Supplementary Information

Young inversion with multiple linked QTLs under selection in a hybrid zone

Cheng-Ruei Lee[1, 2], Baosheng Wang[1,3], Julius Mojica[1], Terezie Mandáková[4], Kasavajhala V. S. K. Prasad[5], Jose Luis Goicoechea[6], Nadeesha Perera[1], Uffe Hellsten[7], Hope N. Hundley[7], Jenifer Johnson[7], Jane Grimwood[8], Kerrie Barry[7], Stephen Fairclough[7], Jerry W. Jenkins[7], Yeisoo Yu[9], Dave Kudrna[6], Jianwei Zhang[6], Jayson Talag[6], Wolfgang Golser[6], Katherine Ghattas[1], M. Eric Schranz[10], Rod Wing[6], Martin A. Lysak[4], Jeremy Schmutz[7], Daniel S. Rokhsar[7], Thomas Mitchell-Olds[1]

[1]Department of Biology, Box 90338, Duke University, Durham, NC, 27708, USA

[2]Institute of Ecology and Evolutionary Biology & Institute of Plant Biology, National Taiwan University, Taipei 10617, Taiwan ROC

[3]Department of Plant Ecology and Genetics, Uppsala University, Norbyvägen 18D, SE-752 36 Uppsala, Sweden

[4]Plant Cytogenomics Group, CEITEC – Central European Institute of Technology, Masaryk University, Kamenice 5, Brno CZ-62500, Czech Republic

[5]Department of Biology, Colorado State University, Fort Collins, Colorado 80523, USA

[6]Arizona Genomics Institute and BIO5 Institute, School of Plant Sciences, University of Arizona, Tucson, AZ, 85721, USA

[7]Department of Energy Joint Genome Institute, Walnut Creek, California, 94598, USA

[8]HudsonAlpha Institute for Biotechnology, Huntsville, Alabama, USA

[9]Phyzen Genomics Institute, Phyzen Inc., Seoul, 151-836, South Korea

[10]Biosystematics Group, Wageningen University & Research Center, Droevendaalsesteeg 1, 6708PB, Wageningen, The Netherlands

This PDF file includes:

Figs. S1 - S9

Tables S1 - S8

Other Supplementary Materials for this manuscript:

All data files is available from http://www.datadryad.org. These include:

Script for power simulation: Script_01-Power simulation.R.txt

Script for ancestral alleles: Script_get_anc_for_variable_sites.py.txt

Script for MANOVA permutation: Script_MANOVA_permutation.R.txt

Script for VCF reformatting: Script_Transf_vcf2fas.py.txt

Script for GATK analysis: Script_Variant_discovery.txt

Boechera library information: Sup Data Boechera Libraries.xlsx

Experimental crosses: Sup Data Experimental Crosses.docx

Gene expression data: Sup Data Expression.xlsx

Genotypes used: Sup Data Genotype_Information.xlsx

RIL genotype matrix: Sup Data Genotype_Matrix_For_RILs.xlsx

Hybrid zone genotype matrix: Sup Data INV122_RP83.snp.geno.xlsx

Annotation of SNPs within the inversion: Sup Data Inversion_SNP_annotation.xlsx

JGI library information: Sup Data JGI_Libraries.xlsx

Breakpoints and primers: Sup Data Locations of breakpoints and related primers.docx

Crossover locations: Sup Data LTMxSAD12_Inferred_Crossovers.xlsx

NIL flowering times: Sup Data LTMxSAD12_NIL_Flowering_time.xlsx

Recombinant inbred genotypes: Sup Data LTMxSAD12_RIL_IDS.xlsx

F2 phenotypes: Sup Data LTMxSDM_Trait_Matrix.xlsx

Linkage map for RILs: Sup Data Map_For_LTMxSAD_RILs.xlsx

Collinear Cross traits_1:

Sup Data ParkerxRuby-KAS6_GH_flowering_20120804-Indiv-Data_3-Blks.xls

Collinear Cross traits_2:

Sup Data ParkerxRuby-KAS6_GH_flowering_20120804-Indiv-Data_9-Blks.xlsx

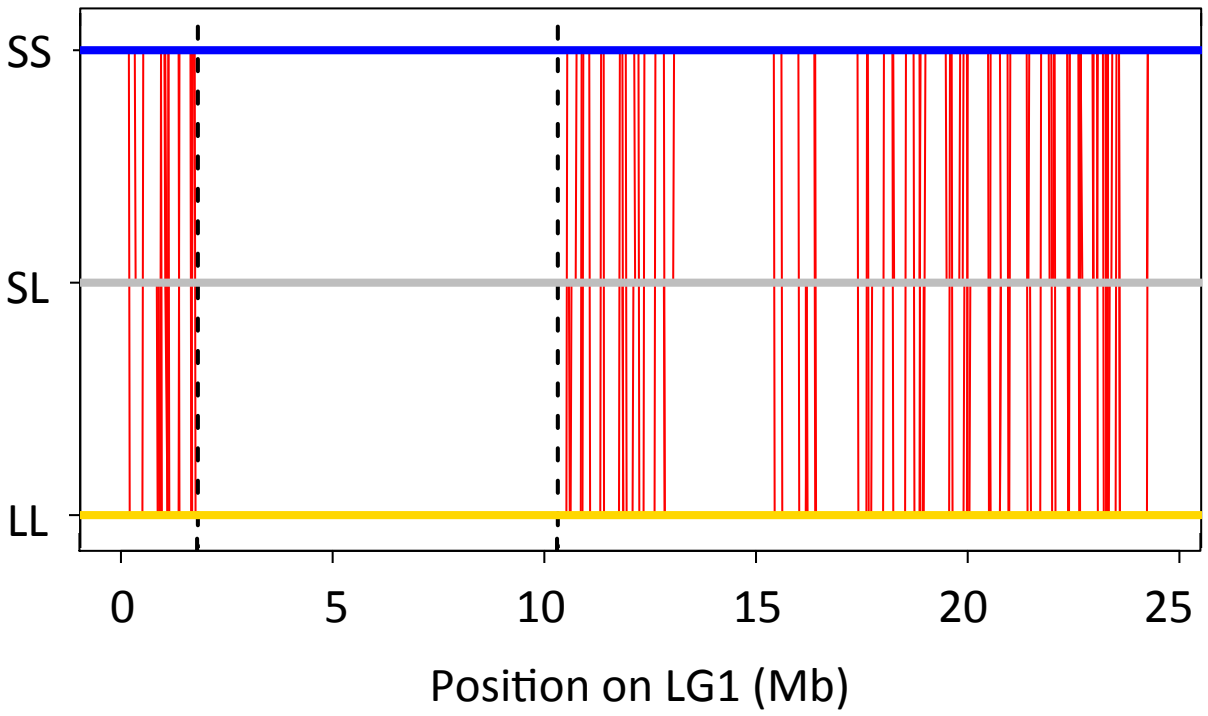
Collinear cross family means:

Sup Data ParkerxRuby-KAS6_GH_Trait_Matrix-Family_Means.xlsx

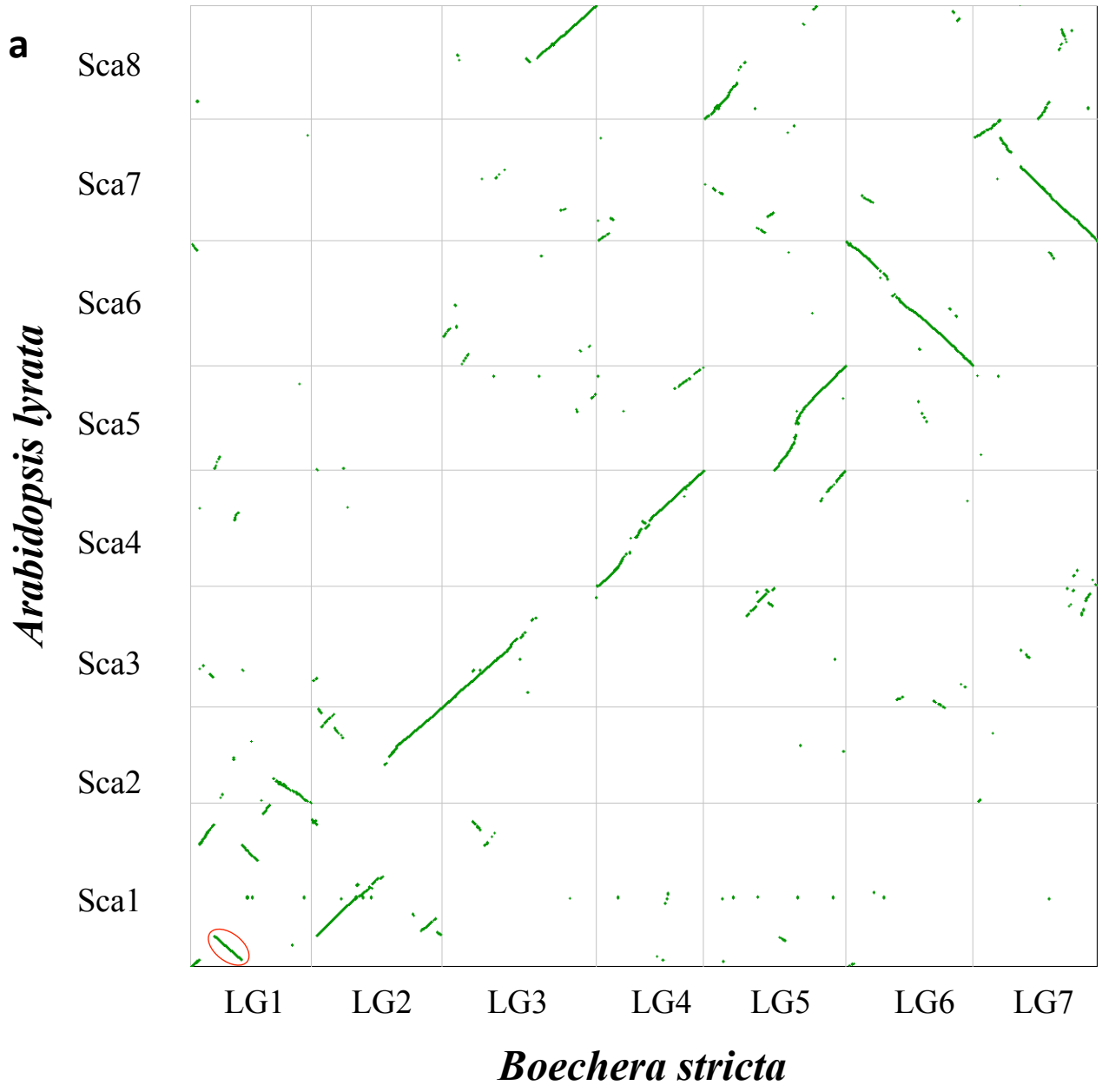
Linkage map data for collinear cross: Sup Data PARxRUBY_Linkage_Map.xlsx

Primers used for gene expression: Sup Data Primers for gene expression.docx

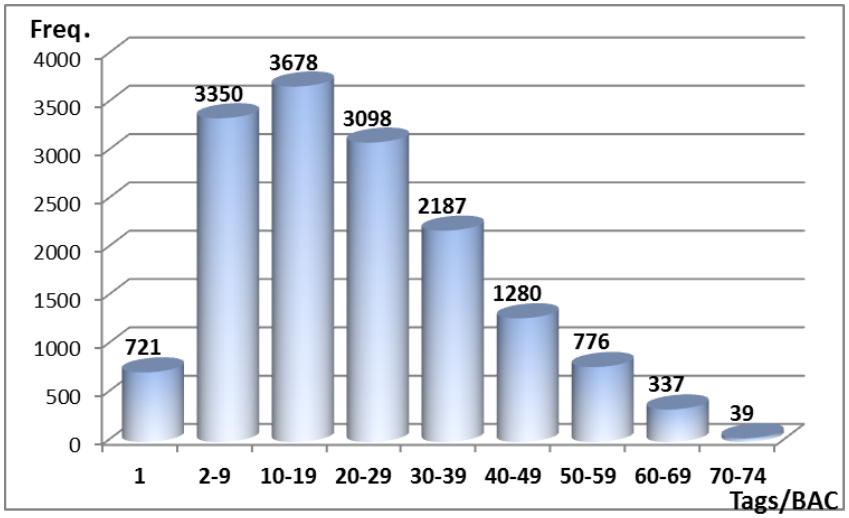
Whole genome profiling: Sup Data Whole Genome Profiling.xlsx



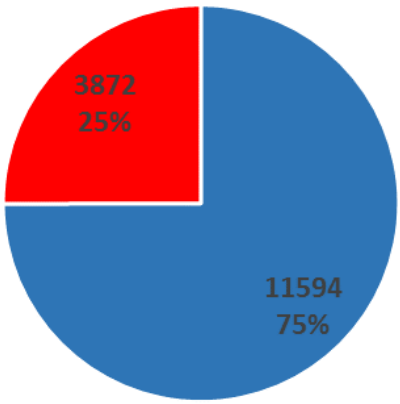
Supplementary Figure 1: Recombination on LG1. No recombination events occurred within the inversion in a recombinant inbred population of 159 F6 individuals. Locations of crossovers are shown along LG1, for F6 individuals from the LTM x SAD12 (WEST-*inv* x EAST-*std*) cross. Horizontal axis shows nucleotide position, and vertical axis shows SAD12 homozygous regions (SS, blue line), heterozygous regions (SL, gray line), and homozygous LTM regions (LL, yellow line). Crossovers are shown by vertical red lines, from one homozygote to another (full length red line) or between homozygous and heterozygous regions (half-height red lines). The inversion region is delimited by vertical dashed lines. These data are compatible with the hypothesis that a pre-existing haplotype was captured by the new inversion, or that an extremely rare double-recombinant was captured by the young inversion before its selective sweep.


a
Sca8
Sca7
Sca6
Sca5
Sca4
Sca3
Sca2
Sca1
Arabidopsis lyrata
LG1
LG2
LG3
LG4
LG5
LG6
LG7
Boechera stricta

b

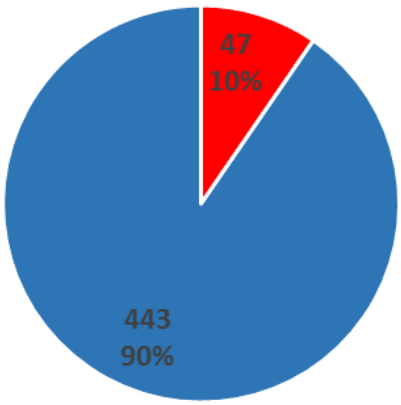


c



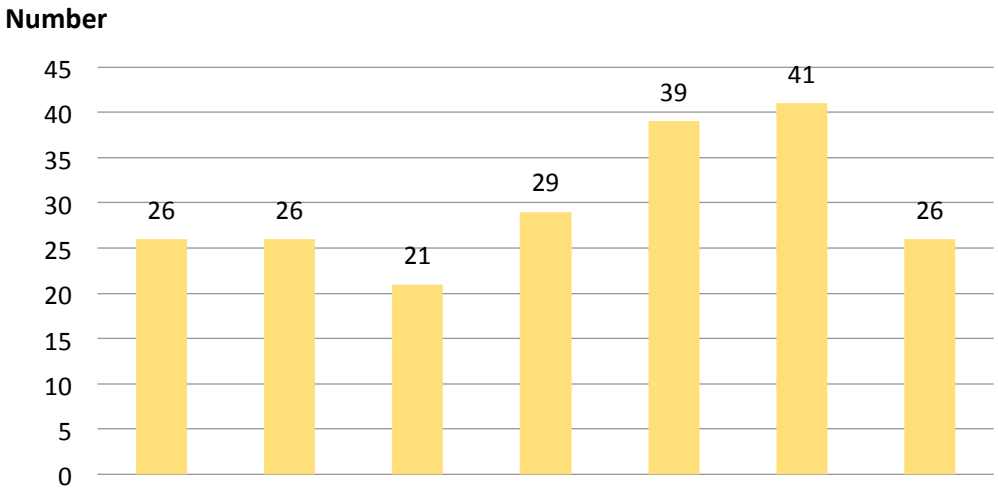
In Contigs Singletons

d

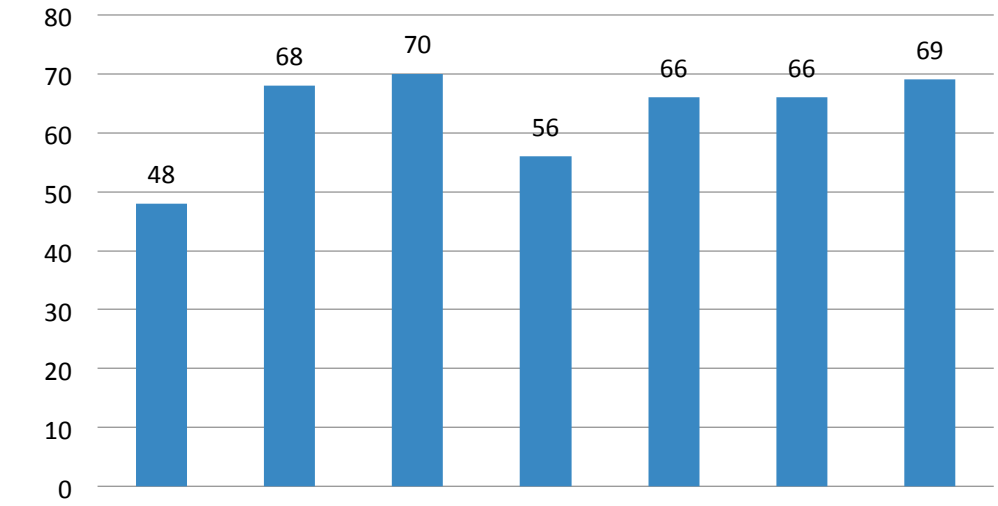


Not anchored Anchored

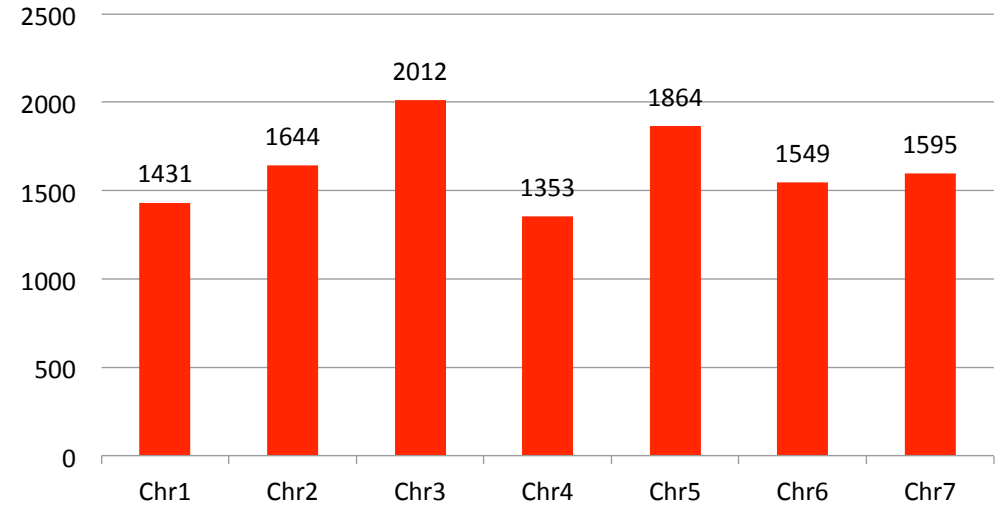
e


Number
45
40
35
30
25
20
15
10
5
0
26
26
21
29
39
41
26

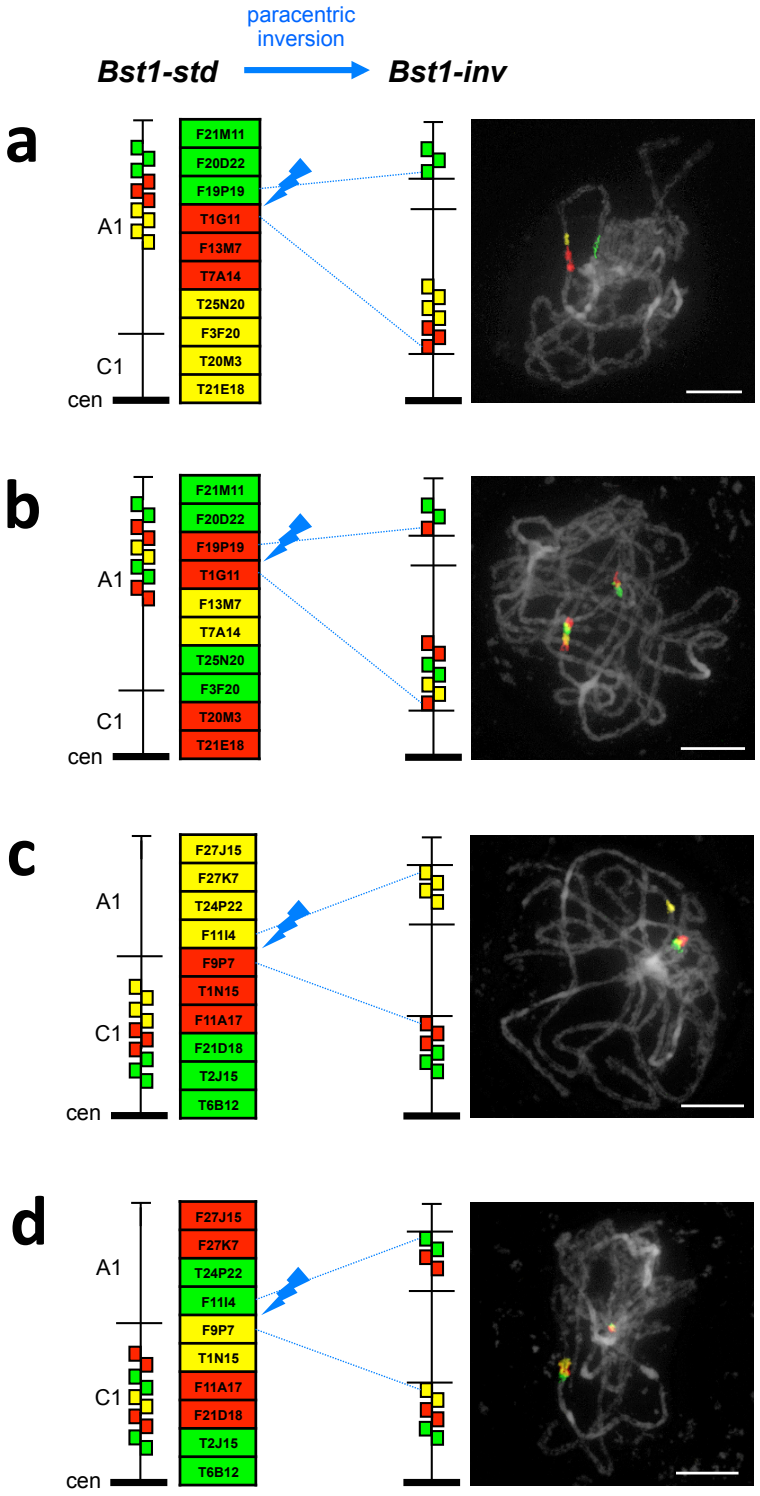
f


80
70
60
50
40
30
20
10
0
48
68
70
56
66
66
69

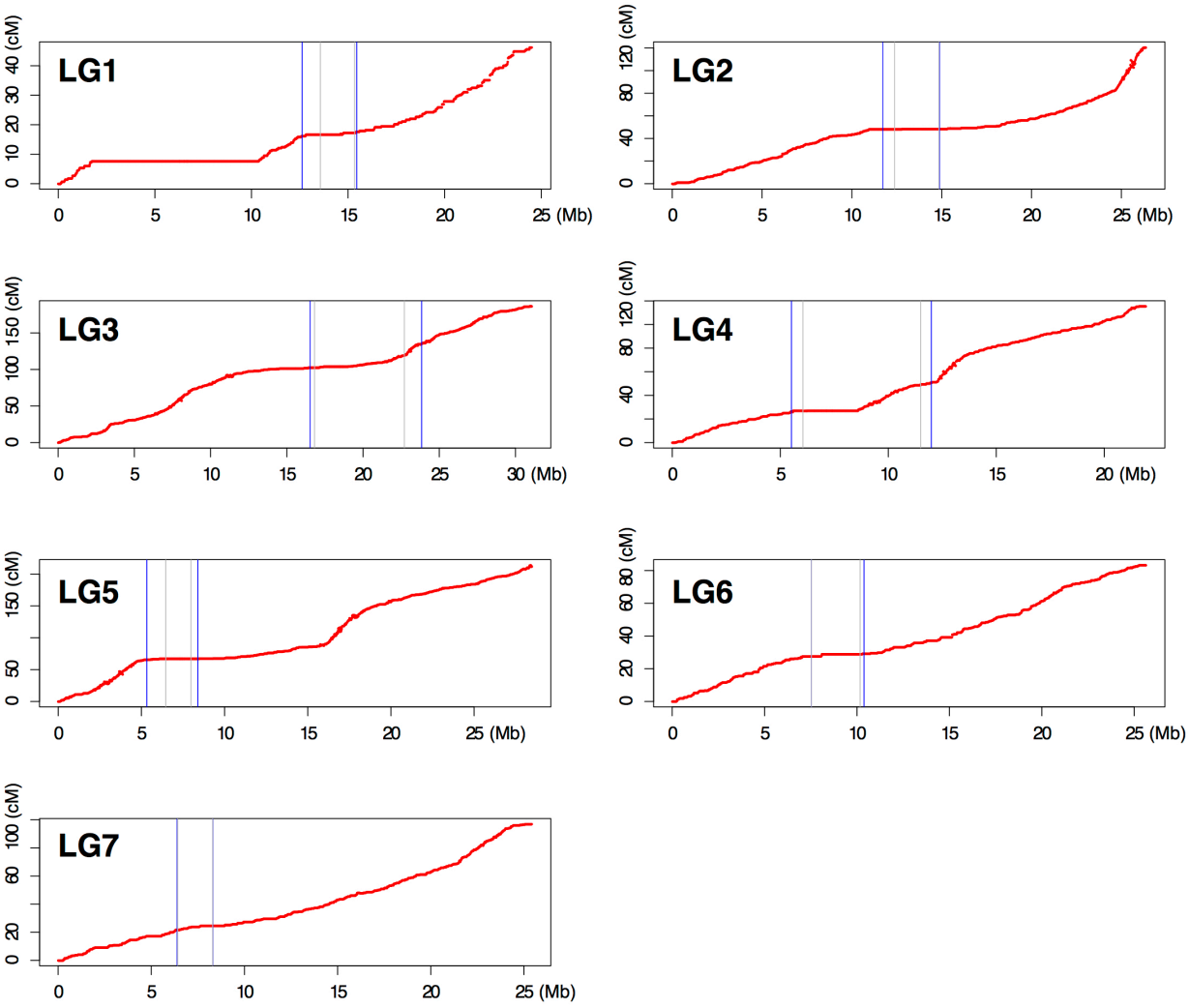
g


2500
2000
1500
1000
500
0
1431
1644
2012
1353
1864
1549
1595
Chr1
Chr2
Chr3
Chr4
Chr5
Chr6
Chr7

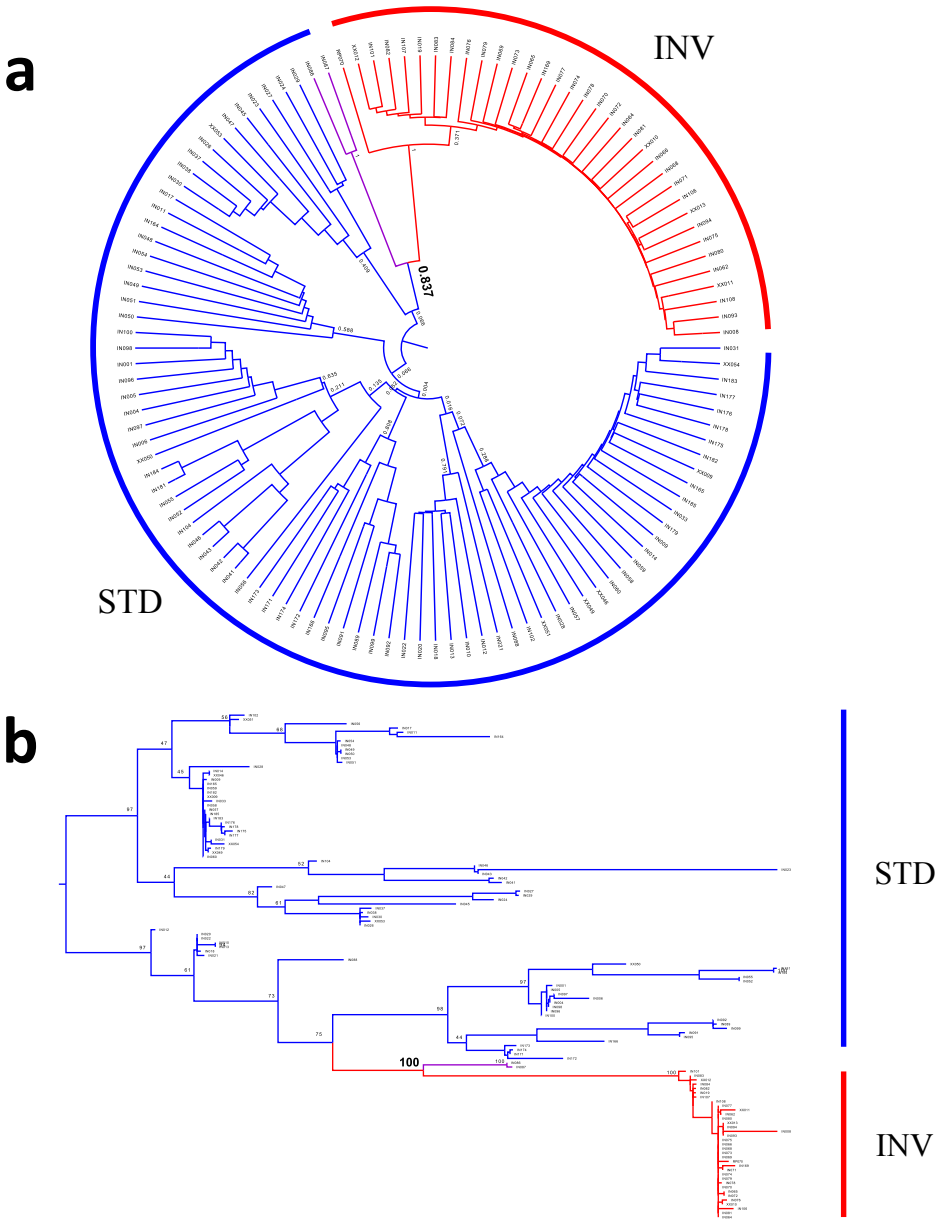
Supplemental Figure 2: Results from genome sequencing. (**a**) Genome-scale synteny between *B. stricta* (n=7) and *A. lyrata* (n=8). The longest eight scaffolds are shown for *A. lyrata*. A complete overview of the most parsimonious model of chromosome evolution of *B. stricta* has been presented elsewhere[53]. The region within the circle represents the paracentric inversion found in the reference genotype for genomic blocks A1b and C1a (Fig. 1). (**b**) Number of Whole Genome Profiling tags / BAC. (**c**) Distribution of clones in contigs. (**d**) Distribution of anchored contigs. (**e**) Number of scaffolds / chromosome. (**f**) Number of contigs / chromosome. (**g**) Number of clones anchored to chromosomes.



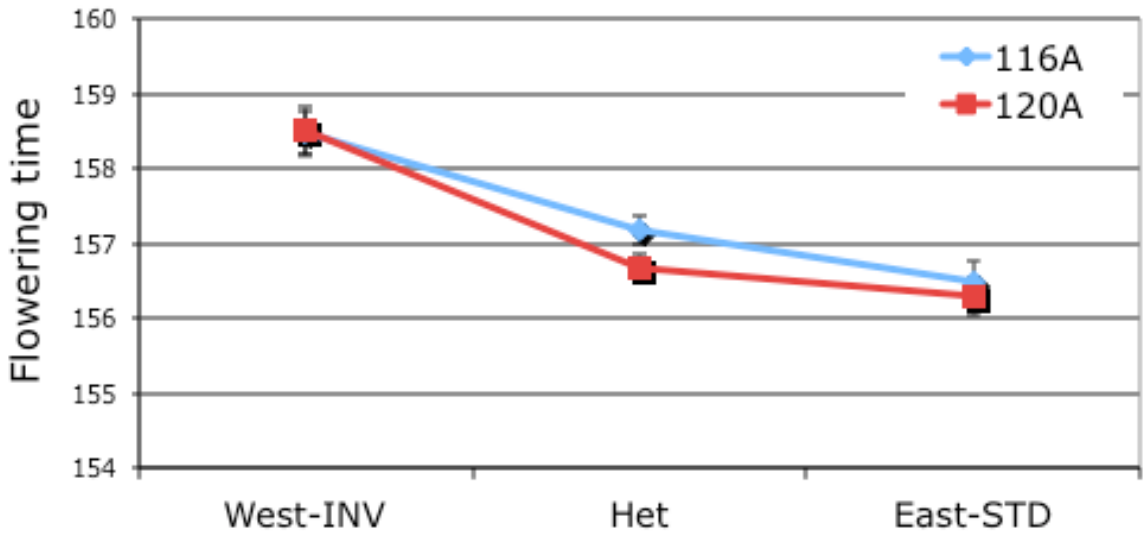
Supplementary Figure 3: Close-up of *Bsi1-inv* breakpoints using comparative chromosome painting on pachytene chromosomes. Shown are diagrams and painting images around the *Bsi1* breakpoints: (**a** and **b**) show the A1 breakpoint between BACs F19P19 and T1G11; (**c** and **d**) show the C1 break-point at F9P7/F11I4.



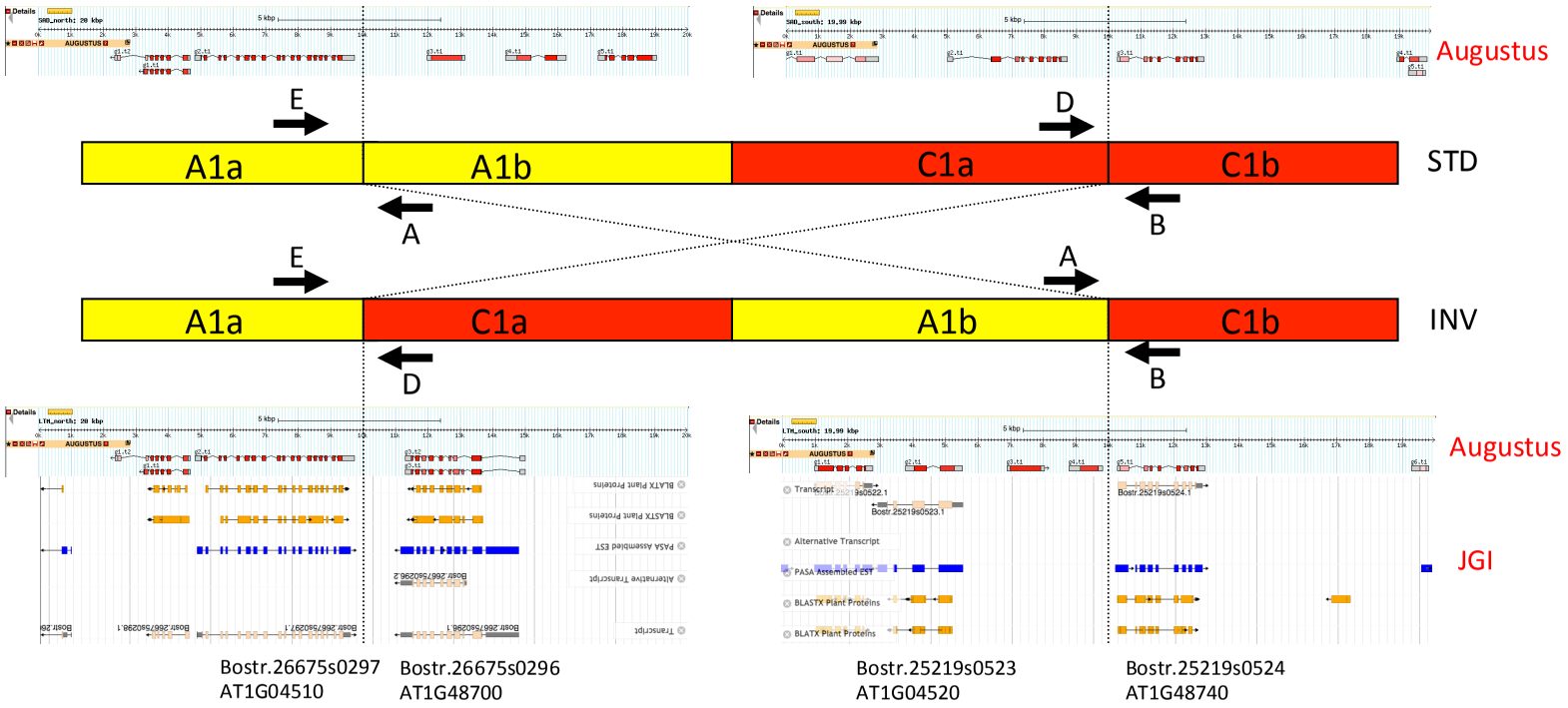
Supplementary Fig. 4: Predicted centromere locations for each linkage group. Shown are plots of physical and linkage map position on seven linkage groups (red lines). Blue vertical lines indicate the inferred margins of centromeres based on homology to *A. thaliana* centromeric BAC clones. Gray vertical lines indicate the margins of *B. stricta* scaffolds that contain these BAC-homologues.



Supplementary Figure 5: Inversion haplotypes form a recent, monophyletic clade. Shown are (**a**) Neighbor-Joining (NJ) and (**b**) Maximum likelihood (ML) trees based on 901 SNPs in the inversion region for 118 *Bsi1-inv* and *Bsi1-std* haplotypes. All *Bsi1-inv* haplotypes form a strongly supported clade (100% bootstrap support on NJ and ML trees) that is sister to a clade formed by two *Bsi1-std* genotypes (IN086 and IN087, purple) from Saddle Mountain. This pair of genotypes cluster together with 84% and 100% bootstrap support on NJ and ML trees, respectively. In turn, these two haplotypes are nested within other *Bsi1-std* genotypes. These results indicate a single origin of *Bsi1-inv* haplotypes from ancestral *Bsi1-std* haplotypes. The most closely related *Bsi1-inv* and *Bsi1-std* genotypes are only 1.7 km away from each other and have very low levels of nucleotide divergence (0.21 SNPs per kb of Genotype By Sequencing reads). The maximum and mean distances between INV haplotypes are 1.230 $\times 10^{-4}$ and 2.904 $\times 10^{-5}$, respectively, allowing a rough calibration for the age of inversion at 2,074 – 8,785 years.

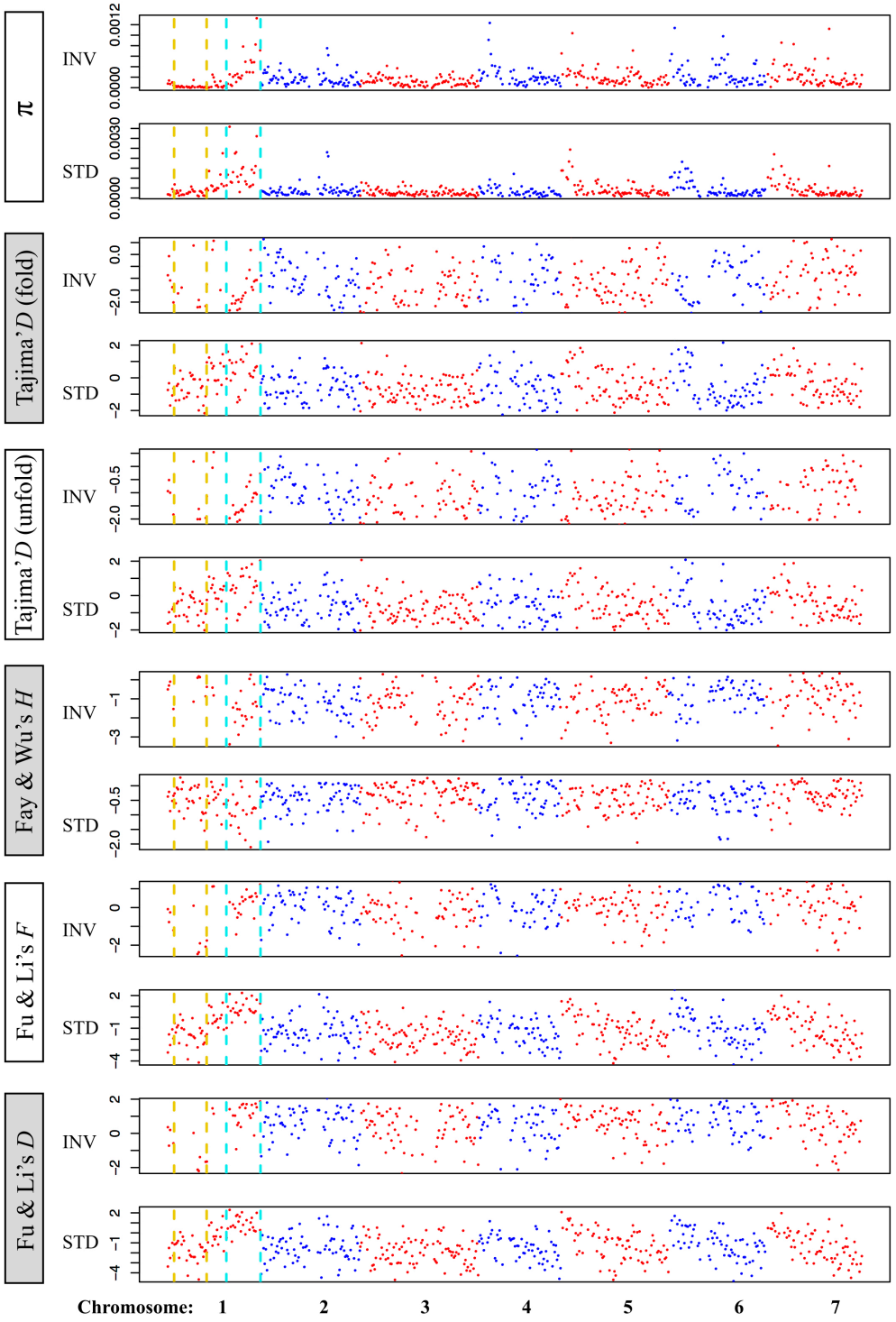


Supplementary Figure 6: Flowering time in two near-isogenic segregating families. Shown is flowering time of two F7 WEST-inv x EAST-std heterogeneous inbred families (HIFs, or segregating near-isogenic lines). Two independent HIFs (116A, blue) and (120A, red) show phenotypic effects of the inversion region segregating within each family, while the rest of the genome is nearly isogenic in each family. Genotypes in families are either homozygous WEST-inv, heterozygous, or homozygous EAST-std. Flowering time is indicated in days; error bars are +/- 1 standard error of the mean.

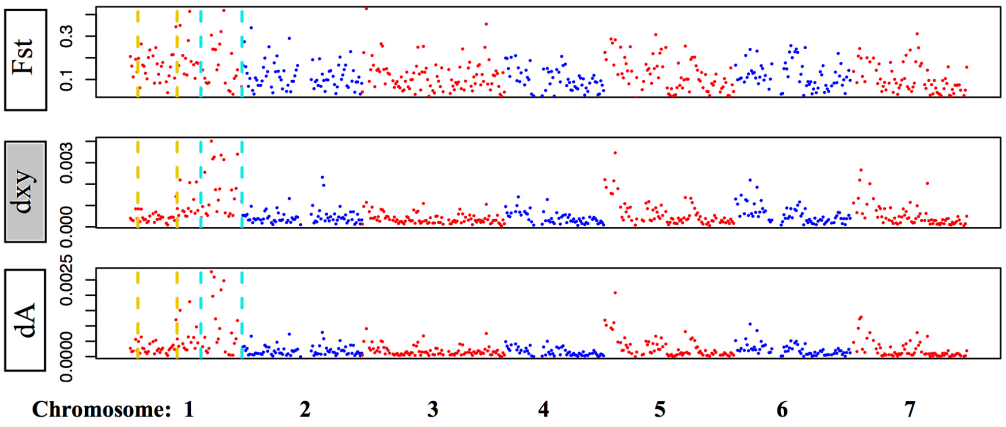


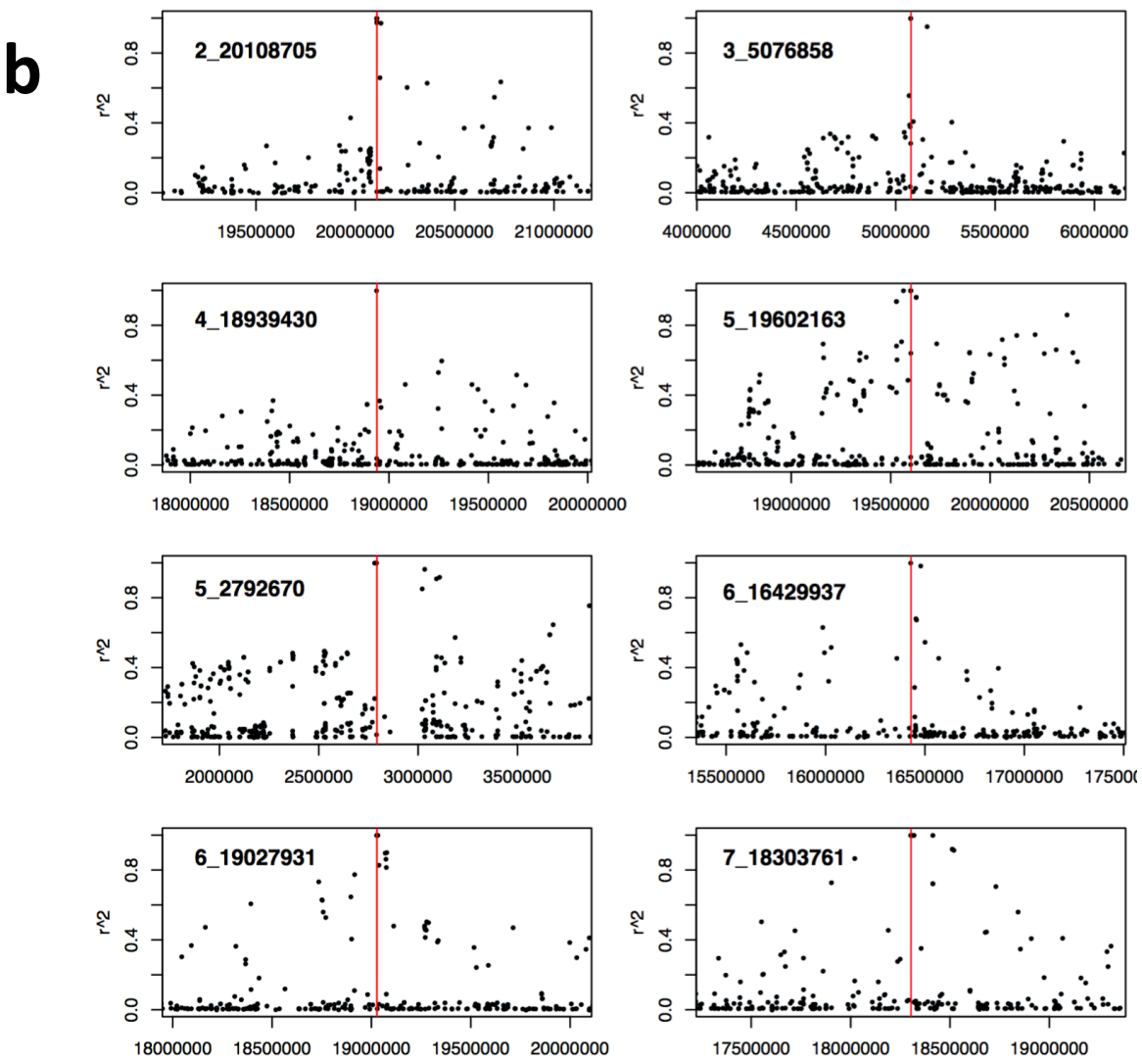
Supplementary Figure 7: Gene models near inversion breakpoints. Gene models syntenic to Arabidopsis are indicated near the inversion breakpoints, based on Augustus analysis of the hard-masked 20kb regions centered on each breakpoint in the reference genome (inversion) and the inferred standard haplotype. Primers used to distinguish inversion vs. standard haplotypes are indicated by lettered arrows. No open reading frames spanning the breakpoints in either the inversion or standard haplotypes were found, hence the chromosomal inversion does not disrupt existing or create new open reading frames. Shown are gene models from Augustus and/or JGI analysis, which largely correspond to the transcriptome-assisted annotation in *B. Stricta* genome v1.2. Different colors here indicate gene models from different annotation methods. Six complete genes are shown: Bostr.26675s0298, Bostr.26675s0297, Bostr.26675s0296, Bostr.25219s0522, Bostr.25219s0523, Bostr.25219s0524. The four genes flanking both breakpoints are labeled.

a

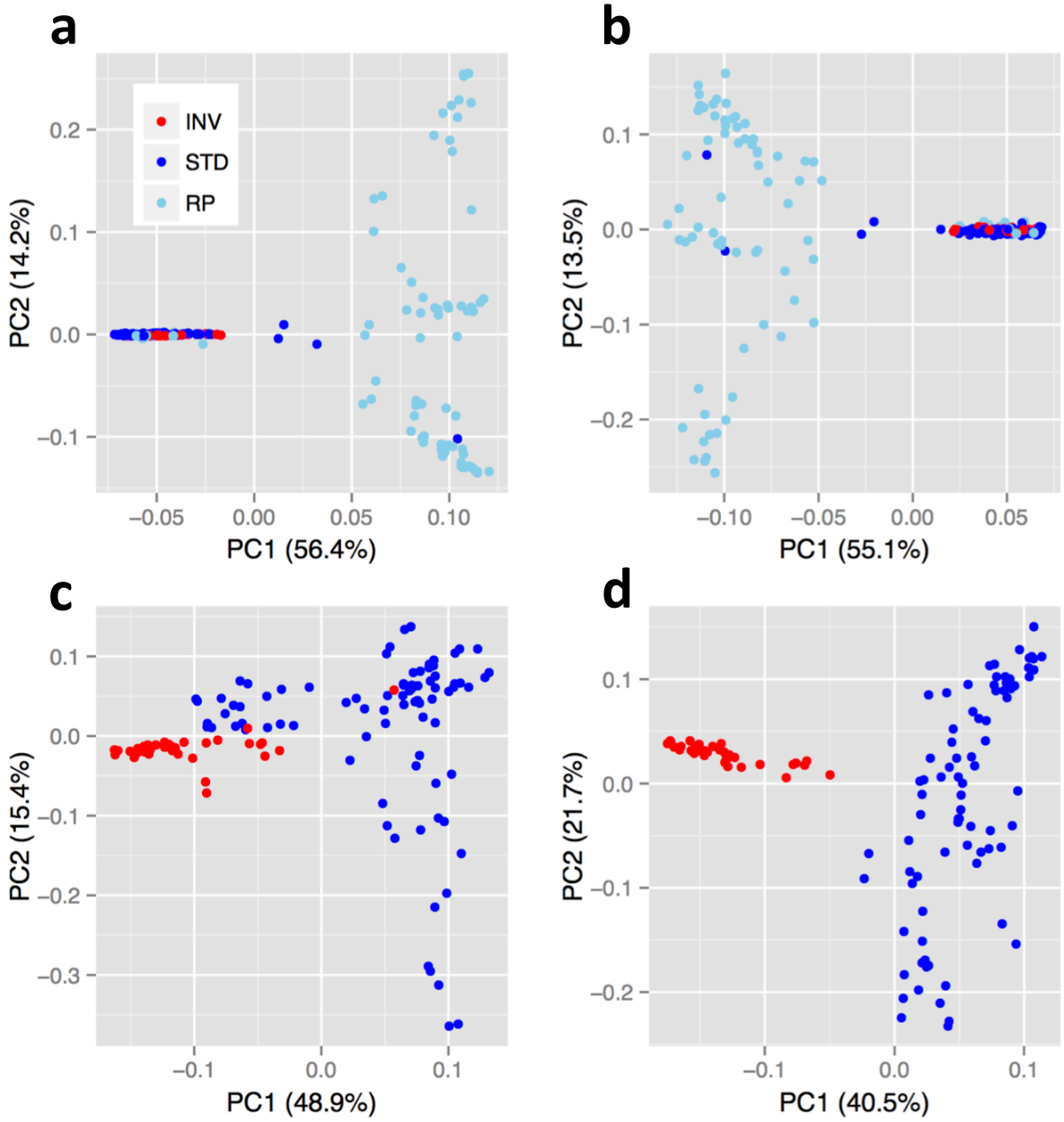

π
INV
STD
0.0012
0.0000
0.0030
Tajima'D (fold)
Tajima'D (unfold)
Fay & Wu's H
Fu & Li's F
Fu & Li's D
0.0
−2.0
2
0
−2
−0.5
−1
−3
−4
Chromosome:
1
2
3
4
5
6
7

a





Supplementary Figure 8: Population genomics analyses. (**a**) Manhattan plots of population genomic parameters estimated in 300 kb non-overlapping windows. Unfolded (or folded) analyses use (or do not use) information on the ancestral allele. Nucleotide diversity (π), Tajima's *D*, Fay & Wu's *H*, Fu & Li's *F* and Fu & Li's *D* were calculated for each of the INV and STD groups. Population genetic differentiation (*Fst*), pairwise nucleotide divergence (*Dxy*), and net pairwise divergence (*dA*) were estimated between INV and STD groups. Alternating colors paint the different chromosomes, and the dashed vertical lines mark the inversion region (golden) and block D region (light blue) on LG1. (**b**) Linkage disequilibrium (r^2) between ten randomly chosen target SNPs having derived allele frequencies similar to *Bsi1-inv*, and their surrounding SNPs within 2 Mb (+/- 1Mb) windows centered on the target SNP, on chromosomes 2-7. Positions of target SNPs are indicated by vertical red lines.


a
INV
STD
RP
PC2 (14.2%)
PC1 (56.4%)
b
PC2 (13.5%)
PC1 (55.1%)
c
PC2 (15.4%)
PC1 (48.9%)
d
PC2 (21.7%)
PC1 (40.5%)

Supplementary Figure 9: Principal Components of species-wide and inversion zone genotypes. Genetic principal component analysis (PCA) of *Boechera stricta* based on genotyping by sequencing SNPs. (**a**) and (**b**) show PCA of 205 accessions based on genome-wide SNPs and the LG1 inverted region, respectively. These results support two diverged subspecies in *B. stricta* (EAST and WEST). All genotypes from the inversion zone were identified as members of the WEST group, except that four *Bsi1-std* genotypes had high levels of admixture from the EAST group (0.44 − 0.92 on PC1), indicating introgression from the EASTERN subspecies in the inversion zone. INV (red): inversion genotypes from inversion zone; STD (dark blue): standard genotypes from the inversion zone; Reference Population (RP, light blue): species-wide samples from outside the inversion zone. (**c)** and (**d**) show PCA of 118 accessions near the inversion zone based on genome-wide SNPs, or the inverted genome region, respectively. For comparison of inversion haplotypes to the most closely related standard haplotypes, we used all genotypes in the left portion of C. (One admixed *inv* genotype was excluded: IN019, shown in red on the right hand side).

Supplementary Table 1. Overview of linkage maps used in these experiments.

Parents	Cross	Generation	Map Length	Number of Bins	Method
EAST-std x WEST-std	Parker x Ruby	F4	1130 cM	1,011	Genotyping by sequencing
EAST-std x WEST-inv	SAD12 x LTM	F6	860 cM	9,273	Light sequencing

Detailed linkage maps are archived at Dryad.

Supplementary Table 2a. Test of segregation distortion in two NIL families between LTM and SAD12. *N* = 144.

Family	INV allele	STD allele	Chi-sq (df=1)	P
116A	142	146	0.0560	0.8140
120A	130	158	2.7220	0.0990

Supplementary Table 2b. Test of Hardy-Weinberg equilibrium (expected genotype frequency calculated from empirical allele frequency) in two NIL families between LTM and SAD12. N = 144.

Family	INV homozygotes	Heterozygotes	STD homozygotes	Chi-sq (df=1)	P
116A	33	76	35	0.4480	0.5033
120A	27	76	41	0.6200	0.4781

Supplementary Table 2c. Analysis of variance for flowering time in the greenhouse for two NIL families between LTM and SAD12. N = 288, R^2 = 14.7%.

Effect	DF	SS	F	P
Family	1	3.0694	0.8851	0.3476
Inversion Genotype	2	151.2487	21.8064	1.60E-09
Inversion Genotype* Family	2	3.1379	0.4524	0.6366

Supplementary Table 2d. Comparison of flowering times between EAST and WEST subspecies. T-test for flowering time in the greenhouse, comparing 12 EAST and 12 WEST genotypes. R^2 = 24.7%. Subspecies means equal 10.07 and 15.06 days for EAST and WEST subspecies, respectively. (Data taken from genotype LSMEANS, Table S1 in ref. 19).

Effect	DF	t-statistic	P
Subspecies	22	2.69	0.0134

Supplementary Table 3a: Test of segregation distortion in a sympatric, reciprocal WEST-*inv* x WEST-*std* F2 cross. *N* = 1,988 alleles = 994 individuals.

Cross	INV allele	STD allele	x^2 (df = 1)	*P*
CL9.1	483	507	0.582	0.446
CL10.1	495	503	0.064	0.800
Combined	978	1010	0.515	0.473

Supplementary Table 3b: Test of Hardy-Weinberg equilibrium (expected genotype frequency calculated from empirical allele frequency) in a sympatric, reciprocal WEST-*inv* x WEST-*std* F2 cross. *N* = 994 individuals.

Cross	INV homo	Het	STD homo	x^2 (df = 1)	*P*
CL9.1	117	249	129	0.022	0.882
CL10.1	127	241	131	0.577	0.447
Combined	244	490	260	0.190	0.663

Note:
CL9.1 = LTM as mother
CL10.1 = Saddle Mountain (SDM) as mother

Supplementary Table 4: Inversion effects on life history traits in a sympatric, reciprocal WEST-*inv* x WEST-*std* F2 cross in the greenhouse [a, b] $N = 1{,}000$.

Trait	Inversion *F*	Cross *F*	Inversion*Cross *F*	Inversion *P*	Inversion*Cross *P*
MULTIVARIATE-TRAITS MANOVA	2.095	1.050	0.779	**0.005**	0.743
Survival after vernalization [c]	0.554	0.656	0.616	0.594	0.536
Flowering or not [c]	0.276	6.867	0.160	0.728	0.855
Rosette width at 4 week	1.701	5.918	0.240	0.173	0.778
Rosette width at 10 week	0.873	6.405	0.164	0.458	0.857
Days to flowering	3.880	0.059	0.265	**0.030**	0.774
Width when flowering	0.077	0.328	0.745	0.926	0.487
Height when flowering	0.543	0.185	0.119	0.562	0.892
Rosette number when flowering	0.124	0.119	2.792	0.902	0.063
Leaf number when flowering	3.726	1.218	0.122	**0.035**	0.881
Fruit number of flowered plants	2.148	4.375	0.996	0.116	0.367
Lifetime fitness (fruit number) [d]	2.172	11.046	0.161	0.104	0.854

a. *P* values were obtained by comparing the true statistic to those from 1,000 permutations. Within each permutation the inversion genotype was randomized within each cross. MANOVA R^2 attributable to the inversion ≈ 4%.

b. Inversion degrees of freedom = 2, Cross degrees of freedom = 1, and Inversion*Cross degrees of freedom = 2

c. 'Survival after vernalization' and 'flowering or not' are binary traits

d. Fruit number of dead or non-flowering plants was set to zero

Supplementary Table 5: Quantitative PCR tests for differential expression between *Bsi1-inv* and *Bsi1-std* homozygotes for seven genes flanking inversion breakpoints and syntenic to *A. thaliana*. *N* = 11 F2 plants homozygous for the inversion. *N* = 12 homozygous for standard haplotypes. Significance was determined by ANOVA with 1000 permutations.

***B. stricta* gene**	***A. thaliana* ortholog**	**$2^{\Delta Ct}$ *F* value (*P* value)**	**ΔCt *F* value (*P* value)**
Bostr.26675s0299	AT1G04500	0.72 (0.412)	3.23 (0.092)
Bostr.26675s0297	AT1G04510	0.61 (0.420)	1.35 (0.213)
Bostr.25219s0523	AT1G04520	0.15 (0.763)	0.08 (0.791)
Bostr.25219s0522	AT1G04530	1.40 (0.279)	2.48 (0.157)
Bostr.25219s0521	AT1G04540	1.83 (0.223)	0.94 (0.381)
Bostr.26675s0296	AT1G48700	2.81 (0.112)	1.65 (0.205)
Bostr.25219s0524	AT1G48740	0.06 (0.817)	0.31 (0.606)

Power analyses were performed with JMP and with simulations in R. For all seven genes, we found that both alleles are expressed, hence major expression differences do not occur for the genes near the breakpoints. Across these genes, our ability to detect a true gene expression difference corresponding to the observed levels had a median power of 19.8%, so our ability to detect individual quantitative expression differences is low. Nevertheless, we can calculate the probability of the observed multi-locus result (zero significant results over seven tests) given the assumption that true expression differences correspond to the observed levels and that these loci are independent. Under these circumstances, power to detect at least one significant result equals 82% without Bonferroni correction for multiple tests, or 40% using a Bonferroni-adjusted significance threshold. In summary, this experiment finds little evidence for moderate to large differences in expression, but subtle, quantitative differences in expression may exist.

Supplementary Table 6. The 16 *Arabidopsis thaliana* BAC clones and positions (start and end) of their homologues on *Boechera stricta* reference genome.

BAC	Length (bp)	Chromo-some	Scaffold	Position on scaffold		Physical position*		Genetic position*	
				Start (bp)	End (bp)	Start (bp)	End (bp)	Start (cM)	End (cM)
F2J6	108061	1	Scaffold12302	308493	945202	945202	12620048	16.215	16.657
F13N6	90698	1	Scaffold8169	269954	611199	15450805	15792050	17.611	17.927
F12K21	117585	2	Scaffold1040	155707	332639	11542806	11719738	48.203	48.203
T23K8	53090	2	Scaffold26959	3955495	4566017	14878627	15489149	48.836	48.519
T25N22	80891	3	Scaffold24399	292205	310014	16506702	16524511	102.54	102.54
MGO3	43570	3	Scaffold26833	1148763	4517264	23843108	27211609	135.649	163.285
T19L18	87080	4	Scaffold26326	531811	810775	5237879	5516843	25.19	26.204
F7F1	89779	4	Scaffold24513	504718	881714	11991703	12368699	50.769	53.988
MGC1	2595	5	Scaffold14419	314149	316758	115224	117833	0.57	0.57
MBD2	79976	5	Scaffold3148	1370932	1780895	4914471	5324434	63.775	65.554
T10F5	102585	5	Scaffold13083	340643	432698	8388774	8480829	67.14	67.14
T4C9	85356	6	Scaffold25463	548445	790858	3981483	4223896	17.171	17.171
T3H13	118737	6	Scaffold9686	59367	119129	7469571	7529333	27.554	27.554
T8M17	114695	6	Scaffold29827	211566	226057	10379396	10393887	29.275	29.275
F5H8	70098	7	Scaffold9638	29598	95279	6319745	6385426	21.732	21.732
T1P17	137519	7	Scaffold2021	8370	19324	8304581	8315535	24.586	24.586

*Position on linkage map

Supplementary Table 7. List of phenotyped traits for the collinear East-std x West-std (Parker x Ruby) QTL mapping experiment

Trait Name	Trait details
GH_ln_RosNum	Rosette number
GH_ln_RosW	Rosette width
GH_ln_RosH	Rosette height
GH_ln_RosVol	Rosette volume
GH_ln_Fresh_Weight	Rosette fresh weight
GH_ln_Dry_Weight	Rosette dry weight
GH_ln_Water_Weight	Rosette water weight = fresh - dry weight
GH_ln_Water_Proportion	Rosette water weight / fresh weight
GH_ln_Leaf_Area	Rosette leaf area by scanner
GH_ln_Leaf_Packing	Rosette leaf area / Rosette volume
GH_ln_Fresh_weight.Area	Rosette fresh weight / leaf area
GH_ln_Dry_weight.Area	Rosette dry weight / leaf area
GH_ln_Water_weight.Area	Rosette water weight / leaf area
GH_ln_Water_proportion.Area	Rosette water proportion / leaf area
GH_ln_Leaf_mean_L	Length of largest leaf
GH_ln_Leaf_mean_W	Width of largest leaf
GH_ln_Leaf_mean_W.L	Leaf width / length
GH_ln_Flower_Julian_day	Flowering time in Julian day
GH_ln_Flower_W	Width when flowering
GH_ln_Flower_H	Height when flowering
GH_ln_Flower_LN	Leaf number when flowering
GH_ln_Flower_Ros	Rosette number when flowering
GH_ln_Flower_Stalk	Stalk number when flowering
GH_ln_fin_W	Final width number when experiment ends
GH_ln_fin_H	Final height
GH_ln_fin_stalk.width	Final main stalk width
GH_ln_fin_FLN	Final remaining flower number
GH_ln_fin_FRN	Final fruit number
GH_ln_fin_longest.silique	Longest silique length
GH_ln_fin_FRNxlongest.silique	Fruie number multiplied by longest silique length
GH_ln_fin_branch_on_mainstalk	Number of reproductive branches on main stalk
GH_ln_fin_length_with_branch	Stalk height in the region with reproductive branches
GH_ln_fin_branch_internode	Internode of the reproductive part of main branch
GH_fin_tip_bud_active	Whether the top youngest flowering bud is still green at the end of expt
GH_fin_have_mature_fruit	Whether it has a dried-up and matured fruit in the end of experiment
GH_fin_central_tip_stop_develop	Whether the top meristem stops develop at the end of experiment

Supplementary Table 8. Flowering time genes in inversion region of Chromosome 1 of *Boechera stricta* . The three genes located in QTL regions (peak +/- 100k bp) are in bold.

Gene_name	Scaffold	Start	End	Arabidopsis thaliana orthology	Putative_function
Bostr.26675s0108	**Scaffold26675**	**659,047**	**661967**	**AT1G50700**	**calcium-dependent protein kinase 33**
Bostr.26675s0088	Scaffold26675	545,606	547858	AT1G50960	gibberellin 2-oxidase 7
Bostr.26675s0074	Scaffold26675	481,495	483840	AT1G51140	basic helix-loop-helix (bHLH) DNA-binding superfamily protein
Bostr.26675s0051	Scaffold26675	334,778	337078	AT1G51450	TRAUCO
Bostr.0124s0064	Scaffold124	350,941	352379	AT1G52740	histone H2A protein 9
Bostr.0124s0099	Scaffold124	558,489	562796	AT1G53090	SPA1-related 4
Bostr.0124s0106	Scaffold124	599,305	600288	AT1G53160	squamosa promoter binding protein-like 4
Bostr.13671s0011	Scaffold13671	56,078	57768	AT1G14400	ubiquitin carrier protein 1
Bostr.13671s0159	**Scaffold13671**	**884,222**	**885478**	**AT1G12910**	**Transducin/WD40 repeat-like superfamily protein**
Bostr.13671s0424	Scaffold13671	2,264,288	2268324	AT1G10570	Cysteine proteinases superfamily protein
Bostr.13671s0476	Scaffold13671	2,522,138	2525008	AT1G10120	basic helix-loop-helix (bHLH) DNA-binding superfamily protein
Bostr.25494s0015	Scaffold25494	68,807	74496	AT1G09570	phytochrome A
Bostr.25219s0051	**Scaffold25219**	**224,347**	**226129**	**AT1G08970**	**nuclear factor Y, subunit C9**
Bostr.25219s0119	Scaffold25219	559,042	575787	AT1G08260	DNA polymerase epsilon catalytic subunit
Bostr.25219s0356	Scaffold25219	1,618,878	1620440	AT1G06040	B-box zinc finger family protein
Bostr.25219s0382	Scaffold25219	1,730,774	1738687	AT1G05830	trithorax-like protein 2
Bostr.25219s0486	Scaffold25219	2,228,447	2230930	AT1G04870	protein arginine methyltransferase 10