNAs and contribute to basically any aspect of RNA metabolism (Linder and Owttrim 2009). Their target specificity largely resides in the N- and C-terminal domains flanking the conserved core domains that are frequently involved in mediating protein interactions (Jankowsky 2011). Based on at least 12 sequence motifs shared, the family is further subdivided in distinct clades, with SF2 containing the DEAD and DExD/H box helicase proteins named after the conserved amino acid motif, the Ski2-like proteins including SUV3 facilitating mitochondrial RNA degradation, RecQ-like proteins involved in regulation of meiotic recombination, and SWI/SNF proteins typically acting in chromatin remodeling (Fairman-Williams et al. 2010; Tang et al. 2010; Higgins et al. 2011; Jankowsky 2011; Khemici and Linder 2018). Members of the *EUKARYOTIC TRANSLATION INITATION FACTORS* (*EIF*) family also belong to the SF2 of RNA helicases (Marintchev 2013). The EIF proteins of the EIF4 family act in multimeric complexes to recruit mRNAs to the small subunits of the ribosomes upon translation initiation (Asano et al. 2000; Hernández and Vazquez-Pianzola 2005). In *A. thaliana*, two isoforms exist, with EIF4A1 being involved in controlling cell cycle progression and ovule development (Bush et al. 2015, 2016).

From the 161 members of the gene family identified in *A. thaliana* (Xu et al. 2013), requirement for germline development has been reported for a few genes only, that is, *MAGATAMA3* (*MAA3*), *MATERNAL EFFECT EMBRYO ARREST29* (*MEE29*), *SLOW WALKER3* (*SWA3*), *CHROMATIN REMODELING 11 (CHR11)*, and *AtSUV3* (Huanca-Mamani

et al. 2005; Pagnussat et al. 2005; Shimizu et al. 2008; Liu et al. 2010). In addition, MNEME (MEM) has previously been described, sharing 43% amino acid identity with SUV3 (Schmidt et al. 2011). *MEM* is predominantly expressed in the MMC, and plants carrying a mutant *mem* allele likely show formation of unreduced gametophytes, representing an element of apomixis (Schmidt et al. 2011). Furthermore, also the homolog of *Hypericum perforatum MEE29* is preferentially expressed in reproductive tissues from apomicts compared with sexual plants (Barcaccia and Albertini 2013). Members of the RNA helicase family might thus be involved in establishing or sustaining apomixis.

Based on investigations on *A. thaliana* and maize, also epigenetic regulatory pathways involving small RNA pathways and DNA methylation to control gene activity or to mediate silencing of transposable elements appear to be relevant for apomixis (Garcia-Aguilar et al. 2010; Olmedo-Monfil et al. 2010; Singh et al. 2011). These pathways involve certain RNA helicases including *DECREASE IN DNA METHYLATION1* (*DDM1*), or genes encoding for DICER LIKE (DCL) proteins (Pikaard and Mittelsten Scheid 2014). However, to date, knowledge about the roles and evolution of RNA helicases in apomicts is largely lacking. Due to the diversity of this gene family, it can be hypothesized that the functional impact of mutations predicted to accumulate in apomicts differs for distinct members. Although detrimental mutations in genes controlling basic cellular or essential developmental processes including reproduction are likely to be deleterious, mutations in others, for example, *MEM* might even be instrumental for apomixis.

To study mutation accumulation of selected germline expressed RNA helicases and to gain insights into aspects of their evolution in sexual and apomictic *Boechera*, we applied in solution hybrid capture also referred to as target enrichment (Gnirke et al. 2009). After enrichment, we used Illumina HighSeq2500 to sequence genomic regions of 58 RNA helicases from 24 *Boechera* accessions representing sexual, facultative apomictic and obligate apomictic reproduction. We also included the gene encoding for the centromere-specific histone variant *CENH3* in this study for comparison, which is characterized by rapid evolution and higher expression levels in the germline of the triploid apomict *Boechera gunnisoniana* as compared with *A. thaliana* (Lysak 2014; Schmidt et al. 2014).

Thereby, we identified different features of mutation accumulation and genetic variation presumably relating more to gene function than to the reproductive modes. Although evidence for purifying selection was found for most of the genes, indications of near neutral evolution for allelic variants from apomictic as compared with sexual accessions were given only for a few helicases. As a prove of concept that high sequence conservation might be correlated with involvement of the proteins in regulating basic cellular or crucial developmental processes, we analyzed lines carrying mutations in a novel RNA helicase AT1G72730 in *A. thaliana*, now named *VASA-LIKE* (*VASL*). Consistent with our expectations, these show defects during reproductive development including alterations in ovule development, gametogenesis, and embryogenesis. Taken together, our study provides new insights into the evolution of RNA helicases in sexual and apomictic *Boechera*, further supporting the relevance of RNA helicases to regulate germline and seed development.

Materials and Methods

Plant Material

Seeds of *Boechera* Á. Löve & D. Löve accessions included in this study were obtained from Timothy F. Sharbel (Global Institute for Food Security, University of Saskatchewan, Canada), Thomas Mitchell-Olds (Duke University, USA), and John Carman (Utah State University; table 1 and Supplementary Material online). B12-663 has been originally been collected from Big Horn Pass Montana, USA (Sharbel T, Saskatoon, Canada, personal communication). Except for three accessions, *Boechera stricta* B12-663, *B. retrofracta* × *stricta* CO11010, and *B. stricta* × *retrofracta* B12-2582, the frequency of apomictic seed formation has previously been determined (Aliyu et al. 2010; Mau et al. 2015; Carman et al. 2019). Therefore, these three accessions were excluded from all analyses summarizing groups by reproductive mode (also see supporting information, Supplementary Material online). In this study, we classified accessions as obligate apomicts when no evidence for seed formation by sexual reproduction is given, and as facultative apomicts when sexual seed formation has been reported even at very low frequencies (Aliyu et al. 2010; Mau et al. 2015).

Seeds were stratified at 4 °C for ~7 days before they were surface sterilized and grown on murashige-skoog plates for 2–3 weeks at 16-h light/darkness at 18 °C night 21 °C day cycle. One leaf per seedling was used for ploidy analysis by flow cytometry using a CyStainUV Precise P Kit (Sysmex Partec, Görlitz, Germany) on a Partec CyFlow Space instrument following manufacturer instructions.

Arabidopsis thaliana (L.) Heynh., ecotype Col-0 was used as wild-type and for generating CRISPR/CAS9 lines (see supporting information, Supplementary Material online). Seeds were stratified overnight at 4 °C before they were surface sterilized and germinated on murashige-skoog plates similar to *Boechera*. Seedlings were transferred to soil (ED73, Einheitserde, Fröndenberg, Germany) and grown at 16-h light/darkness at 22 °C. Lines SAIL_288_C08 and SALK_095627 with T-DNA insertions in the 3′ untranslated region (UTR) were obtained from "The European Arabidopsis Stock Centre" (arabidopsis.info, last accessed January 30, 2018) and referred to as *vasl_1* and *vasl_2*, respectively. Genotyping of plants carrying *vasl_1* or *vasl_2* alleles was done with primers 5′-TGGTACTTGTTGGTGATGCTG-3′, 5′-

Table 1
Summary of *Boechera* Accessions Used, Ploidy Levels, Number of Seedlings Pooled for DNA Extraction, and Read Counts Obtained

Genus	Species	Ploidy	Accession	Reproductive Mode	No of Seedlings Used	Raw Reads Paired	Paired Read Counts after Trimming	Publication of Accession
Boechera	*divaricarpa*	2x	ES517	Facultative	3	9,135,540	7,280,704	Aliyu et al. (2010) and Mau et al. (2015)
Boechera	*retrofracta* × *stricta*	2x	CO11010	Apomictic	2	5,600,566	3,763,336	Carman et al. (2019)
Boechera	*retrofracta*	2x	B12-1131	Apomictic	3	8,803,860	5,936,003	Mau et al. (2015)
Boechera	*williamsii*	2x	B12-1524	Apomictic	3	7,660,767	4,708,395	Mau et al. (2015)
Boechera	*pallidifolia*	2x	B12-1578	Apomictic	3	10,065,290	8,650,998	Mau et al. (2015)
Boechera	*pallidifolia*	3x	B12-1599	Apomictic	3	9,739,570	7,210,650	Mau et al. (2015)
Boechera	*spatifolia*	2x	B12-956	Apomictic	3	9,460,515	7,965,700	Mau et al. (2015)
Boechera	*stricta* × *retrofracta*	4x	B12-2582	Apomictic	2	9,979,526	8,602,569	Mau et al. (2015)
Boechera	*divaricarpa*	2x	ES524	Facultative	3	8,394,949	6,744,721	Aliyu et al. (2010) and Mau et al. (2015)
Boechera	*divaricarpa*	3x	ES598	Apomictic	3	9,442,364	6,404,632	Aliyu et al. (2010) and Mau et al. (2015)
Boechera	*divaricarpa*	3x	ES704	Apomictic	3	9,411,828	6,550,887	Aliyu et al. (2010) and Mau et al. (2015)
Boechera	*holboellii*	2x	ES911	Facultative	2	9,840,426	6,007,448	Aliyu et al. (2010) and Mau et al. (2015)
Boechera	*holboellii*	2x	ES805	Facultative	3	9,686,119	8,430,991	Aliyu et al. (2010) and Mau et al. (2015)
Boechera	*stricta*	2x	B12-693	Facultative	3	10,100,186	7,702,794	Mau et al. (2015)
Boechera	*pallidifolia*	2x	B12-1591	Facultative	3	10,446,466	8,772,196	Mau et al. (2015)
Boechera	*stricta*	2x	B12-663	Sexual	2	10,041,301	8,626,705	
Boechera	*stricta*	2x	B12-1268	Sexual	3	8,809,041	7,094,779	Mau et al. (2015)
Boechera	*crandalii*	2x	B12-1397	Sexual	3	10,464,956	7,190,352	Mau et al. (2015)
Boechera	*crandalii*	2x	ES727	Sexual	3	11,340,638	7,882,158	Aliyu et al. (2010) and Mau et al. (2015)
Boechera	*leisiocarpa*	2x	B12-357	Sexual	3	9,516,505	5,840,745	Mau et al. (2015)
Boechera	*williamsii*	2x	B12-558	Sexual	3	9,441,753	8,116,152	Mau et al. (2015)
Boechera	*holboellii*	2x	ES786	Sexual	2	5,259,327	3,966,639	Aliyu et al. (2010) and Mau et al. (2015)
Boechera	*holboellii*	2x	ES820	Sexual	3	9,699,227	8,126,082	Aliyu et al. (2010) and Mau et al. (2015)
Boechera	*stricta*	2x	LTM	Sexual	3	1,607,3343	13,825,485	Lee et al. (2017)

CGATCTCTACGAGACACTGGC-3′, and 5′-TGGATAAA TAGCCTTGCTTCC-3′, or 5′-GGAAACTCGAGACACTATGC G-3′, 5′-TTTGTCACTTGCGAGTTTTTG-3′, and 5′-ATTTTGCC GATTTCGGAAC-3′, respectively. In addition, two heterozygous lines obtained by CRISPR/Cas9 harboring an insertion of a C after position 982 or T after position 980 of the coding sequence (CDS) are referred to as *vasl_3* and *vasl_4*, respectively. Identification of homozygous lines in the next generation was done by CAPS genotyping (see supporting information, Supplementary Material online).

Phylogenetic Analysis of RNA Helicases with Similarity to VASL

For selection and alignment of protein sequences, see supporting information, Supplementary Material online. Phylogenetic analysis was done using RAxML v8.1.16 with model setting "PROTGAMMAAUTO" and including fast bootstrap analysis (100 replicates) (Stamatakis 2014).

Target Pulldown Probe Design

To define genomic regions of interest, the closest homologs to the *A. thaliana* genes were identified in *B. stricta* LTM (supplementary table S1) (Lee et al. 2017) by BlastN with default settings on Phytozome v11 (https://phytozome.jgi.doe.gov, last accessed February 27, 2020). Homologs were determined based on highest sequence similarities, information on synteny, and on annotations in Phytozome to belong to a gene family. The probe design for the pulldown was based on sequence information from *B. stricta* LTM genomic regions and up to 5-kb upstream sequence downloaded from Phytozome v11, and coding sequences (CDS) from the annotated reference transcriptome of *Boechera gunnisoniana* (Schmidt et al. 2014), depending on availability of sequence data. A total of

120-bp probes at 4× tiling density were designed and synthesized as myBaits 1–20K custom target capture kit by MYcroarray (Ann Arbor, MI) (supplementary table S2, Supplementary Material online).

DNA Preparation, Target Pulldown, and Library Preparation

DNA was isolated using Invisorb Plant Mini Kit (Stratec Molecular GmbH, Berlin, Germany) following manufacturer instructions from two to three seedlings per accession. Genomic DNA was quantified using QUBIT dsDNA HS Assay Kit (Thermo Fisher Scientific) following manufacturer instructions. Genomic DNA was sheared by Covaris before size selection of 300–400-bp fragments for library preparation. Libraries were prepared with the NEBNext Ultra II DNA Library Prep Kit for Illumina and NEBNext Ultra Multiplex Oligos for Illumina (New England Biolabs, Ipswich, MA) and amplified by 12 cycles of polymerase chain reaction prior to the pulldown. Subsequently, the myBaits 1–20K kit was used for target pulldown on the libraries following manufacturer instructions with hybridizations done at 60 °C. Prior to sequencing, libraries were subjected to additional ten cycles of polymerase chain reaction.

Next-Generation Sequencing and Variant Detection

Libraries were sequenced on one lane using 125-bp paired end sequencing on Illumina HighSeq2500 by the Deep Sequencing Core Facility (Heidelberg University). After quality control using FastQC v0.11.04 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/, last accessed June 01, 2017), reads were trimmed with Trimmomatics 0.22 to remove adopters and low-quality bases (Bolger et al. 2014). Reads shorter 40 bp after trimming were discarded. Subsequently, reads were aligned to the *B. stricta* genome using Stampy v1.0.32 with default parameters (Lunter and Goodson 2011; Lee et al. 2017). Samtools 1.6 and picardtools 1.95 (http://broadinstitute.github.io/picard/, last accessed February 02, 2019) were applied for sorting, removal of duplicates, and library indexing before realignment with GATK3.4.46 using RealignerTargetCreator and IndelRealigner (Li et al. 2009; McKenna et al. 2010). Regions with a base coverage >50 were included in further analysis. Variant calling was done with freebayes v.1.1.0-50 (Garrison and Marth 2012) using options for pooled libraries and setting ploidy to the sum of the ploidies of plants and the number of plants used for genomic DNA extraction (table 1). Filtering of variants was applied with VCFtools 0.1.15 with the parameters DP > 10 and QUAL > 30 (Danecek et al. 2011). Raw data are deposited under SRA accession PRJNA563290.

Alignment of Sequence Variants and Identification of CDS

Sequences of coding regions for *B. stricta* and *Arabidopsis lyrata* were downloaded from Phytozome v11 (https://phytozome.jgi.doe.gov). Homologs in *A. lyrata* were identified using BlastN with standard settings. Gene family annotations, highest sequence similarity, and synteny were taken into account for determination of the closest homolog. Allelic variants (haplotypes) were sorted by genomic regions. Subsequently, MAFFT v7.271 was used to align CDS of *B. stricta* and *A. lyrata* to the haplotypes from *Boechera* (Katoh and Standley 2013). Sequence alignments have been deposited on Dryad (doi:10.5061/dryad.xksn02vc6). To compare CDS alignments, all sequences in the alignment were cropped to CDS according to the sequence of the primary transcript annotated for *B. stricta* in Phytozome using Python. All alignments were checked and readjusted manually using PhyDE Phylogenetic Data Editor (http://www.phyde.de/, last accessed August 22, 2018).

Visualization of Usage of Nucleobases along Genomic Regions and Mutation Density along CDS

To visualize the use of different nucleobases across the genomic regions of targets, Python was used to develop a script assessing and plotting the frequencies of distinct nucleobases used at a certain position. Another Python script was developed to plot mutation density and types of mutations along the CDS. Python scripts developed for this study have been deposited on Github (https://github.com/hinz1/gbe).

Assessment of Mutation Load

A Python script was developed to annotate changes in CDS as compared with the *A. lyrata* reference sequence as silent mutations (synonymous single-nucleotide polymorphisms [SNPs]), nonstop mutations (nonsynonymous SNPs), insertions or deletions non causing any stop codons (In_Frame_Ins or In_Frame_Del, respectively), insertions or deletions leading to shifts of the open reading frame (Frame_Shift_Ins or Frame_Shift_Del, respectively), or nonsense mutations introducing premature stop codons. The information was transferred to a custom .maf file further used for visualization of mutation density and of the most deleterious mutation per allele using the Bioconductor R-package GenVisR and the waterfall function (Skidmore et al. 2016). The hierarchical categorization of the most detrimental mutation per allele was applied in the order of mutations leading to premature stop (nonsense mutations), insertion and deletions (INDELs) causing frameshifts, in frame INDELS not affecting the reading frame, nonstop mutations, and synonymous SNPs.

To compare mutation load per gene, the number of nonsynonymous SNPs (here including all SNPs except the class of synonymous SNPs) per CDS was divided by protein length. Two-way ANOVA and Student's *t*-test implemented in R were

used to test for effects of gene identity and reproductive mode (see supporting information, Supplementary Material online).

Assessment of Evolutionary Rates

To estimate evolutionary rates of haplotypes from *Boechera* aligned to their *A lyrata* homologs, we used the yn00 program implemented in PAML (Yang 2007). To allow the analysis, stop codons were replaced by gaps using PHYDE (http://www.phyde.de/). Variants with frameshifts resulting in stop codons were excluded from further analysis. Ratios of nonsynonymous to synonymous SNPs ($\omega = dN/dS$) were estimated using the LWL85m model after (Li et al. 1985; Yang 2007). Ratios from groups of sexual versus apomictic haplotypes of one gene were compared by Student's *t*-test in R 3.5.1 (https://www.r-project.org/, last accessed August 02, 2018). Fishers *F*-test implemented in R was applied to test for higher variance of mean ω-ratios in triploid as compared with diploid accessions. Boxplot visualization was done using the boxplot function implemented in R using default options.

Morphological Investigation of Reproductive Defects

Morphological characterization of seed development was done by clearing in Hoyers solution as described and DIC microscopy with a Zeiss Axio Imager M1 (Zeiss, Oberkochen, Germany) (Schmidt et al. 2011). Pictures captured were cropped and processed in Adobe Photoshop CS2 Version 9.0 (Adobe Systems Inc., San Jose, CA).

Results

Genes with Structural Similarity to *Drosophila* VASA Were Identified among Targets

RNA helicases are a large gene family involved in regulatory processes in any cell and tissue type. To our knowledge, the evolution of members of this family has so far not been studied in related sexual and apomictic plants. For this study, in addition to *CENH3*, we selected 58 RNA helicases that are active in cells of the female germline (supplementary fig. S1 and table S1, Supplementary Material online) (Wuest et al. 2010; Schmidt et al. 2011, 2014). They were chosen to be either enriched, or consistently active at key steps of reproduction in sexual or apomictic plants, or to show differences in activity between the reproductive modes. Hereafter, the *A. thaliana* gene identifier or name is used to refer to *Boechera* homologs.

These helicases comprised members of different clades, including genes with described roles for germline development or meiotic recombination, and in addition such encoding for uncharacterized proteins (supplementary table S1, Supplementary Material online). In order to identify similarities of the latter to described proteins, we modeled their *B. stricta* protein sequences to related protein structures using Phyre2 (Kelley et al. 2015). Interestingly, similarities to the *Drosophila* VASA were predicted with high confidence for the homologs of AT1G72730, now named VASA-LIKE (VASL), and of AT1G51380 (supplementary table S1, Supplementary Material online). To establish their evolutionary relation to other genes in the study, we performed a phylogenetic analysis including 14 additional RNA helicases selected based on sequence similarity (supplementary fig. S2, Supplementary Material online). Interestingly, VASL is grouped together with EIF4A1 (AT3G13920) and EIF4A2 (AT1G54270) and this group forms a cluster with AT1G51380.

Target Enrichment and Sequencing Allowed Identification of Allelic Variants in *Boechera*

To study allelic variation of candidate genes in apomictic versus sexual *Boechera*, target enrichment was performed (Gnirke et al. 2009). Twenty-four *Boechera* accessions representing 12 species or interspecific hybrids of parental species at three ploidy levels were selected to this aim including sexual, facultative, and obligate apomictic accessions (table 1) (Aliyu et al. 2010; Beck et al. 2012; Mau et al. 2015; Carman et al. 2019). Target pulldown was performed on indexed libraries prepared from genomic DNA isolated from two to three pooled seedlings per accession (table 1). Sequencing on the Illumina HighSeq2500 platform (125-bp paired end) resulted in 3,763,336–13,825,485 paired reads per accession after quality control and trimming (table 1). Variant calling resulted in identification of 147 alleles in total for all accessions after mapping of the reads to the *B. stricta* LTM genome, variant calling and filtering (Lee et al. 2017). High-quality data and annotations of allelic variants were achieved for all targeted regions.

Sequence Variation Was Generally Higher in Noncoding than in Coding Regions

Evolutionary forces affect CDS and noncoding regions differently. Whereas changes causing an altered protein sequence can directly impact protein function, upstream regulatory regions are crucial for gene regulation. To visualize the overall sequence variation of target regions, allelic variants (haplotypes) from all accessions were aligned. We used a sliding window approach to plot the number of distinct nucleobases represented along the sequence with a binning of 12 positions separately for the groups of apomictic, facultative apomictic, or sexual accessions (fig. 1 and supplementary fig. S3, Supplementary Material online). Furthermore, we analyzed variation across all accessions to allow to identify positions conserved within the same reproductive mode but different between groups. Most differences in base usage and regions of overall high sequence variation in either sexual accessions or apomicts were identified in noncoding regions including upstream regulatory regions or introns, for example, in

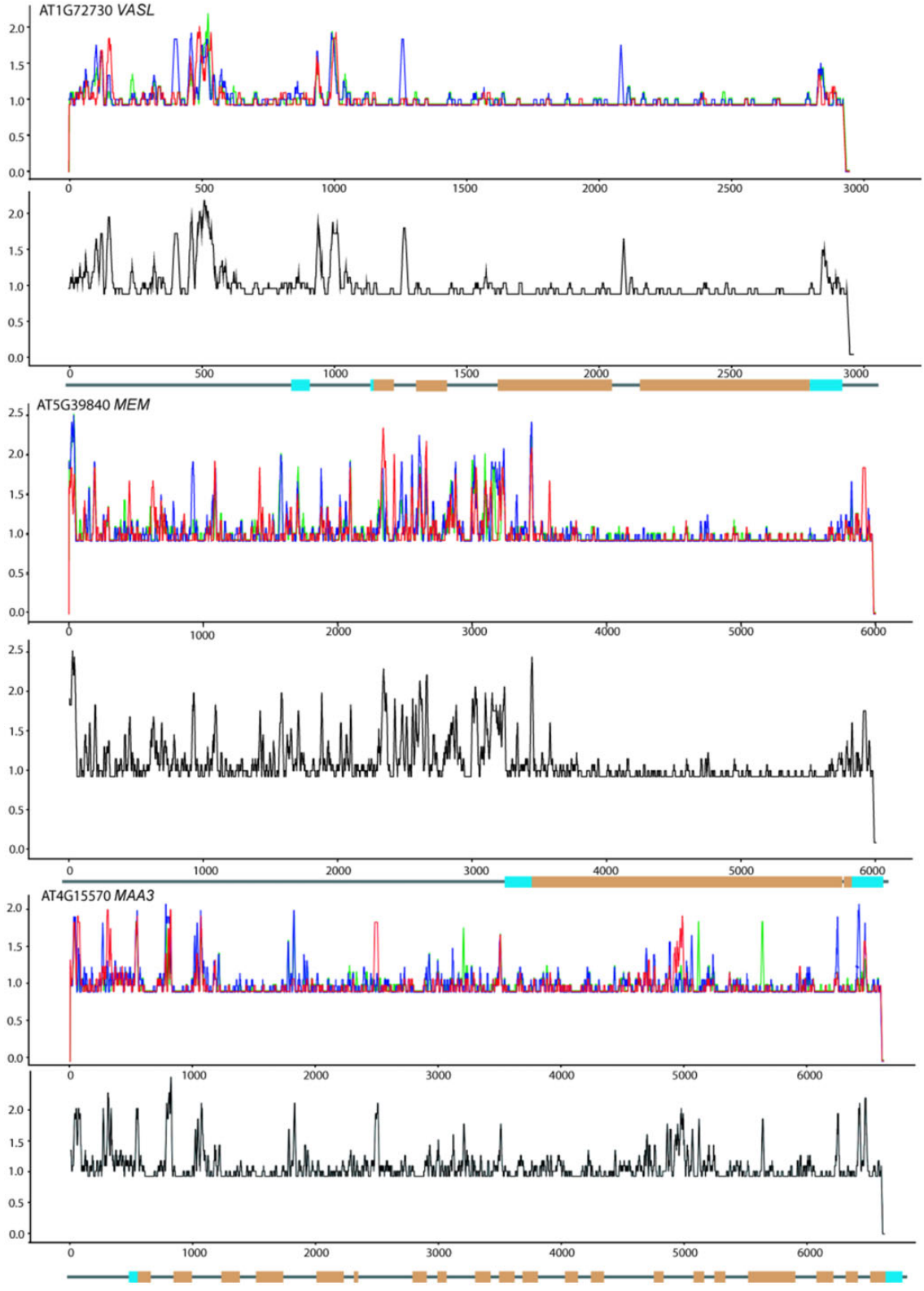

Fig. 1.—Sliding window approach to visualize the usage of nucleobases per position in genomic regions of selected RNA helicase genes. For *VASL*, *MEM*, and *MAA3*, the number of different nucleobases comprised in the allelic variants was plotted for all variants of apomictic (blue), facultative apomictic (green), sexual accessions (red), or all accessions (black) with a binning of 12 bases per datapoint. Orange boxes below symbolize exons, blue boxes UTRs, and gray bars noncoding regions. For all genes, a higher sequence variability was observed in noncoding regions as compared with exonic regions or UTRs.

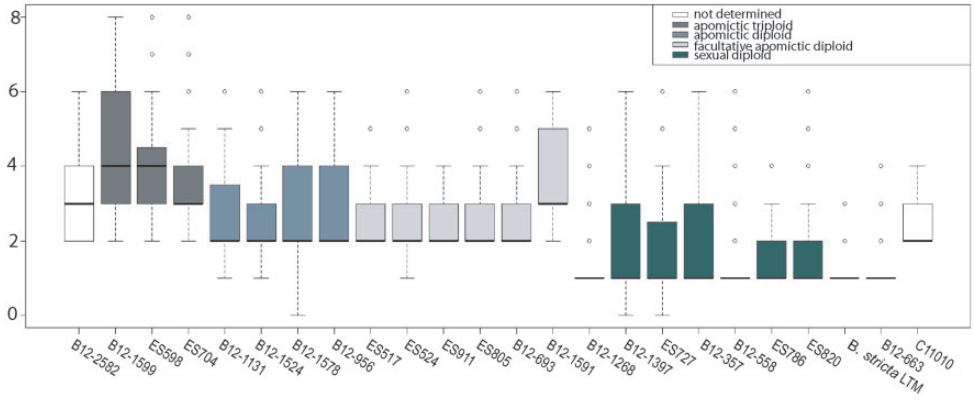

Fig. 2.—Box and whiskers plot depicting the number of allelic variants per accession for all genes. Boxes indicate the lower and upper quartile of numbers of unique CDS variants. Horizontal lines in each box indicate the medium numbers of variants observed. Accession numbers are given below.

AT2G40700 (*RNA HELICASE 17, RH17*) or AT3G13920 (*EIF4A1*) (fig. 1 and supplementary fig. S3, Supplementary Material online). However, even within these regions, stretches of lower sequence variation can be identified, likely allowing to predict positions of conserved motifs. Interestingly, a region of high variability both within and between the groups was observed for example for *MEM* upstream of the start codon (fig. 1). Furthermore, for *VASL*, sequence variation was observable mostly in the putative promoter region and in an intron within the 5′ untranslated region (UTR) (fig. 1). Although in case of *MAA3* differences in sequence variation between reproductive modes were mostly observable in intronic regions, also in the 17th exon high sequence variation was identified for facultative apomicts (fig. 1). Future investigations are needed to elucidate if these regions are functionally important for the regulation of the genes and if differences are relevant for the reproductive modes, or if this high variability mainly results from the absence of selective pressure.

CDS Variants Are Mostly under Purifying Selection in Diploid and Triploid Apomicts

To describe mutation accumulation of targeted genes in more detail, we specifically investigated CDS. Therefore, haplotypes were aligned to CDS from *A. lyrata* as a suitable reference for variant calling in *Boechera* (Lovell et al. 2017). Based on this, synonymous or nonsynonymous SNPs, INDELs, and mutations leading to premature stop codons (nonsense mutations) were identified. Interestingly, the numbers of distinct alleles identified correlated with reproductive mode and ploidy level (fig. 2 and supplementary fig. S4, Supplementary Material online): In all diploid sexual accessions for at least half of all genes, only one haplotype was identified, in contrast to two or more most frequently observed in apomicts (fig. 2). In the same line, the tendency was more pronounced in apomicts than in sexual accessions to show differences in the numbers of mutations per MB for different allelic variants of the same accession (fig. 3*A*). Nevertheless, also differences were observed depending on the gene. The maximum number of allelic variants identified per gene in any accession ranged between 18 and 117 (supplementary table S1, Supplementary Material online), with only few sequence variants identified, for example, for EIFs and higher numbers, for members of chromatin remodelers and others (supplementary fig. S4 and table S1, Supplementary Material online).

To obtain insights into the evolutionary mechanisms shaping diversification of the alleles in *Boechera*, we further estimated the ratios of nonsynonymous (d*N*) versus synonymous (d*S*) SNPs ($\omega = dN/dS$). This can be taken as an indication for the evolutionary rate, with $\omega < 1$ indicating purifying selection, $\omega = 1$ neutral evolution, and $\omega > 1$ diversification (Yang and Bielawski 2000). For asexual reproduction, an accumulation of slightly deleterious mutations under nearly neutral evolution has been hypothesized (Schiffels et al. 2011; Hojsgaard and Hörandl 2015). To uncover if this holds true for the RNA helicases, we estimated ω-values using the yn00 program implemented in PAML (Yang 2007). Although this test indicated purifying selection for CDS both from sexual and apomictic accessions (supplementary fig. S5, Supplementary Material online), differences in ω were still evident between genes, with the homologs of *ESP3* (AT1G32490), *EIF4A1* (AT3G13920), *EIF4A2* (AT1G54270), AT1G51380, AT1G72730 (*VASL*), *CHR12* (AT3G06010), *CHR11* (AT3G06400), *ABA OVERLY SENSITIVE 6* (*ABO6*; AT5G04895), and others being under strong purifying selection, whereas *SNF2-RING-HELICASE LIKE5 (FRG5)* (AT1G11100) and *CHR40* (AT3G24340) showed relatively high ω-values pointing toward near neutral evolution.

Furthermore, ω-values averaged for alleles of obligate apomictic as compared with sexual accessions were only significantly different for six genes (Student's *t*-test; *P* value < 0.05; supplementary fig. S5, Supplementary Material online). Unlike the homologs of *RECQ4A* (AT1G10930), *RH5* (AT1G31970), AT2G35920, and AT3G02060 that even showed lower

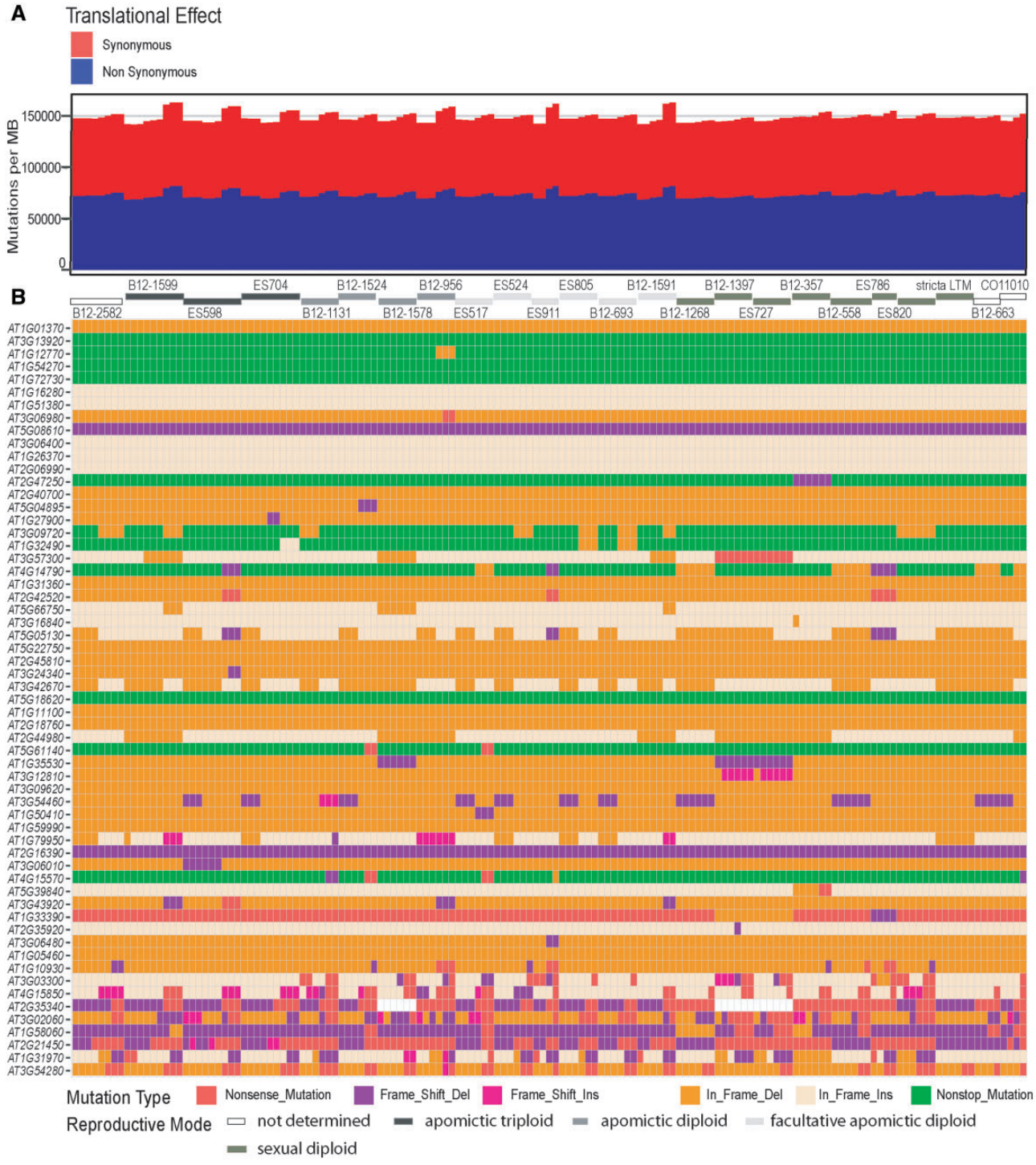

Fig. 3.—Visualization of mutation density per accession and most detrimental mutations per gene and accession as compared with *A. lyrata* as reference. (*A*) Plot of the abundance of synonymous and nonsynonymous SNP per MB of CDS sequences. (*B*) Visualization of the most deleterious mutation per haplotype and accession with a mutation hierarchy from deleterious to harmless from Nonesense_Mutation leading to a premature stop codon, Frame_Shift_Deletion, Frame_Shift_Insertion, In_Frame_Del, and In_Frame_Ins leading to a deletion or insertion without causing a frameshift, and Nonstop_mutation causing a nonsynonymous SNP. Rows in the plot represent different genes, whereas the 147 allelic variants identified in the study per gene are presented in columns. Boxes between (*A*) and (*B*) indicate reproductive mode and ploidy. The plot has been done using the waterfall function from GenVisR (Skidmore et al. 2016).

selection pressure on allelic variants of sexual accessions as compared with apomicts, only for *ESP3* and AT1G58060 higher rates of mutation accumulation were identified for alleles from apomictic accessions.

As higher ploidy is commonly perceived as a strategy to buffer the effect of deleterious mutations, we tested in addition, if different allelic variants accumulate mutations at higher rates in triploid apomicts as compared with diploids. This would result in a larger variance of mean d*N*/d*S* ratios. We tested for this using the *F*-test implemented in R (P value $<$ 0.05; www.r-project.org/). However, this trend was not observed for any of the RNA helicases preselected based on broader distributions of ω-values in all apomicts, except for AT5G61140 at near significant levels ($P = 0.632$) (supplementary fig. S6, Supplementary Material online).

Allelic Variants with Premature Stop Codons Were Identified in More than Half of the Genes

To categorize mutations based on their effect on the protein encoded, a hierarchical categorization of the most detrimental mutation per allele was applied (fig. 3) (Skidmore et al. 2016). As compared with *A. lyrata*, none of the genes showed only synonymous SNPs (fig. 3*B*). Nevertheless, EIF4A1 and EIF4A2 (AT3G13920 and AT1G54270, respectively), as well as the uncharacterized proteins encoded by *VASL* (AT1G72730), and *CHR17* (AT5G18620) showed only nonstop mutations as the most deleterious changes present. For additional 21 genes studied, in frame INDELs were the most deleterious type of mutations represented in any accession (fig. 3*B*). In the CDS of 11 helicases, mutations leading to stop codons (nonsense mutations, INDELs) were exclusively found in allelic variants from apomicts in contrast to alleles from sexual accessions (fig. 3*B*). Interestingly, these comprise genes involved in epigenetic regulatory pathways, maintenance of genome stability, or regulation of reproduction, including *DCL3* (AT3G43920), *CHR40* (AT3G24340), *FRG2* (AT1G50410), *REGULATOR OF TELOMERE ELONGATION HELICASE1 (RTEL1;* AT1G79950), and *MAA3* (AT4G15570) (Shimizu et al. 2008; Ahmad et al. 2010; Groth et al. 2014; Recker et al. 2014; Zhou et al. 2018). In contrast, mutations leading to premature stop codons were only observed for alleles from sexual accessions for *MEM* (AT5G39840), in addition to the *INO80* ortholog AT3G57300, AT2G47250, and AT2G35920 (fig. 3*B*). For other genes, including *DCL2* (AT3G03300), *RH1* (AT4G15850), and *RH5* (AT1G31970), mutations leading to premature stop were identified in accessions of both reproductive modes (fig. 3*B* and supplementary table S1, Supplementary Material online). Moreover, high abundance of alleles with deleterious mutations including absence of start codons and premature stop codons was observed for several genes, including *CHR34* (AT2G21450), *MEE29* (AT2G35340), and *FASCIATED STEM4 (FAS4*, AT1G33390) (fig. 3). The consistent identification of frameshift mutations is *mis*leading for the homolog of *PIGMENT DEFECTIVE* (AT5G08610) and *DECREASE IN DNA METHYLATION1 (DDM1*, AT5G66750), as in these cases reference sequences from *A. lyrata* are much divergent from the sequences in *Boechera*. Taken together, the data indicate that the presence of deleterious mutations is more dependent on the gene and possibly functional implications rather than on the reproductive mode.

The Load of Deleterious Mutations Varies Greatly between Genes

Apart from the most deleterious mutations contained in a CDS, the mutation density and position of amino acid change or INDELs typically impacts protein folding and function. Although the core domains of RNA helicases are evolutionary highly conserved, aspects of gene evolution and functionalization are driven by changes in flanking C- and N-terminal sequences (Sloan and Bohnsack 2018). To gain more detailed insights into sequence changes, we plotted the mutation distribution along the CDS. First, comparisons included all 147 allelic variants from any accession by showing the percentage of alleles holding a mutation when averaging 12 consecutive bp (fig. 4). Here, we focused on *MEM*, *VASL*, and AT3G16840 all previously described as significantly enriched during germline formation in *Arabidopsis* (Schmidt et al. 2011), in addition to *FAS4*, *EIF4A1*, and *RECQ4A*. Overall, mutation density varied greatly between these proteins with mainly synonymous and only few nonsynonymous SNPs present in *EIF4A* and *VASL*, consistent with the strong purifying selection observed (fig. 4 and supplementary fig. S5, Supplementary Material online). While sharing the conserved PFAM domains "DEAD/DEAH box helicase" (PF00270) and "helicase conserved C-terminal domain" (PF00271) with the former, AT3G16840 is characterized by longer N- and C-termini showing a significant load of nonstop mutations and in frame INDELs. *MEM* harbored mainly changes in the N-terminal domain, but for a few alleles also nonsense mutations close to the stop codon. A number of mutations are also represented in the conserved helicase domains, especially in the "mitochondrial degradosome RNA helicase C-terminal" (PF12513) domain. Interestingly, both FAS4 and RECQ4A presented a high mutation load including deleterious frameshift mutations in different sequence contexts along the protein.

Furthermore, mutation load and position might differ between reproductive modes. To investigate this, we compared the mutation plots of *MEM* and *MAA3* between the groups of sexual or obligate apomictic accessions (supplementary fig. S7, Supplementary Material online). Interestingly, most positions of mutations were shared between both groups with only few differences observed between variants from sexual or apomictic accessions (supplementary fig. S7, Supplementary Material online). Nevertheless for *MEM*, a slightly higher amount of synonymous SNPs and nonstop

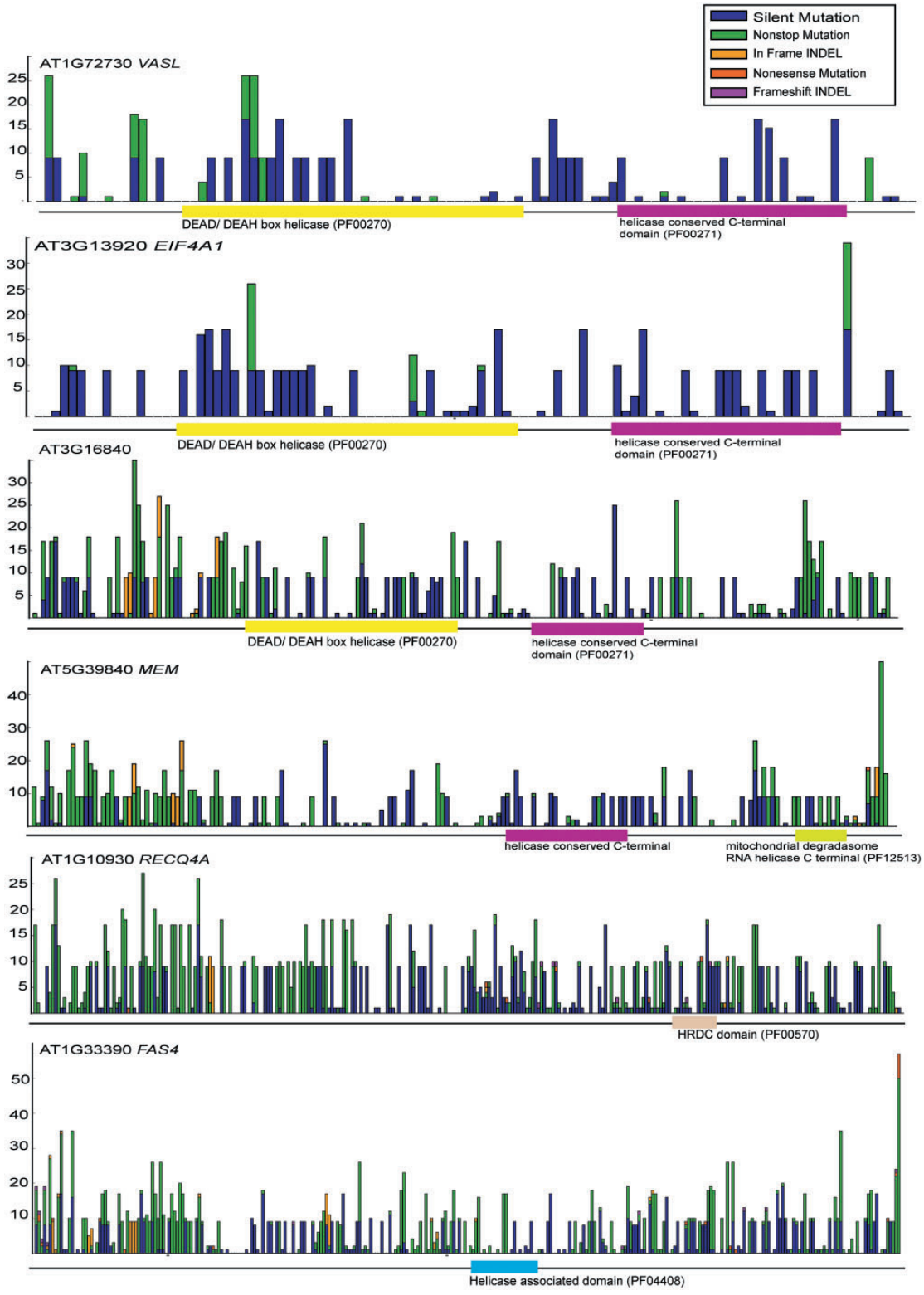

Fig. 4.—Plot of mutation density along the CDS of selected RNA helicase genes. For selected RNA helicases, mutations as compared with *A. lyrata* used as reference were plotted along the CDS in windows of 12 bp. Types of mutations depicted are silent mutations, nonstop mutations (nonsynonymous SNPs), in frame INDELs, nonsense mutations introducing a stop codon, and frameshift INDELs. Depicted are the genes encoding for the conserved proteins EIF4A1 and VASL harboring only silent or nonstop mutations in any accession, MEM and the undescribed helicase AT3G16840 accumulating mutations mainly in the N- and C-terminal regions, and *RecQ4A* and *FAS4* accumulating all types of mutations along the CDS including conserved domains.

mutations was identified in sequences from apomicts as compared with sexual *Boechera* accessions, in particular 5′ of the "helicase conserved C-terminal domain" (supplementary fig. S7, Supplementary Material online). Similarly for *MAA3*, a tendency toward more nonstop mutations was apparent in apomicts. In addition, few deleterious mutations in the conserved "AAA-domain" (PF13086) leading to a premature stop codon (frameshift INDELs and nonsense mutations) were identified only in allelic sequences of apomictic accessions.

Our analyses indicate greater differences in mutation load dependent on the gene identity than on the mode of reproduction. Consistently, gene identity was identified as significant factor using two-way ANOVA ($P < 0.001$). This difference was supported by pairwise comparisons of selected RNA helicases (supplementary fig. S8, Supplementary Material online). Student's *t*-test indicated differences in mutation accumulation between allelic variants of sexual accessions as compared with facultative or obligate apomictics for less than half of the genes ($P < 0.05$), whereas highly significant differences were observed between the mutation accumulation of *VASL* alleles as compared with any other gene tested within sexual accessions ($P < 0.001$).

In summary, differences in mutation load largely differed between genes, with the RNA helicases encoding for VASL and the related evolutionary conserved EIF proteins remaining highly conserved both in sexual and apomictic species. In contrast, for the coding regions of other genes, more divergence between allelic variants was observed, including such with known functions for germline development or candidate genes for elements of apomixis.

High Sequence Conservation of *VASL* Correlates with Roles during Reproductive Development

Our data indicated different evolutionary pressure to act on the RNA helicases both in apomicts and in sexuals. Interestingly, a small number of genes were identified to be highly conserved and under purifying selection independent of the reproductive mode, including the undescribed gene *VASL*.

To gain first insights into the potential roles of *VASL* for reproductive development, we took advantage of the established model species *A. thaliana*. First, we analyzed two independent T-DNA insertion lines, *vasl_1* and *vasl_2*. Furthermore, we generated *vasl_3* and *vasl_4* mutant lines with single-nucleotide insertions leading to premature stop codons using CRISPR/Cas9 (Wang et al. 2015). Interestingly, in all lines, similar defects of reproductive development and pleiotropic phenotypes were observed both in heterozygous and in homozygous plants, although at different frequencies (fig. 5 and supplementary fig. S9 and table S3, Supplementary Material online). In agreement with *VASL* expressed predominantly during germline formation (Schmidt et al. 2011), sporophytic defects of ovule development and protruding gametophytes were observed (fig. 5 and supplementary table S3, Supplementary Material online). At low frequencies, indications were given for the formation of ectopic gametophytes likely not derived from a FMS formed at the right position as in wild-type, in addition to occasional observations of odd numbers of nuclei in gametophytes (fig. 5 and supplementary table S3, Supplementary Material online). Further developmental defects were observed during seed development (supplementary fig. S9 and table S3, Supplementary Material online). Although embryo development usually follows a series of well-defined cell divisions and developmental stages (Jenik et al. 2007), in our lines, overproliferation of the embryo and abort of embryo development was observed at low frequencies (supplementary fig. S9 and table S3, Supplementary Material online). Taken together, the analysis indicates the importance of *VASL* for plant reproduction and supports the hypothesis that genes involved in crucial cellular processes or important for developmental processes including reproduction are conserved also in apomicts.

Discussion

Purifying Selection Observed for Most RNA Helicases Challenges Common Hypothesis on Mutation Accumulation in Apomicts

The coexistence of sexual and asexual reproduction is a feature of reproduction across different kingdoms. From an evolutionary perspective, both sexual and asexual reproductions have their unique advantages and downsides. On the one hand sexual reproduction allows de novo combination of genetic information and genomic elements and thus holds great potential for adaptive evolution, but also favorable allelic combinations can get lost by meiotic recombination. Furthermore, investment of energy and resources is required for the production of male gametes and for finding a suitable mating partner (Barton and Charlesworth 1998; Charlesworth 2006; Stelzer 2015). Asexual reproduction on the other hand can lead to the rapid fixation of hybrid genotypes frequently outperforming their parents. Over time however, mutations are predicted to accumulate in genomes of apomictic lineages leading to their rapid extinction (Muller 1964; Stelzer 2015). Investigations on different asexual species have only partially proven this hypothesis. Confirmation of this prediction has previously been found for *Oenothera*, where higher rates of mutation accumulation have been identified in functionally asexual as compared with sexual taxa based on comparisons of transcriptomes (Hollister et al. 2015). In the functional asexual taxa, recombination is nearly fully suppressed. Combined with self-fertilization, this results in offspring genetically identical to the mother plants by another system than apomixis (Hollister et al. 2015). Moreover, in apomictic *Boechera*, genome-wide higher frequencies of mutations have been

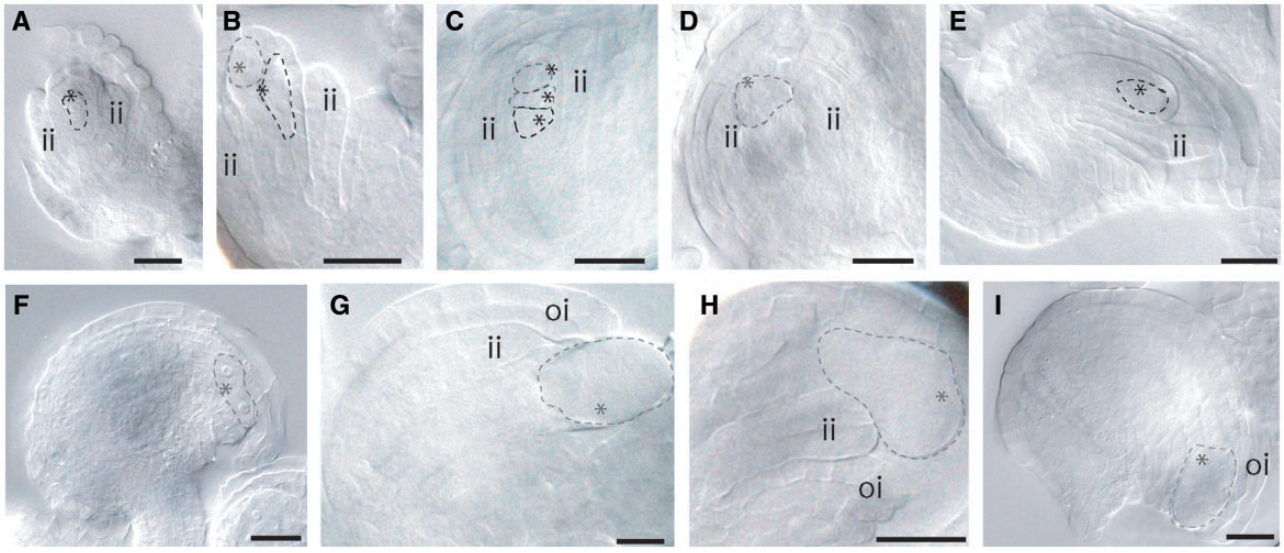

Fig. 5.—Reproductive phenotypes in lines carrying mutant alleles of *VASL* in *A. thaliana*. Seed clearing and DIC microscopy was used to study reproductive development in wild-type and lines carrying mutant alleles. (*A*) FMS in the wild-type. (*B–D*) Additional or ectopic gametophytic cells in *vasl_1/VASL* (*B*, *D*) and *vasl_3/VASL* (*C*). (*E*) Wild-type like gametophyte after the first mitotic division harboring two nuclei in *vasl_3/VASL*. (*F*) Ectopic gametophyte in *vasl_1/VASL*. (*G–I*) Slightly protruding (*I*) or protruding (*G*, *H*) gametophytes in *vasl_2/vasl_2* (*G*), *vasl_1/VASL* (*H*), and *vasl_3/VASL* (*I*), respectively. (ii) Inner integuments; black stars denote FMS; dark gray stars ectopic cells, ectopic, or protruding gametophytes. Scale bars are 20 μm.

described as compared with sexual lineages (Lovell et al. 2017). Although an elevated mutation load in apomicts is even presented at phylogenetically conserved sites, the effect is least observable for coding sites (Lovell et al. 2017). This suggests that also in apomicts mutation accumulation is not stochastically affecting any position in the genome but underlies certain evolutionary constraints. In the same line, from a study of *Ranunculus auricomus*, indications are given that gene evolution in apomicts is not random but related to gene functions (Pellino et al. 2013). Indeed, for CDS, elevated d*N*/d*S* ratios and thus indications either for accumulation of deleterious mutations or for diversifying selection are only given for certain genes, including such functionally important for meiosis and gametogenesis, whereas the majority of genes are under purifying selection (Pellino et al. 2013). To explain this observation, a relatively high degree of persistent facultative sexuality has been postulated to counteract the genome-wide effect of mutation accumulation (Hojsgaard and Hörandl 2015). Occasional processes of sexual reproduction would lead to an efficient unmasking of deleterious mutations and rapid purging from the population (Hojsgaard and Hörandl 2015; Hodač et al. 2019). Interestingly, although obligate apomicts have been defined based on maturing seeds in *Boechera* (Aliyu et al. 2010; Mau et al. 2015), greater developmental plasticity during early stages of germline development has recently been described and also rare reversions from apomictic to sexual reproductions are postulated (Carman et al. 2019). In apomictic *Boechera*, both apospory and diplospory are common. Diplospory is most frequently of *Taraxacum*-type, where a first division restitution of meiosis takes place in the AIC followed by meiosis II, but also *Antennaria*-type of diplospory has been reported where the AIC omits meiosis to directly initiate gametogenesis (Carman et al. 2019). In addition, rare occurrence of second division restitution maintaining meiotic recombination cannot fully be excluded. With respect to the evolutionary young apomicts of *R. auricomus*, it has been proposed that the time of divergence from sexual lineages was too short to allow substantial diversification (Hojsgaard and Hörandl 2015). In general, mutation accumulation in asexual species or lineages is largely influenced by the time of evolution and by population size (Brandt et al. 2017). Large population sizes thereby efficiently counteract the tendency of mutation accumulation. This has recently been shown by comparing ancient asexual to sexual oribatid mites (Brandt et al. 2017). Here, in contrast to common expectations, higher efficiency of purifying selection was observed in asexual rather than in sexual mites (Brandt et al. 2017). Apart from the effect of large populations, it can also not fully be excluded that so far largely overlooked genome repair mechanisms contribute to a certain degree to the conservation of sequences in asexual species, similar to mechanisms acting on nonrecombining genomes of mitochondria and chloroplasts (Khakhlova and Bock 2006; Brandt et al. 2017). Indeed, unexpected mechanisms of evolution have previously been reported for the ameiotic bdelloid rotifer *Adineta vaga*, where gene conversion and horizontal gene transfer counteract the accumulation of deleterious mutations (Flot et al. 2013). Taken together, the

partially contradicting observations from different studies in general suggest versatile mechanisms shaping gene evolution in different asexual systems and populations.

More Allelic Variants per Accession Were Identified from Apomicts than from Sexuals

In our study, rather than investigating a population, we selected accessions to represent part of the genetic diversity and distinct geographic origins in *Boechera* (Aliyu et al. 2010; Mau et al. 2015; Carman et al. 2019). Likewise done previously (Lovell et al. 2017), we defined mutations by comparison of sequences to the homologs of *A. lyrata*. Thereby, the finding of sequence changes largely at similar positions in sexual and apomictic *Boechera* reflects the general divergence of *Boechera* and *Arabidopsis* split about 10–14.5 Ma (timetree.org, last accessed October 01, 2019). Moreover, the number of haplotypes per gene in different accessions was correlated with the reproductive mode and ploidy, in agreement with previous studies indicating highest levels of heterozygosity for triploid apomicts, intermediate levels for diploid apomicts, and lowest levels for diploid sexuals (Lovell et al. 2013). In this line, especially in case of polyploid apomicts, the frequencies of mutation accumulation are likely to differ between alleles (Hojsgaard and Hörandl 2015). In consequence, different evolutionary forces can shape mutation accumulation on different alleles to allow loss of function or even de novo functionalization of one allelic variant. Although polyploidy is perceived as mechanism to buffer deleterious mutations, the situation in diploid apomicts is less well understood as complex genetic interactions might play a role in addition to copy number variations previously reported (Aliyu et al. 2013). Nevertheless, no consistent tendency of rapid diversification between alleles of the same accession was observed in triploid as compared with diploid apomicts. Disentangling the origins of the different allelic variants is also not trivial, as many apomicts are derived by hybridization or intra-specific crosses several times (Lovell et al. 2013).

CDS Regions Are Less Variable than Non-CDS

To gain a better understanding of the diversifications of alleles in *Boechera*, we accessed sequence divergence along the genomic regions. Hereby, increased variability in the sequences was evident particularly for noncoding regions. The identification of conserved stretches in regulatory regions when comparing different taxa, often indicates sequence motives important for regulation, particularly transcription factor binding sites. Our comparison of variance allows to distinguish regions of overall high variability and from such with differences between reproductive modes. These might serve as an indication for changes in regulation, as previously been described for the *APOLLO*-locus showing polymorphisms linked to apomixis in the 5′ UTR leading to changes in transcription factor binding sites (Corral et al. 2013). Therefore, identification of differences in regulatory regions between sexual and apomictic accessions can serve as starting points for investigations of potential functional roles of different alleles.

Gene Evolution Relates to Function in Sexual and Apomictic Lineages

Compared with noncoding regions, less variability was observed in CDS. Based on their allelic variants, *Boechera* RNA helicases could be classified into distinct groups: Genes with high sequence conservation being under strong purifying selection independent of the reproductive mode, genes showing more nonsynonymous sequence changes in the evolutionary less conserved N- and C-terminal domains, and genes harboring a number of allelic variants with deleterious mutations. These differences seemed to be mainly dependent on gene function, with, for example, members of the evolutionary conserved EIFs remaining conserved in apomictic and sexual *Boechera* accessions. In contradiction to common predictions for a subset of helicases, even higher sequence conservation of allelic variants was given in apomicts than in sexual accessions. However, it needs to be taken into account that in our analysis stop codons were replaced by deletions and alleles with frameshift mutations were not taken into account, as yn00 cannot account for stop codons. This likely leads to an underestimation of sequence divergence for a subset of genes.

Genes Less Conserved in Apomicts Involve Members of Epigenetic Regulatory Pathways

Although overall CDS remained under purifying selection in any accession, near neutral evolution was given for certain genes, including the homolog of *FRG5*. As *FRG1* and *FRG2* are involved in RNA-directed DNA methylation (Groth et al. 2014), also the *Boechera* homolog of *FRG5* might serve similar roles. Moreover, *ESP3* as a player in epigenetic regulatory pathways, was identified to be among the few genes with higher mutation load in apomicts or mutations causing stop codons only in alleles from apomicts. These findings are in agreement with the notion of apomixis being related to complex changes in the epigenetic landscape (León-Martínez and Vielle-Calzada 2019). However, as none of the deleterious mutations identified is consistently present in all apomicts investigated, such mutations are more likely to contribute to sustaining apomixis rather than being causal for the reproductive mode. Likewise, it is possible that different genes in similar epigenetic pathways are affected by mutations leading to similar deregulations of gene activity related to expression of components of apomixis. It is also important to note that genetically apomixis is typically linked to one or few recombination repressed loci (Barcaccia and Albertini 2013). In this respect, genes encoded on these loci should remain to be prone to mutation accumulation even at low rates of

facultative sexuality. On the other hand, identification of elevated mutation frequency alone cannot be taken as sufficient indication for a gene to be linked to an apomixis locus.

Interestingly, also for *MEM*, evidence for few more potentially deleterious mutations in core regions of apomictic as compared with sexual alleles was given. As mutant *mem* alleles cause formation of an additional and unreduced gametophyte in the same ovule in *A. thaliana* (Schmidt et al. 2011), such mutations might be relevant for aspects of apomixis in *Boechera*. Furthermore, high sequence variation was observed upstream of the *MEM* coding region, suggesting differences in the regulation of *MEM* activity. To regulate distinct developmental features, the observed diversity of *MEM* allelic variants might correlate to the developmental flexibility during germline development in *Boechera* (Carman et al. 2019). Nevertheless, *MEM* alleles carrying deleterious mutations were identified only from sexual accessions. As these mutations were in the C-terminal domain close to the stop codon, the mutation might not severely affect MEM functionality. Future functional investigations of allelic variants will be required to disentangle their potential relevance for aspects of apomixis.

Interestingly, for a subset of helicases, a high frequency of deleterious mutations was observed in both reproductive modes. Together with a high number of allelic variants observed for these genes, this might indicate cases of redundancy or copy number variations, or exceptionally problems of accurate mapping due to close homologs present in the genome.

High Sequence Conservation of Certain RNA Helicases Indicates Functions in Development

In contrast to the expected mutation accumulation in apomicts, we also identified a number of RNA helicases under strong purifying selection in all accessions. These include *CHR17* and members of the EIF family, involved in controlling fundamental cellular functions of translational control and regulation of nucleosome patterning, respectively, and in turn important to control gene or protein activity (Struhl and Segal 2013; Li et al. 2014; Dutt et al. 2015). From the animal germline, it is well appreciated that EIF proteins have essential functions in the regulation of translation of stored mRNAs (Henderson et al. 2009). Similar to the importance of such mechanisms in animals, evidence suggests that translational regulation of stored mRNAs is an essential mechanism also in the plant germline, in particular during meiosis (Schmidt et al. 2011). Here, we identified the undescribed gene *VASL* not only to be highly conserved in *Boechera* but also to have similarities to animal-specific VASA proteins, while being related to *EIF4A1* and *EIF4A2* based on phylogenetic analysis. Nevertheless, low bootstrap support for certain nodes likely results from the overall high divergence of RNA helicases.

Our functional analysis identifies *VASL* as gene involved in regulating reproductive development. In agreement with a relation to *EIF4A1*, defects in ovule development observed partially resemble phenotypes in *eif4A1* and *eif4A2* mutants (Bush et al. 2015). EIF proteins typically act in multimeric complexes often involving different isoforms, some of which are not essential for translation initiation but rather modify the efficiency of the process (Hinnebusch and Lorsch 2012). Importantly, also VASA has previously been determined to play roles in translation initiation and cell cycle regulation, in particular with respect to genes required for germline development (Yajima and Wessel 2011; Marintchev 2013). Therefore, VASL might serve similar roles in plant reproductive tissues. Strikingly, similar effects were observed in lines with insertions in the 3′ UTR as for the lines with a premature stop codon, as well as in heterozygous and homozygous lines. As the 3′ UTR is relevant for mRNA stability or translational regulation (Srivastava et al. 2018), these lines likely result in altered levels of VASL. Similar to EIF4A1 acting in a level-dependent manner (Vain et al. 2011; Bush et al. 2016), likely also VASL levels are relevant its mode of action.

Our observation of high conservation of genes important for reproductive development appears partially contradicting to the observations in *R. auricomus* where genes with roles for gametophyte development showed higher mutation accumulation in apomicts (Pellino et al. 2013). For such genes, higher rates of mutations in apomicts might be explained by dosage effects or recessive mutations being buffered by unreduced genomes in apomictic gametophytes. Nevertheless, even in sexual plants, genes preferentially expressed in female gametes are prone to higher rates of protein evolution as compared with the genome-wide average (Gossmann et al. 2014). This is in agreement with the finding of accumulation of deleterious mutations in *FAS4* and *CHR34* in alleles of any accession, which are predominantly transcribed in the cells of the female gametophyte in *A. thaliana* (Wuest et al. 2010). High evolutionary rates of genes in female gametogenesis might in general allow for a certain developmental flexibility during reproductive development, potentially also involved in sustaining the differences between sexual and apomictic reproduction.

In conclusion, our data indicate that versatile mechanisms are in place to shape the evolution of genes in apomictic *Boechera*. Thereby, the gene function appears as a key determinant for the acquisition of mutations. Based on the hypothesis that a high degree of purifying selection can be taken as an indication for important functional roles, *VASL* has been identified as a novel gene playing a role to regulate reproductive development. This indicates that the data presented provide an ideal starting basis for future functional investigations on the roles of the genes and allelic variants. Important new insights into aspects of the evolution of RNA helicases in sexual and apomictic *Boechera* are presented. A deeper understanding of aspects of the evolutionary forces acting on

apomicts, in particular with respect to genes important for regulating reproduction, is not only of great interest from a scientific point of view. In longer terms, this understanding can also contribute to allow predictions on behavior of engineered apomicts released to fields for agricultural use.

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Acknowledgments

We thank Timothy Sharbel (Global Institute for Food Security), Thomas Mitchell-Olds (Duke University, USA), and John Carman (Utah State University, USA) for providing *Boechera* seeds and Marcus A. Koch (Centre for Organismal Studies Heidelberg [COS], Germany) for providing instruments and facilities and support through his technical assistant. We also thank Jan U. Lohmann (COS Heidelberg) for contributing instruments and Christiane Kiefer (COS Heidelberg) for critical reading of the manuscript and for helpful discussions. We are grateful to Anne Pfeiffer (COS Heidelberg) for advice concerning CRISPR/Cas9 and to Qi-Jun Chen (College of Biological Sciences, China Agricultural University, Beijing, China) for providing the pHEE401E plasmid. This work was supported by funding of Deutsche Forschungsgemeinschaft (DFG) (SCHM2448/2-1 and SCHM2448/2-2 to A.S.).

Author Contributions

A.S. conceived the project. M.K. and A.S. planned and conducted the bioinformatics data analysis. A.S., D.I., B.H.N., C.V., and A.L. planned and performed the experimental work. A.S., M.K., B.H.N., and D.I. analyzed the data. A.S. wrote the manuscript. All authors approved the manuscript.

Literature Cited

Ahmad A, Zhang Y, Cao X-F. 2010. Decoding the epigenetic language of plant development. Mol Plant 3(4):719–728.

Aliyu OM, Schranz ME, Sharbel TF. 2010. Quantitative variation for apomictic reproduction in the genus *Boechera* (Brassicaceae). Am J Bot. 97(10):1719–1731.

Aliyu OM, Seifert M, Corral JM, Fuchs J, Sharbel TF. 2013. Copy number variation in transcriptionally active regions of sexual and apomictic *Boechera* demonstrates independently derived apomictic lineages. Plant Cell 25(10):3808–3823.

Arkov AL, Ramos A. 2010. Building RNA-protein granules: insight from the germline. Trends Cell Biol. 20(8):482–490.

Asano K, Clayton J, Shalev A, Hinnebusch AG. 2000. A multifactor complex of eukaryotic initiation factors, eIF1, eIF2, eIF3, eIF5, and initiator tRNA(Met) is an important translation initiation intermediate in vivo. Genes Dev. 14(19):2534–2546.

Audiffren J, Pardoux E. 2013. Muller's ratchet clicks in finite time. Stoch Process Their Appl. 123(6):2370–2397.

Barcaccia G, Albertini E. 2013. Apomixis in plant reproduction: a novel perspective on an old dilemma. Plant Reprod. 26(3):159–179.

Barton NH, Charlesworth B. 1998. Why sex and recombination? Science 281(5385):1986–1990.

Beck JB, et al. 2012. Does hybridization drive the transition to asexuality in diploid *Boechera*? Evolution 66(4):985–995.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30(15):2114–2120.

Brandt A, et al. 2017. Effective purifying selection in ancient asexual oribatid mites. Nat Commun. 8(1):873.

Bush MS, Crowe N, Zheng T, Doonan JH. 2015. The RNA helicase, eIF4A-1, is required for ovule development and cell size homeostasis in *Arabidopsis*. Plant J. 84(5):989–1004.

Bush MS, et al. 2016. eIF4A RNA helicase associates with cyclin-dependent protein kinase A in proliferating cells and is modulated by phosphorylation. Plant Physiol. 172(1):128–140.

Carman JG. 1997. Asynchronous expression of duplicate genes in angiosperms may cause apomixis, bispory, tetraspory, and polyembryony. Biol J Linn Soc Lon. 61(1):51–94.

Carman JG, et al. 2019. Apospory and diplospory in diploid *Boechera* (Brassicaceae) may facilitate speciation by recombination-driven apomixis-to-sex reversals. Front Plant Sci. 10:724.

Charlesworth B. 2006. The evolutionary biology of sex. Curr Biol. 16(17):R693–R695.

Conner JA, Ozias-Akins P. 2017. Apomixis: engineering the ability to harness hybrid vigor in crop plants. Methods Mol Biol. 1669:17–34.

Corral JM, et al. 2013. A conserved apomixis-specific polymorphism is correlated with exclusive exonuclease expression in premeiotic ovules of apomictic *Boechera* species. Plant Physiol. 163(4):1660–1672.

Danecek P, et al. 2011. The variant call format and VCFtools. Bioinformatics 27(15):2156–2158.

Dutt S, et al. 2015. Translation initiation in plants: roles and implications beyond protein synthesis. Biol Plant. 59(3):401–412.

Fairman-Williams ME, Guenther U-P, Jankowsky E. 2010. SF1 and SF2 helicases: family matters. Curr Opin Struc Biol. 20(3):313–324.

Flot J-F, et al. 2013. Genomic evidence for ameiotic evolution in the bdelloid rotifer *Adineta vaga*. Nature 500(7463):453–457.

Friday AJ, Keiper BD. 2015. Positive mRNA translational control in germ cells by initiation factor selectivity. BioMed Res. 2015:1–11.

Gao M, Arkov AL. 2013. Next generation organelles: structure and role of germ granules in germline. Mol Reprod Dev. 80(8):610–623.

Garcia-Aguilar M, Michaud C, Leblanc O, Grimanelli D. 2010. Inactivation of a DNA methylation pathway in maize reproductive organs results in apomixis-like phenotypes. Plant Cell 22(10):3249–3267.

Garrison E, Marth G. 2012. Haplotype-based variant detection from short-read sequencing. arXiv: 1207.3907.

Gnirke A, et al. 2009. Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. Nat Biotechnol. 27(2):182–189.

Gossmann TI, Schmid MW, Grossniklaus U, Schmid KJ. 2014. Selection-driven evolution of sex-biased genes is consistent with sexual selection in *Arabidopsis thaliana*. Mol Biol Evol. 31(3):574–583.

Groth M, et al. 2014. SNF2 chromatin remodeler-family proteins FRG1 and -2 are required for RNA-directed DNA methylation. Proc Natl Acad Sci U S A. 111(49):17666–17671.

Gustafson EA, Wessel GM. 2010. Vasa genes: emerging roles in the germ line and in multipotent cells. BioEssays 32(7):626–637.

Hafidh S, Čapková V, Honys D. 2011. Safe keeping the message: mRNP complexes tweaking after transcription. In: Collins LJ, editor. RNA infrastructure and networks. New York: Springer New York. p. 118–136.

Henderson MA, et al. 2009. A germline-specific isoform of eIF4E (IFE-1) is required for efficient translation of stored mRNAs and maturation of both oocytes and sperm. J Cell Sci. 122(10):1529–1539.

Hernández G, Vazquez-Pianzola P. 2005. Functional diversity of the eukaryotic translation initiation factors belonging to eIF4 families. Mech Dev. 122(7–8):865–876.
Higgins JD, Ferdous M, Osman K, Franklin F. 2011. The RecQ helicase AtRECQ4A is required to remove inter-chromosomal telomeric connections that arise during meiotic recombination in *Arabidopsis*. Plant J. 65(3):492–502.
Hinnebusch AG, Lorsch JR. 2012. The mechanism of eukaryotic translation initiation: new insights and challenges. Cold Spring Harb Perspect Biol. 4:pii: a011544.
Hodač L, Klatt S, Hojsgaard D, Sharbel TF, Hörandl E. 2019. A little bit of sex prevents mutation accumulation even in apomictic polyploid plants. BMC Evol Biol. 19(1):170.
Hojsgaard D, Hörandl E. 2015. A little bit of sex matters for genome evolution in asexual plants. Front Plant Sci. 6:82.
Hojsgaard D, Klatt S, Baier R, Carman JG, Hörandl E. 2014. Taxonomy and biogeography of apomixis in angiosperms and associated biodiversity characteristics. Crit Rev Plant Sci. 33(5):414–427.
Hollister J, et al. 2015. Recurrent loss of sex is associated with accumulation of deleterious mutations in *Oenothera*. Mol Biol Evol. 32(4):896–905.
Huanca-Mamani W, Garcia-Aguilar M, Leon-Martinez G, Grossniklaus U, Vielle-Calzada J-P. 2005. CHR11, a chromatin-remodeling factor essential for nuclear proliferation during female gametogenesis in *Arabidopsis thaliana*. Proc Nat Acad Sci U S A. 102(47):17231–17236.
Jankowsky E. 2011. RNA Helicases at work: binding and rearranging. Trends Biochem Sci. 36(1):19–29.
Jenik P, Stewart Gillmor C, Lukowitz W. 2007. Embryonic patterning in *Arabidopsis thaliana*. Annu Rev Cell Dev Biol. 23(1):207–236.
Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 30(4):772–780.
Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg M. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc. 10(6):845–858.
Khakhlova O, Bock R. 2006. Elimination of deleterious mutations in plastid genomes by gene conversion. Plant J. 46(1):85–94.
Khemici V, Linder P. 2018. RNA helicases in RNA decay. Biochem Soc Trans. 46(1):163–172.
Kliver S, et al. 2018. Assembly of the *Boechera retrofracta* genome and evolutionary analysis of apomixis-associated genes. Genes (Basel) 9(4):185.
Kotliński M, et al. 2017. Phylogeny-based systematization of *Arabidopsis* proteins with histone H1 globular domain. Plant Physiol. 174(1):27–34.
Lee C-R, et al. 2017. Young inversion with multiple linked QTLs under selection in a hybrid zone. Nat Ecol Evol. 1(5):119–119.
León-Martínez G, Vielle-Calzada J-P. 2019. Apomixis in flowering plants: Developmental and evolutionary considerations. Curr Top Dev Biol. 131:565–604.
Li F-W, Rushworth CA, Beck JB, Windham MD. 2017. *Boechera* microsatellite website: an online portal for species identification and determination of hybrid parentage. Database (Oxford) 2017:baw169.
Li G, et al. 2014. ISWI proteins participate in the genome-wide nucleosome distribution in *Arabidopsis*. Plant J. 78(4):706–714.
Li H, et al. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25(16):2078–2079.
Li WH, Wu CI, Luo CC. 1985. A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. Mol Biol Evol. 2(2):150–174.
Linder P, Jankowsky E. 2011. From unwinding to clamping—the DEAD box RNA helicase family. Nat Rev Mol Cell Biol. 12(8):505–516.
Linder P, Owttrim GW. 2009. Plant RNA helicases: linking aberrant and silencing RNA. Trends Plant Sci. 14(6):344–352.
Liu M, Shi D-Q, Yuan L, Liu J, Yang W-C. 2010. SLOW WALKER3, encoding a putative DEAD-box RNA helicase, is essential for female gametogenesis in *Arabidopsis*. J Integr Plant Biol. 52(9):817–828.
Lovell JT, Williamson RJ, Wright SI, McKay JK, Sharbel TF. 2017. Mutation accumulation in an asexual relative of *Arabidopsis*. PLoS Genet. 13(1):e1006550.
Lovell JT, et al. 2013. On the origin and evolution of apomixis in *Boechera*. Plant Reprod. 26(4):309–315.
Lunter G, Goodson M. 2011. Stampy: a statistical algorithm for sensitive and fast mapping of Illumina sequence reads. Genome Res. 21(6):936–939.
Lysak MA. 2014. Live and let die: centromere loss during evolution of plant chromosomes. New Phytol. 203(4):1082–1089.
Marintchev A. 2013. Roles of helicases in translation initiation: a mechanistic view. Biochim Biophys Acta 1829(8):799–809.
Mau M, et al. 2015. Hybrid apomicts trapped in the ecological niches of their sexual ancestors. Proc Natl Acad Sci U S A. 112(18):E2357–E2365.
McKenna A, et al. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20(9):1297–1303.
Muller HJ. 1964. The relation of recombination to mutational advance. Mutat Res. 106:2–9.
Nakajima K. 2018. Be my baby: patterning toward plant germ cells. Curr Opin Plant Biol. 41:110–115.
Nonomura KI. 2018. Small RNA pathways responsible for non-cell-autonomous regulation of plant reproduction. Plant Reprod. 31(1):21–29.
Olmedo-Monfil V, et al. 2010. Control of female gamete formation by a small RNA pathway in *Arabidopsis*. Nature 464(7288):628–632.
Pagnussat GC, et al. 2005. Genetic and molecular identification of genes required for female gametophyte development and function in *Arabidopsis*. Development 132(3):603–614.
Pellino M, et al. 2013. Asexual genome evolution in the apomictic *Ranunculus auricomus* complex: examining the effects of hybridization and mutation accumulation. Mol Ecol. 22(23):5908–5921.
Pikaard CS, Mittelsten Scheid O. 2014. Epigenetic regulation in plants. Cold Spring Harb Perspect Biol. 6(12):a019315.
Recker J, Knoll A, Puchta H. 2014. The *Arabidopsis thaliana* homolog of the helicase RTEL1 plays multiple roles in preserving genome stability. Plant Cell 26(12):4889–4902.
Rushworth CA, Song BH, Lee CR, Mitchell-Olds T. 2011. *Boechera*, a model system for ecological genomics. Mol Ecol. 20(23):4843–4857.
Schiffels S, Szöllosi GJ, Mustonen V, Lässig M. 2011. Emergent neutrality in adaptive asexual evolution. Genetics 189(4):1361–1375.
Schilling MP, Gompert Z, Li FW, Windham MD, Wolf PG. 2018. Admixture, evolution, and variation in reproductive isolation in the *Boechera puberula* clade. BMC Evol Biol. 18(1):61.
Schmidt A, Schmid MW, Grossniklaus U. 2015. Plant germline formation: common concepts and developmental flexibility in sexual and asexual reproduction. Development 142(2):229–241.
Schmidt A, et al. 2011. Transcriptome analysis of the *Arabidopsis* megaspore mother cell uncovers the importance of RNA helicases for plant germline development. PLoS Biol. 9(9):e1001155.
Schmidt A, et al. 2014. Apomictic and sexual germline development differ with respect to cell cycle, transcriptional, hormonal and epigenetic regulation. PLoS Genet. 10(7):e1004476.
Shimizu KK, Ito T, Ishiguro S, Okada K. 2008. MAA3 (MAGATAMA3) helicase gene is required for female gametophyte development and pollen tube guidance in *Arabidopsis thaliana*. Plant Cell Physiol. 49(10):1478–1483.
Singh M, et al. 2011. Production of viable gametes without meiosis in maize deficient for an ARGONAUTE protein. Plant Cell 23(2):443–458.

Skidmore ZL, et al. 2016. GenVisR: Genomic Visualizations in R. Bioinformatics 32(19):3012–3014.
Sloan KE, Bohnsack MT. 2018. Unravelling the mechanisms of RNA helicase regulation. Trends Biochem Sci. 43(4):237–250.
Spillane C, Curtis MD, Grossniklaus U. 2004. Apomixis technology development-virgin births in farmers' fields? Nat Biotechnol. 22(6):687–691.
Srivastava A, Lu Y, Zinta G, Lang Z, Zhu J-K. 2018. UTR-dependent control of gene expression in plants. Trends Plant Sci. 23(3):248–259.
Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30(9):1312–1313.
Stelzer C-P. 2015. Does the avoidance of sexual costs increase fitness in asexual invaders? Proc Natl Acad Sci U S A. 112(29):8851–8858.
Struhl K, Segal E. 2013. Determinants of nucleosome positioning. Nat Struct Mol Biol. 20(3):267–273.
Tang L, Nogales E, Ciferri C. 2010. Structure and function of SWI/SNF chromatin remodeling complexes and mechanistic implications for transcription. Prog Biophys Mol Biol. 102(2–3):122–128.
Tekleyohans DG, Nakel T, Gross-Hardt R. 2017. Patterning the female gametophyte of flowering plants. Plant Physiol. 173(1):122–129.
Vain P, et al. 2011. A T-DNA mutation in the RNA helicase eIF4A confers a dose-dependent dwarfing phenotype in *Brachypodium distachyon*. Plant J. 66(6):929–940.
Wang ZP, et al. 2015. Egg cell-specific promoter-controlled CRISPR/Cas9 efficiently generates homozygous mutants for multiple target genes in *Arabidopsis* in a single generation. Genome Biol. 16(1):144.
Warren WC, et al. 2018. Clonal polymorphism and high heterozygosity in the celibate genome of the Amazon molly. Nat Ecol Evol. 2(4):669–679.
Windham MD. 2010. *Boecheria*. In: Flora of North America Editorial Committee, editors. Flora of North America North of Mexico. Vol. 7. New York/Oxford: Oxford University Press. p. 348–412.
Wuest SE, et al. 2010. *Arabidopsis* female gametophyte gene expression map reveals similarities between plant and animal gametes. Curr Biol. 20(6):506–512.
Xu R, Zhang S, Huang J, Zheng C. 2013. Genome-wide comparative in silico analysis of the RNA helicase gene family in *Zea mays* and *Glycine max*: a comparison with *Arabidopsis* and *Oryza sativa*. PLoS One 8(11):e78982.
Yajima M, Wessel GM. 2011. The multiple hats of Vasa function and its regulation of cell cycle progression. Mol Reprod Dev. 78(10–11):861–867.
Yang Z. 2007. PAML 4: Phylogenetic Analysis by Maximum Likelihood. Mol Biol Evol. 24(8):1586–1591.
Yang Z, Bielawski JP. 2000. Statistical methods for detecting molecular adaptation. Trends Ecol Evol. 15(12):496–503.
Zhou LZ, Juranic M, Dresselhaus T. 2017. Germline development and fertilization mechanisms in maize. Mol Plant 10(3):389–401.
Zhou M, Palanca AMS, Law JA. 2018. Locus-specific control of the de novo DNA methylation pathway in *Arabidopsis* by the CLASSY family. Nat Genet. 50(6):865–873.

Associate editor: Yves Van De Peer