Government of Saskatchewan
Ministry of Agriculture

ADF Project Progress Report Format

1. **Project title and ADF file number.**
   Title: "Preserving hybrid vigour through a novel apomixis breeding strategy in *Brassica* crops"
   File-#: 20180141

2. **Name of the Principal Investigator and contact information.**
   Prof. Dr. Timothy Sharbel,
   Department of Plant Sciences, College of Agriculture and Bioresources, University of Saskatchewan, 51 Campus Drive, Saskatoon, SK S7N 5A8 Canada
   Email: tim.sharbel@usask.ca

3. **Name of the collaborators and contact information.**
   Dr. Martin Mau (**author of the progress and the final reports**)
   Department of Plant Sciences, College of Agriculture and Bioresources, University of Saskatchewan, 51 Campus Drive, Saskatoon, SK S7N 5A8 Canada
   Email: martin.mau@usask.ca

4. **Abstract/ Summary**: *An outline on overall project objectives, methods, key findings and conclusions for use in publications and in the Ministry database (**Maximum of 500 words** or one* page*).*

   Abstract Content: The abstract addresses the following (usually 1–2 sentences per topic):
   - Key aspects of the literature review
   - Problem under investigation or research question(s)
   - Clearly stated hypothesis or hypotheses
   - Methods used (including brief descriptions of the study design, sample, and sample size)
   - Study results
   - Conclusions
   Asexual seed formation (*i.e.* apomixis) is rare among crop plants, but its implementation in modern breeding strategies could become a game changer for agriculture due to its potential to fixate the hybrid vigour in clonal crops. This study utilizes diploid natural apomictic and previously developed synthetic diploid apomictic hybrids lines of the apomixis model *Boechera* for a two-pronged approach which will facilitate the implementation of apomixis technology in breeding schemes of canola and other key crops. The diploid synthetic apomictic Boechera hybrids producing haploid pollen are unique in their potential to infectiously transfer apomixis which allow the repeated introgression of the apomictic father into recipient sexual without polyploidization which typically leads to sterility in the progeny. This approach enables to minimize the gene-space from the father which harbors all required genetic factors to trigger the apomixis trait in the backcrosses in each generation and facilitates the simultaneous identification of all factors required to express the trait by sequencing and analysing the segregation patterns of paternal gene space in the generated backcrosses.
   We have successfully executed a backcrossing scheme of using 6 sexual recipient and 4 synthetic hybrid apomictic *Boechera* lines to produce 12 near-isogenic lines (NILs) with reduced genomic proportion from the apomictic father that contain all required genetic factors to trigger the apomixis trait. Overall, we performed 37036 emasculations and 8491 pollinations to produce and geno-/phenotype 3877 hybrid F_1/F_2 and BC_1 to BC_3/BC_4 backcrosses from 12 NILs. We geno- and phenotyped all parents and hybrid progeny to select only diploid unbalanced apomicts as donors in the next backcross generation and show that the paternal apomictic phenotype is transferred across the several backcross generations (BC_1 to BC_4). We have successfully sequenced, assembled, and annotated chromosome-level whole genomes from all parental lines, and we are currently

completing the skim-sequencing of all NILs. Genomes of both, parental lines and NILs, are currently used to find novel and validate known apomixis candidate factors.

In a second approach we were testing whether an apomict such as *Boechera* can be crossed into close relative crops, such as from the mustard family (*Brassicales*) to potentially trigger the apomixis trait in the crop plant directly. We developed and tested different crossing methods in various mustard crops. We demonstrate the first successful introgression of an apomict (*i.e. Boechera sp.*) into Canola (i.e. *B. napus* cv 'Golden') and validated the introgression with GISH and apo-specific PCR markers. The five interspecific canola hybrids host different sets of apomixis candidate factors and whole genome sequencing was used to validate the nature of introgressed *Boechera* fragments. Phenotypic analyses of intergeneric F1 and F2/BC hybrids revealed subtle differences in flower shape as well as plant height and substantial differences in seed size/viability and genome size compared to native Canola.

5. **Extension Messages:** *key outcomes and their importance for producers/processors and the relevant industry sector (**3-5 bullet points** **in lay language).***
    - Proof of stable inheritance of unbalanced apomixis across several backcross generation in the model plant *Boechera* minimizes effort to identify all apomixis factor simultaneously by analysing only least common paternal (*i.e.* apomictic donor) genome fragments in the final backcrosses instead of whole genomes.
    - Generated synthetic apomictic backcrosses enable segregation analyses to refine and validate any independently or previously generated apomixis candidate factor dataset.
    - Chromosome-scale whole genome assemblies of all parental lines used in the backcross array in conjunction with low-coverage sequencing of all final backcrosses allow for generation of independent apomixis factor list.
    - Generation of first intergeneric hybrids between an apomict (*i.e. Boechera*) and *Brassica napus* and introgression of subsets of apomixis candidate factors showcases engineering of apomixis in canola is possible.
    - Various geno-/phenotyping and bioinformatics protocols developed and applied to NIL and intergeneric hybrid datasets (*e.g.* apomixis marker PCR screen, embryo rescue protocols for intergeneric hybrids, genome assembly/annotation & SNP/Indel analyses pipelines).

6. **Introduction:** *Brief project background and rationale (**Maximum of 1500 words or 1.5-3 pages**).*

Engineering asexual seed production (apomixis) in major crop plants, which predominantly reproduce sexually, could provide an enabling technology to rapidly fix and maintain valuable genotypes and associated heterotic traits without inbreeding depression [[1, 2]] and facilitate selection for complex genetic traits not available to current breeding strategies [[3]]. In apomicts three essential developmental events (*i.e.* functional elements) are genetically controlled: *(i)* the formation of a chromosomally-unreduced egg cell via *apomeiosis*, *(ii) parthenogenetic* embryonic development of this egg cell and, *(iii)* fertilization of the polar nuclei with male gametes to produce *chromosomally balanced endosperm* (*i.e.* pseudogamy) [[4]]. A number of early studies have postulated single dominant mendelian factors for apomixis inheritance [[4-7]], which are typically hemizygous or highly divergent in allele composition [[6, 8, 9]]. More recently, segregation of single apomixis elements have been shown in a number of natural-occurring apomicts [[10], [11, 12] and [13]]. Taken together, these results point to a more complex origin of apomixis in natural apomicts, with monogenic control in some species [[14]] and with multiple segregating independent loci in others [[15, 16]].

Therefore, the most straightforward method to engineer apomixis in a crop plant is to introduce a transgenic "apomixis cassette" containing genes conferring the required developmental components of apomixis [[17]] whereby the resulting apomictic lines could be used as pollen donors to create apomictic F_1 individuals [[18, 19]]. Yet, this cassette does not exist, although a plethora of, mostly single and dominant, apomixis candidate loci in various model plants have been reported (*e.g.* in *Poa*, *Taraxacum*, *Tripsacum*, *Hieracium*, *Hypericum*, and *Boechera*; [[20]]). Thus it is clear that not all necessary apomixis factors are known, and is a reflection of apomixis being a species-specific trait which is not controlled by a single apomixis candidate gene [[20]].

A second strategy to introduce apomixis into crop plants is the wide intergeneric cross between a natural apomictic relative into crop plants which has been attempted for example *from Pennisetum squamulatum/purpureum* into pearl millet [[21]], from *Tripsacum dactyloides* into *Zea mays* [[22]] or from *Elymus rectisetus* into wheat [[23]]. All reported attempts demonstrated the transfer of apomixis into the crop plant to some extent but failed to stably generate apomictic progeny which retained the crop plant phenotype. Multiple effects have hindered the successful implementation of apomixis in major crops, such as (1) high male sterility in the progeny which hinders back crossing, (2) polyploidization tendencies in the progeny [[24]], and (3) an increase of facultative seed production or inviable pollen and low seed set [[20]]. A further challenge is linkage between the desired transferred apomixis factors and other factors which negatively affect apomixis (*e.g.* linkage drag and alteration of epigenetic programs; [[20]]).

Our study follows a two-pronged approach, which (1) will generate backcrosses for the simultaneous discovery of all genetic factors for apomixis initiation, and (2) test the transferability of apomixis into multiple *Brassica* crops. The key advantage of our study is the use of a number of well-characterized natural and synthetic apomictic hybrid lines of the North-American Brassicaceae genus *Boechera* Á. Löve & D. Löve [[25-27]] from a 4-year pilot study ([[28]], submitted) which fulfill all necessary requirements for the successful transfer of apomixis via backcrosses and intergeneric crosses without ploidy barriers into a *Brassica* crop.

Boechera is characterized by the presence of diploid sexual, as well as diploid and polyploid (mostly triploid) apomictic lineages [[29]]. In addition to diploid sexuals, diploid and polyploid natural apomictic *Boechera* typically produce seeds with balanced maternal and paternal genome ratios in the endosperm (*i.e.* maternal to paternal genome ratios of $2_m:1_p$ or a multiple thereof; [[29, 30]]), and from here on such apomictic individuals are termed 'balanced apomicts'. Double fertilization with haploid gametes leads to sexual seeds composed of a diploid (2C) embryo and triploid (3C) endosperm (*i.e*. the 2C:3C embryo to endosperm genome content ratio is referred to as 'seed ploidy'), the latter of which is composed of 2 maternal ($2C_m$) and 1 paternal (C_p) haploid genomes. Balanced diploid apomicts produce seeds with 2C:6C ploidy ratio, with a $4C_m:2C_p$ genome ratio in the endosperm. Also triploid apomicts with triploid male gametes can form balanced seeds (3C:9C seed ploidy with a $6C_m:3C_p$ endosperm genome ratio; [44]). In some natural diploid and polyploid *Boechera* deviations from a balanced endosperm are tolerated [[30, 31]] and those individuals are referred to as 'unbalanced apomicts'. For example, some diploid apomicts with haploid male gametes produce progeny with 2C:5C seed ploidy ($4C_m:C_p$ endosperm genome ratio) and triploid apomicts with hexaploid pollen produce progeny with 3C:12C seed ploidy ($6C_m:6C_p$ endosperm genome ratio; [[29, 30]]).

In this study we approach the transfer of the apomixis trait from diploid unbalanced apomictic *Boechera* donor lines directly into multiple *Brassica* crops (*B. napus*, *B. rapa* and *B. juncea*) or via the diploid bridge species *Diplotaxis erucoides*. Thereby, we meet several key prerequisites which can facilitate apomixis transfer [[32]]. First, we demonstrated the successful transfer of apomixis via haploid pollen from these diploid unbalanced *Boechera* donor lines into diploid sexual recipients in inter- and intra-specific crosses of the genus *Boechera,* and we showed the stable inheritance of the apomixis trait in subsequent generation ([[28]], submitted). Second, the pollen in these crosses is unaltered, with basic chromosome number and ploidy being similar between the apomictic donor and the bridge species. Another asset which facilitates the generation of an apomictic crop is

the close phylogenetic relation between *Boechera* and *Brassica* crops like canola. We use genetic markers for female (*APOLLO*; [[12]]) and male apomeiosis (*UPGRADE2*; [[11, 33]]) in addition to flow cytometric analyses of seeds to proof the transfer of apomixis.

The transfer of apomixis from natural diploid unbalanced apomictic *Boechera* lines into *Brassica* crops (mostly diploid or amphidiploid) may be successful but other major crops such wheat or maize might not benefit from our crossing strategy, because of greater phylogenetic distance, ploidy barriers (*e.g.* hexaploid wheat vs. diploid *Boechera*), or larger genome size differences and resulting homeology deficiencies. Therefore, we use the synthetic diploid unbalanced apomictic hybrid lines in a recurrent backcrossing assay which restores most of the sexual genotype, while we will select in each of the generation for the diploid, unbalanced apomictic paternal phenotype to minimize the paternal gene-space containing all the genetic factors for apomixis trait expression. We use multiple diploid unbalanced apomictic independent backcross lines of inter- and intraspecific origin in our whole genome comparison between and across hybrid backcrosses and their parental lines, which will statistically correct for phylogenetic-mediated background noise and facilitate the selection of apomixis-specific regions.

7. Objectives and the progress towards meeting each objective.

Objectives **(Please list the original objectives and/or revised objectives if Ministry-approved revisions have been made to original objectives. A justification is needed for any deviation from original objectives).**	**Status** **(e.g. completed/not completed)**
a) *Generation of diploid, hybrid unbalanced apomictic Boechera backcrosses*	**90% completed (BC1F1 to BC3F1)** – Successful generation, geno- and phenotyping of backcrosses throughout all 12 near-isogenic lines (NILs) with inheritance of selected phenotype (i.e. diploid, unbalanced apomixis). 98% of all BC1 crosses, >90% of all BC2 crosses, ~75% of all BC3 crosses across 12 NILs are completed, and geno-/phenotyping concluded. Reduction of the number of backcross generation from BC5 to BC3 → *see justification in paragraph 1 of section 9a.*
b) *Transfer of apomixis from selected diploid unbalanced apomictic Boechera lines into sexual bridging species via hybrid crosses*	**100% completed** – We tested various crossing methods and generated sufficient crosses to draw conclusions on the potential to introgress the genome of an apomict into the recipient bridge species or *Brassicales* crop (*i.e. B. juncea*, *B. napus* and *B. rapa*). Crosses into the bridge species *D. erucoides*, *B. juncea*, and *B. rapa* did not generate any hybrid progeny or the generated progeny was not of hybrid origin (*i.e.* false positive seed potentially derived through cross contamination with outcrossing seeds from neighbouring plants or own seeds). Nonetheless, apomict was successful introgressed into *Brassica napus*. Results produced together with objectives c) and d).
c) *Testing for stable inheritance of the apomixis trait in the bridging species*	**~80% completed** – We successfully generated BC1F1 and F2 plants from 5 interspecific hybrids between recipient *B. napus* and two unbalanced apomictic *Boechera* donor lines. Embryo rescue protocols were established. For all 5 lines

	~50 F2 plants were planted and phenotyped. For 1 of 5 lines backcrosses into B. napus were successfully planted via embryo rescue and phenotyped. For the remaining 3 BC lines plants have to be planted via embryo rescue. For all BC and F2 hybrids must be genotyping must be completed. Results produced together with objectives b) and d)..
d) *Generating apomictic Brassica crops via intergeneric crosses*	**100% completed –** This is the first time an apomict was successful introgressed into *Brassica napus*. Five independent hybrids from two different donor lines were produced. Genomic in situ hybridizations and apomixis marker PCR confirmed the introgression of Boechera genomic fragments into the *B. napus* genome. Subset of apomixis-specific orthologs from Boechera can be found in the five hybrid lines. Results produced together with objectives b) and c).
e) *Identification of all apomixis factors in the backcrosses of de novo generated hybrid apomictic Boechera and Brassica crops*	**~70% completed** – NIL sequencing & Bioinformatics in progress. All analyses pipelines established. Parental genomes sequenced, assembled to chromosome-level and annotated – ready for parental SNP analyses in NILs and validation of external apomixis factor lists (*e.g.* ADF #20150123).

Please add additional lines as required.

8. **Methodology:** *Specify project activities undertaken during the entire project period (without referring to previous progress report). Include approaches, experimental design, methodology, materials, sites, etc.* ***(Maximum of 5 pages)***

Plant material and cultivation conditions

Boechera seeds were sterilized for five minutes in 70% EtOH and additional fifteen minutes in 10% common household bleach with one intermediate and three final rinses of distilled water. Seeds were cultured on autoclaved MS media (4.3g MS nutrient powder, bioPLUS™, GeneLinx International, Dublin/Ohio, USA; 8 g phytoagar, bioPLUS™, GeneLinx International, Dublin/Ohio, USA; pH 5.7 with 0.5M NaOH) in sealed Petri dishes and vernalized at 4°C in the dark for two weeks until germination. In some accessions phytohormonal inhibitors cause reduced or lack of germination. In order to break this dormancy the seeds can be placed in distilled water - optionally supplemented with 1% hydrogen peroxide - for 24hours prior to their transfer onto media plates. Then, the Petri dishes were transferred to a growth chamber under long-day conditions for 1-2 weeks with 16 h in light and 8h in dark, 22°C. Subsequently, *Boechera* seedlings were transferred into 1-inch pots and grown on Sunshine Mix#8 as a substrate for 28 days (Sunshine® Mix #8/Fafard®-2; Sun Gro Horticulture, Vancouver, Canada). Then, the plants were transferred to four-inch square pots for all parental plants and 2.35-inch pots for the F1 progeny filled with Sunshine® LT5 substrate (Sun Gro Horticulture, Vancouver, Canada) and placed into a cold chamber for six weeks at 4°C for vernalisation. Finally, the vernalized plants were transferred in a temperature-controlled greenhouse and grown under long-day conditions (17h light and 21°C; 7 h dark and 18°C) at 1200 µmol/m^2/sec. Seeds from *Brassica napus, B. rapa,* and *B. juncea* were placed directly into Sunshine Mix#8 in a 6-inch pot. While *Brassica napus* and *B. juncea* did not need vernalization, all *B. rapa* were transferred at seedling stage (approx. 2 weeks in growth chamber under long-day conditions with 16 h in light and 8h atv22°C) into a 4°C chamber with short-day conditions (8 h in light and 16h in dark,) for 6 weeks. All plants matured in a temperature-controlled greenhouse and grown under long-day conditions (17h light and 21°C; 7 h dark and 18°C) at 1200 µmol/m^2/sec.

Crossing pedigree and generation of synthetic hybrid apomicts

Intergeneric crosses were generated using *Diplotaxis erucoides* (*i.e.* 2n=2x=14, diploid) and three Brassica species (*B. rapa, B. juncea* and *B. napus*) as sexual recipients and 5 diploid unbalanced apomictic *Boechera* accessions (all 2n=2x=14, haploid pollen producer) representing four different *Boechera* species (*B. retrofracta, B. crandallii, B. stricta, B. stricta x spatifolia*) were used as pollen donors. In addition, three sexual *Boechera* pollen donors were used for intergeneric crosses into the sexual recipients to evaluate the influence of the mode of reproduction on the crossing success. A minimum of 200 repetition for each cross combination is performed. Thereby, we tested four different crossing methods. In method 1 recipient flower buds with non-dehiscent anthers were emasculated one day prior to pollination using clean forceps. Under optimal conditions, emasculated flower buds were pollinated at 24h after emasculation. For method 2, immature fully closed flower buds were bagged at day 1, emasculated at day 2 after removal of the bag and pollinated at day 3. As third method we have introduced the emasculation and pollination of immature flower buds with yellow petals at the same day. The hormon-enhanced cross pollination (Method 4) is the most strict and minimal invasive crossing method of all tested methods. Anthers were removed from flowers at anthesis leaving the rest of the flower intact before pollination on pistil with or without removal of the carpel above the style. This allowed to remove any self-incompatibility related hormonal inhibitors to effect pollen tube growth. The pollen was either placed on the papillae of the carpel or on top of the cut site and the pollen pellet was then covered with a cotton ball soaked in a hormone pistil treatment solution (12mg/ml GA3, 3mg/ml NAA, 1.5mg/ml Kinetin, 1mg/ml glucose, distilled water adjusted to pH 5.5) to ensure a moist environment for the pollen which allows the pollen to germinate. Siliques were harvested after three weeks for crosses with *Diplotaxis erucoides*, and six weeks for crosses with *Brassica sp.*.

DNA extraction for genotyping procedures

A previously developed high-throughput extraction protocol for plant DNA from young *Boechera* leaves was applied [[28]]. In brief, *Boechera* leaf samples of approx. 1cm^2 size were collected on ice into a deep well (DW) plate and subsequently frozen at -80°C until further usage. All leaf samples were transferred to liquid nitrogen for two minutes and the tissue was ground for 30 seconds at 1400 rpm with a single tungsten carbide bead (4mm diameter) using a tissue homogenizer (2010 Geno/Grinder, SPEX SamplePrep, Metuchen, NJ, USA). The tissue powder was dissolved in 400 µl Edward's buffer (200mM Tris, 250mM NaCl, 25mM EDTA, 0.5%SDS, pH7.5) supplemented with 1.2% RNase A and incubated at 65°C for 15 minutes with occasional shaking. After centrifuging for 30 seconds at 4000 rpm in room temperature 250 µl of sample suspension per well were transferred without touching the pellet into the 96 DW plate containing the suspension of Genomic Bind buffer (FastID, Genomic Bind, #K1-0001-0050) and magnetic beads (Qiagen, MagAttract Suspension G, #1026901). A tip comb (ThermoFisherScientific, #97002534) was placed into the 96 DW plate containing the sample-magnetic buffer suspension. The sample plate, a DW plate with 250ul genomic wash (FastID, Genomic Wash, #K1-0001-0050), a DW plate with 400ul 70% EtOH, and a plate 0.1-fold TE elutionc(ThermoFisherScientific, #5400620) were loaded into a KingFisher™ Apex Purification System (ThermoFisherScientific, #5400610) and a specific 22-minute-program was used for DNA extraction. The extracted *Boechera* DNA samples were stored until further usage at -30°C.

High molecular weight (HMW) DNA extraction for whole genome sequencing (WGS) procedures

The identification of all genetic components sufficient for the expression of the apomixis trait in the backcrosses of de novo generated hybrid apomictic *Boechera* requires the sequencing of whole genomes of the parental accessions assign genomic regions specific to the donor or to the recipient genomes. The parentage assignment will enable *(i)* the comparison between WGS datasets from final hybrid apomictic backcrosses (optimally at least

BC_5F_1 assuming a mendelian inheritance pattern, but we decided conclude with the BC_3F_1 generation due to the enormous amount of work involved in the backcrosses and the upcoming genetic constrains due to non-mendelian inheritance patterns) with the WGS datasets from parental lines to identify sequence regions belonging to the apomictic paternal male donors, and *(ii)* the cross-comparison of those regions across different species to further identify and select for species- or cross-specific versus apomixis-specific male sequence fragments.

High molecular weight (HMW) DNA enables the sequence of very long DNA sequence reads which is crucial for the *de novo* assembly of whole complex genomes. We successfully used a modified protocol of a sedimentation-based HMW DNA extraction method (Wizard® HMW DNA Extraction Kit; Promega, Cat#: A2920) to extract DNA with 50-300+kb ranges from all native parental donors and recipients of our backcrossing assay and *Diplotaxis erucoides*. For each sample 150mg fresh leaf tissue was frozen in liquid nitrogen and ground into a fine powder using a mortar and pestle. The leaf powder was transferred into a 1.5ml LoBind microcentrifuge tube submerged in 500µl of HMW Lysis Buffer A and incubated at 65°C for 30 minutes with occasional inversion of the tubes. RNA was digested by incubation with 6µl of RNase A solution at 37°C for 30 minutes and remaining proteins were digested by incubation of the sample with additional 20µl of Proteinase K Solution at 56°C for 30 minutes. The sample was chilled on ice 1 minute and centrifuged at 13,000–16,000 × g for 3 minutes at room temperature to pellet any insoluble material. Then the lysate was transferred to a clean 1.5ml LoBind microcentrifuge tube. This step was repeated twice before the sample was incubated with additional 200µl of Protein Precipitation Solution on ice for 15 minutes. The precipitate was pelleted by centrifugation at 13,000–16,000 × g for 10 minutes at room temperature. The supernatant was decanted into a clean 1.5ml LoBind microcentrifuge tube and the centrifugation step was repeated three times. Then the supernatant was slowly decanted into a clean 1.5ml microcentrifuge tube containing 600µl of room-temperature isopropanol to precipitated high molecular weight DNA. The sample was inverted and incubated for 1 minute before the DNA was pelleted by 2 minutes of centrifugation at 13,000–16,000 × g in room temperature. Subsequently, the supernatant was discarded and the pellet was washed twice with 600µl 70% ethanol at room temperature. Finally, the ethanol was aspirated for 15 minutes, and the pellet was resuspended in 70µl of 10 mM TRIS (pH=8.0-8.4; e.g. Qiagen EB Buffer) and rehydrated by incubation at 65°C for one hour with periodical mixing. The HMW DNA was quantified with Qubit Fluorometric Quantitation (ThermoFisher) and the quality was spectrophotometrically assessed (NanoDrop™ One, ThermoFisher) prior to storage at 2–8°C until further usage. Occasionally samples were purified using the Genomic DNA Clean & Concentrator-10 kit (Zymo Research, Cat.-#: D4011).

Genotyping of parental plants and hybrid progeny for apomixis proxy markers

A PCR-based analysis for the presence/absence of the candidate marker gene either for male (*UPGRADE2*, [[11]]) or female apomeiosis (*APOLLO*, [[12]]) in combination with an internal reference (*ACTIN2*) was conducted for all crossing parents and backcrossing progeny plants according to Mau *et al*. [[33]] with minor changes. PCR primers for a 645 bp fragment of the gene *UPGRADE2* ("PC1pol1-L": 5'-CTTTTCCGTTGACTTTCCGACAAAT-3' and "PC1pol1-R": 5'-TCGATCAATCTCATTCGGGATCTAT-3') and of a 234 bp fragment spanning the apomixis-specific 5'-UTR polymorphism of the gene *APOLLO* ("Lara5-F": 5'-CCTCATCGTACCGTTGCTTCTCTC-3' and "TSP1-R": 5'-GATAGCCCCAAACTCCAAAATCGC-3') were used, in addition to primers for a 132 bp fragment of the housekeeping gene *ACTIN2* which served as internal template control ("RTAct2T7-L": 5'-GTTCCACCACTGAGCACAATGTTACC-3' and "RTAct2T7-R": 5'-AGTCTTGTTCCAGCCCTCTTTTGTG-3'). Marker-assisted selection was also used for the validation of intergeneric hybrids. Therefore, absence/presence (nuclear RNA polymerase D1B (TAIR #: AT2G40030, *Boechera* gene #: BA4G0233000F-GTGATGCATGGAGGGTTGTGAT/BA4G0233000R-AGAGTTGCTTCTGTGAATGGAGC) and insertion-deletion (InDel) markers (CSC1-like protein (TAIR#: AT3G54510, *Boechera* gene #: BA5G0414900F-TGTTGTACGATGGGAAAGATCTTGA/BA5G0414900R-

ACATCTCTTGGTTCAGGAGCCA, and Plastid division protein (TAIR #: AT3G19180, *Boechera* gene #: BA3G0307900F-TGGGTTGCACTGTGTTCTTCTC/BA3G0307900R-AGCAAGGTTTCTTCTGAAGTTCCC) between the apomictic donor *Boechera* and the sexual recipient crops (*Brassica napus, B. rapa, Diplotaxis erucoides*) were developed which are genic markers located on chromosome #3, #4 and #5 in *Boechera*. The orthologous primers from *Boechera stricta* ES672 (reference genome) and *Brassica napus* were used to test the presence of 49 apomixis candidate factors from ADF #20150123 (see **Table 9**) in intergeneric hybrid progeny by multiplex PCR reactions.

Multiplex PCR reactions were performed in a volume of 10 µL, using 2 µl of ten-fold diluted sample DNA, 2.5 µM of each primer, and the two-fold Phusion U Multiplex PCR Mastermix (Thermo Fisher Scientific, Cat# F562L, Waltham, MA, USA). The amplifications were run on a Mastercycler EP Gradient S (Eppendorf, Hamburg, Germany) under the following conditions: 3 min initial denaturation at 98°C; 40 cycles of amplification with 10 s at 98°C, 30 s at 54°C, and 1 min at 72°C; and 10 min of final elongation at 72°C. Depending on the amplicon length each PCR success was verified with either 1.5% or 3.5% agarose gel electrophoresis.

High-throughput flow cytometric seed screen and measurement of the relative nuclear DNA content

A previously developed high-throughput flow cytometric seed screens (FCSS) protocol [[28]] was applied to measure the mode of reproduction (*i.e.* genome ratio between the embryo and the endosperm) from all parental plants and the F_1 progeny by testing 24 seeds per individual. An external diploid sexual control was used per 96 well plate (*B. stricta*, ES854). For FCSS samples a single seed was placed into each well of a 96-well plate (Masterblock 0.5 ml 96-Well Deep Well Microplates, Greiner Bio-One, Cat#82051-472, VWR, Radnor, PA, USA) and a bottom-covering amount of 0.1mm silica powder (Zirconia Silica, Cat# 11079101z, BioSpec Products, Bartlesville, OK, USA) per well was added in addition to a 4mm borosilicate bead (Cat# 26396-633, VWR, Radnor, PA, USA). Plates were sealed with a foil (Peel Seal Cat# KBS-0601-001, LGC Genomics, Teddington, UK) using a heat sealer at 168°C for 3 seconds (ALPS 50 V, Cat# AB-1443A, Thermo Fisher Scientific, Waltham, MA, USA) and homogenised twice for 15 seconds at 1700 rpm (2010 Geno/Grinder, SPEX SamplePrep, Metuchen, NJ, USA) with intercalating centrifugation for one minute at 1260 rcf. Then, 210 µl nuclei extraction buffer (CyStain UV precise P automate nuclei extraction and DNA staining kit, Cat# 05-5002-a, Sysmex, Kobe, Japan) was added and mixed by inverting the plate 4-6 times. 200 µl of the suspension were filtered through a 96-well filter plate (AcroPrep Advanced 350, 30-40 µm PP/PE, Cat# 8027, PALL Corporation, New York, USA) into a new 96-well elution plate. A ratio of 50 µl nuclei suspension and 200µl DAPI staining suspension (CyStain UV precise P automate nuclei extraction and DNA staining kit, Cat# 05-5002-a, Sysmex, Kobe, Japan) was pipetted into the final 96-well measurement plate (96-Well Microplates, Greiner Bio-One, Cat#82050-636, VWR, Radnor, PA, USA) and placed at 4°C for 30 minutes for saturation of the tissue with the stain before measurement.

The relative nuclear DNA content (*i.e.* "ploidy") per leaf sample from all parental and progeny plants was quantified in reference to a diploid internal control (*Pisum sativum,* 8.84 pg/2C genome size; [34]). Each frozen leaf sample was supplied with a 4 mm borosilicate bead. The 96-well plate was sealed with a foil using a heat sealer and maintained in a -30°C freezer until further usage. Frozen leaf tissue was homogenized for five seconds at 1750 rpm using a 2010 Geno/Grinder and centrifuged for 60 seconds at 4000 rpm to collect the leaf powder at the bottom of the well. Subsequently, the tissue powder was suspended in 20 µl CyStain UV precise P automate nuclei extraction buffer and the suspension was mixed using the GenoGrinder for five seconds at 1750 rpm and spun down. Each sample was supplemented with 100 µl of the DAPI staining suspension. The plate was re-sealed, the stain was mixed into the sample by inverting the plate several times. The sample plate was centrifuged, and the suspension was filtered through the 96-well filter plate into a new 96-well elution plate for flow cytometric measurement.

All ploidy and FCSS measurements were performed on a CytoFLEX flow cytometer with plate loader (BeckmanCoulter, Indianapolis, IN, USA) equipped with a 405 nm (Violet) laser. Data was measured and analyzed using the CytExpert Software v2.3 (BeckmanCoulter). Throughout the manuscript "n" refers to DNA content of a haploid anaphase cell, and "x" to the basic chromosome number.

Pollen tube growth in situ analysis

In situ pollen tube germination for wide intergeneric cross combinations between native diploid unbalanced apomictic *Boechera* donors and various crop recipients (*Diplotaxis erucoides*, *Brassica napus* or *Brassica rapa*) was measured (Lu, 2011) and scored (Kuligowska et al., 2015) with minor changes. Carpels were collected 48 h/76 h and 96 h after pollination and were fixed in a 3:1 solution of absolute EtOH and glacial acetic acid for 3 h and washed for 10 minutes with 70% (v/v) EtOH for two times. Pistils were stored at 4°C until the day of the use. In the day of the use, pistils were softened in 1M NaOH for 10 - 15min at 55°C and washed two times with 50 mM KPO4 buffer (4.17 ml 1M K2HPO4, 0.83 ml 1M KH2PO4, 995 ml H2O, pH 7.5). Then, samples were stained for 1 – 10 min with toluidine blue 0.01% (w/v) in 50 mM KPO4 pH 7.5 buffer and washed with 50 mM KPO4 buffer. Then, pistils were stained for one hour with 0.05% (w/v) aniline blue in 50 mM KPO4 pH 8.5 buffer (4.17 ml 1M K2HPO4, 0.83 ml 1M KH2PO4, 995 ml H2O, pH 8.5) in the dark covered with foil. Softened and stained samples were placed on microscopy slides (ThermoFisherScientific, #12-552, 25 x 75 x 1mm) with 100 µl Prolong Antifade Diamond 10 buffer and squash under cover glass (Marienfeld, #5250, 18 x 18mm or larger). Microscope slides can be stored at 4° until examination. Pollen tubes were examined under the fluorescent microscope (excitation filter BP 340–360 nm) equipped with a digital camera. The pollen grains on stigma were calculated and categorized to proof the pollination event according to the criteria: 1- up to 10 pollen grains on stigma, 2- 10-20 pollen grains on stigma, 3- 20-100 pollen grains on stigma, 4- more than hundred pollen grains on stigma. Pollen tube growth was evaluated for pistils with similar pollination ranks and five categories were scored; 0- no pollen tubes formed 1- up to 10 pollen tubes formed, 2- 10-20 pollen tubes formed, 3- 20-100 pollen tubes formed, 4- more than hundred pollen tubes formed. Then the pollen tube growth was evaluated at three positions: 1- pollen grains germinated on stigma, 2- pollen tubes visible at half the length of the style, 3- pollen tubes in the ovary reaching ovules.

Whole genome sequencing

Due to the recurrent hybrid nature of *Boechera* with large structural genome variation and regions with high levels of repeats our sequencing strategy involves the preparation of three different whole genome sequence libraries and their combination into a single de novo genome sequence assembly data pipeline. We developed three library types from 10 parental *Boechera* genomes used in the backcrossing assay and *Diplotaxis erucoides*: (1) high molecular weight (HMW) DNA for the generation of ultra-long read sequence libraries using the Oxford Nanopore Technologies (ONT) sequencing platform, (2) highly pure DNA for the generation of a high-coverage short-read Illumina sequence libraries, and (3) nuclei samples with crosslinked DNA for the generation of Hi-C sequence libraries. We employed the genome center at UC Davis (https://dnatech.genomecenter.ucdavis.edu/) for preparing & sequencing libraries with Illumina short-read sequencer Illumina NovaSeq 6000 (Nova0453) and the long-read libraries with the Oxford Nanopore PromethION platform. Illumina short read libraries were prepared using the QIAseq X DNA library kit (Qiagen) from samples with >1ug in 50ul EB buffer (i.e. 20ng/ul; min. 100ng total) with 280/260 purity between 1.8-2.0 and 260/230 purity of >2.0. Illumina NovaSeq 6000 produced 150bp paired-end read libraries aiming for a 50-fold genome coverage. DNA samples for long-read sequencing using the ONT were stored in 10 mM TRIS (pH=8.0-8.4; e.g. Qiagen EB buffer) and tested for their quality (280/260 purity between 1.8-2.0 and 260/230 purity of 2.0-2.2) and length using the Femto Pulse System (Agilent) prior to library preparation. The ONT sequencing libraries were prepared using a ligation-based sequencing kit (SQK-

LSK109; https://store.nanoporetech.com/ligation-sequencing-kit.html) for multiplexing the 11 samples into three pools of each three samples and a single pool of two samples. The library preparation method comprises the repair of DNA ends and dA-tailing of the sequence strands using the NEB Next End Repair/dA-tailing module, and then sequencing adapters, supplied in the kit, are ligated onto the prepared ends. Per pool 1000ng (pool1) and 500ng (pools 2-4) DNA was loaded into a different flow cell (FLO-PRO002), respectively. Per flow cell a nuclease flush, library reloads and refuelled was performed to guarantee best outcomes. Sequencing runs lasted 72 hours per pool and were performed using the 'Super-accurate basecalling' model with a minimum read filtering quality score of 10. The use of Hi-C sequencing libraries allows the detection of long-range DNA interactions which enables in combination with other sequencing methods (e.g. long-read libraries, deep short read libraries) the assembly of chromosome-scale genomes. In this method, DNA-protein complexes per 2g of leave sample were crosslinked with formaldehyde and nuclei were isolated using the CelLytic Plant Nuclei Extraction Kit (Sigma, Cat# CELLYTPN1). DNA from aliquoted nuclei samples were fragmented, and the DNA was extracted, unique dual index adapters were ligated using the 2S Turbo Unique Dual Indexing Primer Kit (Swift Biosciences) and digested with restriction enzymes to an average length of 828bps at NRC Saskatoon. The resulting DNA fragments were PCR-amplified and sequenced on an Illumina NovaSeq 6000 which produced 150bp paired-end read libraries aiming for a 50-fold genome coverage.

Bioinformatic analyses

Various analyses pipelines were developed or adapted to assembly, annotate, map, and compare whole genomes and identify single nucleotide polymorphisms (SNPs) and insertion-deletion sites (InDels) that are specific to apomictic *Boechera* and were either integrated into sexual *Boechera* (i.e. in the backcrosses) or into *Brassica* crops (i.e. intergeneric F1 hybrids between *Boechera* and canola). Due to space constraints only two examples are provided here:

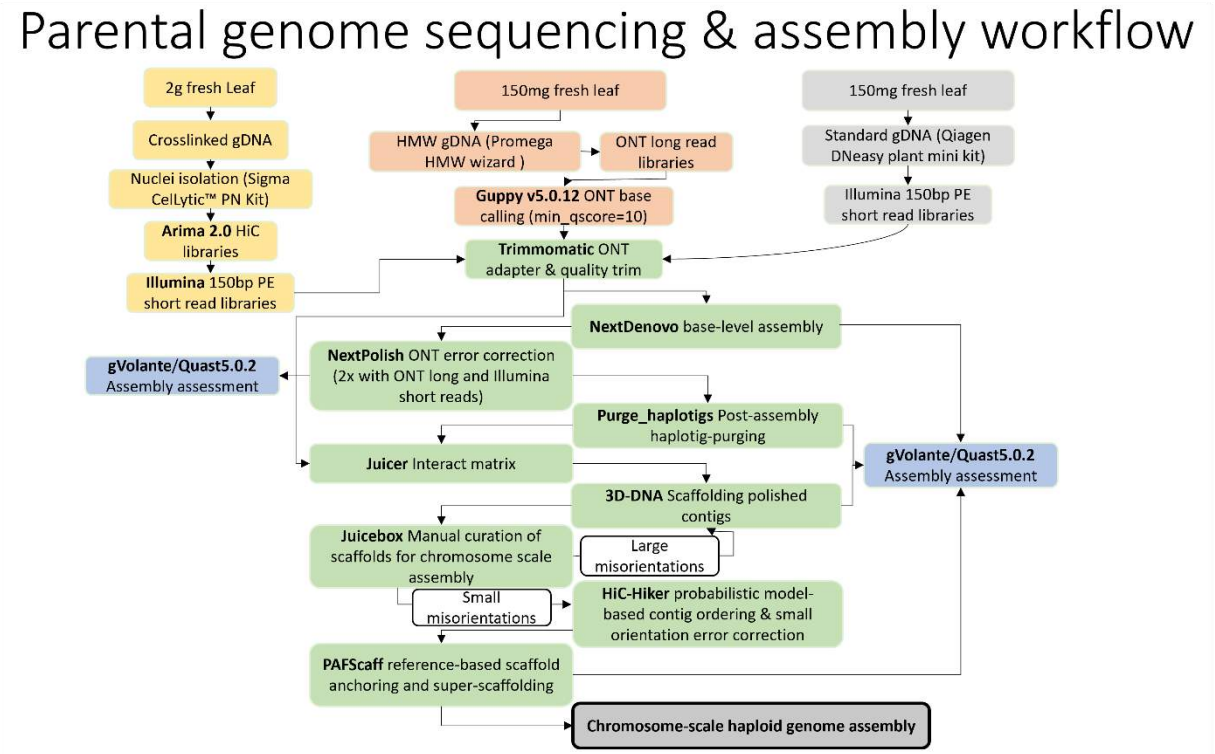

Figure 1. Custom bioinformatic analysis pipeline for chromosome-level haploid genome assembly of 6 sexual recipient and 4 apomictic donor lines from *Boechera sp.* used in the backcrossing array.

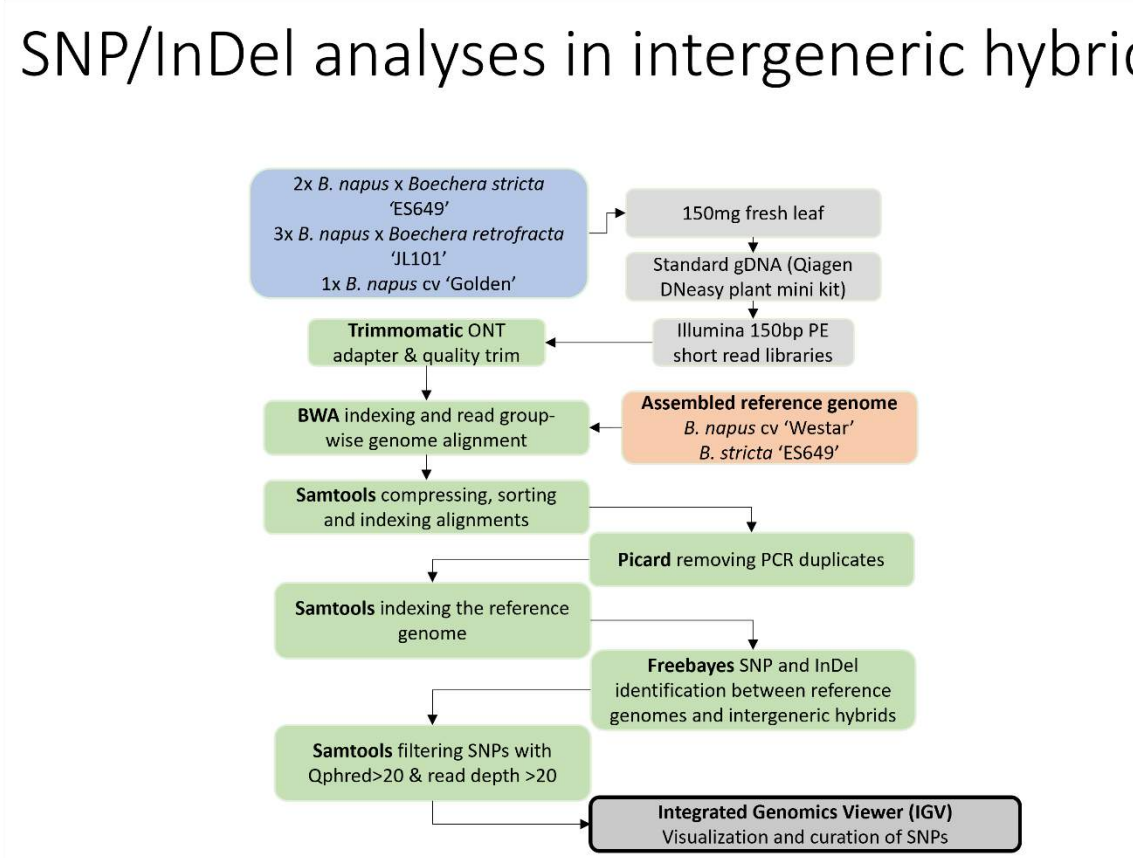

Figure 2. Custom bioinformatic analysis pipeline for single nucleotide polymorphism (SNP) and insertion-deletion (InDel) site identification between reference genomes from *Brassica napus* cv 'Westar' and *Boechera stricta* 'ES649' and the genomes of intergeneric F1 hybrids.

Seed metabolic activity analysis and Embryo rescue protocol

Reduced germination rates can be caused by seed dormancy but also by other factors, such as no or reduced metabolic activity of the seeds leading to their inviability. We observed higher (partial) seed inviabilities in our intergeneric F1 and BC hybrid seeds and analysed the seed metabolic activity to test for seed viability frequencies in our hybrid seeds using a modified protocol according [35]. In brief, 20 non-germinated seeds per test group were soaked in distilled water or 1:4 (v/v) sodium hypochloride for 20 minutes to soften the seed coat. Seeds were washed twice with distilled water and bisected longitudinally to expose the embryo. The seeds were then suspended in a 0.2% solution of 2,3,5 triphenyl tetrazolium chloride and were incubated overnight (24h) at room temperature. Consequently, the seed were then triple-rinsed and stored in distilled water until further use. For analysis the embryos were dissected from seed coat using fine (insuline) syringes and transferred to and empty petri dish placed on a white background under a stereo microscope. The number of stained embryos and the number of total embryos was counted under a stereo microscope and their ratio was used to calculate the frequency of seed viability.

Intergeneric F1 hybrid that could not be propagated directly on soil were subject of the following embryo rescue procedure. Developing grains 14 DAP or older were surface sterilized with 5ml 1% Sodium Hypochlorite solution for 10 minutes in a 15ml Falcon tube under occasional mixing by inversion several times. Sterilized grains were rinsed three times with sterile MilliQ water. Embryos were dissected from the grains under aseptic conditions

and placed on either medium #1 (embryo culture medium (ECM) without hormones: 1L MilliQ water, 4g MS medium, 6% sucrose, 2mg Kinetin, 10 mL NaH2PO4H2O, 10 mL i-inositol, 10 mL Thiamine, 10 mL L-tyrosine, 10 mL L-arginine, 10 mL Glycine, 5.2 g of Phytagel, adjusted to pH 5.7 with NaOH), #2 (ECM with 0.2mg/L 6-Benzylaminopurine (BAP)), #3 (ECM with 1.5mg/L BAP and 0.25mg/L 1-Naphthaleneacetic acid (NAA)), or #4 (ECM with 0.1mg/L NAA). Depending on the age at which the plantlets will be removed from the plate as many as 20-30 embryos can be placed on a single plate with the radical on the media. The plates were wrapped twice around with Parafilm and kept under long-day condition (16h/8h light/dark) between 23 ± 2 ºC until plantlet formation. The plantlets were transferred either directly to soil if roots are properly formed or to medium #3 for plant regeneration for approx. 10 days and if roots still haven't formed to special rooting medium #4 until roots have formed.

9. **Results and discussion:** *Describe and discuss the results accomplished during the entire project period under each objective listed under section 7. The results need to be accompanied with pertinent tables, figures and/or other illustrations. Provide discussion necessary to the full understanding of the results. Where applicable, results should be discussed in the context of existing knowledge and relevant literature. Detail any major concerns or project setbacks. (**Maximum of 30 pages of text not including figures or tables**)*

a) *Generation of diploid, hybrid unbalanced apomictic Boechera backcrosses*

The repeated backcrossing of diploid, hybrid unbalanced apomictic *Boechera* into their sexual mother will stepwise enrich the sexual mother genome while minimizing paternal gene-space in the backcrosses. The goal is to use these backcrosses for the simultaneous discovery of all genetic factors that are sufficient to trigger the apomixis trait via a genetic comparison of multiple parental and hybrid backcross lines. In backcrosses which reproduce purely sexually one can assume mendelian inheritance of traits and a BC_5F_1 line would contain 1.56% of the paternal genome selecting for a single dominant allele and 14.57% selecting for ten dominant alleles. However, apomixis is a non-mendelian trait and the rate at which genes are entering a cross from the donor parent can be influenced by effects from, (1) segregation distortion and suppressed recombination [[14]], (2) epistatic interaction among loci [[36]], and (3) genomic imprinting [[37]] (see **Fig. 1B**). Considering this, a sufficient reduction of the paternal genome contribution in the final backcrosses generation might be impeded by these effects. This fact and other constraints (*i.e*. COVID19 related impedance of greenhouse activities, higher than anticipated labour for some BC lines, moving lab from GIFS to UofS plant sciences BC_5F_1 until end of the project due to seed-to-seed periods of 6 months) led to our decision to reduce the number of backcross generations from BC_5F_1 to BC_3F_1. In addition, the employment of multiple independent backcross lines of inter- and intraspecific origin (**Table 1**) in our genetic comparison will statistically correct for genetic-mediated background noise and facilitate the selection of apomixis-specific regions.

Our 4-year pilot study of the transfer and inheritance of apomixis in the genus *Boechera* revealed the segregation of three apomixis elements, male and female apomeiosis, parthenogenesis and pseudogamy, suggesting their independent genetic control [[28]]. Hence, we can assume a minimum number of 3 to 4 genetic factors although the absolute number of genes triggering the apomixis trait remains unknown. At least, the genetic factors for male (*UPGRADE2*) and female apomeiosis (*APOLLO*) are not genetically linked and segregate in the haploid pollen population of unbalanced apomicts [[28]]. In addition, both apomeiosis marker genes reside in hemizygous regions (*APOLLO*: [[12]], *UPGRADE2*: [[38]]).

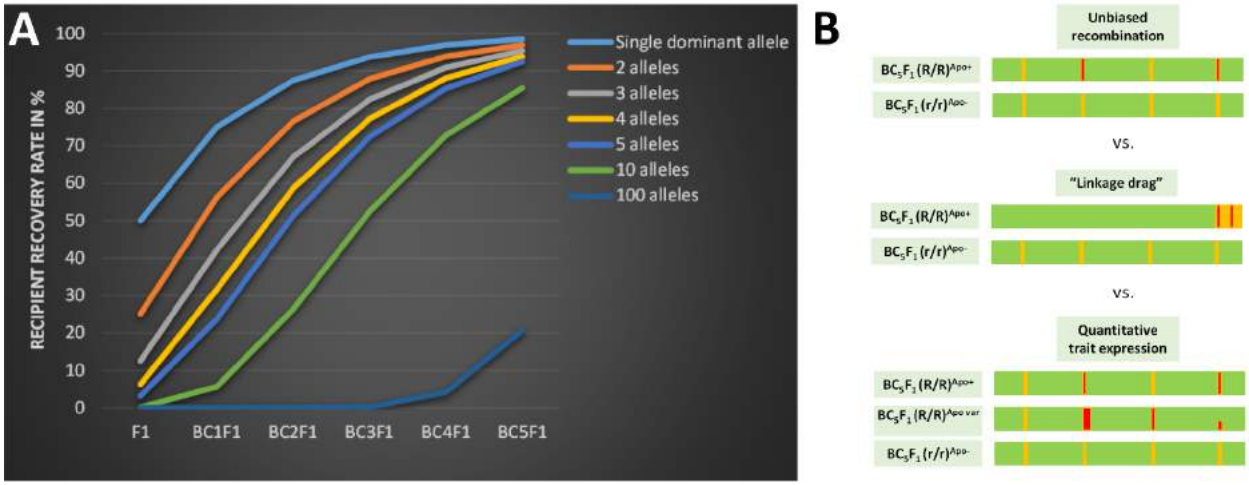

Figure 3. Schematic of the recovery rate of dominant alleles in backcrosses assuming mendelian principles (**A**), and potential deviations from mendelian principles in backcrosses with apomicts (**B**).

Here we utilize 12 newly generated diploid unbalanced apomictic hybrid *Boechera* lines from our 4-year pilot study which produce viable haploid pollen (**Fig. 3** and **Table 1**; see [[28]]) for backcrosses into their diploid sexual *Boechera* mother lines. We aim for the enrichment of the sexual recipient mother genome while selecting for the apomictic phenotype of the apomictic donors. Therefore, we select for diploid unbalanced apomictic *Boechera* lines which produce haploid pollen (*i.e.* vector for apomixis-inducing genetic factors, 2C:5C seed ploidy) in each generated backcrossing generation and discard all polyploid, aneuploid, sexual or diploid balanced apomictic (*i.e.* diploid pollen producers) progeny by using flow cytometry of leaf and seed material in addition to a PCR-based screen for proxy markers for female (*APOLLO*, [[12]]) and male apomeiosis (*UPGRADE2*, [[11]]; **Fig. 4**).

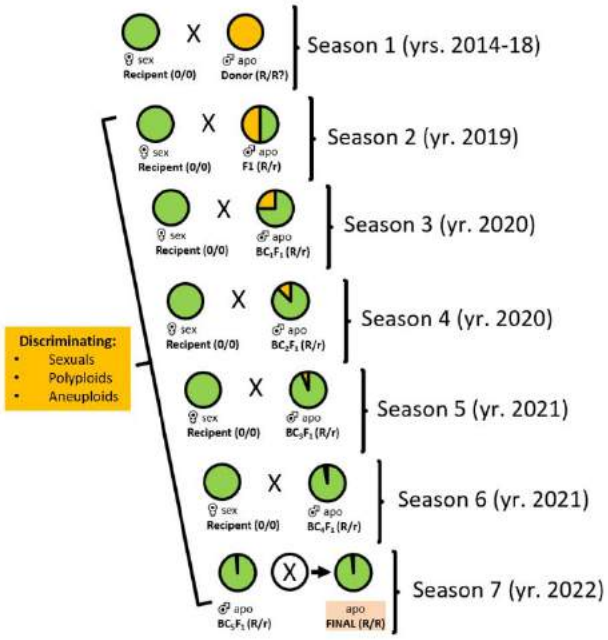

Figure 4. Schematic of the backcross procedure of previously generated diploid unbalanced apomictic hybrid *Boechera* donor lines into diploid sexual *Boechera* recipient lines.

Over the total project time we have sown 5950 diploid sexual recipient individuals (*i.e.* 2C:3C seed producers) and 13741 diploid unbalanced apomictic hybrid donor lines (F1/F2 to BC4F1 2C:5C seed producers, **Table 2**). The germination rates varied between 22.7% and 55.5% for the recipient lines and between 0% and 98% for the donor lines (**Table 2**).

Hitherto, we performed 3290 first generation backcross pollinations (BC1) of all diploid sexual recipient lines with diploid unbalanced apomictic F_1 or F_2 hybrid donor lines which produced several thousand seeds (**Table 2**). We have sown 2993 BC1F1 plants from all 12 backcross lines and have performed 3158 second generation backcross pollinations (BC2). We also have sown 4124 BC2F1 plants from 11 of 12 backcross lines and performed 1974 third generation backcross pollinations (BC3). Over 1954 BC3 plants from 8 of 12 lines were planted of which 270 plants survived (so far, the rest is still in the growth & testing phase) for geno- and phenotyping and skim-sequencing selected BC3F1 diploid unbalanced apomictic haploid pollen producers. A single line (JL12x(JL12xJL103)) was sown for the use in BC5 cross pollinations. Currently, the remaining BC2 (1 of 12 NILs) cross pollinations and BC3 selfings (4 of 12) are concluded.

We tested parental plants and their progeny for presence/absence of proxy markers for female (*APOLLO*, [[12]]) and male apomeiosis (*UPGRADE2*, [[11]]; **Fig. 5** and **Table 3**). Only individuals from sexual lines were the apoallele of *APOLLO* is absent were used as sexual recipients, whereby only individuals from hybrid unbalanced apomictic lines with presence of the apoallele of *APOLLO* gene were considered as potential donors until flow cytometric analyses of the leaf ploidy and the seed set proved the opposite. We tested each potential recipient individual across all six recipient lines using our proxy markers and revealed the sexual mode of reproduction with high frequency (194 individuals tested, 5 negatives: 97.42% sexual). We selected only those recipient individuals without the presence of the *APOLLO* marker. In contrast, from 1034 tested F1/F2 hybrids 82.8% were apomictic. The apomixis frequency substantially decreased in all lines at BCF1 generation (1220 plants tested, 33.4% apomixis frequency, min.: 10%, max.: 66.7%) and stabilized at a similarly reduced levels for the BC2F1 (1169 plants tested, 27.3% apomixis frequency, min.: 9.3%, max.: 51.6%), BC3F1 (232 plants tested, 38.4% apomixis frequency, min.: 36.4%, max.: 40.4%) and BC4F1 generations (112 plants tested, 37.5% apomixis frequency; **Table 3**).

The flow cytometric analysis of leaf tissue and seeds (FCSS) was used to detect the frequency of diploid unbalanced apomixis (*i.e.* our selected phenotype) in all donors and generated backcrosses. Only those hybrid and backcrosses progeny were selected as potential donors for the next backcross generation that displayed a diploid nuclear DNA content of leaf material, and had a genome ratio of the embryo and endosperm in their produced seeds of 2C:5C (*i.e.* individuals which produce more than 50% diploid, unbalanced apomictic seeds: 2C:5C seed ploidy). We tested all six sexual recipient lines and detected that ES854 (3 seed vouchers), JL12 (1 seed vouchers), JL76 (1 seed vouchers), JL78 (1 seed vouchers) contained 100% sexual germplasm. Only for recipient lines ES612 (2 of 5 seed vouchers) and ES865 (1 of 2 seed vouchers) we detected low level of apomixis frequencies in the germplasm in a subset of tested seed vouchers. Only seed vouchers with 100% sexual seed profiles were used for subsequent assays.

We screened 5656 seeds produced by 237 F1 and F2 hybrids (**Table 3**) to select individuals which produced the majority (>50%) of their seeds via diploid, unbalanced apomixis (*i.e.* 2C:5C seed ploidy) which proofs that the phenotype of their apomictic father line is inherited (*i.e.* diploid, unbalanced apomict producing haploid pollen). An average of 63.9% of all tested seeds across 12 hybrid lines displayed the diploid unbalanced apomictic seed phenotype (min: 23.8%, max.: 89.8%). In addition, we tested 4971 seeds from 12 of 12 BC1F1 lines (average 60.5% 2C:5C seed frequency, 212 plants), 1745 seeds from 7 of 12 BC2F1 lines (average 47.8% 2C:5C seed

frequency, 74 plants), and 583 seeds from 2 of 12 BC3F1 lines (average 37.2% 2C:5C seed frequency, 21 plants). We demonstrate a relative stable inheritance of our selected seed phenotype (*i.e.* 2C:5C seed ploidy) from the F1/F2 to the first backcross generation with comparable frequencies (63.9% vs 60.5%, respectively), but found a step-wise reduced inheritance throughout BC2 to BC4 (47.8 vs. 35.9 vs. 37.2, respectively).

We have performed flow cytometric analyses of the nuclear DNA content of leaves (*i.e.* leaf ploidy) from 1753 individuals of all six sexual recipient lines (1487 diploids; with 21 polyploid and 11 mixoploid individuals removed from analysis), 1026 individuals of seven unbalanced apomictic hybrid donor lines (19 polyploid individuals removed from analysis), 1225 BC1F1 individuals of 12/12 backcrossing lines (23 polyploid and a single mixoploid individuals removed from analysis), 949 BC2F1 individuals of 10/12 backcrossing lines (83 polyploid and a single mixoploid individuals removed from analysis), 170 BC3F1 individuals of 2/12 backcrossing lines (all diploid individuals), and 114 BC4F1 individuals of one backcrossing line (one triploid individual) to select for diploid individuals and remove all polyploids from the backcross scheme.

Overall, our monitoring system of leaf ploidy, absence/presence of apomeiosis markers and seed ploidy in parental and hybrid backcross progeny is continuously functional across all backcross generations. Nonetheless, the selection of appropriate donor plants remains a laborious two-step process. First, all new plants are screened for the absence/presence of both apomeiosis markers and leaf ploidy and those that are diploid and have the *APOLLO* marker present will be pre-selected as donor lines and then the crosses are performed. Secondly, all backcrosses are finally categorized by the flow cytometric seed screen which happens usually after the pollination process when the donor plants form seeds. This means that in the first instance often crosses are performed with partially confirmed donor lines of which a fraction is discarded later which is a necessary procedural flaw that produces more crosses than utilized but cannot be avoided.

Importantly, the phenotype and genotype from diploid unbalanced apomictic donor lines can be transmitted through backcrosses into the subsequent generations although their frequencies decreases across generations, which is likely due to the segregation of apomixis elements in hybrids and outcrosses [[39]].

Government of Saskatchewan
Ministry of Agriculture

ADF Project Progress Report Format

Table 1. Native and synthetic *Boechera* accession used as parents in backcrosses. Whole genome sequencing methods will be applied to these samples and *Diplotaxis erucoides* to generate fully assembled and annotated reference genomes for comparative genome analyses with selected backcrosses.

Accession	Taxon	Contribution	Mode of reproduction	Origin
ES612	*B. stricta*	recipient	sexual	native
ES854	*B. stricta*	recipient	sexual	native
ES865	*B. stricta*	recipient	sexual	native
JL12	*B. crandallii*	recipient	sexual	native
JL76	*B. spatifolia/stricta x spatifolia*	recipient	sexual	native
JL78	*B. spatifolia/stricta x spatifolia*	recipient	sexual	native
ES612xES649	*B. stricta x stricta*	donor	unbalanced apomictic	synthetic intraspecific outcross
ES612xJL106	*B. stricta* x *stricta/spatifolia*	donor	unbalanced apomictic	synthetic interspecifc hybrid
ES854xES649	*B. stricta x stricta*	donor	unbalanced apomictic	syntheticintraspecific outcross
ES854xJL103	*B. stricta x crandallii*	donor	unbalanced apomictic	synthetic interspecifc hybrid
ES854xJL107	*B. stricta x stricta*	donor	unbalanced apomictic	synthetic intraspecific outcross
ES865xES649	*B. stricta x stricta*	donor	unbalanced apomictic	synthetic intraspecific outcross
ES865xJL103	*B. stricta x crandallii*	donor	unbalanced apomictic	synthetic interspecifc hybrid
ES865xJL106	*B. stricta x stricta/spatifolia*	donor	unbalanced apomictic	synthetic interspecifc hybrid
ES865xJL107	*B. stricta x stricta*	donor	unbalanced apomictic	synthetic intraspecific outcross
JL12xJL103	*B. crandallii x crandallii*	donor	unbalanced apomictic	synthetic intraspecific outcross
JL76xJL106	*B. stricta/spatifolia* x *stricta/spatifolia*	donor	unbalanced apomictic	synthetic intraspecific outcross
JL78xJL106	*B. stricta/spatifolia* x *stricta/spatifolia*	donor	unbalanced apomictic	synthetic intraspecific outcross
ES649	*B. stricta*	donor	unbalanced apomictic	native
JL106	*B. stricta x spatifolia*	donor	unbalanced apomictic	native
JL103	*B. crandallii*	donor	unbalanced apomictic	native
JL107	*B. stricta*	donor	unbalanced apomictic	native

Table 2. Overview of the performed backcrosses, the sown recipient and donor plants and generated progeny of 12 independent diploid unbalanced apomictic *Boechera* backcross lines in the years 2019 to 2023 (total on 2023-06-25).

						F1/F2 for BC1								BC1F1 for BC2								BC2F1 for BC3								BC3F1 for BC4								BC4F1 for BC5							
Recipient x donor taxa	Apomicic donor line	Sexual recipient line	# Recipients sown	# Recipient survivors	% Germination rate	# Donors sown	# Donor survivors	% Germination rate	# Inflorescences emasculated	# Pistils pollinated	# Siliques harvested	# Seeds harvested	Fecundity	# Donors sown	# Donor survivors	% Germination rate	# Inflorescences emasculated	# Pistils pollinated	# Siliques harvested	# Seeds harvested	Fecundity	# Donors sown	# Donor survivors	% Germination rate	# Inflorescences emasculated	# Pistils pollinated	# Siliques harvested	# Seeds harvested	Fecundity	# Donors sown	# Donor survivors	% Germination rate	# Inflorescences emasculated	# Pistils pollinated	# Siliques harvested	# Seeds harvested	Fecundity	# Donors sown	# Donor survivors	% Germination rate	# Inflorescences emasculated	# Pistils pollinated	# Siliques harvested	# Seeds harvested	Fecundity
B. crandallii x B. crandallii	JL12xJL103	JL12	560	274	48.9	120	112	93.3	115	116	0	0		224	188	83.9	448	246	0	0		150	147	98.0	1355	287	73	1333	18.3	508	178	35.0	392	69	25	515	21	195	114						
B. stricta x B. stricta	ES612xES649	ES612	1120	363	32.4	90	0	0.0	571	201	128	3440	26.9	405	219	54.1	1688	293	199	4277	21.5	505	71	14.1	816	128	109	1507	13.8	264	92	34.8													
B. stricta x B. stricta	ES612xJL106	ES612				94	72	76.6	828	293	195	384	2.0	96	0	0.0	959	238	179	1108	6.2	370	100	27.0	420	120	76	2271	29.9	100															
B. stricta x B. stricta	ES854xES649	ES854				735	116	15.8	219	127	73	824	11.3	275	5	1.8	2621	218	90	1120	12.4	535	157	29.3	2949	345	89	1355	15.2	328															
B. stricta x B. stricta	ES854xJL103	ES854	1365	310	22.7	620	134	21.6	444	118	58	290	5.0	261	20	7.7	224	67	28	1011	36.1	399	125	31.3	1124	250	116	1409	12.1	343															
B. stricta x B. spatifolia	ES854xJL107	ES854				418	105	25.1	1521	290	118	657	5.6	469	98	20.9	1747	409	80	805	0.0	493	274	55.6	377	94	42	801	19.1	200															
B. spatifolia x (B. stricta x spatifolia)	JL76xJL106	JL76	660	366	55.5	166	87	52.4	159	146	54	87	1.6	134	50	37.3	1426	276	167	359	2.1	220	206	93.6	2417	336	162	557	3.44	152															
B. spatifolia x (B. stricta x spatifolia)	JL78xJL106	JL78	760	370	48.7	472	191	40.5	1795	666	298	468	1.6	301	106	35.2	1910	318	144	259	1.8	325	217	66.8	254	60	44	191	4.34																
B. stricta x B. stricta	ES865xES649	ES865				115	63	54.8	420	165	133	474	3.6	126	11	8.7	1022	273	135	1014	7.5	659	73	11.1	1492	354	155	75	0.48	59															
B. stricta x B. crandallii	ES865xJL103	ES865	1485	404	27.2	190	90	47.4	1556	420	233	367	1.6	287	38	13.2	1595	407	246	304	1.2	271	27	10.0																					
B. stricta x (B. stricta x spatifolia)	ES865xJL106	ES865				766	130	17.0	887	304	178	420	2.4	369	110	29.8	1295	290	143	104	1.0																								
B. stricta x B. spatifolia	ES865xJL107	ES865				689	82	11.9	1613	444	246	48	0.2	46	11	23.9	377	123	78	406	5.2	197	58	29.4																					
Total	**Total**		**5950**	**2087**	**39.2**	**4475**	**1182**	**38.0**	**10128**	**3290**	**1714**	**7459**	**5.6**	**2993**	**856**	**26.4**	**15312**	**3158**	**1489**	**10767**	**8.6**	**4124**	**1455**	**42.4**	**11204**	**1974**	**866**	**9499**	**13.0**	**1954**	**270**	**34.9**	**392**	**69**	**25**	**515**	**20.6**	**195**	**114**						

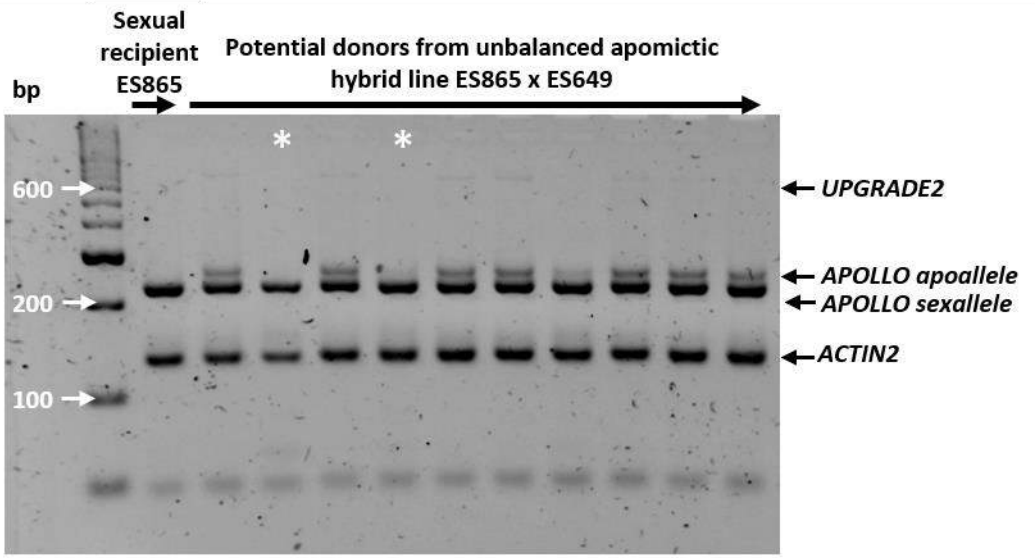

Figure 5. Examples of PCR-based genotyping of a diploid sexual recipient line *B. stricta* 'ES865' and multiple potential donor plants of the unbalanced apomictic hybrid line *B. stricta* 'ES865 x ES649'. The multiplex PCR primers target presence/absence of the apomixis proxy-markers for male (*UPGRADE2*) and female -apomeiosis (*APOLLO* apoallele) in addition to the internal control *ACTIN2*. All sexual lines, including the recipient line *B. stricta* 'ES865', typically harbor only *ACTIN2* and the sexallele of the *APOLLO* gene. All potential unbalanced apomictic donor plants should harbor ACTIN2 and the sex- and apoallele of the *APOLLO* gene. The *UPGRADE2* gene is segregating from the *APOLLO* gene and is therefore not the primary proxy marker for apomixis [[33]]. Therefore, only samples that lack the apoallele of the *APOLLO* gene and which therefore would be discarded (see asterisk).

ADF Project Progress Report Format

Table 3. PCR-based genotyping and flow cytometry-assisted phenotyping of hybrid unbalanced apomictic *Boechera* donor lines and their hybrid backcrosses with native sexual recipient *Boechera* accessions.The apomixis frequencies were calculated based on the transmission of the apo allele of the *APOLLO* gene (female apomeiosis-linked marker, Corral *et al.* 2013). We used flow cytometric seed screen of 24 seeds per individual to selected for diploid, unbalanced apomicts with haploid pollen formation (*i.e.* 2C:5C genome ratio of the embryo and endosperm in seeds). Data points missing in the table are currently analysed (status: 2023-06-25).

F1/F2

BC line	Genotype[a]						Phenotype[b]				
	# plants	# plants-n/a	# n/a	# *APO* pos	# *APO* neg	Apomixis frequency	# plants	# seeds	# seeds-n/a	# 2C:5C	% 2C:5C
ES612xJL106	73	71	2	60	11	84.5	12	279	270	212	78.5
JL12xJL103	99	98	1	77	21	78.6	57	1373	1268	558	44.0
ES612xES649	87	87	0	78	9	89.7	13	294	253	150	59.3
ES854xES649	80	79	1	64	13	81.0	18	420	405	205	50.6
ES854xJL103	88	86	2	70	16	81.4	23	534	530	476	89.8
ES854xJL107	60	59	1	55	4	93.2	6	140	129	92	71.3
ES865xES649	27	23	4	20	3	87.0	10	231	214	74	34.6
ES865xJL103	143	133	10	106	27	79.7	11	255	247	188	76.1
ES865xJL106	128	123	5	80	43	65.0	28	658	544	443	81.4
ES865xJL107	82	77	5	61	16	79.2	18	421	358	249	69.6
JL78xJL106	189	185	4	151	34	81.6	38	886	838	736	87.8
JL76xJL106	13	13	0	12	1	92.3	3	165	160	38	23.8
Total	**1069**	**1034**	**35**	**834**	**198**	**82.8**	**237**	**5656**	**5216**	**3421**	**63.9**

BC1F1

BC line	Genotype[a]						Phenotype[b]				
	# plants	# plants-n/a	# n/a	# *APO* pos	# *APO* neg	Apomixis frequency	# plants	# seeds	# seeds-n/a	# 2C:5C	% 2C:5C
ES612xJL106	45	43	2	23	20	53.5	16	380	324	205	63.3
JL12xJL103	166	155	11	73	82	47.1	42	975	918	181	19.7
ES612xES649	268	264	4	102	162	38.6	18	420	395	231	58.5
ES854xES649	144	143	1	44	99	30.8	54	1251	1002	537	53.6
ES854xJL103	20	15	5	10	5	66.7	10	238	125	99	79.2
ES854xJL107	94	88	6	23	65	26.1	12	279	209	67	32.1
ES865xES649	69	67	2	27	40	40.3	25	594	450	220	48.9
ES865xJL103	42	39	3	12	27	30.8	5	117	112	92	82.1
ES865xJL106	140	140	0	31	109	22.1	6	138	133	118	88.7
ES865xJL107	21	20	1	2	18	10.0	2	48	43	38	88.4
JL78xJL106	177	169	8	22	147	13.0	9	207	184	135	73.4
JL76xJL106	81	77	4	17	47	22.1	13	324	289	110	38.1
Total	**1267**	**1220**	**47**	**386**	**821**	**33.4**	**212**	**4971**	**4184**	**2033**	**60.5**

BC2F1

BC line	Genotype[a]						Phenotype[b]				
	# plants	# plants-n/a	# n/a	# *APO* pos	# *APO* neg	Apomixis frequency	# plants	# seeds	# seeds-n/a	# 2C:5C	% 2C:5C
ES612xJL106	237	237	0	22	136	9.3	5	117	70	33	47.1
JL12xJL103	64	64	0	33	31	51.6	30	696	643	101	15.7
ES612xES649	62	57	5	20	37	35.1	7	165	132	108	81.8
ES854xES649	162	114	48	21	89	18.4	8	188	186	123	66.1
ES854xJL103	78	78	0	32	46	41.0	9	231	173	103	59.5
ES854xJL107	274	250	24	59	191	23.6					
ES865xES649	101	101	0	26	75	25.7	11	255	239	126	52.7
ES865xJL103	27	0	27								
ES865xJL106											
ES865xJL107	58	0	58								
JL78xJL106	162	162	0	47	115	29.0					
JL76xJL106	106	106	0	13	93	12.3	4	93	87	10	11.5
Total	**1331**	**1169**	**162**	**273**	**813**	**27.3**	**74**	**1745**	**1530**	**604**	**47.8**

BC3F1

BC line	Genotype[a]						Phenotype[b]				
	# plants	# plants-n/a	# n/a	# *APO* pos	# *APO* neg	Apomixis frequency	# plants	# seeds	# seeds-n/a	# 2C:5C	% 2C:5C
ES612xJL106											
JL12xJL103	153	143	10	52	91	36.4	36	841	631	163	25.8
ES612xES649	92	89	3	36	53	40.4	30	699	655	301	46.0
ES854xES649											
ES854xJL103											
ES854xJL107											
ES865xES649											
ES865xJL103											
ES865xJL106											
ES865xJL107											
JL78xJL106											
JL76xJL106											
Total	**245**	**232**	**13**	**88**	**144**	**38.4**	**66**	**1540**	**1286**	**464**	**35.9**

BC4F1

BC line	Genotype[a]						Phenotype[b]				
	# plants	# plants-n/a	# n/a	# *APO* pos	# *APO* neg	Apomixis frequency	# plants	# seeds	# seeds-n/a	# 2C:5C	% 2C:5C
ES612xJL106											
JL12xJL103	114	112	2	42	70	37.5	21	583	567	211	37.2
ES612xES649											
ES854xES649											
ES854xJL103											
ES854xJL107											
ES865xES649											
ES865xJL103											
ES865xJL106											
ES865xJL107											
JL78xJL106											
JL76xJL106											
Total	**114**	**112**	**2**	**42**	**70**	**37.5**	**21**	**583**	**567**	**211**	**37.2**

[a] The transfer of apomixis deferred from the presence of a proxy marker for female apomeiosis (*APOLLO*, Corral et al., 2013). In addition, the transfer frequency of a proxy marker for male apomeiosis (*UPGRADE2*, Mau et al., 2013) was monitored.

[b] Hybrid progeny and backcrosses were selected as potential donors for the next backcross generation based on flow cytometric analyses of the nuclear DNA content of leaf material (*i.e.* only diploids selected), the genome ratio of the embryo and endosperm in their produced seeds (*i.e.* individuals which produce more than 50% diploid, unbalanced apomictic seeds: 2C:5C seed ploidy), and the presence of proxy markers for female (*APOLLO*, Corral *et al.*, 2013) and male apomeiosis (*UPGRADE2*, Mau *et al.*, 2013).

b) *Transfer of apomixis from selected diploid unbalanced apomictic Boechera lines into sexual bridging species via hybrid crosses*

This part of the study examines the potential transfer clonal seed production (*i.e.* apomixis) from a natural apomict into a sexual species that functions as bridge between a crop plant and its phylogenetically distant natural apomictic relative with the ultimate goal to introduce clonal seed production in sexual crop plants (see **Fig. 6**). For this purpose a bridging species combines characteristics such as, (1) genome size and ploidy similar to the natural apomictic donor (*Diplotaxis erucoides*: 2n=2x=14, diploid with ~1.02-1.32Mbps/2C genome size, *Boechera sp.*: 2n=2x=14, diploid with ~440Mbps/2C genome size) and (2) crossability with the target crops species (e.g. with *B. rapa*, [[40, 41]], and *B. napus*, [[42]]).

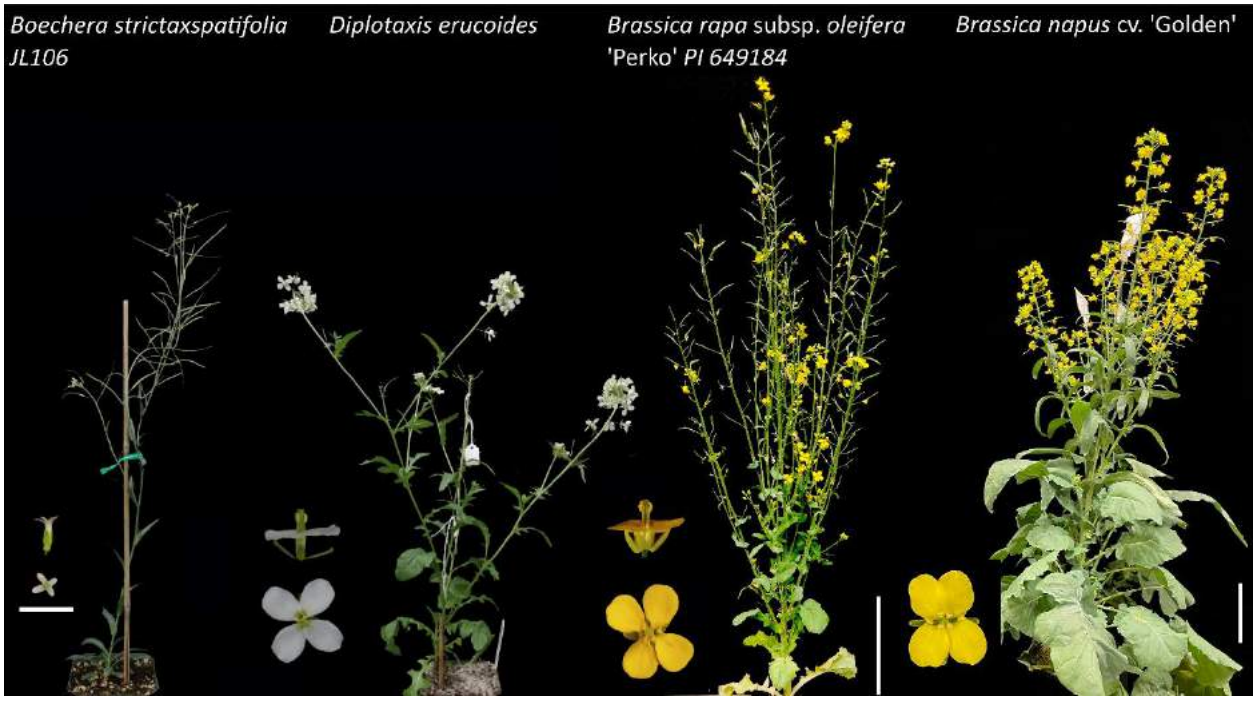

Figure 6. Example graphs of a diploid unbalanced apomictic pollen donor plant from the genus *Boechera*, the diploid sexual recipient bridge species *Diplotaxis erucoides* and the amphidiploid sexual recipient crops *Brassica rapa* and *Brassica napus*. Scale bar on left side for flowers = 10mm, on right side for whole plants = 15cm.

The goal is to generate a hybrid apomictic bridging species that shares the characteristic of its parental apomictic donor (*i.e.* a diploid unbalanced apomict with 2C:5C genome ratio of the embryo and endosperm which produce viable haploid pollen). This will minimize the phylogenetic distance for wide intergeneric crosses between apomictic donor and potential recipient *Brassica* crops (*i.e. B. napus, B. rapa*, refer to section **d)**) and avoid potential mentor effects (e.g. pollen incompatibility, parental imprinting) and ploidy barriers (**Fig. 7**) and ultimately may facilitate the transfer of genetic factors for the expression of apomixis into *Brassica* crops.

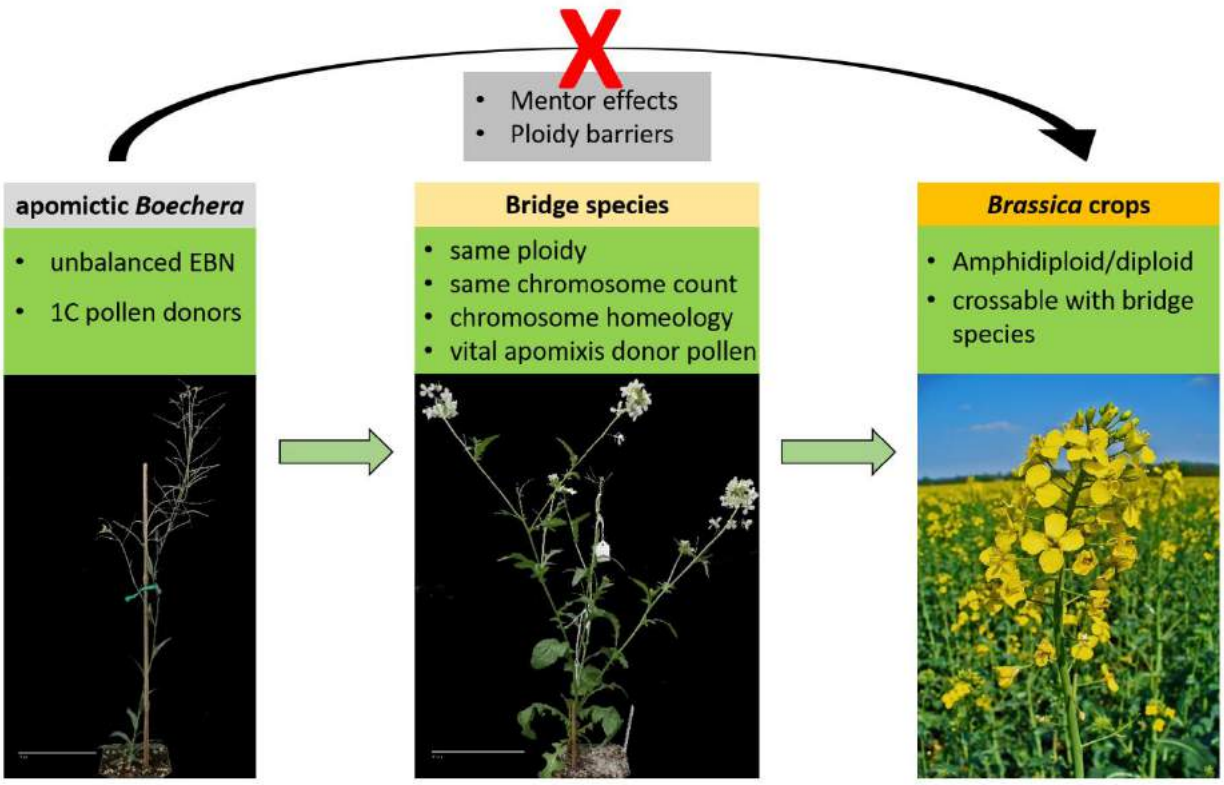

Figure 7. Schematic of the bridge species crossing concept and desired characteristics of the native apomictic relative, the bridge species and the crop plant.

We used five different diploid, unbalanced apomictic *Boechera* lines which produce haploid pollen as donors for crosses into diploid sexual *Diplotaxis erucoides* (**Tables 4** and **5**). In total, 231 *B. napus* cv. "Golden" individuals, and 120 *B. rapa* cv. "Perko" individuals in addition to 2529 individuals from five diploid, unbalanced apomictic *Boechera* lines were planted. As the seed-to-seed period for most of the *Boechera* lines is approx. six months we introduced a planting system with three months cycles for min. 30 seeds from each donor line of the genus *Boechera*. In *Diplotaxis erucoides* the flowering time starts approx. one month and the seed formation approx. 1.5 months after sowing. Hence, the sowing of *Diplotaxis erucoides* was aligned with the planting scheme of *Boechera* and a new set of *Diplotaxis erucoides* was planted approx. 4 weeks before the flowering time of a fresh set of *Boechera* lines. From all tested *Brassica* crops only *B. rapa* needed a stratification period which we optimized to eight weeks at ~4C. Cultivars of *B. napus* and *B. juncea* started flowering approx. one month after sowing and set seeds approx. 2.5 months after sowing.

Table 4. Pollen donor lines of the genus *Boechera (Brassicaceae)* for the generation of intergeneric hybrids with *Diplotaxis erucoides* and *Brassica* species.

Accession	Taxon	FCSS seed count	% C-values Embryo:Endosperm[a]						MOR interpretation	Pollen ploidy	Pollen viability (%)[b]	Apomixis transfer rate (% min-max)[c]	Seeds sown (2019-2020)	Seeds sown (2020-2021)	Seeds sown (2021-2022)	Seeds sown (2022-2023)	Plants survived (2019-2023)	Seed germination rate (%)[d]
			2C:3C	2C:4C	2C:5C	2C:6C	2C:7C	3C:6C										
ES649	*B. stricta*	40		0.20	0.75		0.05		unbalanced apo	1C	-	38.0 - 51.8	100	90	450	320	96	10.0
JL103	*B. crandallii*	182	0.02		0.92	0.01	0.05	0.01	unbalanced apo	1C	36.4	1.9 - 66.7	70	150	309	0	146	27.6
JL106	*B. stricta x spatifolia*	30			0.97	0.03			unbalanced apo	1C	-	13.6 - 35.7	10	270	240	0	161	31.0
JL101	*B. retrofracta*	34			0.97		0.03		unbalanced apo	1C	-	50.7 - 100	70	70	140	0	150	53.6
JL107	*B. stricta*	53	0.02		0.91	0.02	0.06		unbalanced apo	1C	13.0	33.3 - 60.0	10	80	150	0	44	18.3
JL12	*B. crandallii*	36	1.00						sexual	1C	-	0	120	0	0	0	87	72.5
ES865	*B. stricta*	36	1.00						sexual	1C	-	0	90	0	0	0	51	56.7
JL854	*B. stricta*	18	1.00						sexual	1C	49.8	0	90	0	0	0	40	44.4

[a] The C value ratio between the embryo and the endosperm (*i.e.* ploidy) of the mature seed as measured by flow cytometry defines the mode of reproduction (MOR). A ratio of 2C:3C refers to a sexual seed formed by fertilization of the egg cell and the central cell by each one reduced sperm nuclei. A seed ploidy ratio of 2C:5C refers to unbalanced apomictic seeds formed by parthenogenetic development of an unreduced egg cell and fertilization of the central cell by a reduced sperm nuclei (*i.e.* pseudogamy). A seed ploidy ratio of 2C:6C refers to balanced apomictic seeds formed by parthenogenetic development of an unreduced egg cell and fertilization of the central cell by an unreduced sperm nuclei (i.e. pseudogamy).

[b] Pollen viability taken from Mau et al. (manuscript submitted).

[c] The apomixis transfer rate refers to the frequency of a female apomeiosis proxymarker (*i.e.* apoallele of the *APOLLO* gene, Corral *et al.* 2013) measured in a synthetic hybrid formed by the specific donor accession.

[d] Germination counted 14 days after seeding.

(Status of 06-25-2023)

Table 5. Sexual recipient crop species of the Brassicaceae for incrosses with apomictic donor lines from the genus *Boechera* (Brassicaceae).

Taxon	Cultivar	USDA GRIN number	Seeds sown (2019-2020	Seeds sown (2020-2021)	Seeds sown (2021-2022)	Seeds sown (2022-2023)	Plants survived (2019-2023)	Seed germination rate (%)[a]	Genome size (pg/2C)[b]
Diplotaxis erucoides	-	BCN 3463	70	40	40	0	123	82	1.097±0.003
Brassica juncea	Domo	PI 458928	15	0	0	0	5	33	3.1*
Brassica napus	Golden	PI 649126	25	45	123	38	96	50	2.402±0.05
Brassica rapa subsp. *oleifera*	Steinacher Frueher Winterruebsen	PI 649185	5	0	0	0	3	60	1.6*
Brassica rapa subsp. *oleifera*	Perko	PI 649184	15	45	60	0	68	57	1.6*

[a] Germination counted 14 days after seeding.

[b] Genome size inferred from flow cytometric analyses of leaf C-values. Values marked with asterisks inferred from https://cvalues.science.kew.org/search/angiosperm

(Status of 06-25-2023)

We applied a pollen tube growth *in situ* protocol mechanical and natural selfings of the recipient crop cultivars *D. erucoides*, *B. napus* cv. "Golden" and *B. rapa* cv. "Perko" and their intergeneric hybrid progenies and to examine the level of crossing success and self-incompatibility (SI) as these can vary between cultivars and can also be influenced by environmental stressors (*i.e.* heat, **Table 7**). The optimal sampling of pollen tubes in *Boechera* is 48hrs after pollination, as such we applied this protocol also to selfings in the selected crop cultivars (**Fig. 9**). Interestingly, while in *Boechera* high levels of full-length pollen tubes have formed after 48hrs and the pollen tube formation in the crops has been triggered on the stigma but they did not have developed fully into the placenta after 48hrs (**Fig. 9**). In order to apply this test to hybrid progeny we have to optimize the sample collection time. Therefore we have performed mechanical and natural selfing in *B. napus* cv. "Golden" within and across individuals and collected the samples at 48hrs, 72hrs and 96hrs. In total, we have analysed 259 pistils from crosses with *Boechera* lines into *D. erucoides* (165 pistils), *B. rapa* (24) and *B. napus* (57). Overall, it seems that the pollen tube growth is impeded in the carpel in all intergeneric crosses. This is potentially due to SI mechanisms on the recipient carpel or due to (partial) genetic incompatibilities between recipient and donor.

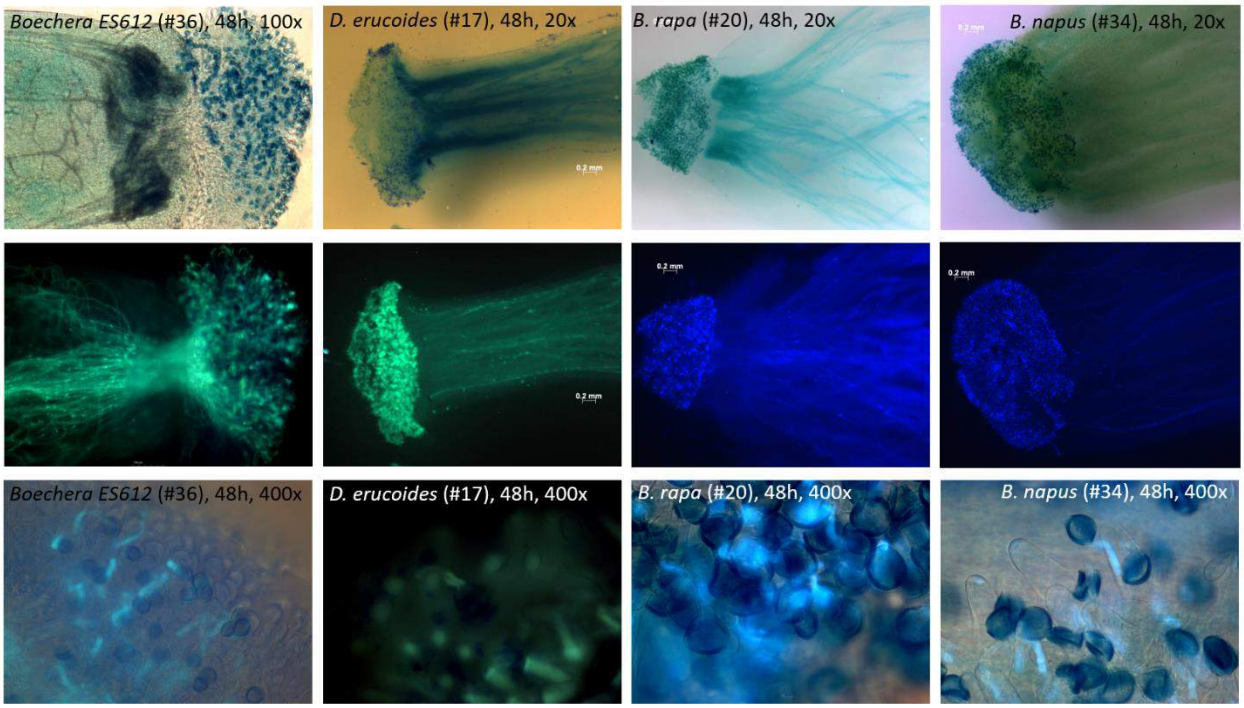

Figure 9. Pollen tube *in situ* analysis with brightfield and fluorescence microscopy in mechanical selfings of sexual *Boechera* ES612, *Diplotaxis erucoides*, *Brassica rapa* cv. "Perko', and *Brassica napus* cv. "Golden".

Therefore, we tested four alternative crossing methods (**Table 6**) to compare their efficiency against our standard method (crossing method 1) for which we emasculate flowers with exposed yellow petals at day 1 and pollinate them at day 2. In method 2 immature fully closed flower buds were bagged at day 1, emasculated at day 2 and pollinated at day 3. In crossing method 3 the emasculation and pollination of immature flowers with yellow petals at the same day (refer to pollen tube analyses part). The crossing method 4 is the most stringent (*i.e.* least error prone) and minimal invasive method in which anthers were removed from flowers at anthesis leaving the rest of the flower intact. Then the pistil was either left intact for pollination or the pistil was cut beneath the style (*i.e.* to circumvent potential pollen incompatibilities) and then covered with fresh pollen from the donor plant. Finally, the pollinated portion of the pistil was covered with cotton ball soaked in a hormone solution with GA3, NAA and Kinetin to enhance proper pollen tube formation (see Method section).

Interestingly, crossing method 2 resulted in higher hybrid seed counts in most of the comparable crosses with lower numbers of pollinated pistils compared to crossing method 1. Crossing methods 3 and 4 resulted in none or low progenies. This result is due to a higher false positive rate among the generated seeds due to cross contamination with own pollen in those methods where the pistil was exposed for extended periods (methods 1 and 2) compared to methods with immediate pollination events (methods 3 and 4). Interestingly, the hormonal treatment and potential circumvention of self incompatibility constrains in method 4 did not lead to a significant production of seeds.

In total, 1408 crosses of four diploid, unbalanced apomictic *Boechera* donor lines into the diploid sexual recipient *Diplotaxis erucoides* were performed and produced 160 seeds in total (all crossing methods together; **Tables 6 and 7**).

Table 6. Crossing methods applied in wide intergeneric cross pollinations between recipient crop cultivars *D. erucoides, B. napus* cv. "Golden" and *B. rapa* cv. "Perko" and five diploid, unbalanced apomictic *Boechera* donor lines.

			Crossing method 1*			Crossing method 2*			Crossing method 3*			Crossing method 4*		
Cross combination	Recipient line	Donor line	N inflorescences	N crossed pistils	N hybrid seeds	N inflorescences	N crossed pistils	N hybrid seeds	N inflorescences	N crossed pistils	N hybrid seeds	N inflorescences	N crossed pistils	N hybrid seeds
1	*Diplotaxis erucoides*	*Boechera stricta* 'ES649'	45	92	61	141	235	49	86	105	0	37	134	0
2	*Diplotaxis erucoides*	*Boechera stricta* 'JL107'	57	72	0									
3	*Diplotaxis erucoides*	*Boechera stricta x spatifolia* 'JL106'	120	188	27				22	31	0	7	32	0
4	*Diplotaxis erucoides*	*Boechera retrofracta* 'JL101'	198	244	8	72	104	10	6	9	0	17	67	0
5	*Diplotaxis erucoides*	*Boechera crandallii* 'JL103'	22	25	0	16	25	5	27	45	0			
6	*B. napus* cv. 'Golden'	*Boechera stricta x spatifolia* 'JL106'	n/a	73	1				19	28	0	58	135	0
7	*B. napus* cv. 'Golden'	*Boechera crandallii* 'JL103'	62	90	30	6	6	12	13	15	0	2	9	0
8	*B. napus* cv. 'Golden'	*Boechera retrofracta* 'JL101'	52	54	208	1	1	0	7	15	0	43	134	0
9	*B. napus* cv. 'Golden'	*Boechera stricta* 'JL107'	6	7	0				1	2	0			
10	*B. napus* cv. 'Golden'	*Boechera stricta* 'ES649'	40	62	25	37	37	288	37	53	0	40	134	4
11	*B. napus* cv. 'Golden'	*Boechera stricta* 'ES612'				1	1	16						
12	*B. rapa* cv. 'Perko'	*Boechera stricta* 'ES649'	3	4	0	43	65	39	43	55	0	15	37	0
13	*B. rapa* cv. 'Perko'	*Boechera crandallii* 'JL103'	9	9	0	6	6	0	8	12	0			
14	*B. rapa* cv. 'Perko'	*Boechera retrofracta* 'JL101'	29	29	0	8	9	0	32	38	0	3	7	0
15	*B. rapa* cv. 'Perko'	*Boechera stricta x spatifolia* 'JL106'	36	39	0				22	33	0	7	19	0
16	*B. rapa* cv. 'Perko'	*Boechera stricta* 'JL107'	9	12	0				1	1	0			
Total			**688**	**1000**	**360**	**331**	**489**	**419**	**324**	**442**	**0**	**229**	**708**	**4**

*Crossing method 1: emasculated flower with exposed yellow petals at day 1 and pollinated at day 2; crossing method 2: immature fully closed flower buds were bagged at day 1, emasculated at day 2 and pollinated at day 3; crossing method 3: closed flower buds with yellow petals were emasculated and pollinated at the same day; crossing method 4: anthers were removed from flowers at anthesis leaving the rest of the flower intact, pollenination on pistil with or with removed pistil, cover pollen pellet with cotton ball soaked in hormone solution.

Only crossing method 1 & 2 produced seeds. In addition, 856 crosses of five diploid, unbalanced apomictic *Boechera* donor lines into the diploid sexual recipient *B. napus* cv. 'Golden' were performed and produced 567 seeds in total (all crossing methods together; **Tables 6 and 7**). The crossings methods 1 and 2 produced similar numbers of progeny (264 and 300 seeds, respectively) while crossing method 3 did not produce any seeds and crossing method 4 produced four seeds with donor accession *Boechera stricta* 'ES649'. In total 39 seeds derived from 369 crosses, solely from method 2, between five diploid, unbalanced apomictic *Boechera* donor lines into the diploid sexual recipient *B. rapa* cv. 'Perko'.

We have planted all harvested seeds from crosses between recipient *D. erucoides*, *B. napus* cv 'Golden and *B. rapa* cv 'Perko' with the apomictic *Boechera* donor lines. The overall germination rate was high (72.2%, min: 20%, max: 100%). We geno- and phenotyped 62 F1 from crosses into *D. erucoides,* 451 F1 from crosses into *B. napus* cv 'Golden, and 34 F1 from crosses with *B. rapa* cv 'Perko'.

The flow cytometric analyses of nuclear DNA contents (*i.e.* ploidy) of leaf material is continuously performed on all parental individuals to assure their correct ploidy levels. We used *Pisum sativum* as internal control (8.84 pg/2C [[34]]) to calculate the genome size. So far, all parental plants were diploid.

We developed PCR markers for marker-assisted selection of true intergeneric hybrids between apomictic donor lines and sexual recipient crops and selected one absence/presence (nuclear RNA polymerase D1B (TAIR#: AT2G40030, *Boechera* gene#: BA4G0233000) and two insertion-deletion (InDel) markers (CSC1-like protein (TAIR#: AT3G54510, *Boechera* gene#: BA5G0414900; Plastid division protein: TAIR#: AT3G19180, *Boechera*

gene#: BA3G0307900) which are genic markers located on chromosome #3, #4 and #5 in *Boechera* (**Figures 10 and 11**).

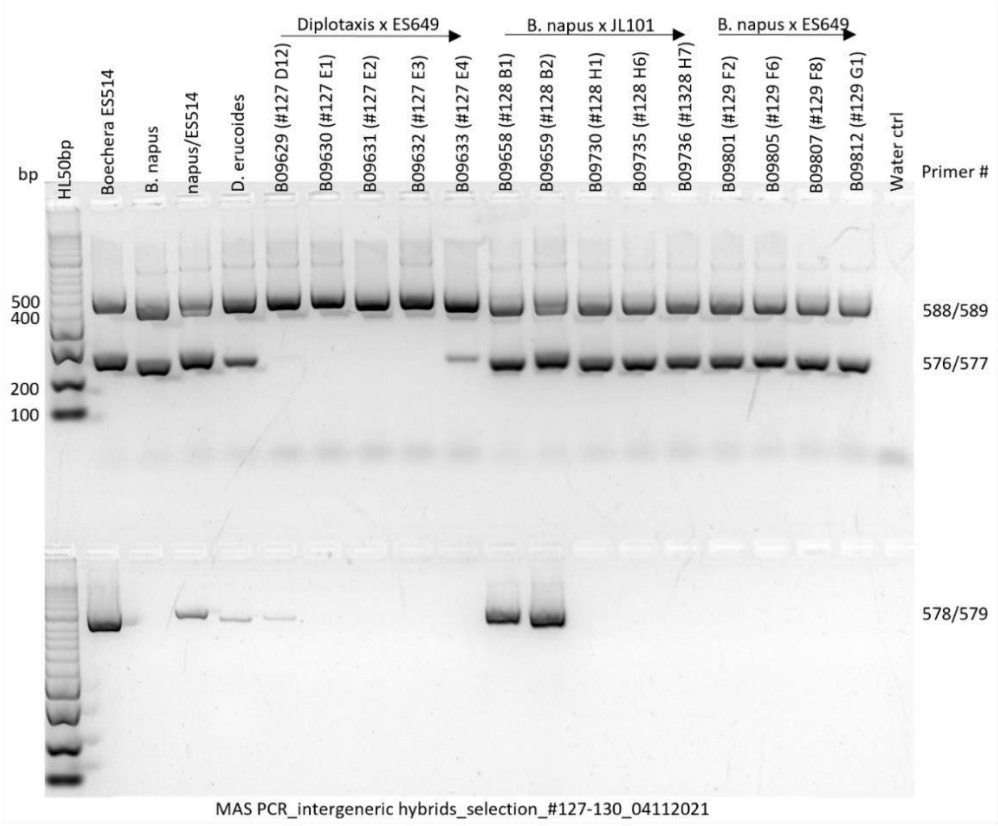

Figure 10. Amplification patterns for the two InDel markers (primer#588/589: nuclear RNA polymerase D1B (NRPD1B, AT2G40030) and primer#576/577: Plastid division protein (CDP1, AT3G19180)) and the presence/absence marker (primer#578/579: CSC1-like protein (AT3G54510)) in the recipient and donor parental lines and potential hybrid progeny.

Government of Saskatchewan
Ministry of Agriculture

ADF Project Progress Report Format

Table 7. Intergeneric hybridizations between diploid apomictic *Boechera* donor lines and sexual diploid recipient *Diplotaxis* or amphidiploid sexual recipient *Brassica* crop lines. The hybrid origin of F1 progeny was verified by marker-assisted selection (MAS) with two insertion-deletion (InDel) markers and a single *Boechera* -specific marker.

				2019-2020			2020-2021			2021-2022			F1 hybrids				F1 MAS					
Cross combination #	Recipient line	Donor line	Donor M	N inflorescences	N crossed pistils	N hybrid seeds	N inflorescences	N crossed pistils	N hybrid seeds	N inflorescences	N crossed pistils	N hybrid seeds	Total N hybrid seeds	N sown	N germinated	Germination rate (%)	N tested (marker 1)	N positive (marker 1)	N tested (marker 2)	N positive (marker 2)	N tested (marker 3)	N positive (marker 3)
1	*Diplotaxis erucoides*	*Boechera stricta* 'ES649'	2C:5C	8	46	4	178	281	106	123	239	0	110	110	75	68.2	47	1	47	0	11	1
2	*Diplotaxis erucoides*	*Boechera stricta* 'JL107'	2C:5C	3	12	0	54	60	0													
3	*Diplotaxis erucoides*	*Boechera stricta x spatifolia* 'JL106'	2C:5C	110	170	4	10	18	23	29	63	0	27	23	23	100.0						
4	*Diplotaxis erucoides*	*Boechera retrofracta* 'JL101'	2C:5C				270	348	18	23	76	0	18	18	5	27.8	10	0	10	0	0	0
5	*Diplotaxis erucoides*	*Boechera crandallii* 'JL103'	2C:5C				38	50	5	27	45	0	5	5	1	20.0	5	0	5	0	0	0
6	*Diplotaxis erucoides*	*Boechera crandallii* 'JL12'	2C:3C				6	6	0													
7	*Diplotaxis erucoides*	*Boechera stricta* 'ES854'	2C:3C				13	15	1				1	0	0							
8	*Diplotaxis erucoides*	*Boechera stricta* 'ES865'	2C:3C				1	1	0													
9	*B. napus* cv. 'Golden'	*Boechera stricta x spatifolia* 'JL106'	2C:5C	n/a	15	0	3	3	0	77	163	0										
10	*B. napus* cv. 'Golden'	*Boechera crandallii* 'JL103'	2C:5C	1	6	0	67	90	42	15	24	0	42	42	38	90.5	11	0	11	0	0	0
11	*B. napus* cv. 'Golden'	*Boechera retrofracta* 'JL101'	2C:5C				55	57	208	50	149	0	208	148	117	79.1	196	2	196	0	61	2
12	*B. napus* cv. 'Golden'	*Boechera stricta* 'JL107'	2C:5C	3	4	0	4	5	0													
13	*B. napus* cv. 'Golden'	*Boechera stricta* 'ES649'	2C:5C				76	93	313	77	187	4	317	272	210	77.2	244	4	244	0	179	1
14	*B. napus* cv. 'Westar'	*Boechera stricta x spatifolia* 'JL106'	2C:5C	n/a	55	1							1	1	1	100.0						
15	*B. rapa* cv. 'Perko'	*Boechera stricta* 'ES649'	2C:5C	3	4	0	61	106	39	58	92	0	39	39	34	87.2	34	0	34	0	7	0
16	*B. rapa* cv. 'Perko'	*Boechera crandallii* 'JL103'	2C:5C				15	15	0	8	12	0										
17	*B. rapa* cv. 'Perko'	*Boechera retrofracta* 'JL101'	2C:5C	11	16	0	26	22	0	35	45	0										
18	*B. rapa* cv. 'Perko'	*Boechera stricta x spatifolia* 'JL106'	2C:5C	36	39	0				29	52	0										
19	*B. rapa* cv. 'Perko'	*Boechera stricta* 'JL107'	2C:5C	9	12	0				1	1	0										
20	*B. rapa* cv. 'Perko'	*Boechera crandallii* 'JL12'	2C:3C	5	5	0																
21	*B. juncea* cv. 'Domo'	*was not used*																				
22	*B. rapa* cv. 'Steinacher Frueher Winterruebsen'	*was not used*																				
Total				**189**	**384**	**9**	**877**	**1170**	**755**	**552**	**1148**	**4**	**768**	**658**	**504**	**72.2**	**547**	**7**	**547**	**0**	**258**	**4**

[a] PCR-based screen of hybrid progeny using two insertion-deletion markers (marker #1: CSC1-like protein (TAIR#: AT3G54510, *Boechera* gene #: BA5G0414900) , marker #2: Plastid division protein (TAIR #: AT3G19180, *Boechera* gene #: BA3G0307900)) and a single *Boechera* -specific marker (marker #3: nuclear RNA polymerase D1B (TAIR #: AT2G40030, *Boechera* gene #: BA4G0233000)).

(Status: 06-25-2023)

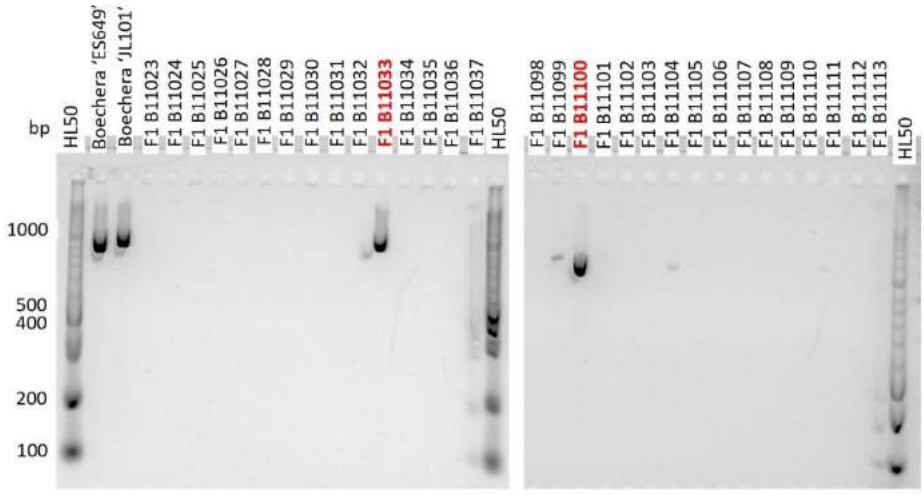

Figure 11. Amplification patterns for the presence/absence marker (primer#578/579: CSC1-like protein (AT3G54510)) in two donor parental lines and the potential hybrid progeny. Red color mark F1 plants carrying a *Boechera* ortholog of the CSC1-like protein (AT3G54510).

c) *Testing for stable inheritance of the apomixis trait in the bridging species*

The resulting F1 progeny from intergeneric crosses between the bridge species *Diplotaxis erucoides* and apomictic *Boechera* donor lines could not be validated using the absence/presence (primer#578/579: CSC1-like protein (AT3G54510)) and InDel markers ((primer#588/589: nuclear RNA polymerase D1B (NRPD1B, AT2G40030) and primer#576/577: Plastid division protein (CDP1, AT3G19180)). Therefore, we were not able to analyse inheritance patterns of introgressions sites for the F1 progeny.

In contrast, we used our validated 5 hybrid lines (F1 plants # B09658, B09659, B09735, B11033, and B11100,) from crosses between recipient sexual *Brassica napus* and apomictic *Boechera* donor lines for the analysis of inheritance patterns of the introgressed genomic portions from apomictic *Boechera* in the progeny of F1 hybrids. We have performed selfing and backcrosses into sexual recipient *B. napus* cv 'Golden' for each of the five F1 hybrid lines and generated at least 50 seeds, respectively. All 243 F1 hybrids progeny derived from selfing of the five confirmed intergeneric hybrid lines were planted and phenotyped for plant height, flower shape, seed size/shape (**Table 8**). The analyses of the collected PCR genotyping and phenotyping data is currently in progress and the exact choice of method depends on our findings from the analysis of the whole genome sequencing of F1 hybrids (**see section 9d**). Backcrosses into *B. napus* cv 'Golden' were performed with 4 of 5 confirmed F1 hybrid plants (individual B09735 could not be used for technical reasons) and resulted in sufficient seeds (Table 8) to direct plant into soil but none of the planted seeds were germinating. Hence, we tested for the metabolic activity of seeds produced by selfed and backcrossed F1 progeny used this data as proxy for seed viability (**Figure 12**). Overall, intergeneric backcross progeny has greater variation in seed size and seed viability compared with progeny derived from selfing. Most of the backcrossed lines show a reduced seed viability (between 0 and 85%) compared to selfed progeny (between 65% and 100%) and native *B. napus* cv 'Golden' (100%). Therefore, we applied a custom embryo rescue protocol to enhance the germination rate of lines with reduced germination. So far, we have performed embryo rescue on 2 of 4 BC lines and showed elevated germination rates for two different embryo rescue media (68.3% and 67.7%) for BC line # B11271 x B11100. The embryo rescue did not

produce viable progeny for the second tested BC line B09853 x B09659. The remaining two BC lines B09858 x B09659 and B11270 x B11033 are currently tested, although the good seed quality and superior seed size of B11270 x B11033 should lead to successful germination directly on soil.

Table 8. Propagation methods for intergeneric hybrid progeny derived from F1 selfing and backcrosses with *B. napus* cv. Golden.

F1 plant #	Taxon	Crossing method	Propagation method*	# BCs	Total N hybrid seeds	N seeds/embryos sown	N germinated	Germination rate (%)
B09658	B. napus x B. retrofracta JL101	backcrossing	direct to soil	58	698	169	0	0.0
		backcrossing	embryo rescue			tbd	tbd	tbd
		selfing	direct to soil		>50	50	47	94.0
B09659	B. napus x B. retrofracta JL101	backcrossing	direct to soil	52	542	tbd	tbd	tbd
		backcrossing	embryo rescue M1			25	0	0.0
		backcrossing	embryo rescue M2			25	0	0.0
		selfing	direct to soil		>50	50	50	100.0
B09735	B. napus x B. retrofracta JL101	backcrossing#	direct to soil	n/a	n/a	n/a	n/a	n/a
		selfing	direct to soil		>50	60	48	80.0
B11033	B. napus x B. stricta ES649	backcrossing	direct to soil	26	147			
		selfing	direct to soil		>50	650	50	7.7
B11100	B. napus x B. stricta ES649	backcrossing	direct to soil	33	412			
		backcrossing	embryo rescue M1			41	28	68.3
		backcrossing	embryo rescue M2			31	21	67.7
		selfing	direct to soil		>50	155	48	31.0

*Various embryo rescue media were tested: M1-embryo culture medium (ECM) without additional hormones, M2-ECM with 0.2mg/L 6-Benzylaminopurine (BAP), M3-ECM with 1.5mg/L BAP and 0.25mg/L 1-Naphthaleneacetic acid (NAA), M4-ECM with with 0.1mg /L 1-Naphthaleneacetic acid (NAA)

Backcrosses were not successful with B09735

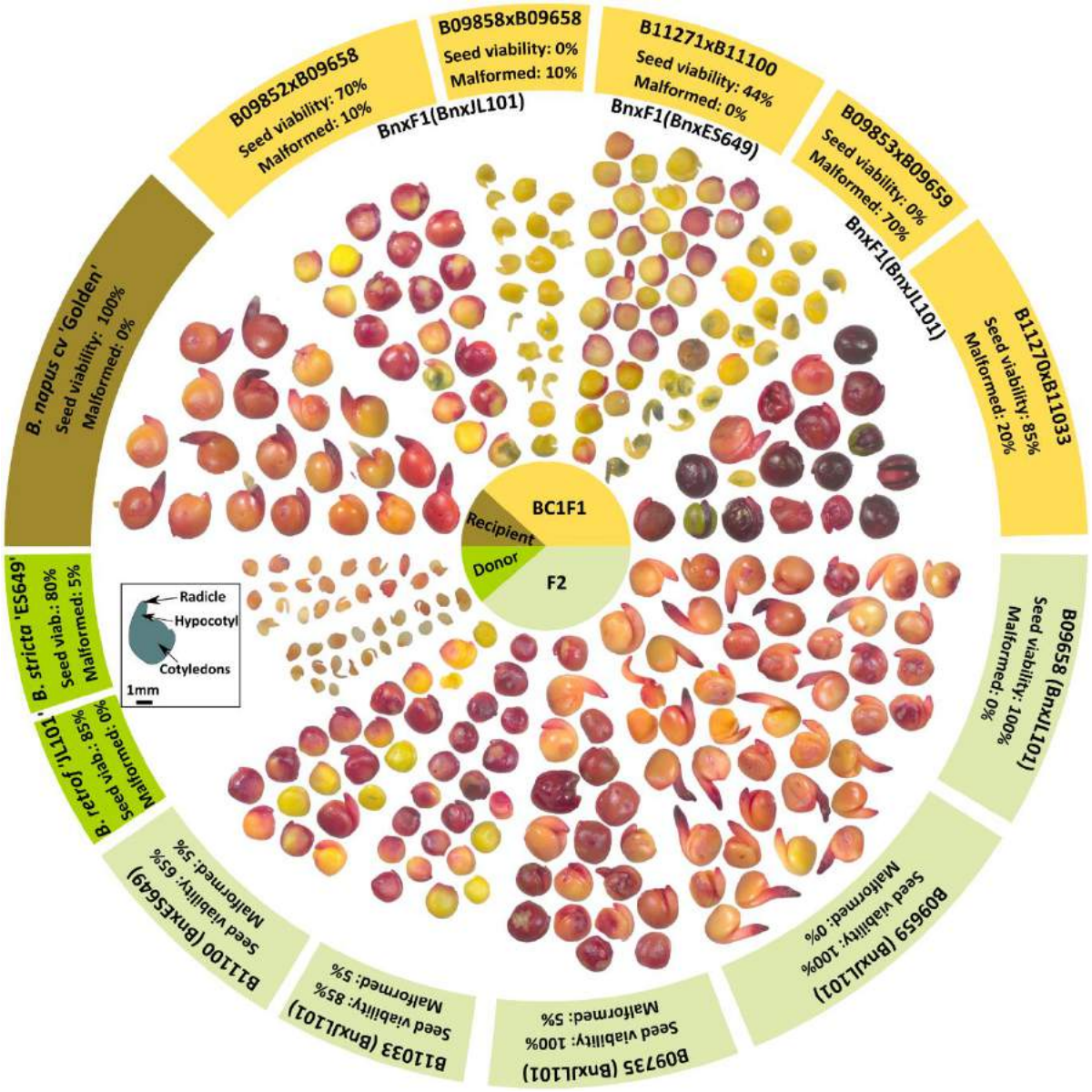

Figure 12. Intergeneric hybrid F2 germplasm viability and embryo size variability. The perforated seeds from F2 hybrids between Canola and apomictic *Boechera* donors were exposed to the colorless tetrazolium solution which reduces hydrogen derived from enzymes in the respiration process of living cells and changes to red. Extracted embryos showing active respiration are considered "viable" and turn red. The darker the color the greater the respiratory activity in the seed. Hybrids derived from selfing show a relative uniform highly viable germplasm set comparable with the native *B. napus* cv. 'Golden'. In contrast germplasm from backcrosses show high variability of embryo size and viability. Some backcrosses show significant reduction in viability and size (*e.g.* B09858xB09658) and others show an increase in embryo size and respiratory activity (*e.g.* B11270xB11033).

d) *Generating apomictic Brassica crops via intergeneric crosses*

The PCR-based screen of 547 F1 plants with the InDel and 258 F1 plants with the presence/absence markers confirmed a hybrid origin for two F1 plants from a cross between *B. napus* cv. 'Golden' and Boechera 'JL101' (B09658 and B09659) and 2 F1 progeny from a cross between *B. napus* cv. 'Golden' and *Boechera* 'ES649' (B11033 and B11100; **Figs. 10** and **11**). From the identified intergeneric hybrids we generated selfed and backrossed F2 progeny (**Figs. 12 and 13, Table 9**).

Figure 13. Pollination bags used for intergeneric F1 hybrid pools from same cross combination to allow outcrossing among individuals of the same cross type and restrict outcrossing between cross types.

Together with our collaborators from the MASARYK UNIVERSITY (Brno, Czech Republic) we applied Genomic in situ hybridization (GISH) as a second validation method (along with the PCR-based method) for the detection of a hybrid origin of the progeny from intergeneric crosses. We validated two F1 plants from a cross between *B. napus* cv. 'Golden' and *Boechera* 'JL101' (B09658 and B09659) and identified a third individual with a single large introgression that has not been identified as a true hybrid by the PCR-based screen (B09735). In addition, we validated two F1 plants from a cross between *B. napus* cv. 'Golden' and *Boechera* 'JL101' (B11033 and B11100). Instead of expected transfer of complete *Boechera* chromosomes into the F1 hybrids, we detected the introgression of different subsets of genomic fragments scattered throughout the chromosomes (**Fig. 12**). This karyotype can be explained by homoploid hybridization were a partial synaptonemal complex formation between *B. napus* and *Boechera sp.* chromosomes and Boechera genome fragments are translocated via crossover and double strand breaks (DSBs). Then, one or several micronuclei carrying aberrant remnant *Boechera*

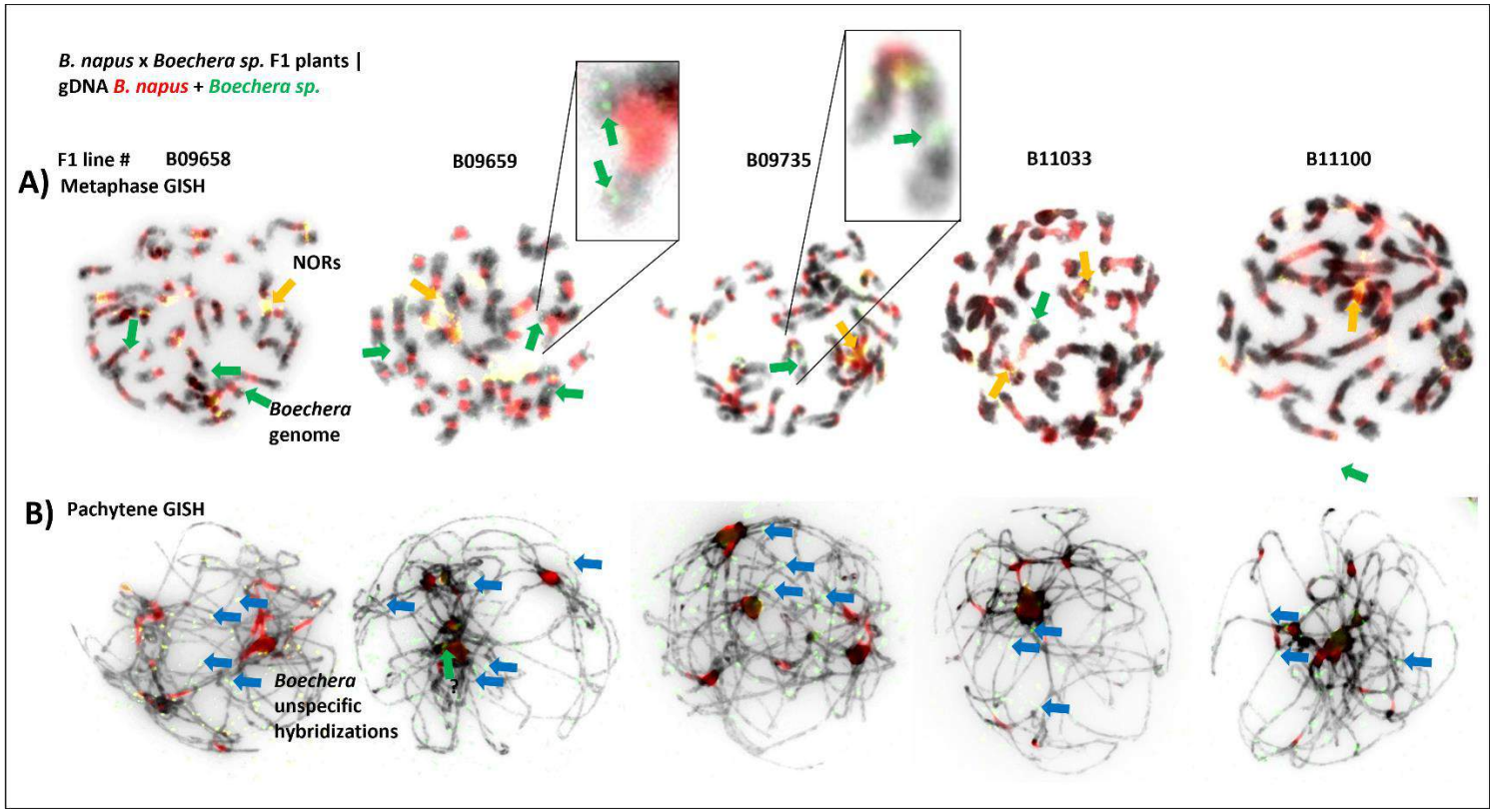

Figure 14. DNA-FISH with *Boechera* and *Brassica napus* specific genomic probes on potential hybrid F1 plants from crosses between *B. napus* and apomictic diploid *Boechera.* Genomic in situ hybridization (GISH) was used to confirm the hybrid nature of F1 lines with *Boechera*-specific InDel markers. GISH results on metaphases show the presence of *Boechera*-specific genomic regions in the *B. napus* genomes (green arrows) but no additional chromosomes from *Boechera* which points to homoploid hybridization. Some regions share probes from *Boechera* and *B. napus* genome which were confirmed to be the conserved Nucleolus Organizer Regions (NOR) between both species (yellow arrows). GISH on thickened but not fully condensed chromosome threads in the pachytene stage was performed to further confirm the presence of chromosome regions from *Boechera* and to measure the total length of all fragments transferred from *Boechera* into the *B. napus* genome. Hybridizations to the pachytene became problematic due to the high frequency of unspecific hybridizations using the whole Boechera genome as a probe and no clear evidence of the *Boechera*-specific chromosome regions on pachytenes was obtained (e.g. green arrow).

sp. chromosomes are formed and break down over time through apoptosis. Finally, spontaneous whole genome duplication (WGD) in the zygote lead to homozygote state that we observe in the GISH metaphase plates.

We tested whether known apomixis factors from the apomixis factor list 1 (AFL1; ADF #20150123, Fig. 15) were transferred into the intergeneric hybrids and developed InDel primers for *B. napus* and *Boechera sp.* orthologs (**Tables 9 and 10**).

Table 1. PCR primers for 49 apomixis candidate factors from ADF #20150123 used to test for introgression of apomixis from diploid unbalanced apomictic *Boechera* donors into the *Brassica napus* cv 'Golden' cultivar.

Primer target	Arabidopsis homolog #	Brassica napus homolog #	Primer #	Primer name	Primer sequence	Tm in °C	Exp. product length (B. napus Z511/ Boechera sp. ortholog)
cysteine-rich RLK (RECEPTOR-like protein kinase) 32 (Boe-gene# 327)	AT4G11480	NC_027764.2	#602/#603	Apogene1F/R	ATTCGCAAACCGGTGTGCTATT ATCTGTTGCCGCTTCGATTGTC	58.4/60.3	411/364 bp
MAP kinase kinase 8 (Boe-gene# 7994)	AT3G06230	NW_019168955.1	#604/#605	Apogene2F/R	AGACATCAAGCCAGCGAATCTG CCCTAACCCTAATCTTGCTGGAG	60.3/62.4	456/440 bp
polyadenylate-binding protein 1-B-binding protein (Boe-gene# 13111)	AT1G31130	NC_027765.2	#606/#607	Apogene3F/R	TGGGTCGAGAGATTCTGCTACT TCCACATTTCCACTTCCTCTCCA	60.3/60.6	108/- bp
Trypsin family protein (Boe-gene# 14081)	AT5G45030	NC_027757.2	#608/#609	Apogene4F/R	GTCCAAATCAACTGCCACATGC GGGTTTGGTGCGATGTAGATGT	60.3/60.3	324/344 bp
TOXICOS EN LEVADURA 4 (Boe-gene# 15759)	AT3G60220	NC_027772.2	#610/#611	Apogene5F/R	AGTACTAGCAGGCCATTGATTTGA CGCTGGGAGAGAGACAGAGAAA	59.3/62.1	99/- bp
HNH endonuclease (Boe-gene# 16390)	AT2G23840	NW_019168938.1	#612/#613	Apogene6F/R	TGCTTCAGAGGCTTGGTTCTTG TGTATTCCAGACAGATTGCGCG	60.3/60.3	276/328 bp
pre-rRNA-processing TSR1-like protein (Boe-gene# 16392)	AT1G42440	NC_027762.2	#614/#615	Apogene7F/R	TGCTAGAGCTGCTCGTCTTCAG CTCACGATAACACGAGGAGCAC	62.1/62.1	203/117 bp
MLO9 (Boe-gene# 16393)	AT1G42560	NC_027770.2	#616/#617	Apogene8F/R	GGCTTTGGAGATTCAAGAGAGACA TCATCAAACACCGATCGCTTCA	61.0/58.4	700/672 bp
BZIP60 (Boe-gene#16395)	AT1G42990	NC_027764.2	#618/#619	Apogene9F/R	GCTATTGCCAAGAAACGAAGAAGG CAGGGATTCCAACAAGAGCACA	61.0/60.3	532/652 bp
LCV2 (Boe-gene#16396, BA1G0146200)	AT1G43130	NW_019168564.1	#620/#621	Apogene10F/R	CGTAGGCACATACACACTACACA CTTCCTCGTGGCTTGGTTCTAC	60.6/62.1	505/620 bp
SET domain-containing protein (Boe-gene#16399)	AT1G43245	NC_027765.2	#622/#623	Apogene11F/R	CCTCCCGCTTATGTGGACTCTA TCAATCGAACGGATTCTGTAAGCA	62.1/59.3	373/591 bp
Sphingomyelin synthetase family protein (Boe-gene#16402)	AT1G43580	NC_027768.2	#624/#625	Apogene12F/R	ATCCTCATCGTTGGGCTCAGAA CAAGCCATAGCGGTTAGGACAG	60.3/62.1	431/355 bp
UGT80B1 (Boe-gene# 16403)	AT1G43620	NW_019169735.1	#626/#627	Apogene13F/R	GCCCAGAAGGAATGAGACCTTT TCTATCCCTTGGGTGGTGATCC	60.3/62.1	180/233 bp
plant/protein (DUF793) (Boe-gene# 16404)	AT1G43630	NW_019168564.1	#628/#629	Apogene14F/R	GGAAGTTTCAGATTCAGTTGTGAGG ACATCTCACAGTCCACAGAGGT	61.3/60.3	211/- bp
TLP1 tubby like protein 1 (Boe-gene# 16405)	AT1G76900	NC_027759.2	#630/#631	Apogene15F/R	TGGGACCATTGTTCTTGGAAGG AAAGCAAGCACACACACCTACA	60.3/58.4	216/- bp
UMAMIT22 nodulin MtN21 /EamA-like transporter family protein (Boe-gene# 16406)	AT1G43650	NC_027764.2	#632/#633	Apogene16F/R	ACATGGCAATGGTGTTTGTACAGA TCCCTTGAGAGATGGCCACTTT	59.3/60.3	146/75 bp
FTSH protease 1 (Boe-gene# 16594)	AT1G50250	NC_027772.2	#634/#635	Apogene17F/R	TCCCTTGAGAGATGGCCACTTT CACACGAGACACCTGCATGAAA	62.4/60.3	483/572 bp
PDH-E1 ALPHA pyruvate dehydrogenase E1 alpha (Boe-gene# 17431)	AT1G01090	NC_027771.2	#636/#637	Apogene18F/R	CGGGTGTTCATGTTGACGGTAT TGGGATCAGCAAACACATTCTCT	60.3/58.9	435/497 bp
ARM repeat superfamily protein (Boe-gene# 19450)	AT5G19820	NC_027766.2	#638/#639	Apogene19F/R	TTGGGAGGGATGATGAATGGGT CCCAATTGACAACTCTTCCCTCC	60.3/62.4	93/- bp
inactive purple acid phosphatase-like protein (Boe-gene# 19452)	AT1G42430	NC_027764.2	#640/#641	Apogene20F/R	GTCCCGTGATACTCCGCCTAAA TTCTCATTCAACAAATCGATGCCC	62.1/59.3	154/237 bp
Patched family protein (Boe-gene# 19453)	AT1G42470	NC_027760.2	#642/#643	Apogene21F/R	GAGAGGATTGTGCCACACAGAG GGGACCCTTAAATGCACTCAAAC	62.1/60.6	360/418 bp
Patched family protein (Boe-gene# 19453)	AT1G42470	NC_027760.2	#644/#645	Apogene22F/R	CTGCTTCATGGCTCGATGACTT AATAAGCACCATGGCCACCTTT	60.3/58.4	500/842 bp
TLR4 regulator/MIR-interacting MSAP protein (Boe-gene# 19454)	AT1G42480	NC_027774.2	#646/#647	Apogene23F/R	GGACGGGCTGTGTGATAGAATG TGTCATTTGCATGCGCTTTAGC	62.1/58.4	433/775 bp
PMI1 plastid movement impaired1 (Boe-gene# 19456)	AT1G42550	NC_027764.2	#648/#649	Apogene24F/R	ATTGCAAAGCAGATCAAAGCGC TCTCCTCCACGTTCCAGATTCC	58.4/62.1	664/735 bp

chloroplast glyceraldehyde-3-phosphate dehydrogenase B subunit gene (Boe-gene# 19457)	n/a	NC_027764.2	#650/#651	Apogene25F/R	GGATCCTAACCATCCAATCAGCA GTCTGCTCCAGCTATAGCCTTG	60.6/62.1	356/493 bp
GAPB glyceraldehyde-3-phosphate dehydrogenase B subunit (Boe-gene# 19458)	AT1G42970	NC_027774.2	#652/#653	Apogene26F/R	ACAGGAGTGTTTGTTGATGGGC AGGTCCCTGTGTGATGCATCTA	60.3/60.3	617/755 bp
electron protein, putative (Protein of unknown function, DUF547) (Boe-gene# 19462)	AT1G43020	NW_019168588.1	#654/#655	Apogene27F/R	TCTCAGCTAAGCACTCCAAGGT ACCGCTCTAACTGTAAGATCTCTTTC	60.3/61.6	407/338 bp
ABR1 Integrase-type DNA-binding superfamily protein (Boe-gene# 19465)	AT5G64750	NC_027762.2	#656/#657	Apogene28F/R	TGCATCAAGTGTATGGTCACATCA ACGACTCTTGAGGCAAATGGAA	60.6/58.4	547/- bp
RP1 ribosomal protein 1 (Boe-gene# 19466)	AT1G43170	NW_019168660.1	#658/#659	Apogene29F/R	CTACAAGAACTGGGCCAAGTCC TGGATCTCCATCATGTGAGCCT	62.1/60.3	201/293 bp
SIRB sirohydrochlorin ferrochelatase B (Boe-gene# 19673, BA1G0192500)	AT1G50170	NC_027759.2	#660/#661	Apogene30F/R	CTCGGCCTCCATAATCTCCTCA GCACTTGTTTGTTCCAGCACAA	62.1/58.1	251/318 bp
PLT2 Integrase-type DNA-binding superfamily protein (Boe-gene# 19713)	AT1G51190	NW_019169374.1	#662/#663	Apogene31F/R	TGCTTGAATCTTGGACTTCTTCGA ACCGGAAAGTTGGTAGTAGTAGAGG	59.3/63.0	231/290 bp
B3 domain protein, putative (DUF313) (Boe-gene# 22991)	AT2G32645	NC_027768.2	#664/#665	Apogene32F/R	CGTCTTCGTTGATGGTTGGAAC GAGGGCAAAGCAGGAAAGACC	60.3/61.8	92/- bp
(Boe-gene# t-RNA_Leu_g_31)		NC_027763.2	#666/#667	Apogene33F/R	TCCTTGAGTACCCGTTGAAGGT TGGAAACAGAAGAACAGGCAGAT	60.3/58.9	133/- bp
PTB3 polypyrimidine tract-binding protein 3 (Boe-gene# BA1G0145300)	AT1G43190	NW_019168564.1	#668/#669	ApoCNV1F/R	TCGTAACAAACCTGACCATGCC ATTGCGGTTGAAGCGGTTCAG	60.3/59.8	269/314 bp
UGT80B1 (Boe-gene# BA1G0146100)	AT1G43620	NC_027773.2	#670/#671	ApoCNV2F/R	TGGGCGGTGATCCTCTTTCTAT TGGTATGGTGGAGCATAAGGACA	60.3/60.6	418/501 bp
n/a (Boe-gene#)	n/a	n/a	n/a	ApoCNV3F/R	n/a	n/a / n/a	n/a / n/a bp
D-aminoacid aminotransferase-like PLP (Boe-gene# BA1G0191700)	AT1G50110	NC_027762.2	#672/#673	ApoCNV4F/R	GGTCGTCTTCTGTCCTGTAAGC ATGTGAAATGGGAAGAGCTCGG	62.1/60.3	297/466 bp
n/a (Boe-gene#)	n/a	n/a	n/a	ApoCNV5F/R	n/a	n/a / n/a	n/a / n/a bp
FU kinase family with ARM repeat domain-containing protein (Boe-gene# BA1G0192900)	AT1G50240	NC_027759.2	#674/#675	ApoCNV6F/R	TGAGCAACCTTGTCCGAAACTC TGCAAATCTGGTGGTTCGAACA	60.3/58.4	392/292 bp
UBC20 ubiquitin-conjugating enzyme 20 (Boe-gene# BA1G0195700)	AT1G50490	NC_027761.2	#676/#677	ApoCNV7F/R	ACAGACTCTCACTCTCTTTCTCCA GTTGCTCCAAAGCTTAGCTGCT	61.0/60.3	458/428 bp
C2H2-like zinc finger protein (Boe-gene# BA1G0230100)	AT1G14580	NC_027774.2	#678/#679	ApoCNV8F/R	GACGGCGATGACGATGATTCAA ACGCTATCACTTCAGCATCTGG	60.3/60.3	406/376 bp
HXXXD-type acyl-transferase family protein (Boe-gene# BA1G0361600)	AT1G03940	NC_027774.2	#680/#681	ApoCNV9F/R	GCTTCAGACGCAATCACATGGA GTTGGATGACTGGACATGACCC	60.3/62.1	113/- bp
(Boe-gene# BA2G0076300)	n/a	NC_027774.2	#682/#683	ApoCNV10F/R	CCATGTTGCAGGAGATGGAGTC CCTCTGCTTCACCAACAAGTGT	62.1/60.3	383/- bp
S-adenosyl-L-methionine-dependent methyltransferases superfamily protein (Boe-gene# BA3G0117400)	AT5G57300	NC_027775.2	#684/#685	ApoCNV11F/R	GCTGAGTCCATTTCCAGGGATG TCCTAAACGCAACATCACCTGC	62.1/60.3	151/182 bp
SUMO2 small ubiquitin-like modifier 2 (Boe-gene# BA3G0141400)	AT5G55160	NC_027762.2	#686/#687	ApoCNV12F/R	CAGGAAGAAGACAAGAAGCCCG GACGCCCATCAAACAAGAAAGC	62.1/60.3	774/845 bp
F-box SKIP23-like protein (DUF295) (Boe-gene# BA3G0141500)	AT5G55150	NC_027759.2	#688/#689	ApoCNV13F/R	CCGTTTCGGTATGATCCTTGCA AGCTTCATGATTTGCTTACTCTCCC	60.3/61.3	323/356 bp
ribosomal protein L30 family protein (Boe-gene# BA3G0141600)	AT5G55140	NC_027761.2	#690/#691	ApoCNV14F/R	AGAAGCTGAAGCCAACCACAAA CCAAACCCATTGATACCAACATTTCT	58.4/60.1	135/246 bp
n/a (Boe-gene# BA3G0218600)	n/a	NC_027768.2	#692/#693	ApoCNV15F/R	CATTGGTCAAACACGGCGAATC ACATGTGTCCTGGGCAAAGAAC	60.3/60.3	94/- bp
Protein kinase family protein (Boe-gene# BA6G0231300)	AT5G18190	NC_027766.2	#694/#695	ApoCNV16F/R	ACCTTACGAGTGGCAAGTGTAC CCAAGCATGTCCATGACCTGAA	60.3/60.3	199/296 bp
Nucleotide-sugar transporter family protein (Boe-gene# BA7G0083800)	AT5G41760	NC_027760.2	#696/#697	ApoCNV17F/R	GAGATACCTACGTATACAACTGCAAGT GTGCGCTTATTTGTTATCCCTTCTATT	61.9/60.4	579/643 bp

Government of Saskatchewan
Ministry of Agriculture

ADF Project Progress Report Format

Table 10. Genotyping intergeneric F1 hybrid between recipient Brassica napus and apomictic Boechera sp. donors for transfer of apomixis candidate genes from apomixis factor list (AFL) 1. Discovered fragments from apomixis candidate genes present in intergeneric hybrid DNA are marked in yellow.

					B09658 fragments sizes in bp			B09659 fragments sizes in bp			B09735 fragments sizes in bp			B11100 fragments sizes in bp			B11033 fragments sizes in bp		
#	Apo marker primer #	Arabidopsis homolog	Brassica napus homolog	Exp. product length (B. napus ZS11/Boechera orthlog)	Boe ctrl	nap	nap extra	Boe ctrl	nap	nap extra	Boe ctrl	nap	nap extra	Boe ctrl	nap	nap extra	Boe ctrl	nap	nap extra
1	#602/#603	AT4G11480	NC_027764.2	411/364 bp	-	364	-	~411	~364	-	-	~411	-	-	~411	-	-	~411	-
2	#604/#605	AT3G06230	NW_019168955.1	456/440 bp	~456	~440	-	~500	~500	-	~440	~456	-	~440	~456	-	~440	~456	-
3	#606/#607	AT1G31130	NC_027765.2	-/108 bp	-	~108	-	-	~108	-	-	~108	-	-	~108	-	-	~108	-
4	#608/#609	AT5G45030	NC_027757.2	324/344 bp	324	344	-	~400	~300	-	~450	~350	-	~400	~324	-	~344	~324	-
5	#610/#611	AT3G60220	NC_027772.2	-/99 BP	1500	99	-	-	~99	-	-	~99	-	-	~99	-	-	~99	-
6	#612/#613	AT2G23840	NW_019168938.1	276/328 bp	328	276	~700	~300	~300	~700 (faint)	~328	~276	-	~328	~276	-	~328	~276	-
7	#614/#615	AT1G42440	NC_027762.2	203/117 bp	117	203	-	~117	~203	~117	~117	~203	-	~117	~203	-	~117	~203	-
8	#616/#617	AT1G42560	NC_027770.2	700/672 bp	672	700	-	~672	~700	-	~672	~700	-	~672	~700	-	~672	~700	-
9	#618/#619	AT1G42990	NC_027764.2	532/652 bp	-	652	-	-	~400	~500	~652	~532	-	-	~532	-	-	~532	(UNSPECIFIC
10	#620/#621	AT1G43130	NW_019168564.1	505/620 bp	620	505		~620	~500	-	~620	-	-	~620	~505	-	~620	~505	-
11	#622/#623	AT1G43245	NC_027765.2	373/591 bp	-	591	-	-	~373	-	-	~373	-	-	~373	-	-	~373	-
12	#624/#625	AT1G43580	NC_027768.2	431/355 bp	355	431		~355	~431	-	~355 + unspecific	~431	-	~355	~431	-	~355	~431	-
13	#626/#627	AT1G43620	NW_019169735.1	180/233 bp	233	180	-	~233	~180	~233	~233	~180	-	~233	~180	-	~233	~180	-
14	#628/#629	AT1G43630	NW_019168564.1	-/211 bp	-	211		-	~211	-	-	211	-	-	~211	-	-	~211	-
15	#630/#631	AT1G76900	NC_027759.2	-/216 bp	-	216		-	~216	-	-	216	-	-	~300	-	-	~216	-
16	#632/#633	AT1G43650	NC_027764.2	146/75 bp	~100	~100	-	~100	~100	-	~100	~100	-	~100	~100	-	~100	~100	-
17	#634/#635	AT1G50250	NC_027772.2	483/572 bp	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
18	#636/#637	AT1G01090	NC_027771.2	435/497 bp	~497	~435	-	~500	~500	-	~497	~435	-	~500	~500	-	~497	~435	-
19	#638/#639	AT5G19820	NC_027766.2	-/93 bp	-	~93	-	-	~93	-	~100	~100	-	-	~100	-	-	~93	-
20	#640/#641	AT1G42430	NC_027764.2	154/237 bp	-	-	-	~237	~154	~210	~237 + ~350 (unspecific?)	~154	-	~237	~154	-	-	~154	-
21	#642/#643	AT1G42470	NC_027760.2	360/418 bp	~418	~360	-	~400	~300	-	~418	~360	-	~418	~360	-	-	~400	-
22	#644/#645	AT1G42470	NC_027760.2	500/842 bp	~842	~500	-	~842	~500	~842	~842	~500	-	~842	~500	-	-	-	-
23	#646/#647	AT1G42480	NC_027774.2	433/775 bp	~775	~433	-	~775	~433	~775	~775	~433	-	~800	~433	-	~775	~500	-
24	#648/#649	AT1G42550	NC_027764.2	664/735 bp	~735	~664	-	~735	~664	-	~735	~664	-	~730	~670	-	~735	~664	-
25	#650/#651	n/a	NC_027764.2	356/493 bp	493	~400	-	-	-	-	~600	~400	-	~600-700	~500	~600-700	~700 (unspecific?)	-	-
26	#652/#653	AT1G42970	NC_027774.2	617/755 bp	~755	~617	-	-	-	-	-	~617	-	~755	~1000	-	~800	~700	-
27	#654/#655	AT1G43020	NW_019168588.1	407/338 bp	~700 (double band)	~550	-	-	~600	-	~600	~500	-	~400 (+~500,	~400	-	-	~500	-
28	#656/#657	AT5G64750	NC_027762.2	547/- bp	-	~700	-	-	-	-	~600 (unspecific?)	~547	-	-	-	-	-	~700	-
29	#658/#659	AT1G43170	NW_019168660.1	201/293 bp	~500	~300	-	-	~300	-	~293	~201	-	~400 (unspecific?)	~201	-	-	~200	-
30	#660/#661	AT1G50170	NC_027759.2	251/318 bp	~318	~755	-	-	~318	-	~318	~251	-	~350	~250-300	~600 (unspecific?)	~318	~251	~500 (unspecific?)
31	#662/#663	AT1G51190	NW_019169374.1	231/290 bp	~290	~231	-	-	-	-	~290	~231	-	~400	~350	-	~290	~231	-
32	#664/#665	AT2G32645	NC_027768.2	~-/92 bp	-	~90	-	-	~90	-	-	~92	-	-	~92	-	-	~92	-
33	#666/#667		NC_027763.2	~-/133 bp	-	~133	-	-	~133	-	-	~133	-	-	~133	-	unspecific	~133	-
34	#668/#669	AT1G43190	NW_019168564.1	269/314 bp	~314	~269	-	-	~269	-	~314	~269	-	~314	~269	-	~314	~269	-
35	#670/#671	AT1G43620	NC_027773.2	418/501 bp	~501	~418	-	-	~501	~800+1200 (unspecific)	~501	~418	-	~501	~418	-	~501	~418	-
36	n/a	n/a	n/a	n/a / n/a bp	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
37	#672/#673	AT1G50110	NC_027762.2	297/466 bp	~466	~297	-	-	unspecific	unspecific	~500	~400		~600	~500	-	~466	~297	~400
38	n/a	n/a	n/a	n/a / n/a bp	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
39	#674/#675	AT1G50240	NC_027759.2	392/292 bp	~292	~400	~440	~400	~500	-	~292	~392	-	~292	~450	-	~292	~392	-
40	#676/#677	AT1G50490	NC_027761.2	458/428 bp	~428	~458	-	-	~520	unspecific	~428	~458	-	~428	~458	-	~428	~458	-
41	#678/#679	AT1G14580	NC_027774.2	406/376 bp	~376	~406		-	~406	-	~376	~406	-	~450	-	-	~500	~500	-
42	#680/#681	AT1G03940	NC_027774.2	113/- bp	-	~113	-	-	~113	-	-	~113	-	-	~113	-	-	~113	-
43	#682/#683	n/a	NC_027774.2	383/- bp	-	~400	-	-	~400	-	-	~383	-	-	~600 (unspecific?)	-	-	~500	-
44	#684/#685	AT5G57300	NC_027775.2	151/182 bp	~200	~200	-	-	~200	-	~182	~151	-	~182	~151	-	~182	~151	-
45	#686/#687	AT5G55160	NC_027762.2	774/845 bp	~845	~774	~600	-	~774	-	~845	~774	~800 (faint)	~845	~774	-	-	-	-
46	#688/#689	AT5G55150	NC_027759.2	323/356 bp	~356	~323	-	-	~400	-	~356	~323	-	~356	~323	-	-	-	-
47	#690/#691	AT5G55140	NC_027761.2	135/246 bp	~246	~135	-	-	~246	-	~300 (+~135, unspecific?)	~135	-	~400 (unspecific?)	~135	-	-	-	-
48	#692/#693	n/a	NC_027768.2	94/- bp	~100	~100	-	~100	~1100 (faint)	-	~94	-	-	~100	~101	-	-	-	-
49	#694/#695	AT5G18190	NC_027766.2	199/296 bp	~296	~199	~300	~450 (faint/unspecific?)	~200	~300	~296	~199	-	~296	~199	~300	-	-	-
50	#696/#697	AT5G41760	NC_027760.2	579/643 bp	~800	~600	~900	-	~643	-	-	~579	-	~800 (unspecific)	~600	~1000 (unspecific?)	-	-	-

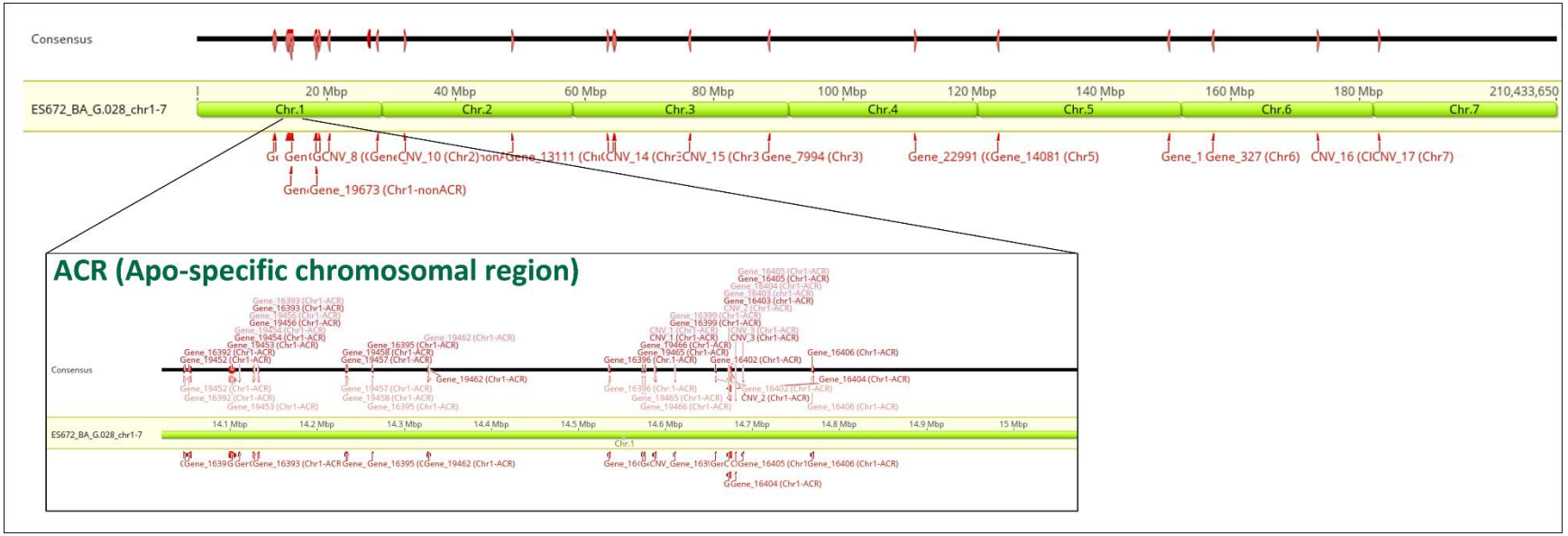

Figure 15. Distribution of candidate apomixis factors such as 17 genes with apo-specific CNVs and 32 genes with apo-specific InDels, from AFL1 (#ADF #20150123) on the reference genome *B. stricta* ES672. The apo-specific chromosomal region on chromosome 1 carries most apo-specific factors (6 CNV genes and 19 genes with 148 of 182 genic InDels). The AFL1 candidate factors were used to design InDel primers for *Boechera* and *B. napus* orthologs to test their introgression into intergeneric hybrids (Tables 9 and 10).

Interestingly, we can demonstrate that different subsets of AFL1 candidate factors (between 1 of 49 and 7 of 49) were introgression into the five different intergeneric F1 hybrid. To validate which genomic fragments (and if which apomictic factors) are present in the intergeneric F1 hybrid genomes we skim-sequenced the whole genomes of all five intergeneric F1 hybrid plants along with one native *B. napus* cv 'Golden' using a low coverage shotgun sequencing pipeline from Illumina (UC Davies DNAtech center; Table 11).

Table 11. Low coverage shotgun sequencing using Illumina NovaSeq 6000 platform and SHT-libraries prepared from five intergeneric F1 hybrids *B. napus* x *Boecher sp.* .

Sample	Taxon	Yield (Mbases)	% >= Q30 bases	Mean Quality Score	*B. napus* genome size (Davies et al. 2023)	Coverage (yield/total *B. napus* genome size)
B09658	*B. napus* x *Boechera retrofracta* JL101	11672	92.71	35.73	1120	10.4
B09659	*B. napus* x *Boechera retrofracta* JL101	12382	92.32	35.65	1120	11.1
B09735	*B. napus* x *Boechera retrofracta* JL101	11349	92.48	35.69	1120	10.1
B11033	*B. napus* x *Boechera stricta* ES649	13471	92.08	35.62	1120	12.0
B11100	*B. napus* x *Boechera stricta* ES649	13024	92.27	35.64	1120	11.6
B12854	*Brassica napus* cv 'Golden'	10765	91.76	35.55	1120	9.6

We have developed an SNP and InDel analyses pipeline (see **Fig. 2** in the Method section) using a mapping strategy of the genomic reads from the hybrids along with the reads form the native *B. napus* cv 'Golden' onto the published genome assembly of *B. napus* cv. 'Westar' (cv 'Golden' is not sequenced yet but is as well a spring-type oilseed rape such as cv 'Westar'; [43]). Currently, we are identifying the Boechera-specific sequence introgression sites in B. napus by mapping the sequenced hybrids along with the sequenced genome of *B. napus* cv 'Golden' onto two references: 1) the annotated whole-genome assembly of *B. napus* cv. 'Westar', and 2) the chromosome-level assembly of *Boechera stricta* ES649 (see section 9e), to retrieve sequence variant calls between references and each of the individual intergeneric hybrids (**see Fig. 16**).

Figure 16. IGV browser view showing a 17kb segment of mate-pair reads from intergeneric hybrids and recipient Canola cv 'Golden' mapping on annotated whole-genome assembly of *B. napus* cv. 'Westar' and sequence variant calls between reference and each of the individual intergeneric hybrids. A similar mapping is performed onto the chromosome-level assembly of *Boechera stricta* ES649 (**see section 9e**). Comparing both mappings is then used to identify *Boechera*-specific SNPs and Indels in the interspecific F1 hybrids.

Subsequently, we scored several phenotype characteristics of the F1 such as seed size and shape variation, flower shape, silique and leaf formation. We could not detect any specific deviations in the formation of siliques or leaves when comparing the validated F1 intergeneric hybrids to their native parental recipient. We discovered a significant difference in flower shape in two validated F1 intergeneric hybrids (B09658 and B09659) compared to negative F1 plants and the native cultivar (**Fig. 17**). Also, seed size of selfed validated F1 intergeneric hybrids (*i.e.* B09658 and B09659 together) does not differ significantly from native cultivar *B. napus* cv. 'Golden' after correction, but when seed size of selfed validated F1 intergeneric hybrid B09659 alone was analysed (*i.e.* most *Boechera* insertions) the seed size is significantly smaller compared to native cultivar *B. napus* cv. 'Golden' (**Figs. 18-20**, **Table 12**). Seed size of backcrosses between validated F1 intergeneric hybrids is significantly smaller compared with selfed seeds of the native *B. napus* cultivar 'Golden' and is comparable with seed size of *Boechera* (**Figs. 18-20**, **Table 12**). Similarly, we also concluded the collection of phenotype data from all backcrossed (with exception of 3/4 backcross lines) and selfed F2 plants. For example, the flow cytometric study of relative genome size variation between selfed intergeneric F2 hybrid progeny and recipient *B. napus* cv. 'Golden' revealed genome size variation within and across selfed F2 progeny of five intergeneric hybrid F2 lines between *Brassica napus* and apomictic diploid *Boechera sp.* (**Fig. 21**). Flow cytometric analysis of the nuclear DNA of leaves from selfed F1 progeny (*i.e.* F2 plants) shows a large spread of genome sizes within each line (between 32 and 70%; **Fig. 21**)). This variation can be caused by a presence/absence variation of chromosomal structures such as (a) insertion–deletion polymorphisms (indels), (b) copy number variation in protein-coding genes or rDNA copies, (c) or of other genomic repeats. These structural variations were likely caused by the hybridization event. Also, the average genome sizes differ significantly between the hybrid lines (between 1.062 and 1.302pg/1C) and spread between 0.818pg/1C and 1.629pg/1C across the lines deviating significantly from the native *B. napus* genome size of 1.157pg/1C; **Fig. 21**).

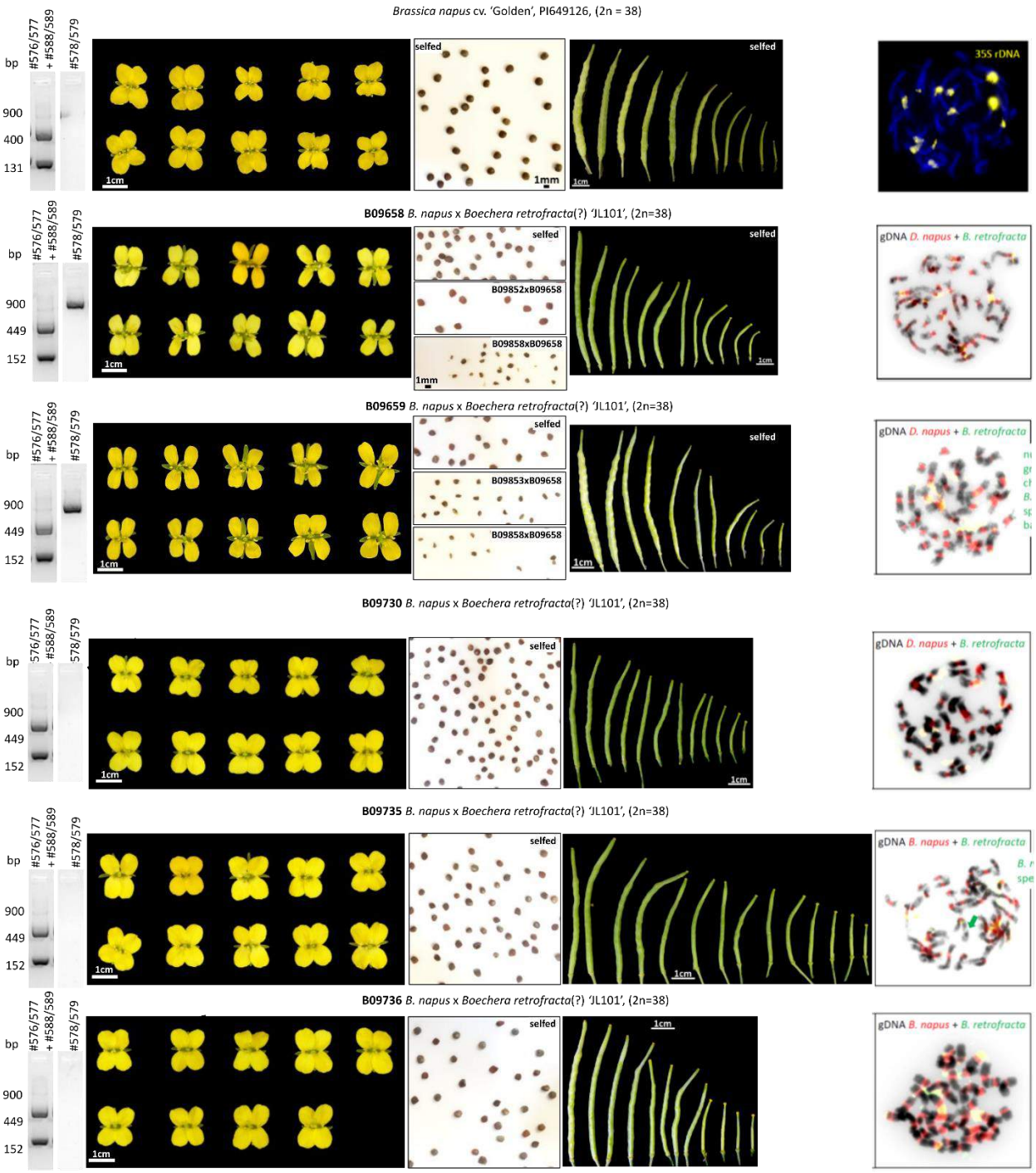

Figure 17 continued next page.

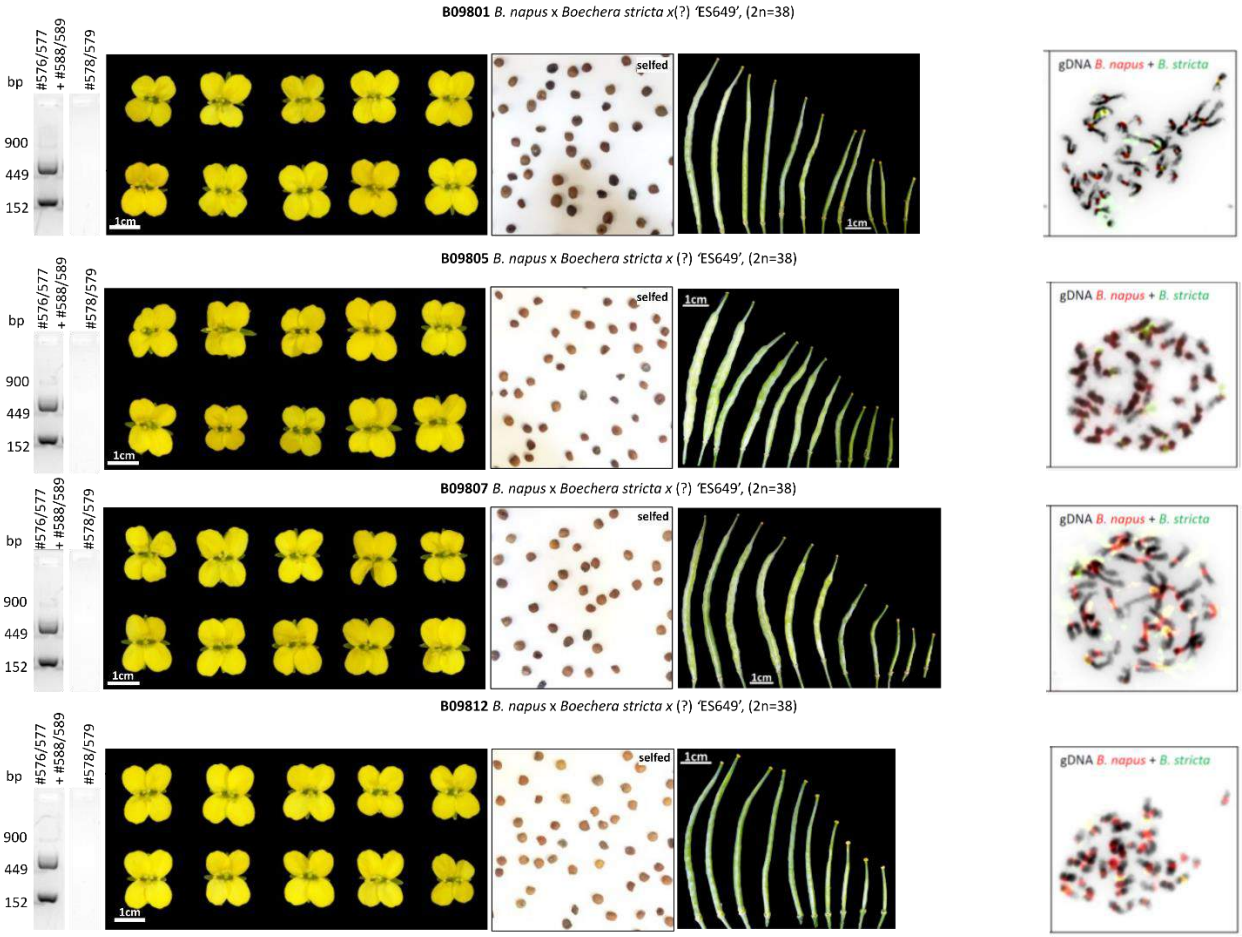

Figure 17. Phenotype scoring of parental recipient *Brassica napus* and hybrid F1 progeny with apomictic *Boechera retrofracta* 'JL101'. The flower morphology, seed size and shape, silique formation was analysed for the native cultivar Brassica napus cv 'Golden', three confirmed F1 hybrids (B09658, B09659, B09735) and six negative F1 progeny (B09730, B09736, B09801, B09805, B09807 and B09812). Validation results from the PCR-based screen and the genomic *in situ* hybridization assays are also given for each sample.

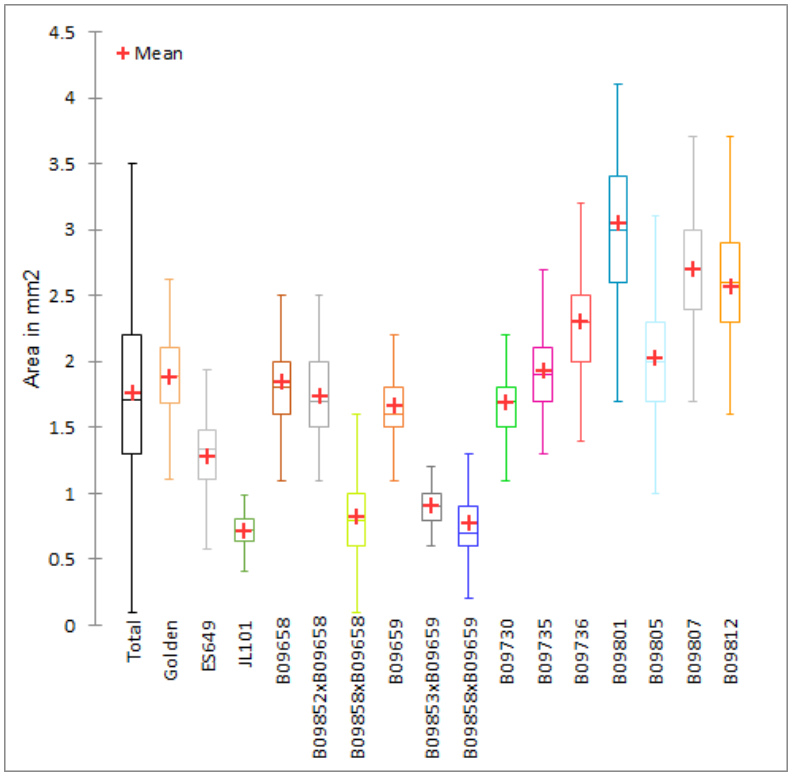

Figure 18. Seed size variation in parental recipient and donor lines, in their selfed and backcrossed progeny from validated and falsified intergeneric F1.

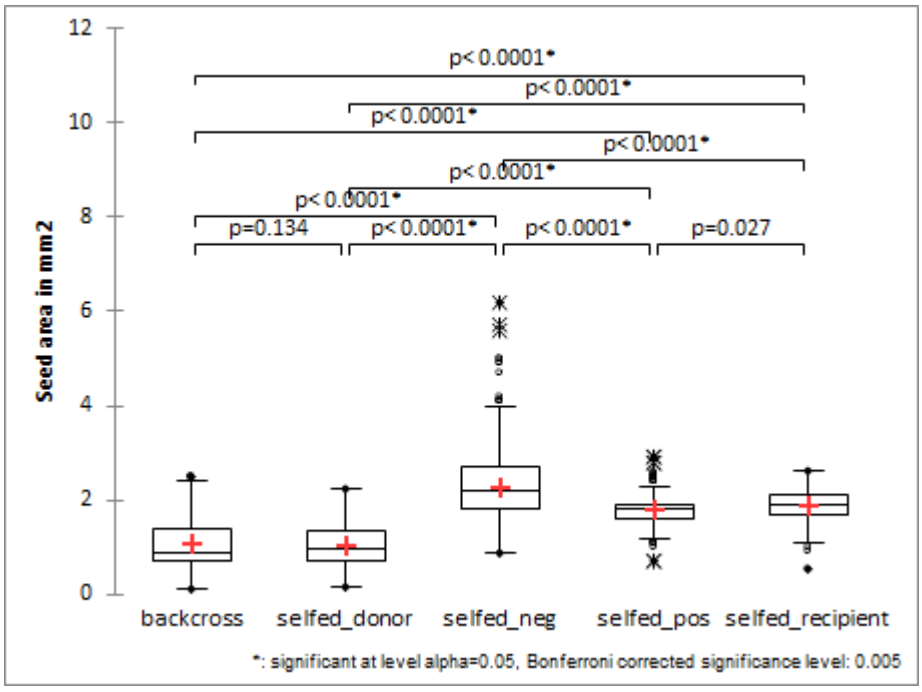

Figure 19. Multiple pairwise comparisons of seed size variation between cross types of recipients, donors and validated or falsified F1 hybrids using Dunn's procedure and Two-tailed test.

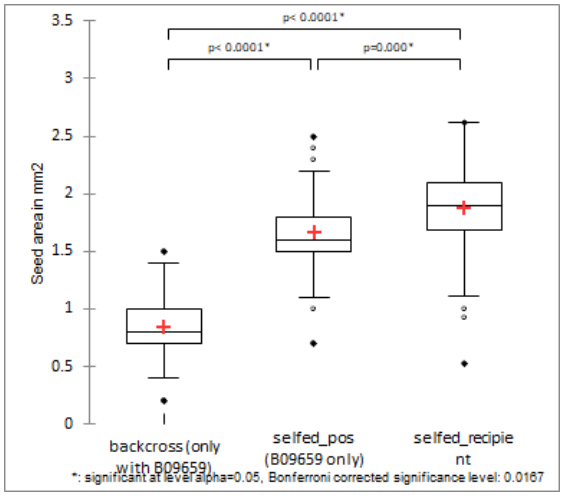

Figure 20. Multiple pairwise comparisons of seed size variation between cross types of hybrid B09659 using Dunn's procedure and Two-tailed test.

Table 12. Summary of significant seed size variation between different cross types.

Significant differences:					
	backcross (napusxMAS positive F1)	selfed_donor (ES649,JL101)	selfed_neg (MAS negative F1)	selfed_pos (MAS positive F1)	selfed_recipient (napus)
backcross (napusxMAS positive F1)	No	No	**Yes**	**Yes**	**Yes**
selfed_donor (ES649,JL101)	No	No	**Yes**	**Yes**	**Yes**
selfed_neg (MAS negative F1)	**Yes**	**Yes**	No	**Yes**	**Yes**
selfed_pos (MAS positive F1)	**Yes**	**Yes**	**Yes**	No	No
selfed_recipient (napus)	**Yes**	**Yes**	**Yes**	No	No

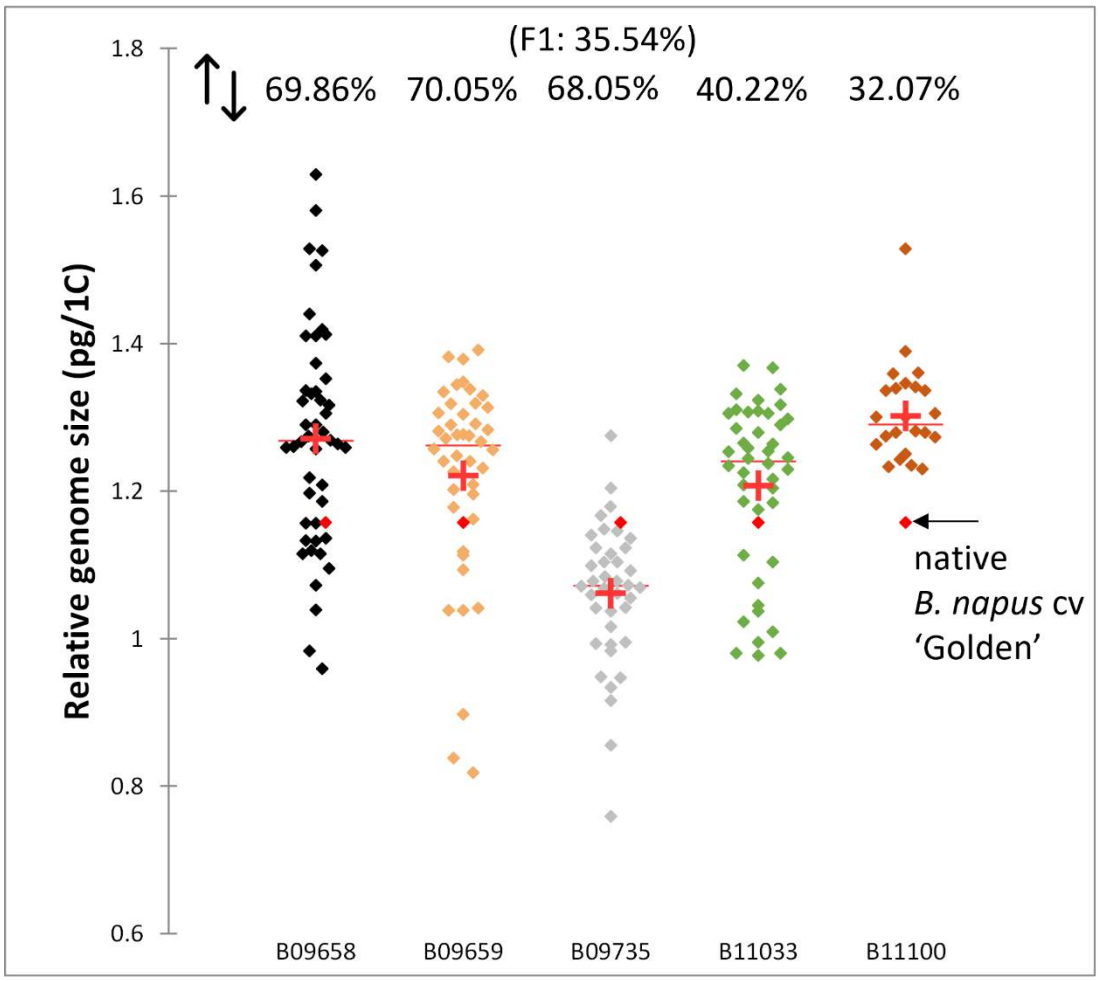

Figure 21. Genome size variation within and across selfed F2 progeny of five intergeneric hybrid F2 lines between *Brassica napus* and apomictic diploid *Boechera sp.*. Flow cytometric analysis of the nuclear DNA of leaves from selfed F1 progeny (*i.e.* F2 plants) shows a large spread of genome sizes within each line (between 32 and 70%). This variation can be caused by a presence/absence variation of chromosomal structures such as (a) insertion–deletion polymorphisms (indels), (b) copy number variation in protein-coding genes or rDNA copies, (c) or of other genomic repeats. These structural variations were likely caused by the hybridization event. Also, the average genome sizes differ significantly between the hybrid lines (between 1.062 and 1.302pg/1C) and spread between 0.818pg/1C and 1.629pg/1C across the lines deviating significantly from the native B. napus genome size of 1.157pg/1C.

e) *Identification of all apomixis factors in the backcrosses of de novo generated hybrid apomictic Boechera and Brassica crops*

For the identification of all apomixis factors in the produced backcrosses we require the mapping of their parental genome contribution and localization of parental genome fragments. Hence, the generation of *de novo* assembled and annotated genomes from all native parental *Boechera* accessions employed in the backcrossing assay is required (**Table 13**). Additionally, we included our bridge species from the wide intergeneric cross assay, *Diplotaxis erucoides*, in our sequencing sample set, which will facilitate the search for optimal genotyping marker

in our screens of potential hybrids with native apomictic *Boechera* donors as a publicly available genome of *Diplotaxis erucoides* is not available yet.

Due to the recurrent hybrid nature of *Boechera* (karyotype: 2n=7x=14, *i.e.* 7 bivalent chromosomes in sexuals and 14 univalents in apomicts) with large structural genome variation and regions with high levels of repeats our sequencing strategy involves the preparation of three different whole genome sequence libraries and their combination into a single de novo genome sequence assembly data pipeline (**see Method section Fig. 1**).

We prepared, (1) high molecular weight (HMW) DNA for the generation of ultra-long read sequence libraries using the Oxford Nanopore Technologies (ONT) sequencing platform, (2) highly pure DNA for the generation of a high-coverage Illumina sequence libraries, and (3) nuclei samples with crosslinked DNA for the generation of Hi-C sequence libraries. The ultra-long read ONT sequence libraries enable the correct alignment irrespective genome complexities, while the high-coverage Illumina sequence libraries was used for error correction of the ONT assemblies. The Hi-C libraries preserve long-range sequence and structure information prior to sequencing of sequences that are closer to each other in the 3D space but distal to each in linear distance and were used for Hi-C scaffolding of contigs for our genome assemblies to define chromosomes *de novo*.

We established a modified protocol of the Wizard® HMW DNA Extraction Kit which resulted in successful HMW DNA extractions with sufficient amounts and quality for all samples (**Fig. 22** and **Table 14**). Genomic DNA samples (approx. 20ng/ul) for the generation of a high-coverage Illumina sequence libraries were prepared from the same individuals used for HMW DNA extractions employing a slightly modified protocol for the DNAesy plant mini Kit (Qiagen, Cat#: 69106). All ONT and Illumina high-coverage DNA samples (**Table 13**) were sent to the DNAtech genome center at UC Davis and we have received the Oxford Nanopore Technology long-read and the Illumina 150bp paired-end read sequencing libraries for 10 parental backcrossing genomes and *Diplotaxis erucoides* (**Tables 14** and **15**).

We decided to apply the proximity ligation method from Arima Genomics Inc. (CA, USA) to our parental accessions from the backcross assay to generate Hi-C sequence libraries that capture the three-dimensional (3D) organizational structure and long-range, gene regulatory interactions of chromatin. This enables the preservation of long-range sequence and structure information prior to sequencing. We have established aliquots of plant nuclei with crosslinked DNA from all samples using the CelLytic Plant Nuclei Extraction Kit (Sigma-Aldrich). All samples were successfully extracted and delivered to our partner at NRC Saskatoon for the library construction using the Arima-HiC2.0 PlantTissue and Accel-NGS 2S Plus DNA Library Kit (Swift Biosciences) protocols. All libraries were successfully sequenced and integrated into our chromosome-level genome assembly pipeline (**see Method section Fig. 1**).

Table 13. High molecular weight DNA from *Boechera sp.* and *Diplotaxis erucoides* accessions for the use in Oxford Nanopore Technology (ONT) sequencing assays, Illumina deep sequencing of short reads and chromatine conformation capture sequencing (HiC).

					ONT			Illumina PE 150bp			HI-C (Illumina PE 150bp)			
Sample #*	**Plant I#**	**Accession #**	**Taxon**	**Mode of reproduction**	**Qubit avg of four measurements (ng/ul)**	**260/280**	**260/230**	**Qubit avg of four measurements (ng/ul)**	**260/280**	**260/230**	**Avg DNA size in bp**	**Amount pg/ul**	**Molarity (nmol/l)**	**qPCR (Stock nM)**
3	B07057	ES612	*Boechera stricta*	sexual	**55.1**	1.91	1.92	19.7	1.94	1.97	803	1990.1	4.40	3.87
3R	B07057	ES612	*Boechera stricta*	sexual	**62.7**	1.93	1.98							
6	B07024	ES854	*Boechera stricta*	sexual	**72.7**	1.87	2.04	22.2	1.84	2.04	879	1386.7	2.84	2.66
6R	B07024	ES854	*Boechera stricta*	sexual	**85.5**	1.91	1.95							
8	B07032	ES865	*Boechera stricta*	sexual	**50.5**	1.91	2	26.9	1.9	2.07	808	3158.5	7.00	7.41
8R	B07032	ES865	*Boechera stricta*	sexual	**40**	2.02	1.98							
19	B07178	ES649	*Boechera stricta*	Unbalanced apomictic	**45.5**	1.92	2.22	24.2	1.87	1.8	848	4590.3	9.66	13.13
19R	B07178	ES649	*Boechera stricta*	Unbalanced apomictic	**77.3**	1.9	1.85							
25	B08686	wildtype	*Diplotaxis erucoides*	sexual	**93.6**	1.92	2.28	21.4	1.86	2.08	885	4470.4	9.01	8.25
26	B08662	JL107	*B. stricta*	Unbalanced apomictic	**24.5**	1.92	2.28				811	2914.3	6.30	5.49
26R	B08662	JL107	*B. stricta*	Unbalanced apomictic	**152.3**	1.83	1.95	17.9	1.84	2.36				
29	B07608	JL106	*B. stricta x spatifolia*	Unbalanced apomictic	**61.7**	1.89	1.91	22.3	1.83	2.1	826	2899.2	6.18	7.07
32	B07435	JL103	*B. crandallii*	Unbalanced apomictic	**30.6**	1.93	2.08	27.5	1.82	2.34	842	3094.8	6.50	6.55
32R	B07435	JL103	*B. crandallii*	Unbalanced apomictic	**563.5**	1.93	2							
38	B07713	JL76	*B. spatifolia*	sexual	**45.4**	1.92	2.24	16.9	1.84	1.87	824	1542.7	3.31	3.47
38R	B07713	JL76	*B. spatifolia*	sexual	**36.4**	1.91	2.17							
44	B07736	JL78	*B. spatifolia*	sexual	**tbd**	1.98	2.13	11.2	1.87	1.66	803	1259.8	2.87	3.55
45	B07692	JL12	*B. crandallii*	sexual	**tbd**	1.89	1.93	21.5	1.86	1.88	780	2096.6	4.71	4.86

*Samples names marked with "R" exemplify repeat samples reserved for second flush.

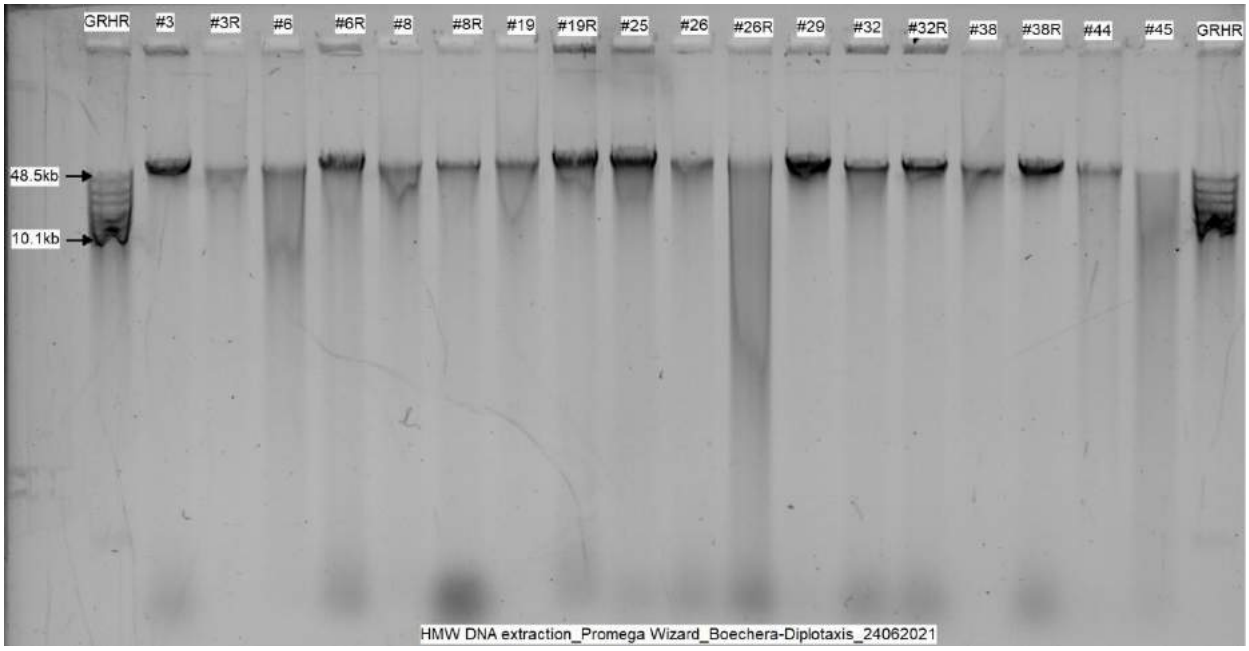

Figure 22. High molecular weight genomic DNA extracts of parental sexual and apomictic *Boechera* accessions used in the backcrossing assay and from *Diplotaxis erucoides*.

Table 14. Oxford Nanopore Technology sequencing libraries of native parental *Boechera* backcrossing lines and *Diplotaxis erucoides* .

Sample #	Accession	Barcode #	Pool #	Reads in mio.	Passed bases	Failed bases	Total bases in Gb	N50 in kb
26	JL107	1	1	23.69	51.77	40.22	101.02	13
32	JL103	2	1					
44	JL78	3	1					
3	ES612	4	2	29.98	51.19	32.89	90.8	7.96
6R	ES854	5	2					
9R	ES865	6	2					
19	ES649	7	3	31.25	64.15	40.52	112.57	8.86
38	JL76	8	3					
45R	JL12	9	3					
25	Diplotaxis	10	4	7.95	32.75	15.54	47.67	14.5
29	JL106	11	4					

Table 11. Illumina 150bp paired-end deep-coverage sequencing libraries of native parental *Boechera* backcrossing lines and *Diplotaxis erucoides* .

Sample	Accession	Approx. genome size in Mbp/1C	PF Clusters	% of the lane	Yield (Mbases)	% >= Q30 bases	Mean Quality Score	fold physical coverage estimate
IL19	ES649	220	59,268,942	1.8	17,899	91.58	35.52	81.4
IL25	Diplotaxis	500	66,564,389	2.03	20,102	92.22	35.66	40.2
IL26	JL107	220	70,228,592	2.14	21,209	91.96	35.6	96.4
IL29	JL106	220	60,976,699	1.86	18,415	91.76	35.57	83.7
IL3	ES612	220	62,458,893	1.9	18,863	92	35.61	85.7
IL32	JL103	220	63,823,317	1.94	19,275	91.76	35.57	87.6
IL38	JL76	220	66,720,897	2.03	20,150	91.84	35.58	91.6
IL44	JL78	220	68,867,192	2.1	20,798	91.33	35.48	94.5
IL45	JL12	220	69,094,412	2.1	20,867	91.42	35.48	94.9
IL6	ES854	220	72,788,603	2.22	21,982	91.81	35.58	99.9
IL9	ES865	220	55,640,952	1.69	16,804	91.86	35.59	76.4

We have employed all three library types (Illumina deep-coverage , ONT and Hi-C) for the successful generation of chromosome-level genome assemblies of all native parental *Boechera* lines (6 sexual recipient and 4 unbalanced apomictic donor lines) used in the backcrossing assay and are currently completing their genome feature annotation (*i.e.* genes, transposable elements, small RNAs, etc.).

Overall, the genome assemblies are of excellent quality and comparable with our reference genome assembly of *Boechera stricta* ES672 (ADF #20150123; **Table 16 and Fig. 23**). Currently, we are performing low coverage shotgun sequencing (Novaseq 6000 on Illumina SHT libraries, UC Davies DNAtech center) on 21 BC_3F1 individuals from 2 of 12 NILs.

Concluding remarks

Now that all data have been collected and the genomes of all parental lines are gold-standard assembled and annotated, we are in the process of performing high-powered bioinformatics and statistical analyses of the remaining NIL genomes to trace the parental genome contribution in our NILs and to detect genomic regions associated with the apomixis trait as well as confirmation of candidate factors from the AFL1 (**Fig. 24**).

Government of Saskatchewan
Ministry of Agriculture

ADF Project Progress Report Format

Table 16. Assessment of completeness of whole genome assemblies from native sexual recipient and apomictic donor lines of Boechera species used for generating near-siogenic lines (NILs).

BUSCO v5 *	Reference genome	Sexual lines						Apomictic lines			
	B. stricta ES672	*B. stricta* ES612	*B. stricta* ES854	*B. stricta* ES865	*B. crandallii* JL12#	*B. stricta x spatifolia* JL76	*B. stricta x spatifolia* JL78#	*B. stricta* ES649#	*B. crandallii* JL103	*B. stricta x spatifolia* JL106	*B. stricta* JL107#
Total # of core genes queried (i.e. Brassicales orthologs)	4596	4596	4596	4596	4596	4596	4596	4596	4596	4596	4596
# of core genes detected											
Complete	4539 (98.76%)	4494 (97.78%)	4477 (97.41%)	4472 (97.30%)	4538 (98.74%)	4522 (98.39%)	4403 (95.80%)	4531 (98.59%)	4530 (98.56%)	4517 (98.28%)	4524 (98.43%)
Complete + Partial	4550 (99.00%)	4507 (98.06%)	4491 (97.72%)	4484 (97.56%)	4547 (98.93%)	4534 (98.65%)	4419 (96.15%)	4538 (98.74%)	4540 (98.78%)	4528 (98.52%)	4538 (98.74%)
# of missing core genes	46 (1.00%)	89 (1.94%)	105 (2.28%)	112 (2.44%)	49 (1.07%)	62 (1.35%)	177 (3.85%)	58 (1.26%)	56 (1.22%)	68 (1.48%)	58 (1.26%)
Average # of orthologs per core genes	1.03	1.03	1.03	1.03	1.03	1.03	1.03	1.04	1.03	1.07	1.07
% of detected core genes that have more than 1 ortholog	2.78	2.69	2.7	2.93	2.71	2.7	2.61	2.98	2.85	6.15	6.83
Scores in BUSCO format:								0.985			
completed BUSCOs (C:)	98.70%	97.70%	97.40%	97.40%	98.80%	98.40%	95.80%	95.60%	98.60%	98.20%	98.40%
single copy BUSCOs (S:)	96.00%	95.10%	94.80%	94.50%	96.10%	95.70%	93.30%	2.90%	95.80%	92.20%	91.70%
duplicated BUSCOs (D:)	2.70%	2.60%	2.60%	2.90%	2.70%	2.70%	2.50%	0.20%	2.80%	6.00%	6.70%
fragmented BUSCOs (F:)	0.20%	0.30%	0.30%	0.30%	0.20%	0.30%	0.30%	1.30%	0.20%	0.20%	0.30%
missing BUSCOs (M:)	1.10%	2.00%	2.30%	2.30%	1.00%	1.30%	3.90%	459600.00%	1.20%	1.60%	1.30%
Total # of core genes queried (n:)	4596	4596	4596	4596	4596	4596	4596		4596	4596	4596
Length Statistics and Composition:											
# of sequences	7	7	7	7	7	7	7	7	7	7	7
Total length (nt)	210433650	191408699	184128410	183146272	192620646	180015912	166269074	196773446	197611246	214446332	215256627
Longest sequence (nt)	33561924	31049017	30534151	30365500	31851292	28082500	27931047	32072999	31473529	34027462	34615755
Shortest sequence (nt)	28423811	22835705	21548024	21886443	23764703	21964905	19739609	22760224	24300107	25672536	23891764
Mean sequence length (nt)	30061950	27344100	26304059	26163753	27517235	25716559	23752725	28110492	28230178	30635190	30750947
Median sequence length (nt)	29531193	27265270	26128158	25951836	27572208	26101000	23576209	28252849	28324657	30718300	30278307
N50 sequence length (nt)	29531193	27265270	26128158	25951836	27572208	26101000	23576209	28252849	28324657	30718300	30278307
L50 sequence count	4	4	4	4	4	4	4	4	4	4	4
Base composition (%)	A:29.64,T:29.62,G:,C:16.75	A:31.95,T:31.92,G:18.01,C:18.03	A:32.00,T:31.99,G:17.93,C:17.96	A:32.03,T:32.01,G:17.91,C:17.94	A:32.03,T:32.01,G:17.95,C:17.95	A:32.09,T:32.07,G:17.89,C:17.93	A:32.16,T:32.13,G:17.78,C:17.77	A: 31.99, T: 31.98, G: 18.00, C: 17.99	A:32.06,T:32.03,G:17.93,C:17.93	A:31.94,T:31.97,G:17.99,C:18.00	A:32.01,T:31.98,G:17.95,C:17.96
# of gaps (>=5 N's)	97	348	427	441	214	86	521	139	188	371	407
GC-content (%)	36.11	36.08	35.93	35.88	35.92	35.82	35.61	36	35.88	36.03	35.95
# of sequences containing non-ACGTN (nt)	0	0	0	0	0	0	0	0	0	0	0

*The completeness of genome assemblies was assessed using the BUSCO v5 pipeline of the gVolante software (Nishimura O, Hara Y, Kuraku S. Bioinformatics. 2017 Nov 15;33(22):3635-3637.). The ortholog sets of 4596 Brassicales genes from OrthoDB v10 were used for these BUSCOs.

BUSCOs were calculated from incompleteletly curated assemblies using XYZ_curated_HiC.fasta files.

HiC contact maps of *Boechera* genome assemblies

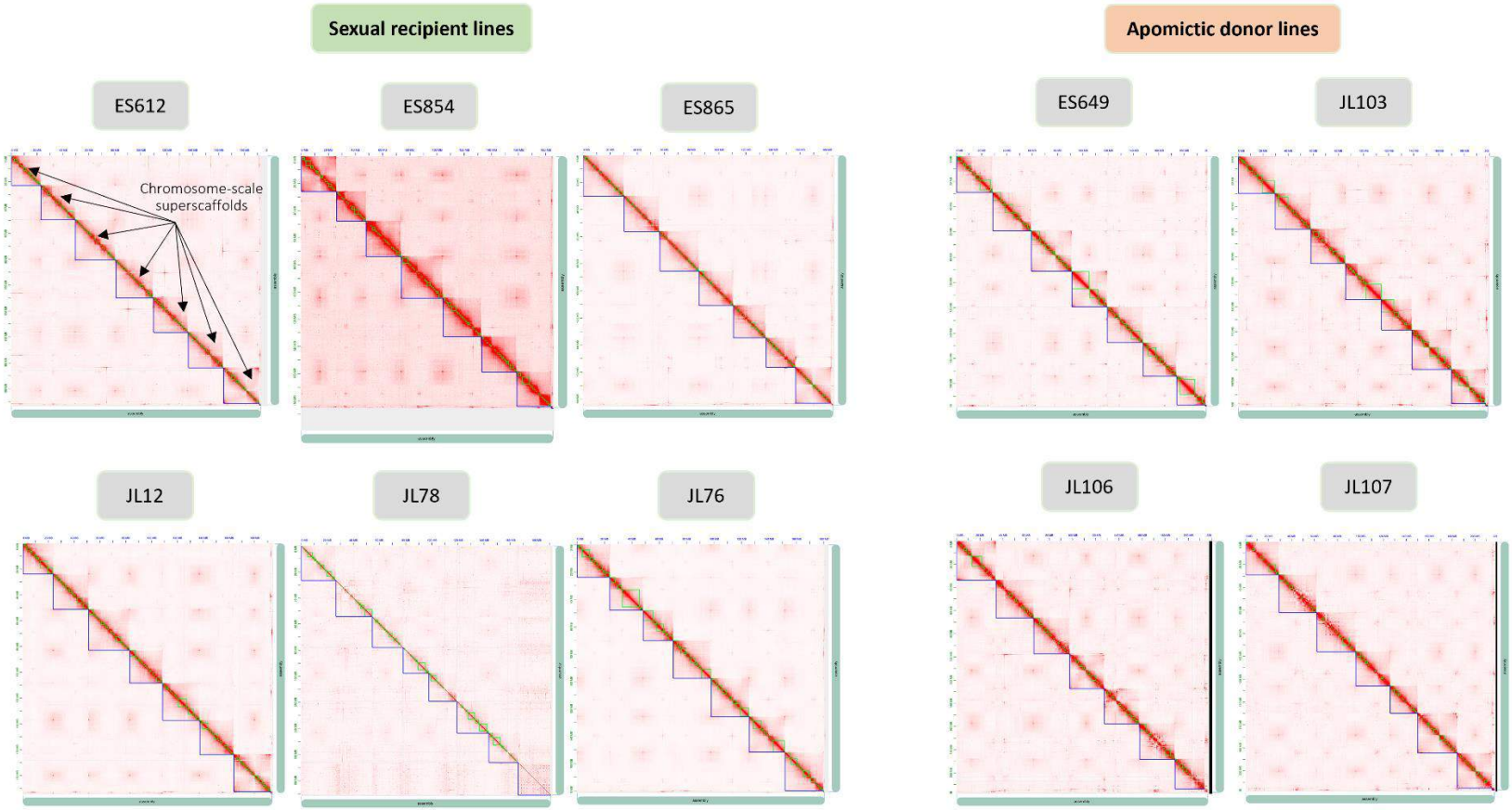

Figure 23. HiC contact maps of chromosome-level *Boechera* genome assemblies from six native sexual recipient and four apomictic donor lines. The is the first chromosome-level assembly of an apomictic *Boechera* genome and the first set of multiple chromosome-level *Boechera* genome assemblies.

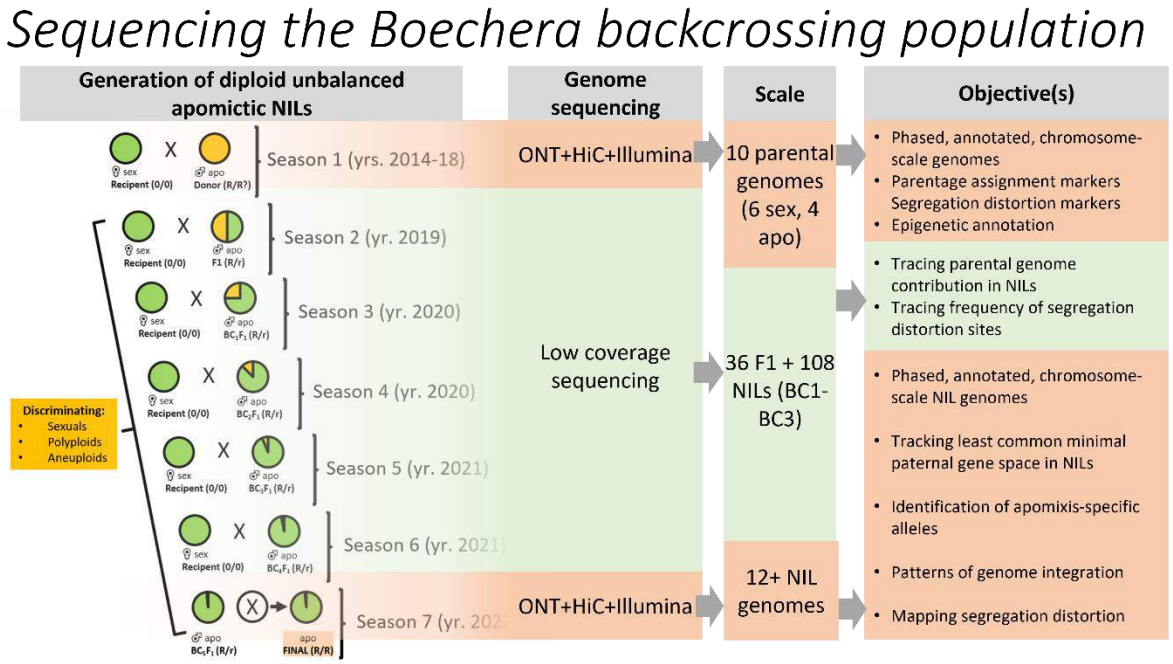

Figure 24. Concept for the identification of genetic factors for apomixis in near-isogenic lines of the apomictic model plant *Boechera.*

10. Conclusions and Recommendations: *Highlight significant conclusions based on the findings of this project, with emphasis on the project objectives specified above. Provide recommendations for the application and adoption of the project findings. (**Maximum of 500 words**)*

Utilizing our previously synthesized intra- and interspecific unbalanced apomictic hybrids for a backcrossing array to build a "magnifying glass" for paternal genomic regions that are least common in the final backcrosses was a full success. In some NILs genomic drag and other genetic constraints impede the crossing success to a certain degree which strongly increases the crossing effort necessary to retrieve sufficient progeny for their use as validated donors in the next generation. Nonetheless, the bulk of all crosses is concluded (8 of 12 NILs in BC3) leaving only 4 of 12 NILs for the remaining crossing to generate BC3F1 progeny. Not all NILs are eventually necessary to retrieve conclusive results when combining the NIL and parental genomes for the analysis on discovering least common paternal gene spaces in the NILs, but with more lines present the precision of the apomixis marker identification will increase. Therefore, we are currently already sequencing the first 2 of 12 NILs that are completed and funnel them into our established SNP and InDel analysis pipeline while the remaining NILs will be added successively.

A prime achievement of this work is the first chromosome-level assembly of an apomictic genome and the generation of a full set of 10 *Boechera* genomes (6 sexual recipient and 4 unbalanced apomictic) which has not been published by any competitor yet. Apomictic genomes are extremely difficult to assembly due to increased levels of heterozygosity and large fractions of repeats and structural variation in their genome. Hence, the selected state-of-the-art sequencing strategy utilizing three library types (Illumina deep-coverage, ONT and Hi-C) was the correct choice. These genomes (from parents and NILs) are extremely valuable as they allow us not only to identify parental SNP analyses in NILs for the establishment of an independent apomixis factor list (validation list =AFL2) but also, they are required to cross-validate any external apomixis factor lists (*e.g.* ADF #20150123) through segregation analysis towards establishing the refined apomixis factor list. This refined apomixis factor list should be used in a study to

provide functional proof of the identified genetic factors directly in Canola. We have provided a letter of intent with a concept how the result of the current study can by applied in this direction (ADF LOI #20230285).
A surprising and exceptional achievement of the present study is the generation of hybrid Canola with two different apomictic *Boechera* species (*B. stricta* ES649 and *B. retrofracta* JL101). The importance of these 5 hybrid lines generated here cannot be overstated as they indicate that apomixis factors can be indeed integrated into this crop. Direct incrosses of close relative apomicts such as *Boechera* may not directly lead to the expression of the apomixis trait, but these hybrids should be prime candidates for studying the geno- and phenotypic effects of the apomictic and non-apomictic sequences in the hybrid progeny in relation to the number, size, and location of these introgressions. In addition, the present study provides various bioinformatic and wet lab protocols that can be applied in similar studies.

11. Is there a need to conduct follow up research? *Detail any further research, development and/or communication needs arising from this project*.

Yes. The data generated here, provide the crucial fundament to refine independently developed apomixis factor lists into a single minimum core set of apomixis factors by combining the candidates from the current approach with those from ADF #20150123. We envisage a follow up research (see our submission ADF LOI #20230285) that will apply our previously gained knowledge on the genetic control of apomixis (ADF#20150123 and ADF#20180141) to provide functional proof of the identified genetic factors directly in Canola. This will be the last step before the synthesis of a modular species-independent "apomixis factor cassette" to advance breeding efforts in Canada's major crops beyond canola.

12. Patents/ IP generated/ commercialized products: *List any products developed from this research.*

None so far.

13. List technology transfer activities: *Include presentations to conferences, producer groups or articles published in science journals or other magazines*.

The work described here is currently in preparation for publication in which ADF will clearly be acknowledged. Furthermore, this work was presented with acknowledgment of the support of ADF as follows:

Posters

1. Mau. M., *et al. "Preserving hybrid vigour through a novel apomixis breeding strategy in Brassica crops"*, Science Day, 15th November 2019, University of Saskatchewan, Sk, Canada

Presentation

1. Mau. M., *"Understanding apomixis in the model Boechera (Brassicaceae): From functional analyses of candidate genes to tracing its mode of inheritance"*, PAG30 January 13-18 2023, invited presentation at the Apomixis workshop, 13th January 2023, San Diego, CA, USA

Media Interviews and Stories on Apomixis Breeding Research:

1. Hein, T., (2019) If Canola were an App, it Would Need an Update, *Germination Magazine*, April 19
2. Pratt, S. (2019) Effort to lock in hybrid vigour makes progress, *The Western Producer*, February 28
3. Isaacs, J. (2019) Apomixis: is the 'Holy Grail' of planting breeding within reach? *Top Crop Manager* WTCM6.7 July 12

Publications related to this work that were published during the term of the grant:

1. Mau, M., et al. (2021) "The spread of infectious asexuality through haploid pollen." New Phytologist **230**: 804-820.

2. Mau M*, et al.* (2022) "Evolution of an Apomixis-Specific Allele Class in Supernumerary Chromatin of Apomictic *Boechera*." Frontiers in Plant Science **13**, doi: 10.3389/fpls.2022.890038.

14. List any industry contributions or support received.

SaskCanola has supported this research with 62'400CAD plus overhead at a maximum rate of 15%.

15. Acknowledgements. *Include actions taken to acknowledge support by the Ministry of Agriculture and the Canada-Saskatchewan Growing Forward 2 bilateral agreement (for projects approved during 2013-2017) or Canadian Agriculture Partnership (For projects approved beyond 2017).*

The authors thank Rick Goertzen, Parvaneh Hashemi, Md. Nazmul Hasan and Hong Wang for their technical support. This work was cofunded by a grant from the Global Institute of Food Security at the University of Saskatchewan to Timothy F. Sharbel.

16. Appendices: *Include any additional materials supporting the previous sections, e.g. detailed data tables, maps, graphs, specifications, literature cited (Use a consistent reference style throughout).*

References

1. Hanna WW, Bashaw EC: **Apomixis: its identification and use in plant breeding.** *Crop Science* 1987, **27**(6):1136-1139.
2. Koltunow AM, Bicknell RA, Chaudhury AM: **Apomixis: molecular strategies for the generation of genetically identical seeds without fertilization.** *Plant Physiology* 1995, **108**:1345-1341 1352.
3. Ma X, Mau M, Sharbel TF: **Genome Editing for Global Food Security**. *Trends in Biotechnology* 2018, **36**(2):123-127.
4. Nogler GA: **Gametophytic apomixis**. In: *In Embryology of Angiosperms.* Edited by In: Johri BM. Berlin: Springer-Verlag; 1984: 475-518.
5. Savidan Y: **Embryological Analysis of Facultative Apomixis in Panicum maximum Jacq.** *Crop Science* 1982, **22**:467-469.
6. Ozias-Akins P, Roche D, Hanna WW: **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.** *Proceedings of the National Academy of Sciences* 1998, **95**:5127–5132.
7. Bicknell RA, Borst NK, Koltunow AM: **Monogenic inheritance of apomixis in two *Hieracium* species with distinct developmental mechanisms.** *Heredity* 2000, **84**(2):228-237.
8. Akiyama Y, Hanna WW, Ozias-Akins P: **High-resolution physical mapping reveals that the apospory-specific genomic region (ASGR) in *Cenchrus ciliaris* is located on a heterochromatic and hemizygous region of a single chromosome**. *Theoretical and Applied Genetics* 2005, **111**(6):1042-1051.
9. Calderini O, Chang SB, De Jong H, Busti A, Paolocci F, Arcioni S, De Vries SC, Abma-Henkens MH, Lankhorst RMK, Donnison IS: **Molecular cytogenetics and DNA sequence analysis of an apomixis-linked BAC in *Paspalum simplex* reveal a non pericentromere location and partial microcolinearity with rice**. *Theoretical and Applied Genetics* 2006, **112**(6):1179-1191.
10. Catanach AS, Erasmuson SK, Podivinsky E, Jordan BR, Bicknell R: **Deletion mapping of genetic regions associated with apomixis in *Hieracium***. *Proceedings of the National Academy of Sciences of the United States of America* 2006, **103**(49):18650-18655.
11. Mau M, Corral JM, Vogel H, Melzer M, Fuchs J, Kuhlmann M, de Storme N, Geelen D, Sharbel TF: **The conserved chimeric transcript *UPGRADE-2* is associated with unreduced pollen formation and is exclusively found in apomictic *Boechera* species.** *Plant Physiology* 2013, **163**(4):1640-1659.
12. Corral JM, Vogel H, Aliyu OM, Hensel G, Thiel T, Kumlehn J, Sharbel TF: **A conserved apomixis-specific polymorphism is correlated with exclusive DEDDh exonuclease expression in premeiotic ovules of apomictic *Boechera*.** *Plant Physiology* 2013, **163**(4):1660 - 1672.

13. Vijverberg K, Milanovic-Ivanovic S, Bakx-Schotman T, van Dijk PJ: **Genetic fine-mapping of DIPLOSPOROUS in *Taraxacum* (dandelion; Asteraceae) indicates a duplicated DIP-gene**. *BMC plant biology* 2010, **10**(1):154.
14. Ozias-Akins P, van Dijk P: **Mendelian genetics of apomixis in plants.** *Annu Rev Genetics* 2007, **41**:509–537.
15. Matzk F, Meister A, Schubert I: **FC analysis on seeds of mono & dicots.** *The Plant Journal* 2000, **21**(1):97-108.
16. Matzk F, Prodanovic S, Bäumlein H, Schubert I: **The inheritance of apomixis in *Poa pratensis* confirms a five locus model with differences in gene expressivity and penetrance**. *The Plant Cell* 2005, **17**:13-24.
17. Mogie M: **The Evolution of Asexual Reproduction in Plants.** London: Chapman & Hall; 1992.
18. Spillane C, Steimer A, Grossniklaus U: **Apomixis in agriculture: the quest for clonal seeds.** *Sexual Plant Reproduction* 2001, **14**:179–187.
19. Grossniklaus U, Nogler GA, van Dijk PJ: **How to avoid sex: the genetic control of gametophytic apomixis.** *The Plant Cell* 2001:1491-1497.
20. Conner JA, Ozias-Akins P: **Apomixis: engineering the ability to harness hybrid vigor in crop plants**. In: *Plant Germline Development: Methods and Protocols, Methods in Molecular Biology.* Edited by Schmidt A, vol. 1669: Springer Science+Business Media LLC; 2017.
21. Dujardin M, Hanna WW: **Crossability of pearl millet with wild *Pennisetum* species**. *Crop Science* 1989, **29**(1).
22. Leblanc O, Grimanelli D, Gonzfilez-de-Leon D, Savidan Y: **Detection of the apomictic mode of reproduction in maize-*Tripsacum* hybrids using maize RFLP markers.** *Theoretical and applied genetics* 1995, **90**(7/8):1198-1203.
23. Liu ZW, Wang RRC, Carman JG: **Hybrids and backcross progenies between wheat (*Triticum aestivum* L.) and apomictic Australian wheatgrass (*Elymus rectisetus* (Nees in Lehm.) A. Loeve and Connor): karyotypic and genomic analyses.** *Theoretical and Applied Genetics* 1994, **89** (5):599-605.
24. Morgan RN, Ozias-Akins P, Hanna WW: **Seed set in an apomictic BC3 pearl millet**. *International Journal of Plant Sciences* 1998, **159**(1):89-97.
25. Böcher TW: **Cytological and embryologal studies in the amphiapomictic *Arabis holboellii* complex.** *Biologiske Skrifter / Kongelige Danske Videnskabernes Selskab* 1951, **6**(7):1-59.
26. Böcher TW: **Experimental taxonomical studies in the *Arabis holboellii* complex.** *Svensk Botanisk Tidskrift* 1954, **48**:31-44.
27. Al-Shehbaz IA, Windham M: **Brassicaceae**. In: *Flora of North America North of Mexico.* Edited by Committee FoNAE, vol. 7. New York and Oxford: Oxford University Press; 2010: 224–746.
28. Mau M, Liiving T, Fomenko L, Goertzen R, Paczesniak D, Boettner L, Sharbel TF: **The spread of infectious asexuality through haploid pollen** *New Phytologist* submitted.
29. Aliyu OM, Schranz ME, Sharbel TF: **Quantitative variation for apomictic reproduction in the genus *Boechera* (Brassicaceae)**. *American Journal of Botany* 2010, **97**(10):1719-1731.
30. Voigt M-L, Melzer M, Rutten T, Mitchell-Olds T, Sharbel TF: **Gametogenesis in the apomictic *Boechera holboellii* complex: the male perspective**. In: *In Apomixis: Evolution, Mechanisms and Perspectives* Edited by Hörandl E, Grossniklaus U, van Dijk PJ, Sharbel TF. Ruggell: A.R.G. Gantner Verlag; 2007: 235-257.
31. Johnston SA, den Nijs TPM, Peloquin SJ, Hanneman JRE: **The significance of genic balance to endosperm development in interspecific crosses**. *Theoretical and Applied Genetics* 1980, **57**:5-9
32. Dresselhaus T, Carman JG, Savidan Y: **Genetic engineering of apomixis in sexual crops: a critical assessment of the apomixis technology.** In: *The flowering of apomixis: from mechanisms to genetic engineering.* Edited by Savidan Y, Carman JG, Dresselhaus T. Mexico, D.F.: CIMMYT,IRD, European Commission DG VI (FAIR); 2001.
33. Mau M, Lovell JT, Corral JM, Kiefer C, Koch MA, Aliyu OM, Sharbel TF: **Hybrid apomicts trapped in the ecological niches of their sexual ancestors.** *Proceedings of the National Academy of Sciences* 2015, **112**(18):2357-2365.
34. Greilhuber J, Ebert I: **Genome sized variation in *Pisum sativum***. *Genome* 1994, **37**:646-655.
35. França-Neto J, Krzyzanowski F: **Tetrazolium: an important test for physiological seed quality evaluation**. *Journal of Seed Science* 2019, **41**:359-366.

36. MacCarthy T, Bergman A: **Coevolution of robustness, epistasis, and recombination favors asexual reproduction**. *Proceedings of the National Academy of Sciences of the United States of America* 2007, **104**(31):12801-12806.
37. Haig D, Westoby M: **Genomic imprinting in endosperm: its effect on seed development in crosses between species, and between different ploidies of the same species, and its implications for the evolution of apomixis.** *Philosophical Transactions of the Royal Society B: Biological Sciences* 1991, **333**:1–13.
38. Mau M, Mandáková TM, Ma X, Ebersbach J, Zou L, Lysak MA, Sharbel TF: **Evolution of an Apomixis-Specific Allele Class in Supernumerary Chromatin of Apomictic Boechera**. *Frontiers in Plant Science* 2022, **13**.
39. Mau M, Liiving T, Fomenko L, Goertzen R, Paczesniak D, Böttner L, Sharbel TF: **The spread of infectious asexuality through haploid pollen**. *New Phytologist* 2021, **230**:804-820.
40. Vyas P, Prakash S, Shivanna KR: **Production of wide hybrids and backcross progenies between Diplotaxis erucoides and crop brassicas**. *Theoretical and Applied Genetics* 1995, **90**(3):549-553.
41. Garg H, Banga S, Bansal P, Atri C, Banga SS: **Hybridizing Brassica rapa with wild crucifers Diplotaxis erucoides and Brassica maurorum**. *Euphytica* 2007, **156**:417–424.
42. Ringdahl EA, McVetty PBE, Sernyk JL: **INTERGENERIC HYBRIDIZATION OF DIPLOTAXIS SPP. WITH Brassica napus: A SOURCE OF NEW CMS SYSTEMS?** *Canadian Journal of Plant Science* 1987, **67**(1):239-243.
43. Song J-M, Guan Z, Hu J, Guo C, Yang Z, Wang S, Liu D, Wang B, Lu S, Zhou R *et al*: **Eight high-quality genomes reveal pan-genome architecture and ecotype differentiation of Brassica napus**. *Nature Plants* 2020, **6**(1):34-45.