CORRECTIONS

Vol. 144: 286–298, 2007

Schranz M.E., Windsor A.J., Song B., Lawton-Rauh A., and Mitchell-Olds T. Comparative Genetic Mapping in *Boechera stricta*, a Close Relative of Arabidopsis.

In Figure 1 of this article (p. 287), the base chromosome number for Halimolobodeae is incorrectly listed as $x = 7$. The correct base chromosome number is $x = 8$.

Comparative Genetic Mapping in *Boechera stricta*, a Close Relative of Arabidopsis[1][C][W][OA]

M. Eric Schranz[2]*, Aaron J. Windsor, Bao-hua Song, Amy Lawton-Rauh, and Thomas Mitchell-Olds

Department of Biology, Duke University, Durham, North Carolina 27708 (M.E.S., A.J.W., B.-h.S., T.M.-O.); and Department of Genetics and Biochemistry, Clemson University, Clemson, South Carolina 29634 (A.L.-R.)

The angiosperm family Brassicaceae contains both the research model Arabidopsis (*Arabidopsis thaliana*) and the agricultural genus *Brassica*. Comparative genomics in the Brassicaceae has largely focused on direct comparisons between Arabidopsis and the species of interest. However, the reduced genome size and chromosome number ($n = 5$) of Arabidopsis complicates comparisons. Arabidopsis shows extensive genome and chromosome reshuffling compared to its close relatives *Arabidopsis lyrata* and *Capsella rubella*, both with $n = 8$. To facilitate comparative genomics across the Brassicaceae we recently outlined a system of 24 conserved chromosomal blocks based on their positions in an ancestral karyotype of $n = 8$, rather than by their position in Arabidopsis. In this report we use this system as a tool to understand genome structure and evolution in *Boechera stricta* ($n = 7$). *B. stricta* is a diploid, sexual, and highly self-fertilizing species occurring in mostly montane regions of western North America. We have created an F_2 genetic map of *B. stricta* based on 192 individuals scored at 196 microsatellite and candidate gene loci. Single-nucleotide polymorphism genotyping of 94 of the loci was done simultaneously using an Illumina bead array. The total map length is 725.8 cM, with an average marker spacing of 3.9 cM. There are no gaps greater than 19.3 cM. The chromosomal reduction from $n = 8$ to $n = 7$ and other genomic changes in *B. stricta* likely involved a pericentric inversion, a chromosomal fusion, and two reciprocal translocations that are easily visualized using the genomic blocks. Our genetic map will facilitate the analysis of ecologically relevant quantitative variation in *Boechera*.

Comparative genetic mapping between related organisms within a phylogenetic framework is a powerful method for understanding genome evolution. Comparative mapping in the grass family (Poaceae) has been successful in detecting collinear genomic regions between a number of domesticated cereal and forage crops, leading to the formulation of the crop circle with rice (*Oryza sativa*) at the center (Moore et al., 1995; Devos, 2005). Rice was selected as the reference point because of its small genome and vast genomic resources, and not because it was phylogenetically well positioned to facilitate comparisons within the family. An analogous situation occurs in the dicot family Brassicaceae, which contains both the model species Arabidopsis (*Arabidopsis thaliana*) as well as the domesticated *Brassica* species. To date, most comparative genomics in the Brassicaceae has largely focused on direct comparisons between Arabidopsis and the species of interest. However, several of the factors that made Arabidopsis ideal for genome sequencing, particularly its reduced genome size and chromosome number (157 Mb, $n = 5$; AGI, 2000; Johnston et al., 2005), complicate its use as a standard in comparative genomics. Recent phylogenetic results have demonstrated that genome and chromosome reduction in Arabidopsis are derived characteristics from its close relatives with $n = 8$.

In a recent taxonomic reclassification of the Brassicaceae based on molecular phylogenetic results (Fig. 1), Arabidopsis and its closest relatives, including the $n = 8$ species *Arabidopsis lyrata* and *Capsella rubella*, were placed within the tribe Camelineae (Al-Shehbaz et al., 2006; Beilstein et al., 2006). Comparative genetic mapping between Arabidopsis, *A. lyrata*, and *C. rubella* found that genome organization of *A. lyrata* and *C. rubella* is largely conserved, and that the $n = 5$ genome of Arabidopsis is the derived state due to a minimum of three inversions, two reciprocal translocations, and three reciprocal translocation/fusion events (Boivin et al., 2004; Kuittinen et al., 2004; Koch and Kiefer, 2005; Yogeeswaran et al., 2005; Lysak et al., 2006). Hence, the $n = 8$ karyotype is likely the ancestral state for the Camelinae, and potentially for much of the Brassicaceae since 38% of the family has a base-chromosome number of $x = 8$ (Al-Shehbaz et al., 2006; Warwick and Al-Shehbaz, 2006). Using the hypothesis of an ancestral karyotype of $n = 8$ similar to the genome structure of *A. lyrata* and *C. rubella*, Lysak et al. (2006) investigated a number of chromosomal reductions in the

[1] This work was supported by Duke University and the Max Planck Society.

[2] Present address: Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, P.O. Box 94062, Amsterdam 1090 GB, The Netherlands.

* Corresponding author; e-mail schranz@science.uva.nl; fax 31–020–525–7832.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) is: M. Eric Schranz (schranz@science.uva.nl).

[C] Some figures in this article are displayed in color online but in black and white in print.

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

[OA] Open Access articles can be viewed online without a subscription.

www.plantphysiol.org/cgi/doi/10.1104/pp.107.096685

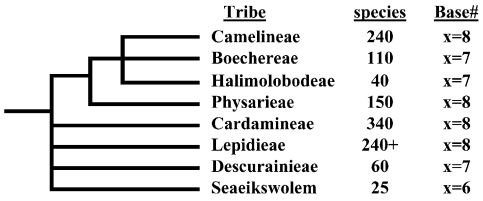

Figure 1. Relationships, number of species, and base-chromosome numbers for the eight tribes of lineage I of the Brassicaceae. A recent taxonomic classification for the Brassicaceae (Al-Shehbaz et al., 2006) proposed 24 tribes and recognized three monophyletic lineages (I–III). Within lineage I there is additional support for the monophyly of the three tribes Camelineae (including Arabidopsis, *A. lyrata*, and *C. rubella*), Boechereae (including *B. stricta*), and Halimolobodeae. Most species of the eight tribes within lineage I have a base-chromosome number of $x = 8$. Karyotype number reduction has occurred several times, including in the Boechereae ($x = 7$).

Brassicaceae using comparative chromosome painting (CCP). To do so they applied multicolor chromosome painting using contiguous bacterial artificial chromosome (BAC) pools of Arabidopsis arranged according to the genetic map of *A. lyrata* and *C. rubella*. The results revealed that karyotypes with reduced chromosome number ($n = 6, 7$) of two taxa from the tribe Camelineae (*Neslia*, *Turritis*) and one taxon from Descurainieae (*Hornungia*) shared conserved chromosome segments that can be related to the ancestral karyotype. Furthermore, the results suggested a common mechanism for chromosome number reduction via a pericentric inversion followed by a reciprocal translocation/fusion event (Lysak et al., 2006; Schranz et al., 2006b).

In addition to comparative mapping done within the Camelineae, there is also a wealth of comparative analyses between the economically important *Brassica* species from the tribes Brassiceae and Arabidopsis. There has been some difficulty in establishing syntenic relationships between *Brassica* and Arabidopsis due to the derived nature of the Arabidopsis genome, the paleopolyploid nature of Brassiceae genomes (Lagercrantz, 1998; Lysak et al., 2005; Parkin et al., 2005; Kim et al., 2006), and the relatively large phylogenetic distance between the two genera (Bailey et al., 2006; Beilstein et al., 2006; Koch et al., 2007). Despite these difficulties, a comprehensive comparison between *Brassica napus* and Arabidopsis identified 21 conserved syntenic blocks shared by *B. napus* and Arabidopsis genomes, which could be duplicated and rearranged to represent 90% of the *B. napus* genome. These conserved blocks represent collinear regions that have been maintained since the divergence of the Arabidopsis and *Brassica* lineages some 20 million years ago (Yang et al., 1999; Koch et al., 2003; Lysak et al., 2005).

An important step toward a unified comparative genomics system across the Brassicaceae can be accomplished by integrating the genomic block system used to show the collinear regions identified between *B. napus* and Arabidopsis (Parkin et al., 2005) with the concept of the $n = 8$ ancestral karyotype shared by *A. lyrata* and *Capsella* (Lysak et al., 2006). We recently proposed a set of 24 genomic blocks (A–X; Schranz et al., 2006b). The order, orientation, and color coding of these blocks are based on their positions in the ancestral karyotype, but using Arabidopsis locus names to define the interval of each block (Fig. 2A). Often the end points of the blocks correspond to pericentromeric or telomeric regions found in one or more species. This suggests that there may be common mechanisms involved in genome evolution across the crucifers.

Using our system of conserved blocks we can further explore the genomic organization of crucifers within a phylogenetic framework (Schranz et al., 2007). In the aforementioned tribal reclassification of the Brassicaceae (Al-Shehbaz et al., 2006), the tribe Camelineae (containing Arabidopsis) is most closely related to the small tribe Halimolobodeae (containing about 40 species) and the Boechereae (with about 110 species; Fig. 1). The Boechereae, made up mostly of species in the genus *Boechera*, are supported as being a monophyletic assemblage based on the fact that they are almost exclusively a North American group and differ from the Camelineae and Halimolobodeae in having a base-chromosome number of $x = 7$ rather than $x = 8$. Thus, they represent an excellent system to examine karyotype evolution accompanying chromosome number reduction.

The genus *Boechera* contains an array of morphologically and ecologically diverse taxa that have mainly radiated in alpine, montane, and desert regions of western North America. The group has great potential for studies of ecology and evolution (for review, see Mitchell-Olds, 2001; Dobeš et al., 2007). Numerous studies have analyzed molecular and phenotypic diversity of *Boechera* species. This includes the molecular evolutionary analysis of gene families (Bishop et al., 2000; Schein et al., 2004; Benderoth et al., 2006), the phylogeography of haplotypes (Dobeš et al., 2004b; Song et al., 2006), the occurrence of supernumary B chromosomes (Böcher, 1951; Sharbel et al., 2004), variation in breeding systems and ploidy (Sharbel and Mitchell-Olds, 2001; Schranz et al., 2005; Schranz et al., 2006a), drought tolerance (Knight et al., 2006), morphological and taxonomical diversity (Rollins, 1993; Al-Shehbaz, 2003; Windham and Al-Shehbaz, 2006), and the evolved responses to pathogen or insect pests (Roy, 1993; Roy and Kirchner, 2000; Jones et al., 2006). The further elucidation of these and other patterns of variation would be greatly aided by the creation and analysis of segregating genetic stocks.

Boechera stricta is one of the most morphologically and molecularly well defined *Boechera* species, making it a good candidate for genetic and genomic studies. Genetic and molecular analyses indicate that *B. stricta* is predominantly inbreeding, diploid, and sexual and most accessions form a monophyletic group (referred to as lineage II in Dobeš et al., 2004b; Schranz et al., 2005), whereas many other species in the genus are apomictic, of hybrid origin, and/or triploid (Sharbel and

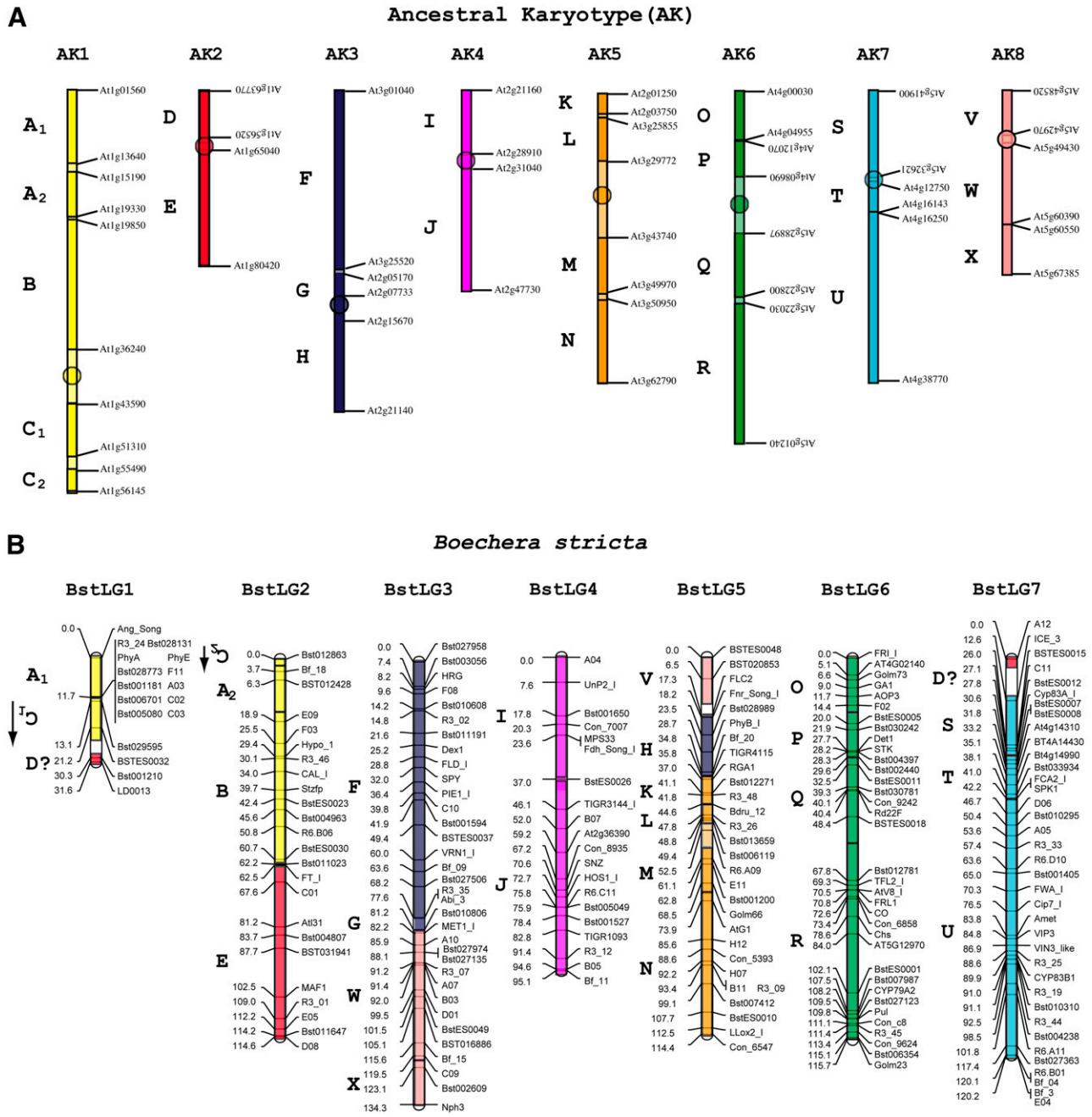

Figure 2. Comparative genome organization of the inferred ancestral karyotype ($n = 8$) based on published genetic maps of *A. lyrata* and *C. rubella* (A), and *B. stricta* using our F_2 genetic mapping results (B). A, Genome blocks of the ancestral karyotype (AK) are labeled by the letters A to X. Each block is one of eight colors, corresponding to each chromosome. Centromeric positions are indicated by the colored circles. Since only the Arabidopsis genome is currently sequenced, the boundaries of the blocks are defined by their flanking Arabidopsis Genome Initiative At locus names. Each block is considered to be in the upright orientation in the ancestral karyotype. Blocks that are inverted relative to Arabidopsis are indicated by upside-down text of the At locus names. B, Genetic map and genomic blocks for *B. stricta*. The seven LGs are labeled as BstLG1 to BstLG7. Marker positions (in cM) are shown on the left hand and the corresponding marker name shown on the right hand of each LG. Genomic blocks, as defined above, are arranged onto the LGs based on sequence similarity of the markers to Arabidopsis. Three LGs are completely conserved (Bst4 = AK4, BstLG6 = AK6, and BstLG7 = AK7). Incongruity of color/letter order of the blocks indicates genomic rearrangements in *Boechera* relative to the ancestral karyotype. Two blocks (A and C) are subdivided in *Boechera* (A_1 and A_2; C_1 and C_2). Blocks that are inverted in *Boechera* (blocks C_1 and C_2) are represented by their names being inverted and by a downward pointing arrow.

Mitchell-Olds, 2001; Schranz et al., 2005). Previously we made reciprocal crosses between 92 different *Boechera* lines from 19 different species to a common *B. stricta* genetic tester (the line SAD12; Schranz et al., 2005). Most of the intraspecific *B. stricta* crosses resulted in successful and highly fertile F_1 progeny. Our common crossing parent (SAD12) also was the focus of a partial genomic sequencing project (Windsor et al., 2006). By analyzing over 39,000 paired end sequences from a collection of medium-sized insert clones, we found that genic and intergenic regions are very similar to Arabidopsis, and conserved microsynteny can be used for rapid identification and cloning of *Boechera* genes (Windsor et al., 2006). We have also sequenced several *Boechera* BACs (Schein et al., 2004; Benderoth et al., 2006), showing that gene order is nearly identical to Arabidopsis. Such conserved microsyteny should facilitate the generation of molecular markers needed to examine patterns of macrosynteny.

In this report, we explore the macrosynteny between *B. stricta* and other crucifer species by comparative genetic mapping using the genomic block system based on the ancestral karyotype (Schranz et al., 2006b). In doing so, we investigate the karyotype reduction from $n = 8$ to $n = 7$ that occurred within the tribe Boechereae. Our construction of an F_2 linkage map will further our understanding of patterns of genome evolution in crucifers and facilitate future ecological genomics work in *Boechera*.

RESULTS

Genetic Markers

By analysis of a large collection of paired end-sequenced clones from *B. stricta* (Windsor et al., 2006) we have designed nearly 200 molecular markers with high similarity to Arabidopsis to facilitate comparative analyses. A summary of the markers developed and method of scoring is presented in Supplemental Table S1. For microsatellite markers we preferentially selected long repeats (with an average total repeat length of 33.8 bp made up from an average of 13.7 repeat units) to expedite genotyping (Supplemental Table S2). We designed primers to approximately 250 simple sequence repeat (SSR) loci, of which 58 were placed onto our genetic map (Supplemental Table S2).

The end-sequenced clones were also used to design primers for genes and/or regions of interest, often in conjunction with other ongoing research projects and objectives. These sequences can be divided into three categories: (1) candidate genes of interest (e.g. flowering time and glucosinolate production); (2) random nuclear loci selected as part of a project looking at the patterns of polymorphism in *Boechera* and other relatives; and (3) targeted regions necessary for synteny comparisons to Arabidopsis. Analysis of these sequences was done to identify single-nucleotide polymorphisms (SNPs) between the two mapping parents. In general, the level of polymorphism was very low, with an average of three polymorphisms per 1,000 bp (B. Song, unpublished data). SNPs were scored using four different methodologies. The majority were scored using a Custom 96-plex GoldenGate Genotyping BeadArray from Illumina. Of the 96 selected SNPs, we obtained genotypic data from 94 of the loci (a 96.8% success rate). This is a much higher success rate than typically seen with custom arrays designed for analysis of human SNPs. We were also very successful in converting SNPs to TaqMan markers as well as cleaved amplified polymorphisms/derived cleaved amplified polymorphisms markers. In addition, nearly all of our molecular markers (193) were scored as codominant, and only three markers were scored as dominant (Con_5393, Pul, and R6.D10).

Linkage Groups

All genotypic data were analyzed with JoinMap 4 using the regression mapping algorithm and Kosambi cM units for genetic linkage analysis (Stam, 1993). Using the most stringent log of the odds (LOD) score threshold of 10.0, all 196 markers resolved into the expected seven linkage groups (LGs). We designate these seven groups as BstLG1 to 7. The genetic map covers a total of 725.8 cM, with an average LG length of 104 cM. The average spacing between markers is 3.9 cM, and no gaps exceed 19.3 cM. The genome size of *B. stricta* has been estimated to be 264 Mb (approximately 1.7× the genome size of Arabidopsis; Schranz et al., 2006a), giving us a ratio of approximately 362 kb/1 cM.

Comparative Genome Analysis and Evolution

The genetic position and comparative genetic data for each marker is summarized in Table I and includes the position of the marker within the *B. stricta* genetic map (LG and relative position in cM), homology of the marker (or based on the similarity of the paired end sequence) to Arabidopsis (given to the At gene it matches or the nearest At gene based on genomic homology), and its assignment to a chromosomal block compared to the ancestral karyotype (as defined by letter and color in Schranz et al., 2006b and summarized in Fig. 2).

The genetic map of *B. stricta* is largely collinear to the ancestral karyotype based on the genetic maps of *A. lyrata* and *Capsella*. Furthermore, almost all markers occur and are collinear within their expected genomic blocks (Fig. 2). Only two genomic blocks (blocks A and C) had to be subdivided based on our *B. stricta* mapping results due to a single pericentric inversion (see below). In the ancestral karyotype block A was defined as the interval from At1g01560 to At1g19330. In *B. stricta* this interval is subdivided into block A_1 on Bst LG1 (At1g01560–At1g13640) and block A_2 on Bst LG2 (AT1G15190–At1g19330). The boundary between block A and block B was delineated based on mapping results from *B. napus* (Parkin et al., 2005) and is one of the

Table I. *Genetic mapping and molecular marker information for B. stricta F_2 map*

∼, Most significant BLAST score to an intergenic region, position of nearest gene is then given.

Group[a]	Position[b]	Marker[c]	Homology[d]	Block[e]	Distortion[f]
BstLG1	0.0	Ang_Song	At1g01510	A_1	–
BstLG1	11.7	R3_24	At1g04550∼	A_1	–
BstLG1	11.7	Bst028131	At1g05670	A_1	–
BstLG1	11.7	PhyA	At1G09570	A_1	–
BstLG1	11.7	PhyE	At4G18130	U	–
BstLG1	11.7	Bst028773	At1g10170	A_1	–
BstLG1	11.7	Bst001181	At1g10760	A_1	–
BstLG1	11.7	F11	At1g12980∼	A_1	–
BstLG1	11.7	A03	At1g13640∼	A_1	–
BstLG1	11.7	Bst006701	At1g51310	C_1	–
BstLG1	11.7	C02	At1g49600∼	C_1	–
BstLG1	11.7	C03	At1g49600	C_1	–
BstLG1	11.7	Bst005080	At1g49270	C_1	–
BstLG1	13.1	Bst029595	At1g03950	A_1	–
BstLG1	21.2	BSTES0032	At1g43245	C_1	–
BstLG1	30.3	Bst001210	At2g44470	J	–
BstLG1	31.6	LD0013	At1G58260	D	–
BstLG2	0.0	Bst012863	At1g55490	C_2	–
BstLG2	3.7	Bf_18	At1G15190	A_2	–
BstLG2	6.3	BST012428	At1g16080	A_2	–
BstLG2	18.9	E09	At1g20050	B	–
BstLG2	25.5	F03	At1g21830	B	–
BstLG2	29.4	Hypo_1	At1g23230	B	–
BstLG2	30.1	R3_46	At1g23530	B	–
BstLG2	34.0	CAL_I	At1g26320	B	–
BstLG2	39.7	Stzfp	At1g27730	B	–
BstLG2	42.4	BstES0023	At1g28130	B	–
BstLG2	45.6	Bst004963	At1g29230	B	–
BstLG2	50.8	R6.B06	At1g30410∼	B	–
BstLG2	60.7	BstES0030	At1g33490	B	–
BstLG2	62.2	Bst011023	None	?	–
BstLG2	62.5	FT_I	At1g65470	E	–
BstLG2	67.6	C01	At1g67420	E	–
BstLG2	81.2	Atl31	At1g71696	E	–
BstLG2	83.7	Bst004807	At1g72250	E	–
BstLG2	87.7	BST031941	At1g73390	E	–
BstLG2	102.5	MAF1	At1g77130	E	–
BstLG2	109.0	R3_01	At1g78690∼	E	–
BstLG2	112.2	E05	At1g79700	E	–
BstLG2	114.2	Bst011647	At1g80740	E	–
BstLG2	114.6	D08	At1g79990∼	E	–
BstLG3	0.0	Bst027958	At3G01760	F	–
BstLG3	7.4	Bst003056	At3g03460	F	–
BstLG3	8.2	HRG	At3g03480	F	–
eBstLG3	9.6	F08	At3G04470∼	F	–
BstLG3	14.2	Bst010608	At3g05310	F	–
BstLG3	14.8	R3_02	At3g05510∼	F	–
BstLG3	21.6	Bst011191	At3g07530	F	–
BstLG3	25.2	Dex1	None	?	–
BstLG3	28.8	FLD_I	At3g10380	F	–
BstLG3	32.0	SPY	At3G11540	F	–
BstLG3	36.4	PIE1_I	At3G12810	F	–
BstLG3	39.8	C10	At3g13670	F	–
BstLG3	41.9	Bst001594	At3g14370	F	–
BstLG3	49.4	BSTES0037	At3g16800	F	–
BstLG3	60.0	VRN1_I	At3g18900	F	*
BstLG3	63.6	Bf_09	At3G21055	F	*
BstLG3	68.2	Bst027506	At3g22760	F	–
BstLG3	77.6	R3_35	At3g24630∼	F	–

Table I. (*Continued.*)

Group[a]	Position[b]	Marker[c]	Homology[d]	Block[e]	Distortion[f]
BstLG3	77.6	Abi_3	At3g24650	F	–
BstLG3	81.2	Bst010806	At2g07690	G	–
BstLG3	82.2	MET1_I	At5G49160	W	–
BstLG3	85.9	A10	At5g49920	W	–
BstLG3	88.1	Bst027135	None	?	–
BstLG3	88.1	Bst027974	At5G50115	W	–
BstLG3	91.2	R3_07	At5g51600∼	W	–
BstLG3	91.4	A07	At5g51490∼	W	–
BstLG3	92.0	B03	At5g51870	W	–
BstLG3	99.5	D01	At5g55550∼	W	–
BstLG3	101.5	BstES0049	At5g53620	W	–
BstLG3	105.1	BST016886	At5G54820	W	–
BstLG3	115.6	Bf_15	At5G58410	W	–
BstLG3	119.5	C09	At5g59490	W	–
BstLG3	123.1	Bst002609	At5g60470	X	–
BstLG3	134.3	Nph3	At5g64330	X	–
BstLG4	0.0	A04	At2g21110	I	–
BstLG4	7.6	UnP2_I	At2g22300	I	*
BstLG4	17.8	Bst001650	At2G25220	I	*
BstLG4	20.3	Con_7007	At3g03100	F	–
BstLG4	23.6	Fdh_Song_I	At2g26250	I	–
BstLG4	23.6	MPS33	At2g26300	I	–
BstLG4	37.0	BstES0026	At2g29690	I/J	–
BstLG4	46.1	TIGR3144_I	At2g33420	J	–
BstLG4	52.0	B07	At2g34890	J	–
BstLG4	59.2	At2g36390	At2g36390	J	–
BstLG4	67.2	Con_8935	At2g38790∼	J	–
BstLG4	70.6	SNZ	At2G39250	J	–
BstLG4	72.7	HOS1_I	At2G39810	J	*
BstLG4	75.8	R6.C11	At2g40940∼	J	*****
BstLG4	75.9	Bst005049	At2g40650	J	*
BstLG4	78.4	Bst001527	At2g42350	J	**
BstLG4	82.8	TIGR1093	At2g43330	J	–
BstLG4	91.4	R3_12	At2g45710∼	J	–
BstLG4	94.6	B05	At2g47460	J	–
BstLG4	95.1	Bf_11	At2G47990	J	–
BstLG5	0.0	BSTES0048	At5g45470	V	–
BstLG5	6.5	BST020853	At5g43900	V	–
BstLG5	17.3	FLC2	At5g10140	R	–
BstLG5	18.2	Fnr_Song_I	At2g15620	H	–
BstLG5	23.5	Bst028989	At2g17420	H	*
BstLG5	28.7	PhyB_I	At2G18790	H	*
BstLG5	34.8	Bf_20	At2G20050	H	–
BstLG5	35.8	TIGR4115	At2g20580	H	–
BstLG5	37.0	RGA1	At2G01570	K	–
BstLG5	41.1	Bst012271	At2g02790	K	–
BstLG5	41.8	R3_48	At2g02860	K	–
BstLG5	44.6	Bdru_12	None	?	–
BstLG5	47.8	R3_26	At3g27400∼	L	–
BstLG5	48.8	Bst013659	At3g27550	L	–
BstLG5	49.4	Bst006119	At4g05460	O	–
BstLG5	52.5	R6.A09	None	?	–
BstLG5	61.1	E11	At3g47600	M	–
BstLG5	62.8	Bst001200	At3g48120	M	–
BstLG5	68.5	Golm66	At3g49900	M	*
BstLG5	73.9	AtG1	At3g51790	N	–
BstLG5	85.6	H12	At3g54670	N	*
BstLG5	88.6	Con_5393	At4g22970∼	U	–
BstLG5	92.2	H07	At3g56760∼	N	–
BstLG5	93.4	B11	At3g57060	N	–
BstLG5	93.4	R3_09	At3g57080∼	N	–

(*Table continues on following page.*)

Table I. (*Continued from previous page.*)

Group[a]	Position[b]	Marker[c]	Homology[d]	Block[e]	Distortion[f]
BstLG5	99.1	Bst007412	At3g59090	N	–
BstLG5	107.7	BstES0010	At3g61520	N	–
BstLG5	112.5	LLox2_I	At3g45140	M	–
BstLG5	114.4	Con_6547	At3g63110~	N	–
BstLG6	0.0	FRI_I	At4g00640	O	*
BstLG6	5.1	AT4G02140	At4G02140	O	–
BstLG6	6.6	Golm73	At4g02485	O	–
BstLG6	9.0	GA1	At4g02750	O	–
BstLG6	11.7	AOP3	At4g03050	O	*
BstLG6	14.4	F02	At4g03540~	O	**
BstLG6	20.0	BstES0005	At4g12300	P	–
BstLG6	21.9	Bst030242	At4g11160	P	*
BstLG6	27.7	Det1	At4g10180	P	–
BstLG6	28.2	STK	At4G09960	P	–
BstLG6	28.3	Bst004397	At4g09980	P	–
BstLG6	29.6	Bst002440	None	?	*
BstLG6	32.5	BstES0011	At5g28690	Q	–
BstLG6	39.3	Bst030781	None	?	–
BstLG6	40.1	Con_9242	At5g25820~	Q	–
BstLG6	40.4	Rd22F	At5g25610	Q	–
BstLG6	48.4	BSTES0018	At5g23320	Q	–
BstLG6	67.8	Bst012781	At5g18230	R	–
BstLG6	69.3	TFL2_I	At5g17670	R	–
BstLG6	70.5	AtV8_I	At5g16570	R	–
BstLG6	70.8	FRL1	At5g16280	R	–
BstLG6	72.6	CO	At5g15780	R	*
BstLG6	73.4	Con_6858	At5g15680~	R	*
BstLG6	78.6	Chs	At5g13930	R	****
BstLG6	84.0	AT5G12970	At5G12970	R	*******
BstLG6	102.1	BstES0001	At5g07440	R	*******
BstLG6	107.5	Bst007987	At5g05570	R	*******
BstLG6	108.2	CYP79A2	At5g05380	R	*******
BstLG6	109.5	Bst027123	At5g04590	R	*******
BstLG6	109.8	Pul	At5g04360	R	***
BstLG6	111.1	R3_45	At5g03905	R	******
BstLG6	111.4	Con_c8	At5G03940	R	****
BstLG6	113.4	Con_9624	At5g03610~	R	******
BstLG6	115.1	Bst006354	At5g02910	R	******
BstLG6	115.7	Golm23	None	?	*******
BstLG7	0.0	A12	At1g58602~	D	*******
BstLG7	12.6	ICE_3	At5g40330~	S	***
BstLG7	26.0	BSTES0015	At5g35870	S	***
BstLG7	27.1	C11	At5g35360~	S	****
BstLG7	27.8	BstES0012	At5g33340	S	****
BstLG7	30.6	Cyp83A_I	At4g13770	T	****
BstLG7	31.8	BstES0007	At4g13970	T	***
BstLG7	31.8	BstES0008	At4g14200	T	***
BstLG7	33.2	At4g14310	At4g14310	T	****
BstLG7	35.1	BT4A14430	At4g14430	T	****
BstLG7	38.1	Bt4g14990	At4g14990	T	**
BstLG7	41.0	Bst033934	At4g15960	T	**
BstLG7	42.2	FCA2_I	At4G16280	U	*
BstLG7	42.2	SPK1	At4G16340	U	*
BstLG7	46.7	D06	At4g17660	U	**
BstLG7	50.4	Bst010295	At4g18490	U	**
BstLG7	53.6	A05	At4g19190~	U	**
BstLG7	57.4	R3_33	At4g20160	U	**
BstLG7	63.6	R6.D10	At4g23280~	U	–
BstLG7	65.0	Bst001405	At4G23540	U	–
BstLG7	70.3	FWA_I	At4g25540	U	–
BstLG7	76.5	Cip7_I	At4g27430	U	–
BstLG7	83.8	Amet	At4g29510	U	–

Table I. (*Continued.*)

Group[a]	Position[b]	Marker[c]	Homology[d]	Block[e]	Distortion[f]
BstLG7	84.8	VIP3	At4g29840	U	–
BstLG7	86.9	VIN3_like	At4G30200	U	–
BstLG7	88.6	R3_25	At4g31170	U	–
BstLG7	89.9	CYP83B1	At4g31490	U	–
BstLG7	91.0	R3_19	At4g31710~	U	–
BstLG7	91.1	Bst010310	At4g31900	U	–
BstLG7	92.5	R3_44	At4g32551	U	–
BstLG7	98.5	Bst004238	At4g33790	U	–
BstLG7	101.8	R6.A11	At4g34180~	U	–
BstLG7	117.4	Bst027363	At4g39800	U	–
BstLG7	120.1	R6.B01	At4g38230~	U	–
BstLG7	120.1	Bf_04	At4G38550	U	–
BstLG7	120.2	Bf_3	At4G38550	U	–
BstLG7	120.2	E04	At4g38690~	U	–

[a]Genetic LG as shown in Figure 2B. [b]Map position (in cM) on LG. [c]Genetic marker name. [d]Homology of *Boechera* genetic marker to gene in Arabidopsis. [e]Position of marker within a genomic block (as defined in Fig. 1A) based on homology to Arabidopsis. [f]Significance of segregation distortion. *, $\alpha = 0.05$; **, $\alpha = 0.01$; ***, $\alpha = 0.001$; ****, $\alpha = 0.0001$; *****, $\alpha = 0.00001$.

few genomic block boundaries that is not defined by a centromeric or telomeric region in Arabidopsis or the ancestral karyotype. Block C must also be divided into two. Block C_1 on BstLG1 spans the interval from AT1g43245 to At1g51310 and block C_2 on BstC2 spans the interval from At1g55490 to At1g56145. The only block that could not be definitively placed on the map was block D, with one marker with homology to block D on BstLG1 and one marker on BstLG7.

There were a total of eight markers whose placement on the map did not agree with expectations based on Arabidopsis. One of these, *PhyE*, mapped to the same position as a close homolog, *PhyA*, on LG1. Hence, this may represent a case of cross hybridization of the primers to a related sequence. Another gene, *FLC2*, mapped to LG5 rather than to its expected position on LG6. However, we have cloned two copies of the *FLC* gene in *Boechera* (M.E. Schranz, unpublished data) and hence the duplicate copy of this locus has likely been integrated at this new genomic position.

We specifically designed a number of probes from the *B. stricta* end-sequence collection that had no significant similarity to Arabidopsis. Interestingly, several of these (Bst011023 on BstLG2, R6.A09 on BstLG5, and Bst002440 on BstLG6) map to the likely pericentromeric regions based on comparison to the ancestral karyotype. The observation that these markers occur within regions with low levels of recombination on their respective LGs supports the possibility that these markers are specific to pericentromeric regions.

Based on comparison to the ancestral karyotype of $n = 8$, we can infer several details of *B. stricta* ($n = 7$) genome structure and evolution (Figs. 2 and 3). First, three of the *B. stricta* LGs are almost completely collinear and unchanged when compared to those in the ancestral karyotype (BstLG4 = AK4, BstLG6 = AK6,

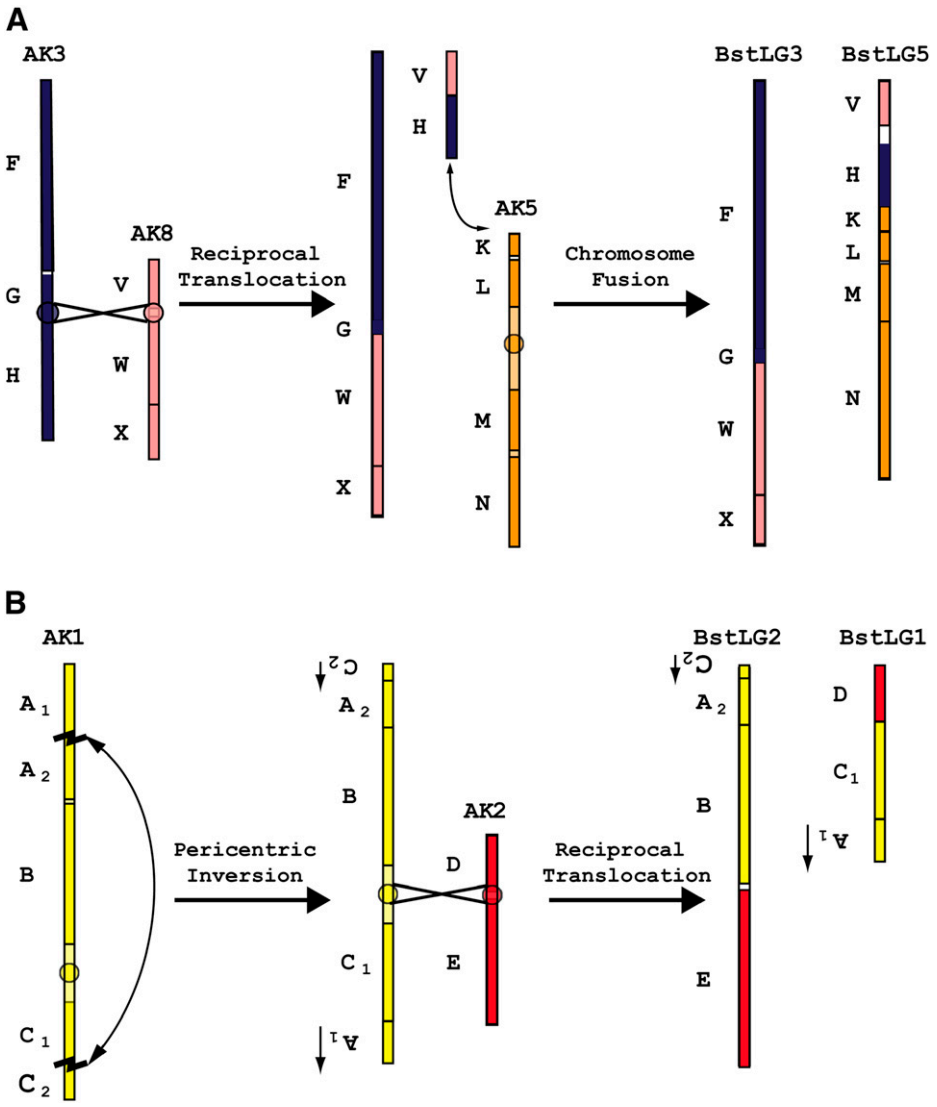

Figure 3. A parsimonious model of genomic changes describing the evolution of the *B. stricta* ($n = 7$) genome from the ancestral karyotype ($n = 8$). A, A reciprocal translocation occurring between the centromeric regions of AK3 and AK8 would result in the formation of BstLG3 (with blocks F and G now fused with W and X) and a second chromosome (made up of blocks V and H). A chromosomal fusion between the telomeric regions of block H and block K (of AK5) would result in the formation of BstLG5. It is equally possible that the chromosomal fusion between block H (of AK3) and block K (of AK5) occurred before the reciprocal translocation event. The combination of the reciprocal translocation and chromosomal fusion events would account for the karyotype reduction from $n = 8$ to $n = 7$. B, A pericentric inversion of AK1 with one breakpoint between blocks A_1 and A_2 and the other between blocks C_1 and C_2 would result in a rearranged chromosome. A reciprocal translocation between the centromeric regions of the rearranged AK1 and AK2 would result in BstLG1 (including blocks D, C_1, and A_1) and BstLG2 (including blocks C_2, A_2, B, and E). [See online article for color version of this figure.]

and BstLG7 = AK7; Fig. 2). This result further validates the derived nature of the Arabidopsis karyotype ($n = 5$) and confirms the existence of a common ancestral karyotype ($n = 8$), as many of the blocks that are collinear in *A. lyrata*, *Capsella*, and now *Boechera* are on different chromosomes in Arabidopsis. For example, blocks S and T are collinear in *A. lyrata*, *Capsella*, and *Boechera*, but not in Arabidopsis where they are found on chromosomes At4 and At5, respectively (Lysak et al., 2006). Second, we can surmise that the other four *B. stricta* LGs were derived by a series of rearrangements that can account for the karyotype reduction from $n = 8$ to $n = 7$. In total, there likely was a pericentric inversion, a chromosomal fusion, and two reciprocal translocations involving five of the eight chromosomes in the ancestral karyotype (Fig. 3).

While we don't know the evolutionary sequence of these events, we can hypothesize a parsimonious series of changes. A reciprocal translocation occurring between the centromeric regions of AK3 and AK8 (Fig. 3A) would result in the formation of BstLG3 (with block F and G now fused with W and X) and a second chromosome (made up of blocks V and H). A chromosomal fusion between the telomeric regions of block H and block K (of AK5) would result in the formation of BstLG5 (Fig. 3A). The combination of the reciprocal translocation and chromosomal fusion events would account for the karyotype reduction from $n = 8$ to $n = 7$. However, our mapping data do not allow us to definitively resolve which centromere was lost (AK3, AK8, or AK5), or by what mechanism this loss occurred. It may well be that there were additional chromosomal rearrangements (such as one or more pericentric inversions).

A pericentric inversion of AK1 with one breakpoint between blocks A_1 and A_2 and the other between blocks C_1 and C_2 would result in a rearranged chromosome (Fig. 3B). A reciprocal translocation between the centromeric regions of the rearranged AK1 and AK2 would result in BstLG1 (including blocks D, C_1, and

A_1) and BstLG2 (including blocks C_2, A_2, B, and E). Again, we cannot rule out the possibility of additional changes.

Based on comparison to the ancestral karyotype, we can also conjecture the centromeric positions in *Boechera* (Fig. 2). Centromeres with conserved positions could lie between the following sets of blocks: I and J on BstLG4, L and M on BstLG5, P and Q on BstLG6, and S and T on BstLG7. Centromeres that are potentially rearranged in *Boechera* relative to the ancestral karyotype might lie between blocks B and E on BstLG2 and between blocks G and W on BstLG3.

Segregation Distortion

By testing for deviation from the expected Mendelian 1:2:1 ratio of markers we found 25% were significant at the 0.05 significance level. This is less than the approximately 35% segregation distortion seen in *A. lyrata* (Kuittinen et al., 2004), but higher than that observed in many intraspecific crosses done in crop plants (Jenczewski et al., 1997). Most distorted loci (42/49) were due to a low frequency of observed homozygous SAD12 genotypes, despite the fact that the SAD12 line was the maternal parent. Segregation distortion was nonrandom, with most distorted loci observed in blocks of multiple colinear markers. In total, there were seven regions in which more than one marker was distorted. There were two particularly large blocks of distortion (on the bottom of BstLG6 and the top of BstLG7) together containing 32 of the 49 distorted loci. We did not find any markers showing segregation distortion to significantly deviate between the observed and expected genotypic frequencies at the 0.10 significance level, suggesting selection on the gametic stage.

Recombination Suppression

BstLG1 showed very low levels of recombination over much of its length, with most recombination only occurring near the ends of the LG. Most of the markers on this LG have homology to Arabidopsis chromosome 1. The equivalent region in Arabidopsis covers at a minimum 6.7 Mb. There is more recombination seen within these syntenic intervals in the published maps of both Arabidopsis and in *A. lyrata*. The low levels of recombination seen in BstLG1 makes marker order and placement very ambiguous. Such a lack of recombination across the interval in *B. stricta* could be due to an intraspecific inversion present between the two parental types or some other mechanism suppressing recombination, such as heterochromatization.

DISCUSSION

The dicot family Brassicaceae is an excellent group in which to examine patterns of genome and sequence evolution (Schranz et al., 2007). Not only does it contain the model species Arabidopsis and the domesticated *Brassica* crops, but it is also the focus of several complete genome-sequencing projects, including *A. lyrata*, *C. rubella*, *Thellungiella halophila* (=*Eutrema halophila*), and *B. rapa*. Added to this is the partial genome sequencing data of *Arabis alpina* (N. Warthmann and D. Weigel, personal communication) and *B. stricta* (approximately 0.15× genome coverage; Windsor et al., 2006). Comparative genetic mapping provides the framework for analysis of these genomic sequences. Comparative mapping has found significant synteny in genomic blocks conserved across a number of species, suggesting that there are common mechanisms involved in genome evolution across the family (for review, see Schranz et al., 2006b).

In this study, we contribute to our understanding of crucifer genome evolution by comparative analysis of our F_2 linkage map for *B. stricta*. Overall, our genetic map is highly collinear with the $n = 8$ genetic maps from *A. lyrata* and *Capsella*. However, we found that the genome evolution and karyotype reduction to $n = 7$ in *B. stricta* did not occur by one simple chromosomal fusion event, but rather involved several chromosomal changes including a pericentric inversion, a chromosomal fusion, and two reciprocal translocations. Most of these chromosomal changes can simply be represented with our genomic block system. Although the evolution from $n = 8$ to $n = 7$ is more complex than one simple fusion it is much easier to interpret than if compared to Arabidopsis (Schranz et al., 2006b). Our results support the hypothesis that there are common mechanisms involved in crucifer genome evolution such that changes tend to occur at certain points while maintaining most of the genome in large syntenic blocks. In addition, we found one LG that has a large degree of recombination suppression as well as two regions showing significant segregation distortion. Our genetic mapping will also facilitate future analysis of quantitative trait variation within *Boechera*.

Markers and Genetic Map Construction

Previously, we had end sequenced a large number of *B. stricta* λ clones to investigate sequence similarity and microsynteny with Arabidopsis (Windsor et al., 2006). In this study, we utilized these sequences to develop nearly 200 SSR and SNP molecular markers for genetic mapping. We successfully genotyped 94 SNPs simultaneously using a custom Illumina bead array. The Illumina bead array technology has been extensively used in studies of human polymorphism, including the Human HapMap project (Altshuler et al., 2005). Only recently has it been used for genetic mapping in plants (Rostoks et al., 2006). Nearly all of our SSR and SNP markers were developed from sequences with high similarity to Arabidopsis, allowing us to investigate patterns of macrosynteny between these two species and other crucifer genomes.

All of our markers could be unambiguously placed into one of the seven LGs using the most stringent LOD score threshold of 10. Our overall sample size, number

of markers, marker density, and genome coverage is higher than that in genetic maps constructed for either *C. rubella* or *A. lyrata* (Boivin et al., 2004; Kuittinen et al., 2004; Yogeeswaran et al., 2005). With an average distance between markers of less than 4 cM and with no gaps greater than 20 cM, our genetic map provides excellent opportunities for precise quantitative trait locus studies, positional cloning, and comparative genomic analyses.

Comparative Genomics and Genome Evolution of *Boechera*

In a recent review, we proposed a framework for comparative genomics for the Brassicaceae based on a set of 24 conserved syntenic blocks (Schranz et al., 2006b). Herein, we demonstrate the utility of these genomic blocks for illustrating and understanding the genome evolution of *B. stricta* from the tribe Boechereae.

Recent molecular phylogenetic analyses have shown that the tribe Boechereae is closely related to the potentially polypheletic tribe Camelineae, containing both the genus *Arabidopsis* and the genus *Capsella* (Bailey et al., 2006; Beilstein et al., 2006; Koch et al., 2007). The two tribes along with several other tribes are strongly supported as a monophyletic assemblage within the Brassicaceae (referred to as lineage I; Fig. 1; Bailey et al., 2006; Beilstein et al., 2006; Koch et al., 2007). The base-chromosome number for most of the tribes within lineage I is $x = 8$ (Fig. 1; Al-Shehbaz et al., 2006). There are several examples of karyotype reduction (Lysak et al., 2006), including the reduction of Arabidopsis to $n = 5$, and increases in chromosome number often due to polyploidy within lineage I (Marhold and Lihova, 2006; Warwick and Al-Shehbaz, 2006). Interestingly, the entire tribe Boechereae shares the base-chromosome number of $x = 7$, presumably from a reduction from $x = 8$ seen in other members of lineage I. A priori, one might expect that this reduction occurred via a single chromosomal fusion event. However, our results demonstrate that the karyotype evolution of *B. stricta* was more complex, involving at a minimum a pericentric inversion, a chromosomal fusion, and two reciprocal translocations (Fig. 3). The proposed types and likely mechanisms of changes are very similar to other examples of karyotype reduction in the Brassicaceae (Lysak et al., 2006). This study suggests that rearrangements often occur at pericentromeres or at telomeres containing nucleolus organizer regions (NORs).

In *A. lyrata* there are a total of five NORs. NORs on different chromosomes associate nonrandomly in interphase nuclei of both Arabidopsis and *A. lyrata* (Armstrong et al., 2001; Pecinka et al., 2004; Berr et al., 2006). The pairing of NORs of nonhomologous chromosomes has been hypothesized to play an important role in facilitating rearrangements between chromosomes (Lysak et al., 2006). In *B. stricta*, there is only a single terminal NOR (Kantama, 2005), hence, if the ancestor of *B. stricta* also contained five NORs than four have been lost. By comparison of the maps of *A. lyrata* and *B. stricta* we can hypothesize that a conserved NOR could be positioned at the top of BstLG4 (above block I). Intriguingly, the other NORs seen in *A. lyrata* are all at the sites of rearrangements in *B. stricta*. Of particular importance are the NORs at the ends of block H (on AL3) and block K (on AL5). This is the site of chromosomal fusion seen on BstLG5, suggesting a possible mechanism for chromosomal fusion by recombination of the two NORs. This process would also lead to the elimination of these two NORs. Another NOR at the bottom of block C (on AL1) could have been lost during the pericentromeric inversion event occurring in *B. stricta*. Finally, there is a NOR at the top of block S (on AL 7). On BstLG7 we have a single marker coming from block D above block S, suggesting this NOR also was lost. Support for the hypothesis that chromosomal rearrangements often occur at NORs comes from the analysis of telomerase-deficient Arabidopsis lines where frequent chromosomal fusions between NORs have been observed (Siroky et al., 2003).

We also have detected a number of rearrangements that likely occurred between pericentromeric regions of the ancestral karyotype during the evolution of the *B. stricta* genome. Specifically, we hypothesize reciprocal translocations between the pericentromeric regions of AK1 and AK2 and between AK3 and AK8. Rearrangements involving pericentromeric regions have also been observed in rearrangements of other crucifer species, including Arabidopsis (Berr et al., 2006; Kawabe et al., 2006; Lysak et al., 2006).

NORs and pericentromeric regions are known to be highly dynamic genomic regions, characterized by long stretches of repetitive DNA (Hall et al., 2004; Fajkus et al., 2005). It is plausible that nonhomologous chromosome association via pairing of repetitive elements in either pericentromeric regions or NORs could be involved in recombination events leading to chromosomal rearrangements. Our results further confirm a generalized set of conserved mechanisms contributing to Brassicaceae genome evolution. One of the major conclusions is that only limited regions of the genome are involved in rearrangements, whereas most of the genome is maintained in discrete and conserved blocks.

Recombination Suppression of BstLG1

BstLG1 has suppressed recombination within the central region of the chromosome. Most of the markers correspond to genic, and not pericentromeric, regions found in Arabidopsis chromosome 1. One possibility to explain the recombination suppression is an intraspecific inversion within BstLG1 between the two parental lines. It is important to note that this would be a separate inversion event than that shown in Figure 3B. Additionally, we do not know whether such intraspecific chromosomal polymorphisms exist within *B. stricta*. Another possibility is that recombination is suppressed because one of the parental BstLG1 chromosomes contains large tracts of heterochromatin.

Heterochromatic chromosomes (*Het*) and supernumary heterochromatic-B chromosomes have been detected in asexually reproducing (apomictic) *Boechera* lines (Sharbel et al., 2004; Kantama, 2005; Sharbel et al., 2005). The *Het* chromosome is derived from *B. stricta* as based on hybridization of pericentromeric repeats (Kantama, 2005; Schranz et al., 2006a) and genetic analysis has demonstrated that the *Het* chromosome can be crossed into *B. stricta* without conferring a dominant apomictic phenotype (Schranz et al., 2006a). These two results suggest that *Het* chromosomes might have originated and been present within sexually reproducing *B. stricta* lineages. From cytological studies we know that the maternal *B. stricta* parent, SAD12, does not carry the *Het* chromosome (Schranz et al., 2006a), but, potentially the parental line, LTM, might bear a *Het* chromosome that was transmitted to the F_1 plant. Future cytogenetic studies will be needed to clarify the reason(s) for the recombination suppression of the BstLG1.

Segregation Distortion

A total of 25% of our markers showed significant transmission ratio distortions. This is only slightly higher than the average ratio for intraspecific crosses of agricultural species (18.4% ± 11.0%; Jenczewski et al., 1997). Several recent analyses of interspecific crosses of wild species have reported much higher levels of distortion (Hall and Willis, 2005; Bratteler et al., 2006), including *A. lyrata* with an average of 35% (Kuittinen et al., 2004). Hence, the degree of segregation distortion seen for *B. stricta* is within expectations. Most markers were distorted because of an underrepresentation of SAD12 homozygous genotypes. This was surprising since SAD12 was the maternal parent, suggesting that cytoplasmic incompatibility is not responsible for the distortion.

Segregation distortion can occur at two levels, reflecting selection at either gametic or genotypic levels. Prezygotic selection at the gametic level causes deviation from 1:1 allelic ratios, producing genotypic ratios of p^2:$2pq$:q^2, where p and q indicate allele frequencies after selection. In contrast, postzygotic selection causes deviation from this predicted genotypic ratio, for example if heterozygotes have higher survival or one homozygote has reduced viability. At each distorted locus we calculated allele frequencies and compared observed and predicted genotypic ratios. Results fit a model of prezygotic gametic selection that favors LTM alleles in several genomic regions. We found no evidence for postzygotic selection favoring particular genotypes. Thus, there is no indication of heterotic influences on plant viability.

Gametic selection can occur via competition between pollen and/or ovules with different genotypes caused by self-incompatability loci, inbreeding depression due to genetic load, or genetic isolation evolved between the parental populations. *B. stricta* is a highly self-compatible species with very high levels of homozygosity (Song et al., 2006), thus, we do not expect either self-incompatibility or inbreeding depression to be the cause of the segregation distortion. However, there could be significant genetic isolation between these two populations. The two populations are geographically isolated, being approximately 1,000 km apart. Also, the habitats of the two populations are quite different (Fig. 4); the SAD12 locality is a sagebrush grassland in a river valley, whereas the LTM locality is a subalpine meadow. The two sites differ in levels of precipitation and temperature. Additionally, the two parental genotypes cluster into different groups based on STRUCTURE analyses of 229 single-copy nucleotide loci and the Wright's fixation index value between the two groups is 0.4 (B. Song and T. Mitchell-Olds, unpublished data). Finally, we also know that the two populations have different chloroplast haplotypes. SAD12 has chloroplast haplotype DG and LTM has haplotype AH based on analysis of the *trn*L intron-*trn*L/F IGS region (Schranz et al., 2005). The LTM haplotype (AH) is one of two main haplotypes found to have likely colonized formerly glaciated areas of North America, with the AH haplotype spreading into the Northwest and AS spreading in the Southwest, Northcentral, and Northeast (Dobeš et al., 2004b). The SAD12 haplotype (DG) is rare and is a geographically restricted haplotype that may represent an isolated glacial refugial population (C. Dobeš and M. Koch, personal communication).

Future Directions

CCP using multicolored BACs arranged according to the ancestral karyotype (Lysak et al., 2006) would be invaluable for resolving several remaining questions about genome evolution in *Boechera*. First, CCP would confirm the order and orientation of the genomic blocks in *B. stricta* and help resolve the uncertain

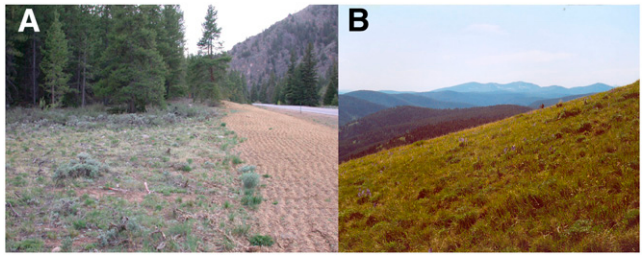

Figure 4. The collection sites of the maternal and paternal parents of our mapping population of *B. stricta* differ substantially in abiotic environments. A, The maternal SAD12 locality in Colorado is a sagebrush grassland in a river valley occurring at an elevation of 2,530 m, at a latitude of 38.7° N, has an average monthly precipitation of 43.0 mm, an average monthly high temperature of 8.5°C, and an average monthly low temperature of −6.9°C. B, The paternal LTM locality in Idaho is a subalpine meadow occurring at an elevation of 2,390 m, at a latitude of 45.7° N, has an average monthly precipitation of 94.7 mm, an average monthly high temperature of 9.2°C, and an average monthly low temperature of −3.5°C. The two sites are approximately 1,000 km apart. [See online article for color version of this figure.]

placement of block D. Second, CCP could identify rearrangements that were not detected by our mapping due to low marker density or that occur in regions of low recombination (such as centromeres or in BstLG1). This will help us resolve whether or not there is recombination suppression in BstLG1 due to an intraspecific chromosomal inversion. Third, CCP could identify the location of centromeres, NORs, and heterochromatic regions of our *Boechera* genetic map. This information will be important in future quantitative trait analyses. Locating centromeres would also resolve which centromere has been lost (potentially one between blocks V and H) and by what mechanism this occurred. Finally, CCP could be used to address whether any of the genomic rearrangements identified in our genetic mapping of *B. stricta* are unique to this species or if they are shared with all members of the Boechereae tribe (with $x = 7$) and/or with members of the closely related tribe Halimolobodeae (with $x = 8$). We also could test for shared breakpoints by tracking a handful of markers spanning the block boundaries in F_2 families in crosses made in other taxa.

Comparative genetic mapping and CCP could also help resolve conflicting phylogenetic signals by tracking specific and rarely occurring genomic changes (Lysak and Lexer, 2006). For example, such approaches could be used to identify monophyletic groups within the polyphyletic Camelineae (Bailey et al., 2006; Koch et al., 2007) and help resolve ambiguous relationships within the genus *Boechera* (Schranz et al., 2005).

Our construction of a *B. stricta* genetic map will also greatly facilitate our analyses of ecologically important quantitative variation and positional cloning of the underlying genes segregating in our cross. There is substantial variation for a number of important traits, including glucosinolate content, flowering time, and drought tolerance that should be amenable to quantitative analyses. Our placement of many candidate gene markers onto our genetic map will facilitate these analyses. Furthermore, we have already advanced our mapping population to the F_5 generation and will soon have recombinant inbred lines for this perennial species. Finally, our success in exploiting microsynteny between *Boechera* and Arabidopsis for sequencing genes of interest should allow for successful positional cloning of quantitative trait loci (Schein et al., 2004; Benderoth et al., 2006; Windsor et al., 2006).

MATERIALS AND METHODS

Plant Materials and DNA Isolations

A genetic cross was made between two highly inbred lines of *Boechera stricta* (Graham) Al-Shehbaz. The maternal line SAD12 was collected in Gunninson County, Colorado by Dr. Bitty Roy and the paternal line LTM was collected in Lemhi County, Idaho (Fig. 4). Details about the plant populations, locations, and the genetic crossing have been described previously (Schranz et al., 2005). A total of 192 F_2 lines were grown, with seeds placed on moist filter paper in sealed petri dishes and cold treated at 4°C for 3 weeks in the dark. The petri dishes were then transferred to a growth chamber until seed germination. The germinated seedlings were then transferred to 96-well flats. Seedlings were grown for 4 weeks and then transplanted to pots ($11 \times 11 \times 13$ cm). The plants were grown in a controlled growth room under long-day conditions (16 h light and 8 h dark).

DNA from each F_2 line was isolated using the Qiagen DNeasy Plant Mini kit and the Qiagen Genomic-tip 100/G kit (Qiagen) was used for the two parental genotypes.

Genetic Marker Development and Analysis

We previously reported our analysis of approximately 39,000 paired end sequences from the SAD12 genotype of *B. stricta* (Windsor et al., 2006) that was used as the maternal parent in our genetic cross. We utilized these end-sequenced clones to develop a number of different genetic marker systems. Foremost, we used sequences that had strong similarity/homology to Arabidopsis (*Arabidopsis thaliana*) to facilitate comparative analyses. In several instances, we used the similarity data from the paired end sequences to help infer homology. A summary of the markers developed is presented in Supplemental Table S1. This includes the name of the marker, the type of polymorphism (SSRs or SNPs), the method by which they were analyzed (sequencer, gel based, cleaved amplified polymorphism, derived cleaved amplified polymorphism, Illumina bead array, or TaqMan probe), the primers used to detect and/or score the polymorphism, and the λ clone from which it was derived.

To identify SSRs to use as molecular markers we screened the end-sequenced clones for SSRs (microsatellites) using the SPUTNIK program (Abajian, 1994). We allowed for no errors (insertions, mismatches, and deletions) to the repeat. Primers that flanked the repeat element and had an amplicon <300 bp were designed using PRIMER 3 software (Rozen and Skaletsky, 2000; Supplemental Table S2). In addition, we used 10 microsatellite loci reported in previous studies (Clauss et al., 2002; Dobeš et al., 2004a; Schranz et al., 2005; Song et al., 2006). Most microsatellites were scored by analysis on 4% MetaPhor Tris-acetate EDTA agarose gels (Cambrix Bio Science). Primers used in previous studies were run as described previously (Schranz et al., 2005; Song et al., 2006).

Second, we used the end-sequenced clones to design primers for genes and/or regions of interest, often in conjunction with other ongoing research projects and objectives. Primer pairs were designed from the end-sequenced clones using either PRIMER 3 software (Rozen and Skaletsky, 2000) or with PRIMACLADE software when the end sequences were aligned to Arabidopsis genomic sequence (Gadberry et al., 2005). The primers were used to amplify and sequence genomic DNA of the mapping parents (SAD12 and LTM) to identify SNPs. PCR and DNA sequencing was done as previously described (Windsor et al., 2006) or by Genaissance Pharmaceuticals Inc.

Sequences were quality trimmed and assembled into contigs using either phred-phrap-consed (Ewing and Green, 1998; Ewing et al., 1998; Gordon, 2004) or SeqMan 5.0 (DNAStar Inc.) at stringent quality thresholds.

SNPs were identified and formatted en masse using *SnpDetector* (A.J. Windsor, unpublished data), a script written in Python (http://www.python.org). To identify candidate SNPs, *SnpDetector* performs pairwise alignments between orthologous sequences by calling the National Center for Biotechnology Information's *bl2seq* (Altschul et al., 1990; Zhang and Madden, 1997) and/or EMBOSS' *water* (Rice et al., 2000) programs. Subsequently, *SnpDetector* parses the alignments generated, detects candidate SNPs, and screens the candidate SNPs relative to the requirements needed for genotyping.

Identified SNPs were scored using four different methodologies. A total of 94 SNPs were analyzed using a Custom 96-plex GoldenGate Genotyping BeadArray from Illumina Inc. Second, 16 SNPs were scored using TaqMan probes (Applied Biosystems). Cleaved amplified polymorphisms were identified using both the SNP2CAPS (Thiel et al., 2004) and BlastDigester (Ilic et al., 2004) programs. Finally, derived cleaved amplified polymorphisms (Michaels and Amasino, 1998) were designed using the derived cleaved amplified polymorphism Finder 2.0 (Neff et al., 2002).

Genetic Map Construction

We used the JoinMap v4 program using the weighted least-squares method for map construction (Stam, 1993). A LOD score threshold of 10 was used to assign all markers to one of seven LGs.

Segregation Distortion

To test for segregation distortion, we examined the fit of each marker to the expected 1:2:1 ratio with χ^2 tests using the JoinMap program (Stam, 1993). For

those markers that showed significant deviations in segregation at the 0.05 level, we used χ^2 tests to investigate whether the observed genotypic frequencies differed from the expected genotypic frequencies (e.g. are the marker classes in Hardy-Weinberg equilibrium). If the allele frequencies significantly deviate from the expected $p^2 + 2pq + q^2$ this is taken as evidence for selection on the zygotic stage. If the results are nonsignificant this could be due to either selection at the gametic stage or on the zygotic stage with additive fitness values.

Sequence data from this article can be found in the GenBank/EMBL data libraries under accession numbers DU667459 to DU708532.

Supplemental Data

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

Supplemental Table S1. Marker data.

Supplemental Table S2. Microsatellite primer information.

ACKNOWLEDGMENTS

Thank you to Martin Lysak, Hans de Jong, and two anonymous reviewers for comments on the manuscript. We also thank Christoph Dobeš, Marcus Koch, J. Chris Pires, and Tim Sharbel for discussions. Additionally, we thank Kathy Springer, Petra Eissmann, Steffi Gebauer-Jung, Domenica Schnabelrauch, Kerstin Weniger, and Kevin Shianna for technical assistance. We also acknowledge Andrew Heidel for providing primer pairs. Thank also to Dr. Bitty Roy for the photograph of the SAD12 field site.

Received January 29, 2007; accepted February 28, 2007; published March 16, 2007.

LITERATURE CITED

Abajian C (1994) Sputnik. http://espressosoftware.com/pages/sputnik.jsp

AGI (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature **408:** 796–815

Al-Shehbaz IA (2003) Transfer of most North American species of *Arabis* to *Boechera* (Brassicaceae). Novon **13:** 381–391

Al-Shehbaz IA, Beilstein MA, Kellogg EA (2006) Systematics and phylogeny of the Brassicaceae (Cruciferae): an overview. Plant Syst Evol **259:** 89–120

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol **215:** 403–410

Altshuler D, Brooks LD, Chakravarti A, Collins FS, Daly MJ, Donnelly P (2005) A haplotype map of the human genome. Nature **437:** 1299–1320

Armstrong SJ, Franklin FCH, Jones GH (2001) Nucleolus-associated telomere clustering and pairing precede meiotic chromosome synapsis in *Arabidopsis thaliana*. J Cell Sci **114:** 4207–4217

Bailey CD, Koch MA, Mayer M, Mummenhoff K, O'Kane SL Jr, Warwick SI, Windham MD, Al-Shehbaz IA (2006) Toward a global phylogeny of the Brassicaceae. Mol Biol Evol **23:** 2142–2160

Beilstein MA, Al-Shehbaz IA, Kellogg EA (2006) Brassicaceae phylogeny and trichome evolution. Am J Bot **93:** 607–619

Benderoth M, Textor S, Windsor AJ, Mitchell-Olds T, Gershenzon J, Kroymann J (2006) Positive selection driving diversification in plant secondary metabolism. Proc Natl Acad Sci USA **103:** 9118–9123

Berr A, Pecinka A, Meister A, Kreth G, Fuchs J, Blattner FR, Lysak MA, Schubert I (2006) Chromosome arrangement and nuclear architecture but not centromeric sequences are conserved between *Arabidopsis thaliana* and *Arabidopsis lyrata*. Plant J **48:** 771–783

Bishop JG, Dean AM, Mitchell-Olds T (2000) Rapid evolution in plant chitinases: molecular targets of selection in plant-pathogen coevolution. Proc Natl Acad Sci USA **97:** 5322–5327

Böcher TW (1951) Cytological and embryological studies in the amphi-apomictic *Arabis holboellii* complex. Kong Danske Vidensk Selsk Biol Skr **6:** 1–59

Boivin K, Acarkan A, Mbulu RS, Clarenz O, Schmidt R (2004) The Arabidopsis genome sequence as a tool for genome analysis in Brassicaceae: a comparison of the Arabidopsis and *Capsella rubella* genomes. Plant Physiol **135:** 735–744

Bratteler M, Lexer C, Widmer A (2006) A genetic linkage map of *Silene vulgaris* based on AFLP markers. Genome **49:** 320–327

Clauss MJ, Cobban H, Mitchell-Olds T (2002) Cross-species microsatellite markers for elucidating population genetic structure in *Arabidopsis* and *Arabis* (Brassicaeae). Mol Ecol **11:** 591–601

Devos KM (2005) Updating the ''crop circle''. Curr Opin Plant Biol **8:** 155–162

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

Dobeš CH, Mitchell-Olds T, Koch MA (2004b) Extensive chloroplast haplotype variation indicates Pleistocene hybridization and radiation of North American *Arabis drummondii*, *A. x divaricarpa*, and *A. holboellii* (Brassicaceae). Mol Ecol **13:** 349–370

Dobeš CH, Sharbel TF, Koch MA (2007) Towards understanding the dynamics of hybridization and apomixis in the evolution of genus *Boechera* (Brassicaceae). System Biodivers (in press)

Ewing B, Green P (1998) Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res **8:** 186–194

Ewing B, Hillier L, Wendl MC, Green P (1998) Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res **8:** 175–185

Fajkus J, Sykorova E, Leitch AR (2005) Telomeres in evolution and evolution of telomeres. Chromosome Res **13:** 469–479

Gadberry MD, Malcomber ST, Doust AN, Kellogg EA (2005) Primaclade—a flexible tool to find conserved PCR primers across multiple species. Bioinformatics **21:** 1263–1264

Gordon D (2004) Viewing and editing assembled sequences using Consed. *In* AD Baxevanis, DB Davison, eds, Current Protocols in Bioinformatics. John Wiley & Co., New York, pp 11.12.11–11.12.43

Hall AE, Keith KC, Hall SE, Copenhaver GP, Preuss D (2004) The rapidly evolving field of plant centromeres. Curr Opin Plant Biol **7:** 108–114

Hall MC, Willis JH (2005) Transmission ratio distortion in intraspecific hybrids of *Mimulus guttatus*: implications for genomic divergence. Genetics **170:** 375–386

Ilic K, Berleth T, Provart NJ (2004) BlastDigester—a web-based program for efficient CAPS marker design. Trends Genet **20:** 280–283

Jenczewski E, Gherardi M, Bonnin I, Prosperi JM, Olivieri I, Huguet T (1997) Insight on segregation distortions in two intraspecific crosses between annual species of *Medicago* (Leguminosae). Theor Appl Genet **94:** 682–691

Johnston JS, Pepper AE, Hall AE, Chen ZJ, Hodnett G, Drabek J, Lopez R, Price HJ (2005) Evolution of genome size in Brassicaceae. Ann Bot (Lond) **95:** 229–235

Jones T, Kulseth S, Mechtenberg K, Jorgenson C, Zehfus M, Brown P, Siemens DH (2006) Simultaneous evolution of competitiveness and defense: induced switching in *Arabis drummondii*. Plant Ecol **184:** 245–257

Kantama L (2005) Chromosome studies and genetic analyses of natural and synthetic apomictic model species. PhD thesis. Wageningen University, Wageningen, The Netherlands

Kawabe A, Hansson B, Hagenblad J, Forrest A, Charlesworth D (2006) Centromere locations and associated chromosome rearrangements in *Arabidopsis lyrata* and *A. thaliana*. Genetics **173:** 1613–1619

Kim JS, Chung TY, King GJ, Jin M, Yang TJ, Jin YM, Kim HI, Park BS (2006) A sequence-tagged linkage map of *Brassica rapa*. Genetics **174:** 29–39

Knight CA, Vogel H, Kroymann J, Shumate A, Witsenboer H, Mitchell-Olds T (2006) Expression profiling and local adaptation of *Boechera holboellii* populations for water use efficiency across a naturally occurring water stress gradient. Mol Ecol **15:** 1229–1237

Koch M, Al-Shehbaz IA, Mummenhoff K (2003) Molecular systematics, evolution, and population biology in the mustard family (Brassicaceae). Ann Mo Bot Gard **90:** 151–171

Koch MA, Dobeš C, Kiefer C, Schmickl R, Klimes L, Lysak MA (2007) Supernetwork identifies multiple events of plastid trnF((GAA)) pseudogene evolution in the Brassicaceae. Mol Biol Evol **24:** 63–73

Koch MA, Kiefer M (2005) Genome evolution among cruciferous plants: a lecture from the comparison of the genetic maps of three diploid species—*Capsella rubella*, *Arabidopsis lyrata* subsp *Petraea*, and *A. thaliana*. Am J Bot **92:** 761–767

Kuittinen H, de Haan AA, Vogl C, Oikarinen S, Leppala J, Koch M, Mitchell-Olds T, Langley CH, Savolainen O (2004) Comparing the linkage maps of the close relatives *Arabidopsis lyrata* and *A. thaliana*. Genetics **168:** 1575–1584

Lagercrantz U (1998) Comparative mapping between *Arabidopsis thaliana* and *Brassica nigra* indicates that *Brassica* genomes have evolved through extensive genome replication accompanied by chromosome fusions and frequent rearrangements. Genetics **150:** 1217–1228

Lysak M, Berr A, Pecinka A, Schmidt R, McBreen K, Schubert I (2006) Mechanisms of chromosome number reduction in *Arabidopsis thaliana* and related Brassicaceae species. Proc Natl Acad Sci USA **103:** 5224–5229

Lysak MA, Koch MA, Pecinka A, Schubert I (2005) Chromosome triplication found across the tribe Brassiceae. Genome Res **15:** 516–525

Lysak MA, Lexer C (2006) Towards the era of comparative evolutionary genomics in Brassicaceae. Plant Syst Evol **259:** 175–198

Marhold K, Lihova J (2006) Polyploidy, hybridization and reticulate evolution: lessons from the Brassicaceae. Plant Syst Evol **259:** 143–174

Michaels SD, Amasino RM (1998) A robust method for detecting single-nucleotide changes as polymorphic markers by PCR. Plant J **14:** 381–385

Mitchell-Olds T (2001) *Arabidopsis thaliana* and its wild relatives: a model system for ecology and evolution. Trends Ecol Evol **16:** 693–700

Moore G, Devos KM, Wang Z, Gale MD (1995) Cereal genome evolution—grasses, line up and form a circle. Curr Biol **5:** 737–739

Neff MM, Turk E, Kalishman M (2002) Web-based primer design for single nucleotide polymorphism analysis. Trends Genet **18:** 613–615

Parkin IAP, Gulden SM, Sharpe AG, Lukens L, Trick M, Osborn TC, Lydiate DJ (2005) Segmental structure of the *Brassica napus* genome based on comparative analysis with *Arabidopsis thaliana*. Genetics **171:** 765–781

Pecinka A, Schubert V, Meister A, Kreth G, Klatte M, Lysak MA, Fuchs J, Schubert I (2004) Chromosome territory arrangement and homologous pairing in nuclei of *Arabidopsis thaliana* are predominantly random except for NOR-bearing chromosomes. Chromosoma **113:** 258–269

Rice P, Longden I, Bleasby A (2000) EMBOSS: the European molecular biology open software suite. Trends Genet **16:** 276–277

Rollins RC (1993) The Cruciferae of Continental North America. Stanford University Press, Stanford, CA

Rostoks N, Ramsay L, MacKenzie K, Cardle L, Bhat PR, Roose ML, Svensson JT, Stein N, Varshney RK, Marshall DF, et al (2006) Recent history of artificial outcrossing facilitates whole-genome association mapping in elite inbred crop varieties. Proc Natl Acad Sci USA **103:** 18656–18661

Roy BA (1993) Floral mimicry by a plant pathogen. Nature **362:** 56–58

Roy BA, Kirchner JW (2000) Evolutionary dynamics of pathogen resistance and tolerance. Evolution Int J Org Evolution **54:** 51–63

Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. *In* S Krawetz, S Misener, eds, Bioinformatics Methods and Protocols: Methods in Molecular Biology. Humana Press, Totowa, NJ, pp 365–386

Schein M, Yang ZH, Mitchell-Olds T, Schmid KJ (2004) Rapid evolution of a pollen-specific oleosin-like gene family from *Arabidopsis thaliana* and closely related species. Mol Biol Evol **21:** 659–669

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

Schranz ME, Kantama L, de Jong H, Mitchell-Olds T (2006a) Asexual reproduction in a close relative of *Arabidopsis*: a genetic investigation of apornixis in *Boechera* (Brassicaceae). New Phytol **171:** 425–438

Schranz ME, Lysak MA, Mitchell-Olds T (2006b) The ABC's of comparative genomics in the Brassicaceae: building blocks of crucifer genomes. Trends Plant Sci **11:** 535–542

Schranz ME, Song BH, Windsor AJ, Mitchell-Olds T (2007) Comparative genomics in the Brassicaceae: a family-wide perspective. Curr Opin Plant Biol **10:** 168–175

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

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

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

Siroky J, Zluvova J, Riha K, Shippen DE, Vyskot B (2003) Rearrangements of ribosomal DNA clusters in late generation telomerase-deficient Arabidopsis. Chromosoma **112:** 116–123

Song BH, Clauss MJ, Pepper A, Mitchell-Olds T (2006) Geographic patterns of microsatellite variation in *Boechera stricta*, a close relative of *Arabidopsis*. Mol Ecol **15:** 357–369

Stam P (1993) Construction of integrated genetic-linkage maps by means of a new computer package—Joinmap. Plant J **3:** 739–744

Thiel T, Kota R, Grosse I, Stein N, Graner A (2004) SNP2CAPS: a SNP and INDEL analysis tool for CAPS marker development. Nucleic Acids Res **32:** e5

Warwick SI, Al-Shehbaz IA (2006) Brassicaceae: chromosome number index and database on CD-Rom. Plant Syst Evol **259:** 237–248

Windham MD, Al-Shehbaz IA (2006) New and noteworthy species of *Boechera* (Brassicaceae) I: sexual diploids. Harv Pap Bot **11:** 61–88

Windsor AJ, Schranz ME, Formanova N, Gebauer-Jung S, Bishop JG, Schnabelrauch D, Kroymann J, Mitchell-Olds T (2006) Partial shotgun sequencing of the *Boechera stricta* genome reveals extensive microsynteny and promoter conservation with Arabidopsis. Plant Physiol **140:** 1169–1182

Yang YW, Lai KN, Tai PY, Li WH (1999) Rates of nucleotide substitution in angiosperm mitochondrial DNA sequences and dates of divergence between *Brassica* and other angiosperm lineages. J Mol Evol **48:** 597–604

Yogeeswaran K, Frary A, York TL, Amenta A, Lesser AH, Nasrallah JB, Tanksley SD, Nasrallah ME (2005) Comparative genome analyses of *Arabidopsis* spp.: inferring chromosomal rearrangement events in the evolutionary history of *A. thaliana*. Genome Res **15:** 505–515

Zhang JH, Madden TL (1997) PowerBLAST: a new network BLAST application for interactive or automated sequence analysis and annotation. Genome Res **7:** 649–656