File S1

Supporting Methods

F2 population generation and sequencing: An F2 population from an initial cross between plants S. aethnensis Pro2-21 (collected from Piano Provenzana, Mt. Etna, altitude 2036 m.; male parent) and S. chrysanthemifolius Ran1-5 (collected near Randazzo, Mt. Etna, altitude 763 m. see table 1 in (Chapman *et al.* 2013); female parent) was generated and grown in a growth room with the same temperature and light/dark cycle as the plants analysed in the initial transcriptomic investigation (Chapman *et al.* 2013). The parent plants were morphologically typical of the species and showed no evidence for introgression in a previous analysis (Chapman *et al.* 2013). Some admixture, however, cannot be ruled out, especially for the S. aethnensis parent, which was collected from a population close to the high altitude edge of the hybrid zone.

In total 147 F2 plants were grown, 64 of which were selected for RNA-seq analysis. F2 hybrid breakdown was observed in this cross, with 69 plants developing signs of necrosis or stunted growth, and not reaching maturity (collectively termed 'necrotic'; Figure 1 A-D) and 78 healthy plants that grew to maturity (Figure 1 E-F) in the population. The 64 sequenced plants comprised 48 healthy and 16 necrotic F2s and apart from this designation were selected randomly.

RNA was extracted from leaf tissue of 3 week old seedlings using an RNeasy kit (Qiagen, Manchester, UK) with additional DNA removal using DNase (Qiagen), and from each an individual RNA-seq library was constructed. Sampling from seedlings reduces the variability associated with developmental and morphological differences that are more pronounced as the plants mature. These 64 libraries were then run (as a multiplex) across two lanes of a HiSeq2000 at the Wellcome Trust Centre for Human Genetics.

Expression Analysis: Comparisons were then made between the 48 healthy and 16 necrotic plants using t-tests, with significant differences being detected after applying an FDR of 5%. To determine if

certain categories of Gene Ontology (GO) were over-represented in the differentially expressed loci we carried out Fisher's Exact Tests in Blast2GO (Conesa *et al.* 2005; Gotz *et al.* 2008), applying a 5% FDR.

Linkage Mapping: SNPs from the loci that fulfilled the above criteria were identified using Proseq and imported into Microsoft Excel where consensus genotypes (per locus per individual) were produced. Up to 20% inconsistent genotyping within an individual was considered tolerable as this was likely a result of mis-mapping of a read by CLC-GW, or mis-scoring of a SNP by *SAM*tools. If more than 20% of SNPs showed an erroneous pattern then this locus was scored as missing data for this individual. Because a major aim of this portion of the analysis was to map highly differentiated outlier loci, we preferentially added to this list several outlier loci that did not fulfil criterion (2) above, i.e. less than five (but more than two) SNPs differentiated the parental alleles. For these loci, inconsistent SNP genotypes within a locus were not tolerated.

Manual genotype verification: Several markers on linkage group (LG) 4 were never homozygous in the F2 for alleles inherited from the S. chrysanthemifolius parent and so we sequenced twelve loci in or close to this region from the two parents and the two F1 plants to ensure that they were homozygous (for alternate alleles), and heterozygous, respectively. PCR amplification and sequencing followed the protocol outlined in (Chapman *et al.* 2013).

For six of the above loci and ten other loci exhibiting segregation distortion length differences caused by indels were identified between the parental alleles (either directly from the F2 transcriptome alignments, or from sequencing intron-containing regions of the loci from the parental individuals). For genotyping, loci were PCR amplified using a fluorescent '3rd primer' using protocols established for microsatellites (Chapman *et al.* 2008) and resolved on an ABI 3730 in the Department of Zoology, University of Oxford. Primer sequences are given in Table S2. Alleles were scored with GeneMarker (ver. 2.4.0; www.softgenetcs.com).
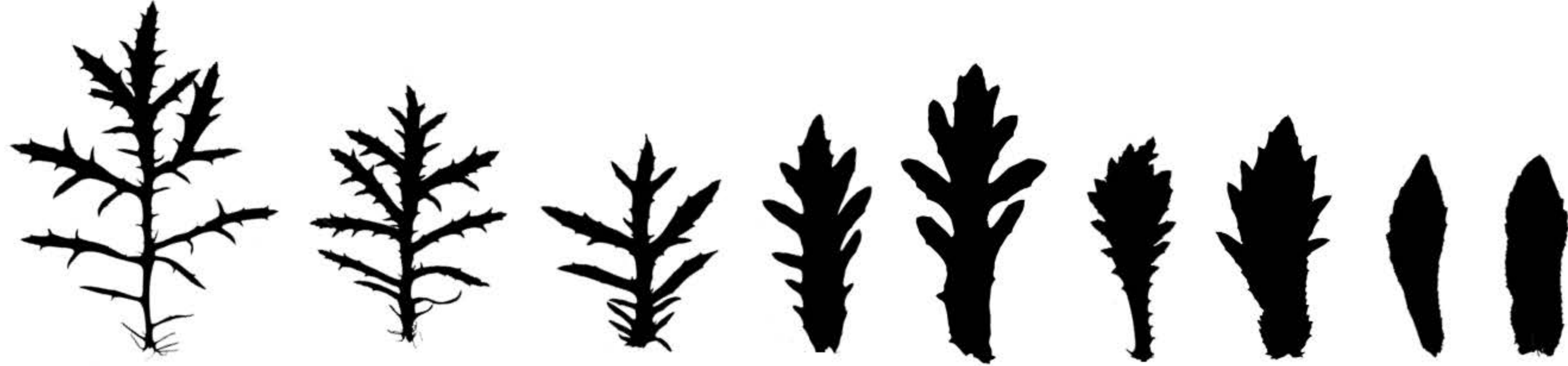
M. A. Chapman *et al.*

Analysis of Polymorphism and Divergence: Values were compared between LGs and between

species, where appropriate, using Mann-Whitney U-tests, and the distribution of outlier markers per

bin (5 and 10 cM bins) was compared to that expected based on marker density alone using a

Fisher's Exact Test, in PAST (http://folk.uio.no/ohammer/past/). Analysing the data based on bins of

a certain number of loci was precluded due to inability to resolve map positions uniquely (i.e.
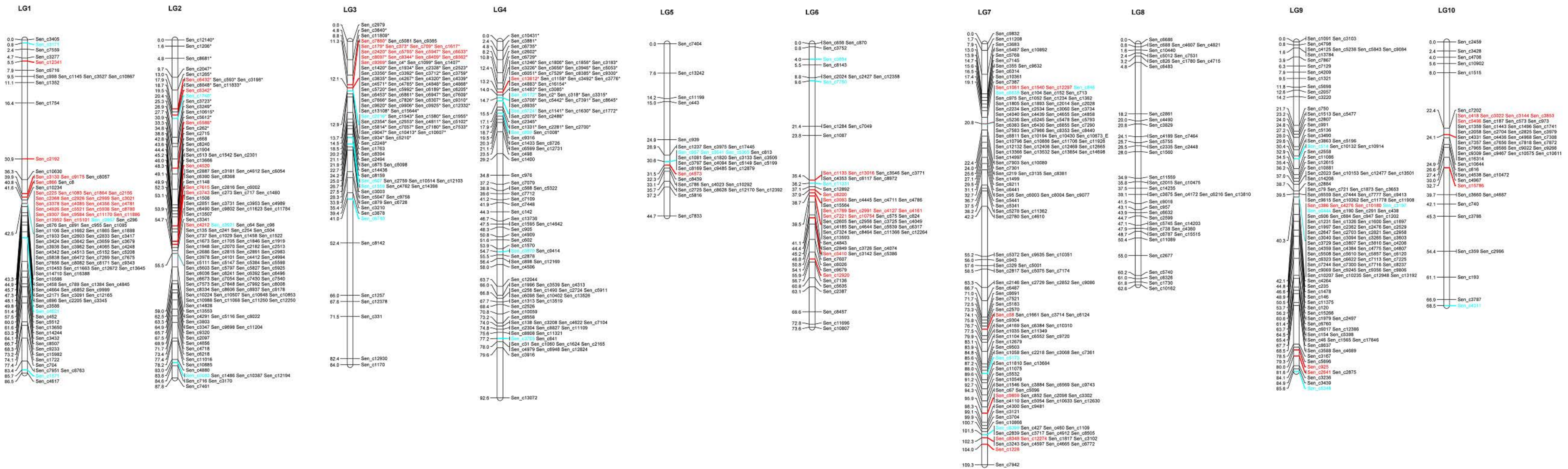
clustering of many loci).

Chapman, M. A., S. J. Hiscock and D. A. Filatov, 2013 Genomic divergence during speciation driven by adaptation to altitude. Molecular Biology and Evolution **30:** 2553-2567.
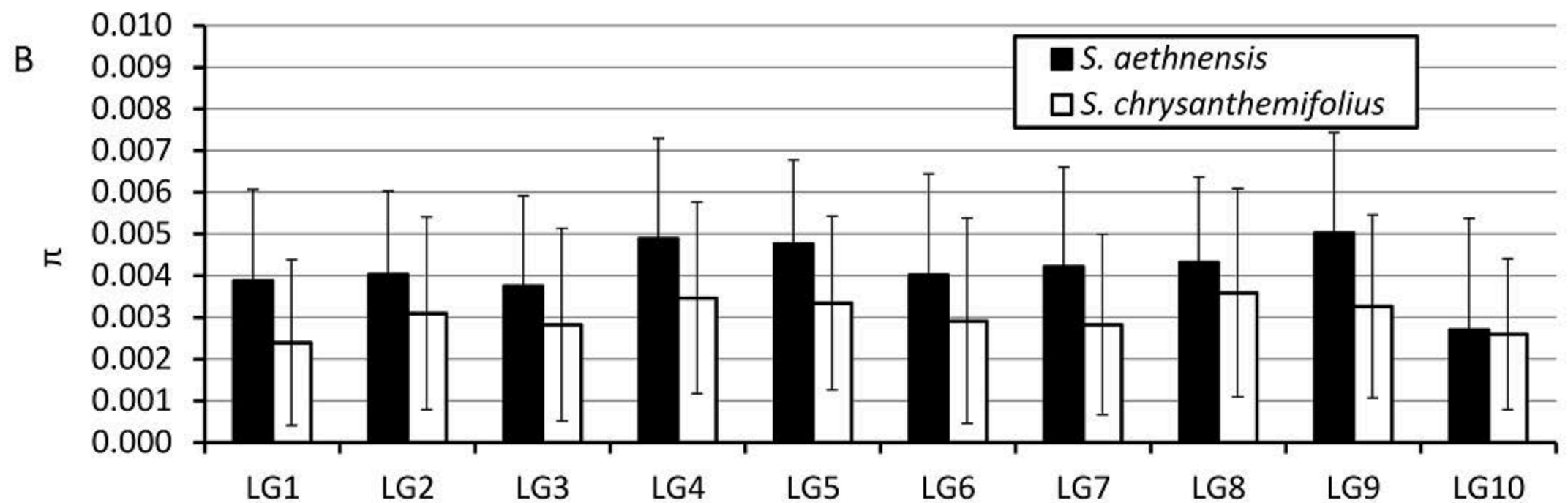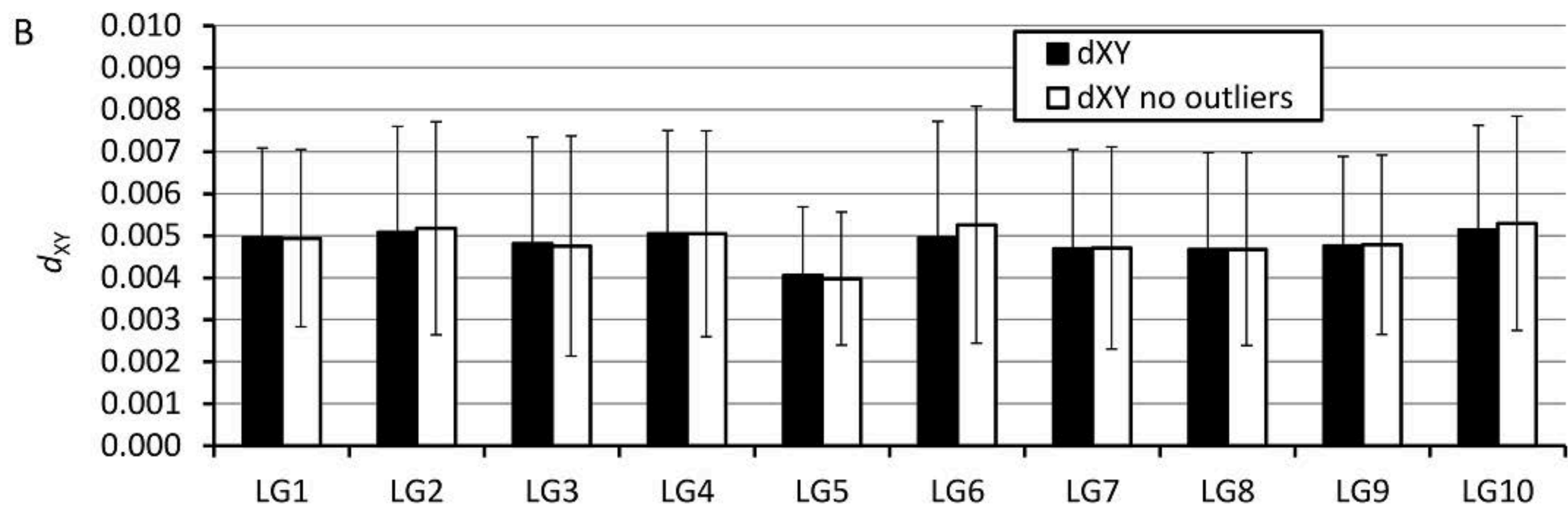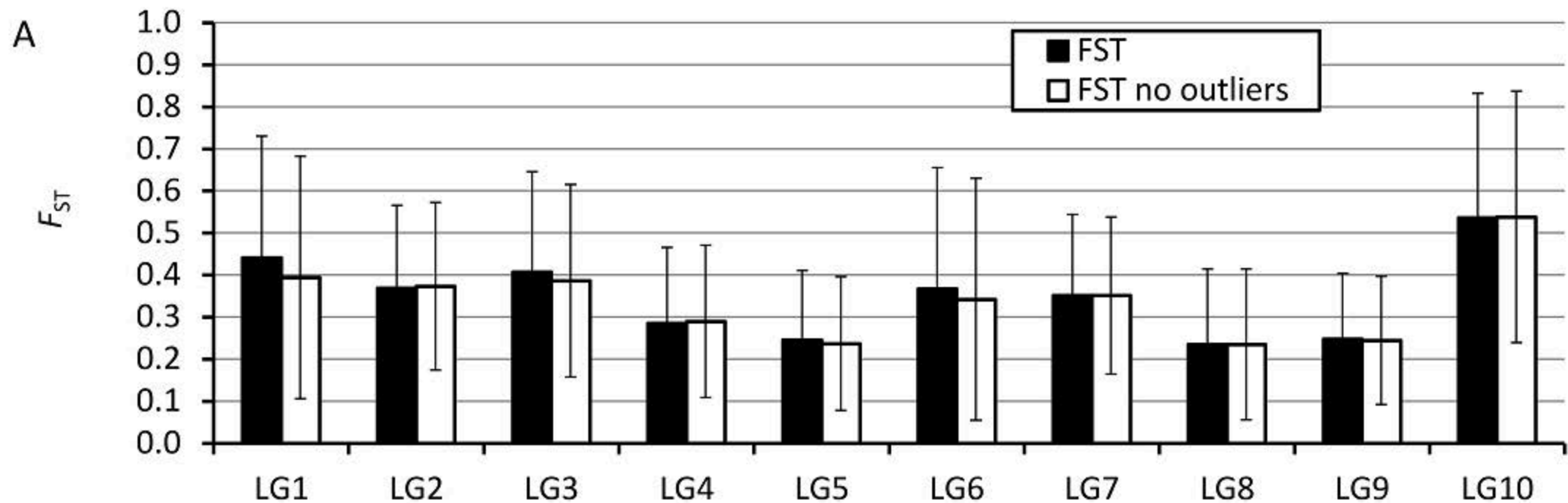
Chapman, M. A., C. H. Pashley, J. Wenzler, J. Hvala, S. Tang *et al.*, 2008 A genomic scan for selection reveals candidates for genes involved in the evolution of cultivated sunflower (*Helianthus annuus*). The Plant Cell **20:** 2931-2945.

Conesa, A., S. Gotz, J. M. Garcia-Gomez, J. Terol, M. Talon *et al.*, 2005 Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics **21:** 3674-3676.

Gotz, S., J. M. Garcia-Gomez, J. Terol, T. D. Williams, S. H. Nagaraj *et al.*, 2008 High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids Research **36:** 3420-3435.

Genetic linkage map showing linkage groups LG1 through LG10 with marker positions and names.
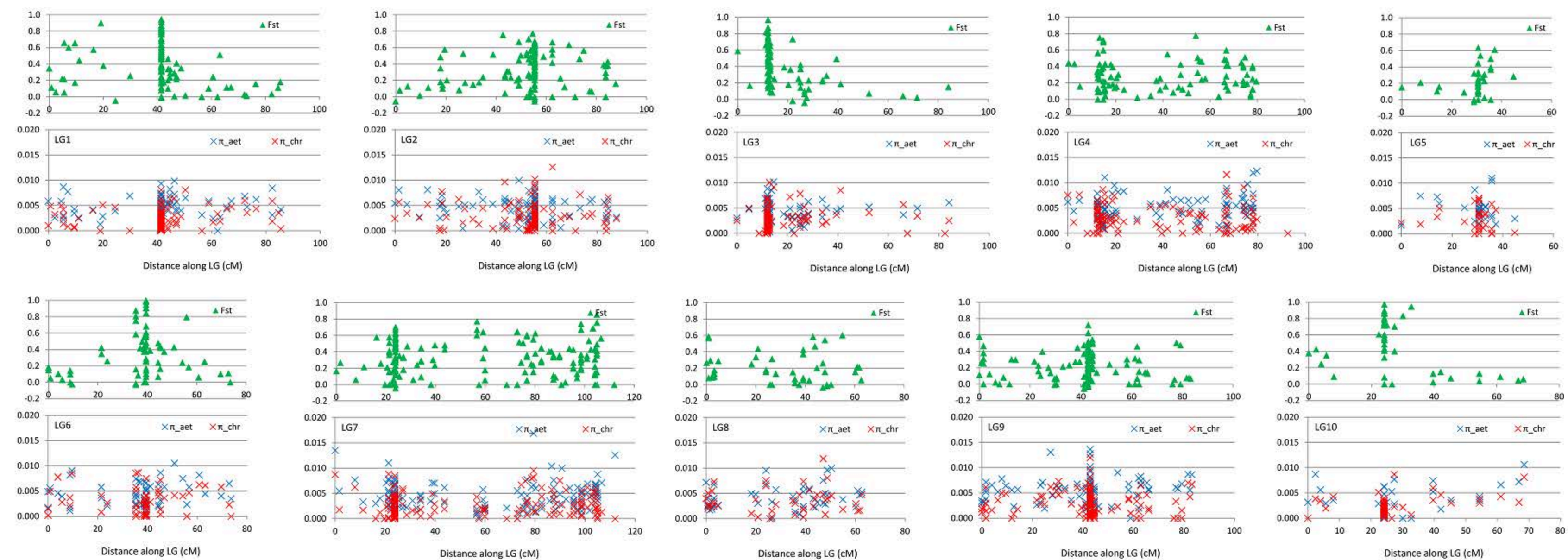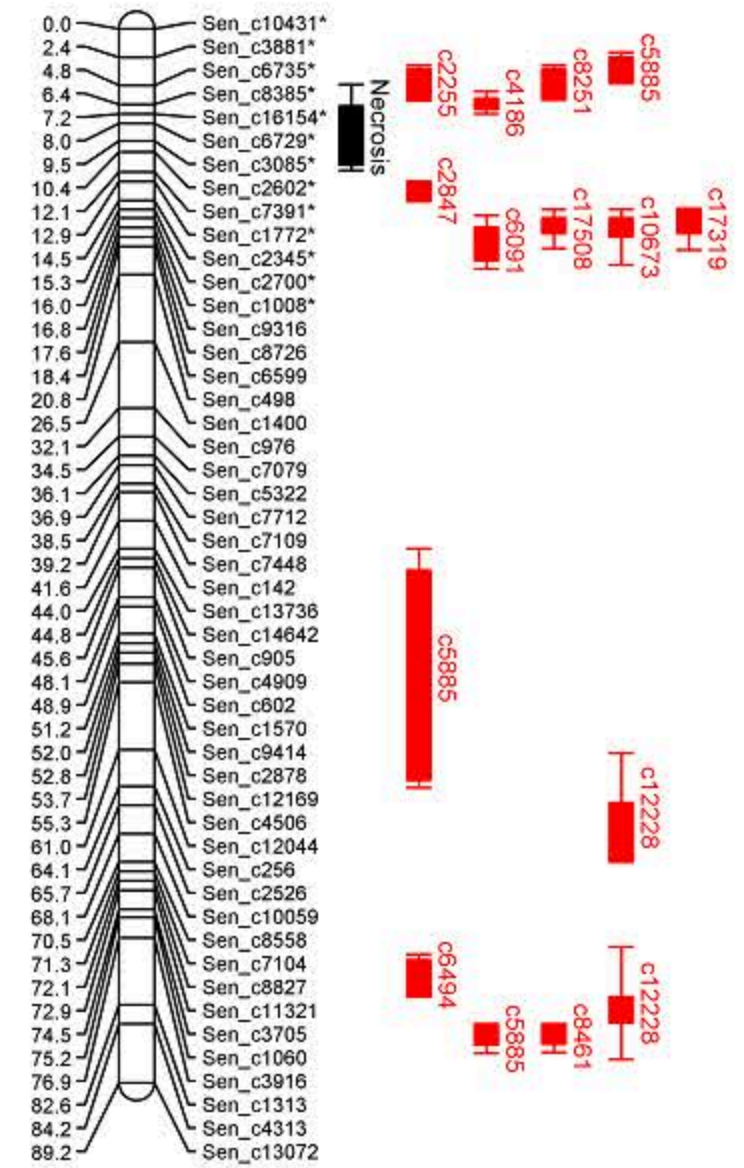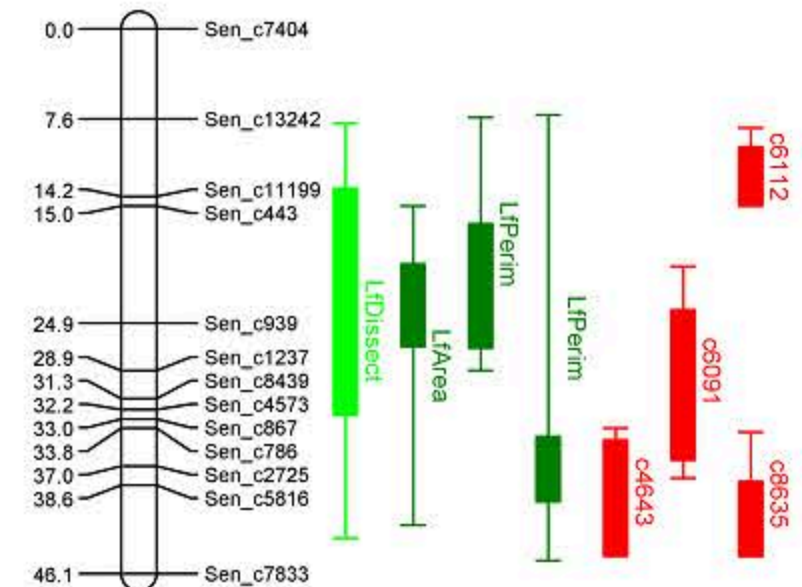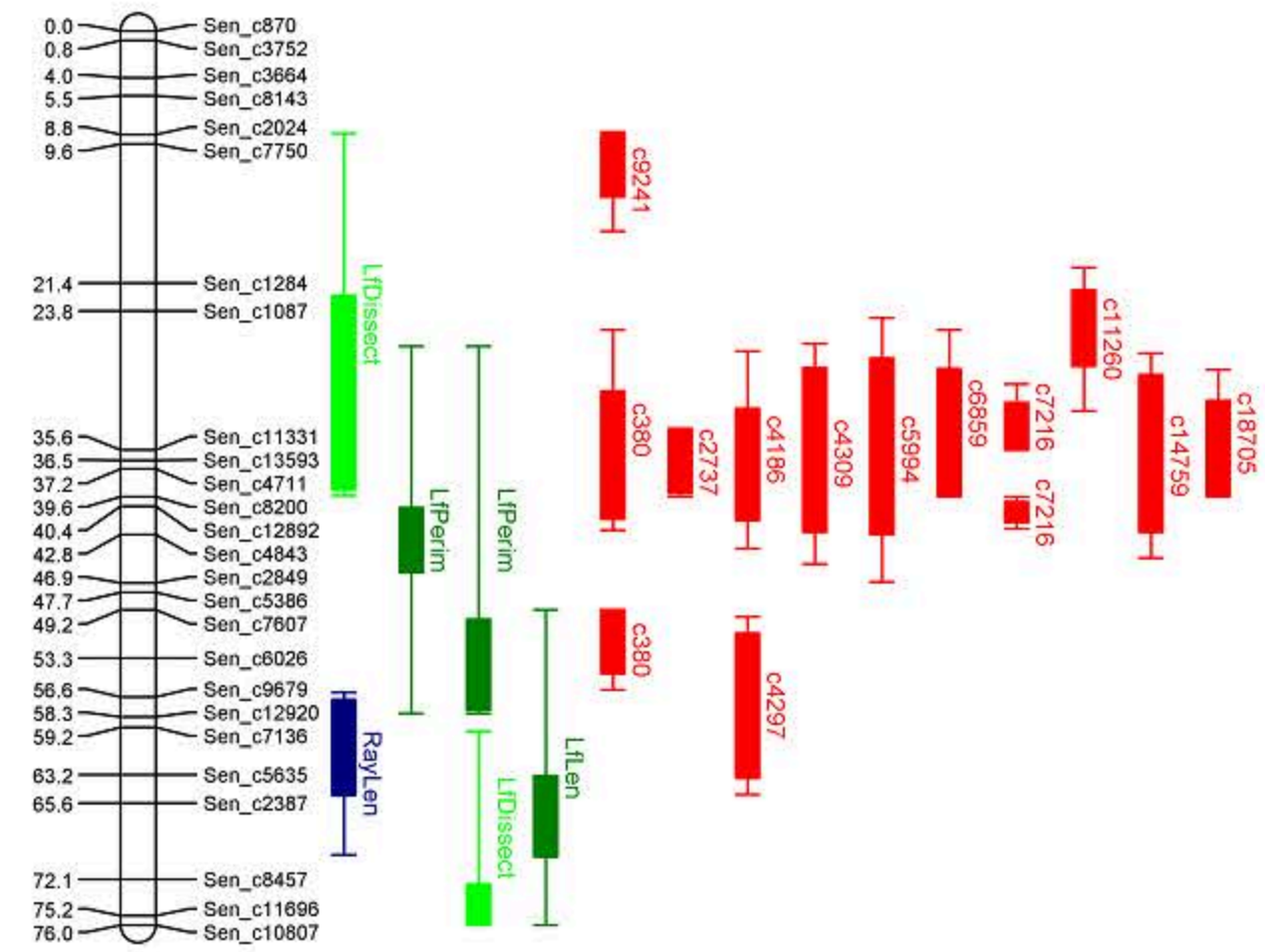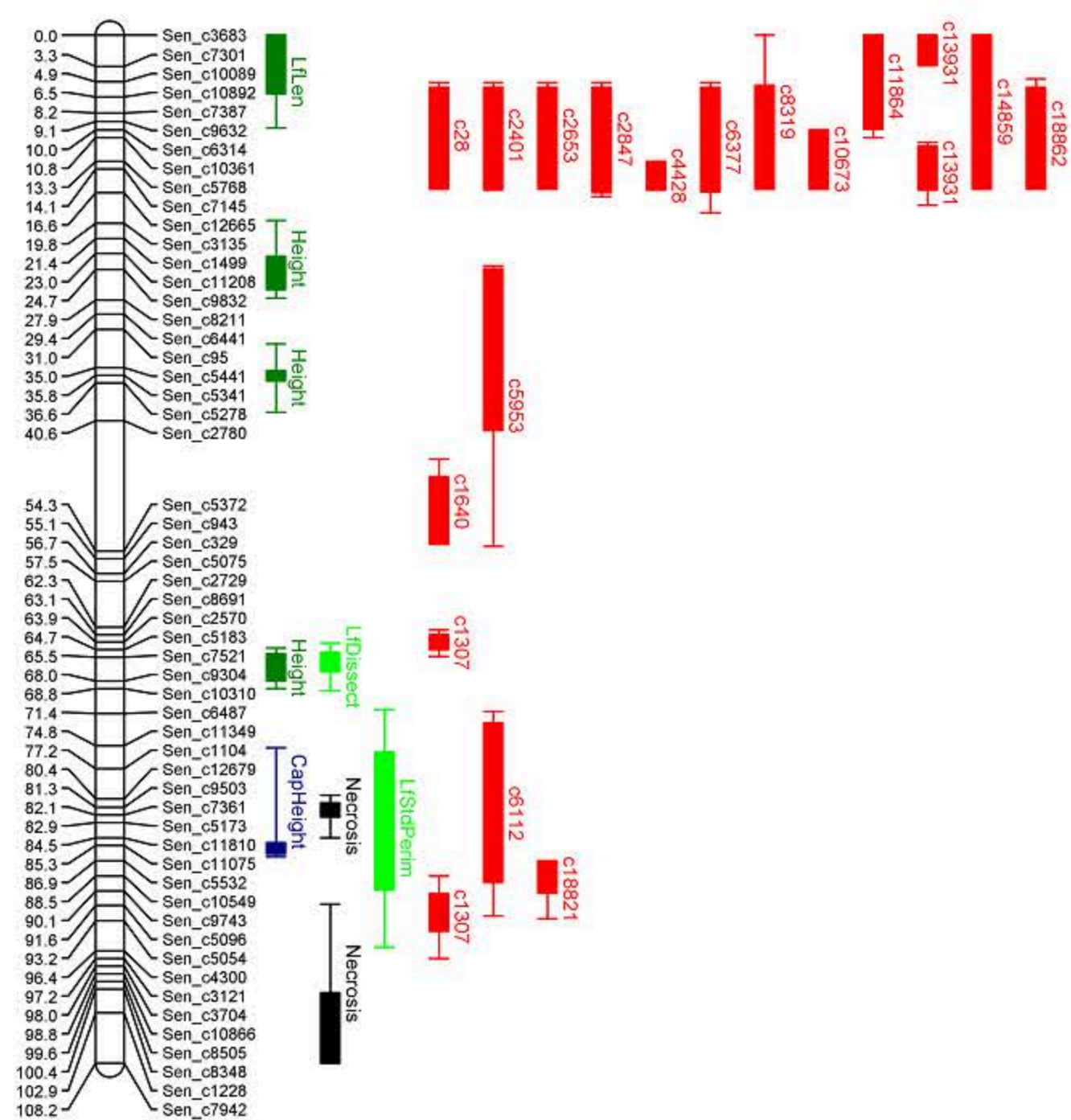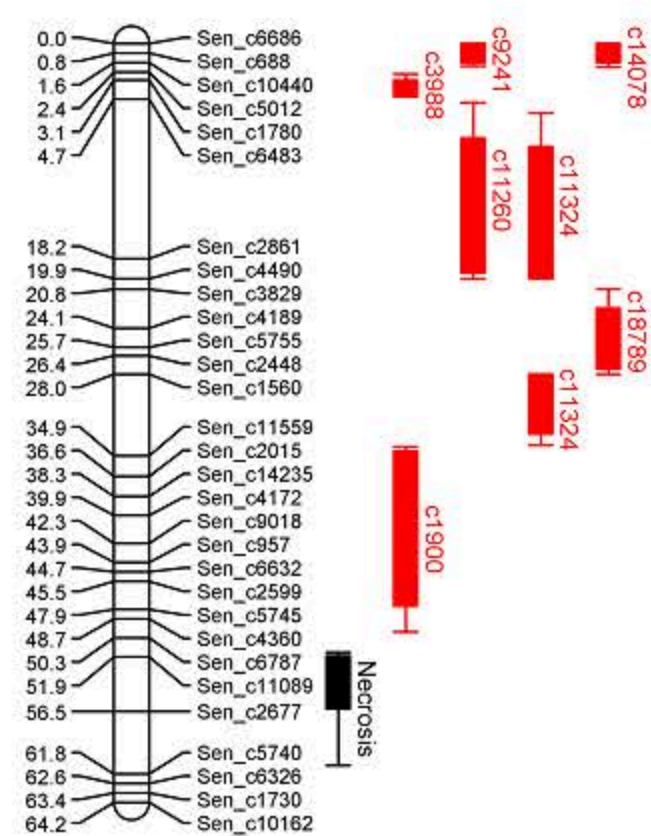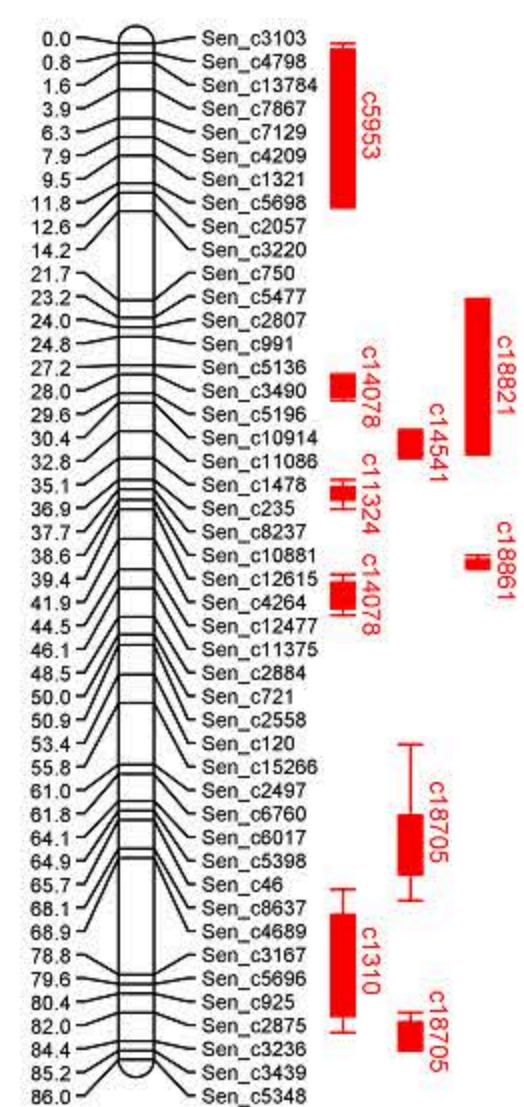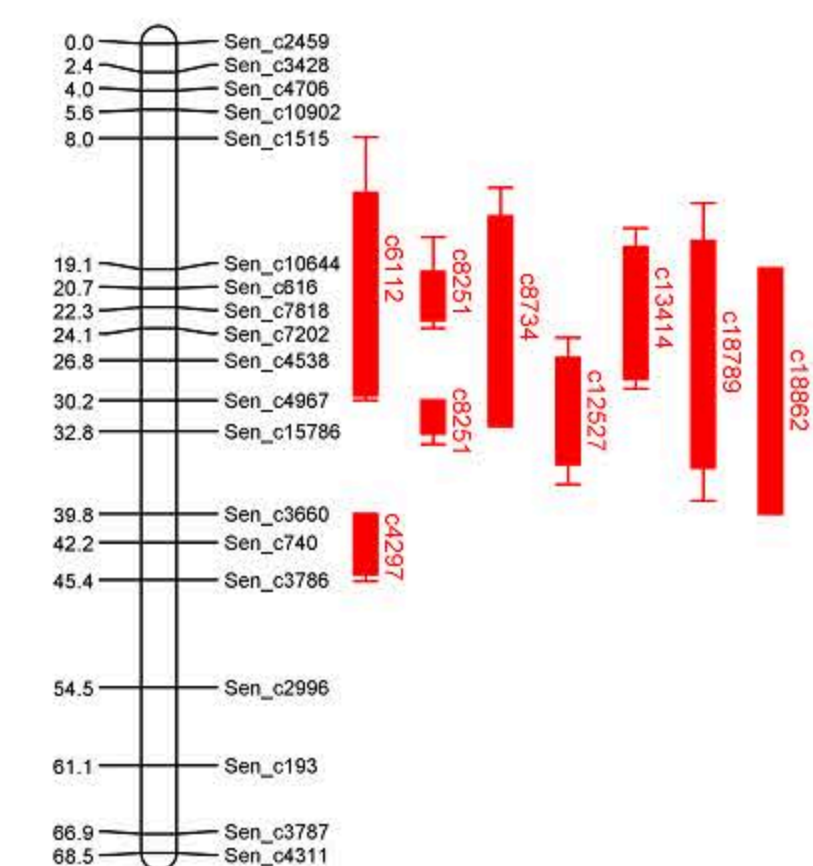
Table S1 – Numbers of sequence reads produced per individual.

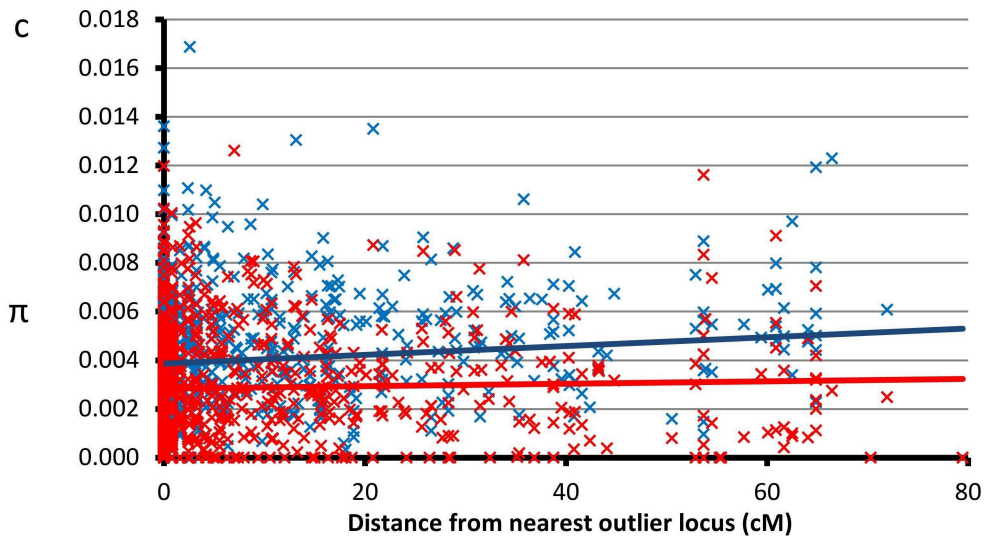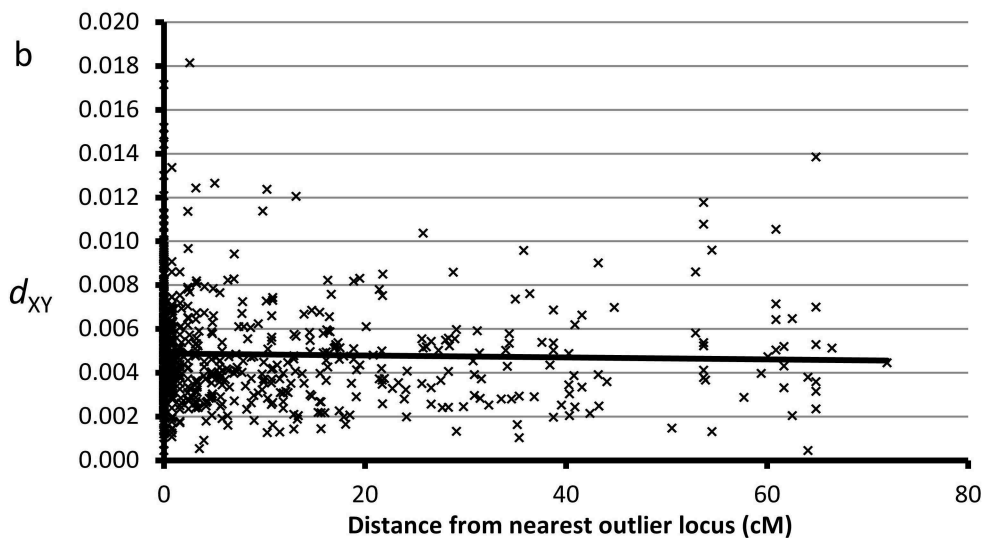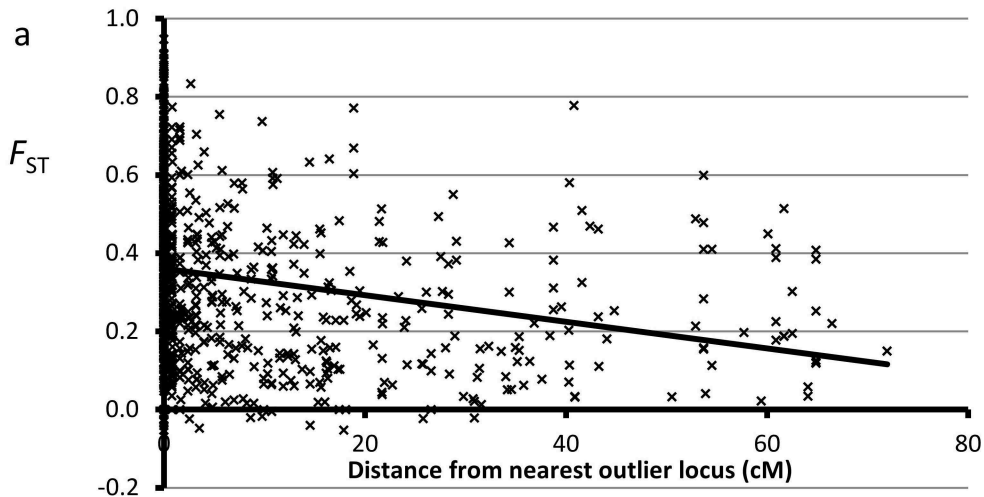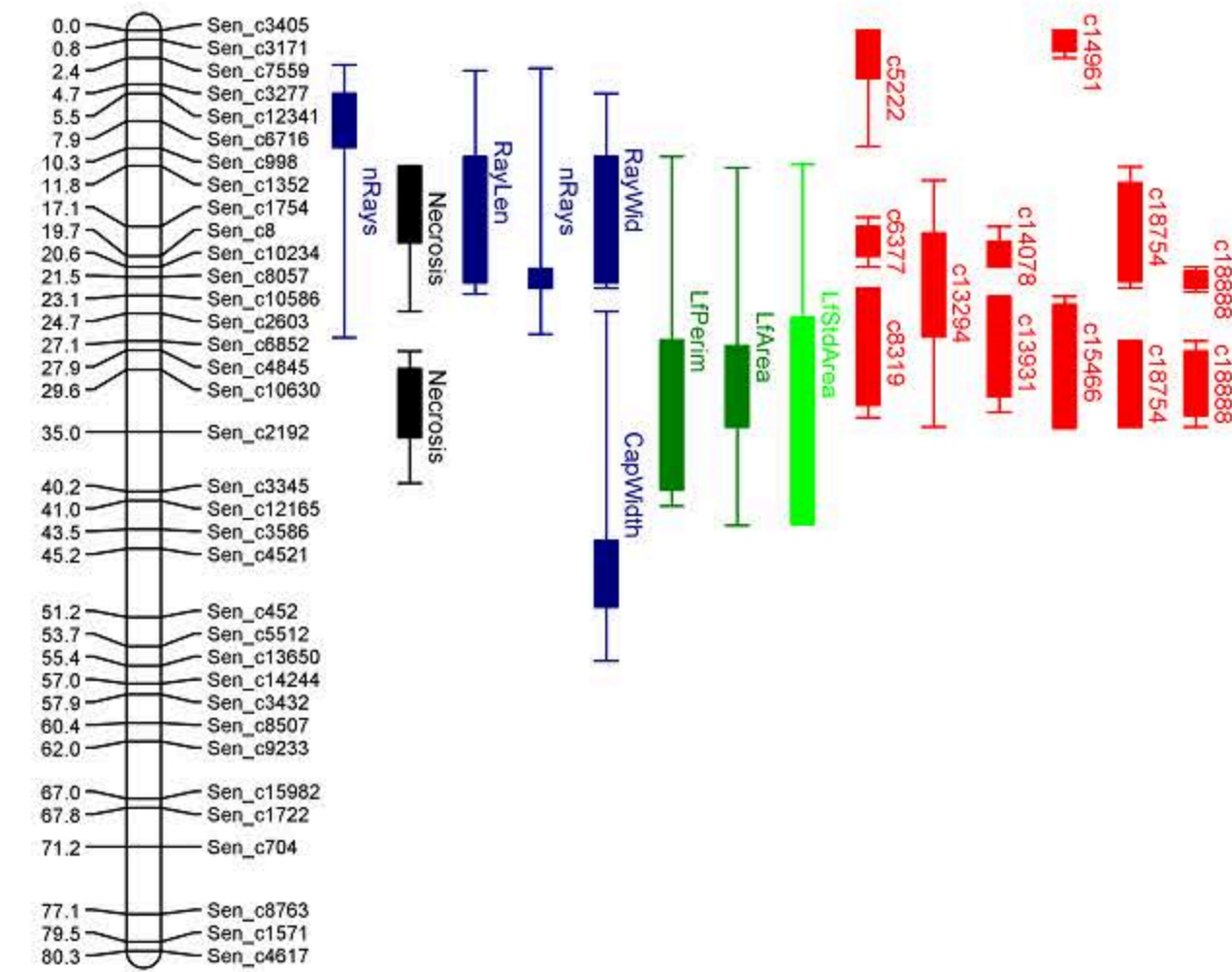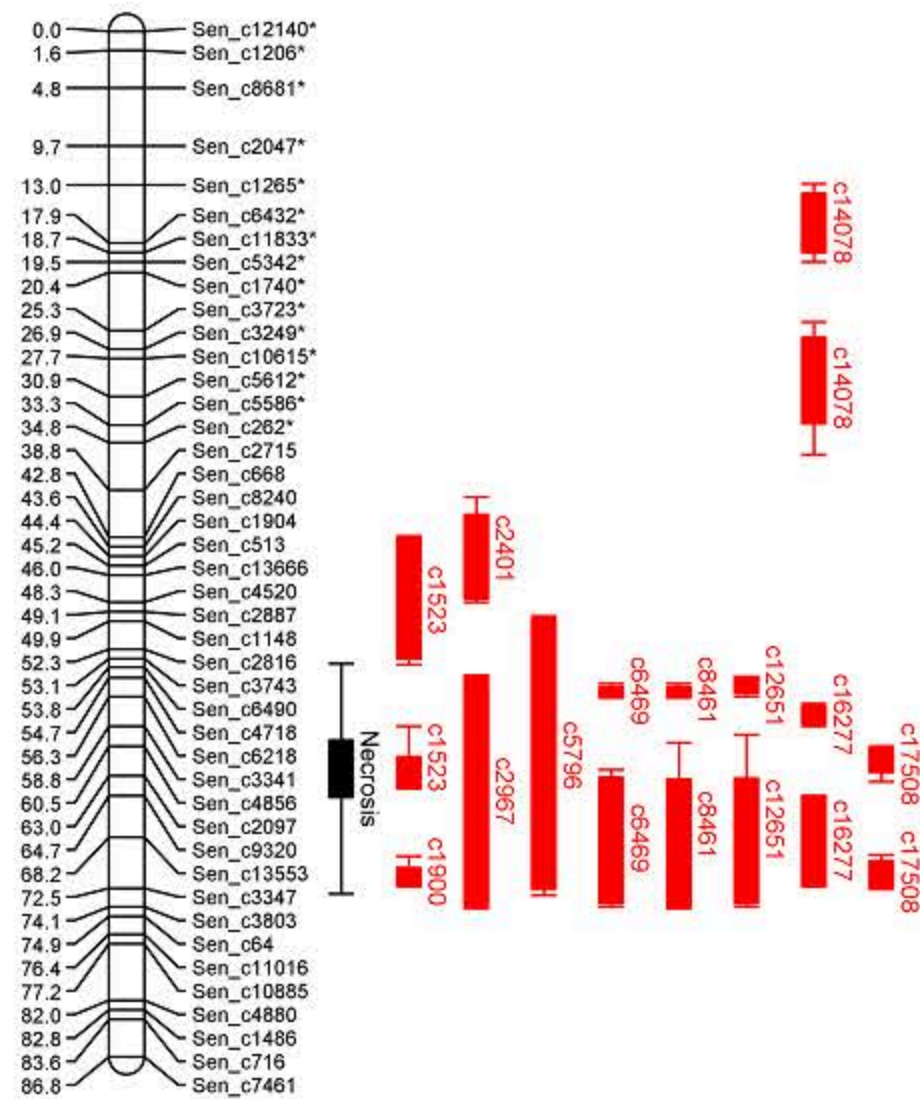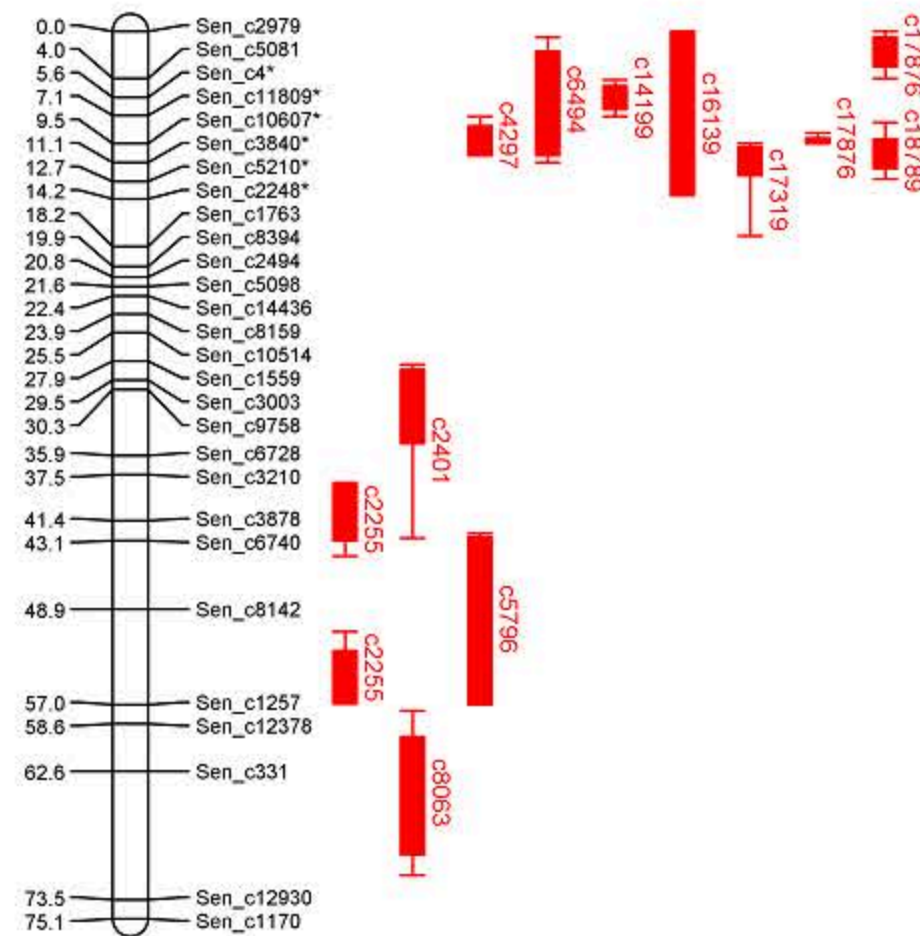| Species	Individual	N reads
S. chrysanthemifolius	R1-5	12591356
S. aethnensis	P2-21	25237815
F2	F2_2	5946413
F2	F2_6	5858567
F2	F2_7	5649616
F2	F2_8	5410637
F2	F2_9	5953172
F2	F2_12	6549678
F2	F2_14	5993052
F2	F2_15	6149542
F2	F2_16	5715691
F2	F2_17	5779442
F2	F2_22	5492286
F2	F2_24	6129351
F2	F2_25	6231166
F2	F2_27	6137433
F2	F2_29	5671652
F2	F2_30	5411550
F2	F2_32	5859679
F2	F2_35	5680176
F2	F2_42	5650940
F2	F2_44	5750924
F2	F2_45	5639667
F2	F2_47	6621453
F2	F2_49	6342642
F2	F2_52	5077913
F2	F2_53	6065291
F2	F2_54	5534249
F2	F2_64	5809735
F2	F2_65	5710319
F2	F2_66	5822554
F2	F2_69	6315459
F2	F2_71	5960569

Species	Individual	N reads
F2	F2_76	5944820
F2	F2_77	5796127
F2	F2_78	6150781
F2	F2_80	6010262
F2	F2_81	5676708
F2	F2_82	6486196
F2	F2_83	6405013
F2	F2_85	5264016
F2	F2_86	5636553
F2	F2_89	6318809
F2	F2_90	5959283
F2	F2_91	5453964
F2	F2_93	6068437
F2	F2_100	6176081
F2	F2_104	6627566
F2	F2_108	5124406
F2	F2_111	5765542
F2	F2_115	6041251
F2	F2_116	5753443
F2	F2_119	5445717
F2	F2_120	5512703
F2	F2_122	5779572
F2	F2_124	6381769
F2	F2_126	6413537
F2	F2_131	5892944
F2	F2_134	6291519
F2	F2_136	6476101
F2	F2_137	5346393
F2	F2_138	6328657
F2	F2_142	6111139
F2	F2_143	7008111
F2	F2_145	5585352
F2	F2_147	6021070

Table S2 – Primers used in the targeted genotyping and sequencing.

A. PCR primers used to amplify and sequence loci in and near the region of extreme transmission ratio distortion on LG4

Locus*	F primer	R primer	
4_c10431	GGCTTCCTCGAGAACTTGC	CCAATGTCTGACCCGAACTT	
4_c6735	CCGGTAGTGGGAAGTTGAAA	TCCATTCAGATGGGGATGAT	length difference uncovered, so was genotyped on a subset of the F2 (see below)
4_c1856†	GGAACCTGTGGCTCCAGTAG	GGGTTCCCATAAGTTGATCG	length difference uncovered, so was genotyped on a subset of the F2 (see below)
4_c6172†	AAAGCAAGCAAGTGCCATCT	GGTTCTGGCTTGAAGCTTTG	length difference uncovered, so was genotyped on a subset of the F2 (see below)
4_c2	ACAACACCAGTTCCATGCAC	TGTCACCATCAAGTTGTTCCA	
4_c2075§	TGGTTGATTTTACRATGCACA	CCAAACTGCAGCTGTTCGTA	
4_c2700§	CAATTGGGGCAGCTTGTAAT	AACCTCCAWCGTGCTTGAAC	
4_c1331	TGGTCATATGGAAGCGGTTT	TGAAGCCAGGGAGAAGGAAT	
4_c809	ATTTTTACACCGCCGTTGAC	GGAAGCCATTTCTGACCAAA	length difference uncovered, so was genotyped on a subset of the F2 (see below)
4_c8726	ACGCTTACCAAGCTCCATTG	ACCGGGGAAGAAATTGAAAC	length difference uncovered, so was genotyped on a subset of the F2 (see below)
4_c6599	TCAATTGTGTTAGTTGGAGTTGG	CGTGCGCTCTAYATTAGCTG	
4_c12731	ATAACCCGAGGCACAATGAG	AATGGCATTTAACGCTCTGG	length difference uncovered, so was genotyped on a subset of the F2 (see below)

† locus was not homozygous in the parents but was heterozygous in the F1
§ locus was not heterozygous in the F1

B. PCR primers used to amplify and genotype loci showing segregation distortion

Locus*	F primer**	R primer	
2_c5342	CACGACGTTGTAAAACGACATCGACTCGKAATCCCTCTT	CGCTCCATCGTACAACAAAA	length difference from RNAseq data
2_c6432	CACGACGTTGTAAAACGACGGAGATGGAGCTCTGGAATG	TCCACCTCTTTGCTCCTTTG	length difference from RNAseq data
3_c4848	CACGACGTTGTAAAACGACCAAACACGTCACACCTCACC	CCTAACCAGCTCATGGGTTT	length difference from RNAseq data
3_c5992	CACGACGTTGTAAAACGACGCAGATGCACCTAAAGCTACC	TGCAATACTAATACGCAGTTCCA	length difference from RNAseq data
3_c6205	CACGACGTTGTAAAACGACGATGCCAGCTCTCTTGAAGG	ACCACAAACCCCTCAAAACC	length difference from RNAseq data
3_c6633	CACGACGTTGTAAAACGACGGATCTTCGAATTGATTTGATT	CGGATTCATAGGCGGTAAWA	length difference from RNAseq data
3_c7609	CACGACGTTGTAAAACGACTTTGTCTCAATTTGGCAGSA	CCCCTACTGCTACTGCTTTTC	length difference from RNAseq data
3_c8097	CACGACGTTGTAAAACGACTGAAAGTTTGGCAGTGGTTG	GGGGCACACTTCCACCTAAT	length difference from RNAseq data
3_c8562	CACGACGTTGTAAAACGACATCCAGCCCCCAAAAATAAC	TTGGTGGGTGTGAAGTGAAA	length difference from RNAseq data
3_c13171	CACGACGTTGTAAAACGACGCTGTTTGGGTATTTTGGTGA	TGTGAAGAATTTTCATGAGATAGGG	length difference from RNAseq data
4_c809	CACGACGTTGTAAAACGACATTTTTACACCGCCGTTGAC	CCATACACTTGAAACTGCGTTC	length difference from PCR sequencing (above)
4_c1856	CACGACGTTGTAAAACGACGGAACCTGTGGCTCCAGTAG	CGAACAACATTAAACGAACTTATG	length difference from PCR sequencing (above)
4_c6172	CACGACGTTGTAAAACGACAAAGCAAGCAAGTGCCATCT	ATAGAACAGTCCGAAACGGTTG	length difference from PCR sequencing (above)
4_c6735	CACGACGTTGTAAAACGACCCGGTAGTGGGAAGTTGAAA	TTTTTGTTCACGCACCAATATG	length difference from PCR sequencing (above)
4_c8726	CACGACGTTGTAAAACGACACGCTTACCAAGCTCCATTG	GAAAATGAGTGTGCAGGTCAAA	length difference from PCR sequencing (above)
4_c12731	CACGACGTTGTAAAACGACATAACCCGAGGCACAATGAG	CCCAAATYAACCCATTCATTA	length difference from PCR sequencing (above)

*LG is given at the start of primer name
**note the universal addition to the F primer (CACGACGTTGTAAAACGAC)

2301 loci differentially-expressed in necrotic F2 (reduced to most specific terms)

GO Term	Name	Type	FDR	single test p-Value	# in test group	# in reference group	# non annot test	# non annot reference group	Over/Under
	Fisher's Exact Test								
	Test-Set: 2301DE_ttest.txt								
	Test all Gene Ontology terms if they are enriched in a test group when compared to a reference group using Fisher's Exact Test with Multiple Testing Correction of FDR (Benjamini and Hochberg).								
GO:0003735	structural constituent of ribosome	F	4.20E-27	1.70E-30	90	103	1811	12607	over
GO:0022627	cytosolic small ribosomal subunit	C	4.30E-17	4.80E-20	32	11	1869	12699	over
GO:0005730	nucleolus	C	3.20E-16	4.00E-19	107	235	1794	12475	over
GO:0022625	cytosolic large ribosomal subunit	C	1.80E-14	2.50E-17	32	17	1869	12693	over
GO:0001510	RNA methylation	P	2.30E-07	7.20E-10	40	72	1861	12638	over
GO:0005886	plasma membrane	C	9.50E-06	3.60E-08	440	2268	1461	10442	over
GO:0010103	stomatal complex morphogenesis	P	1.00E-05	4.10E-08	44	102	1857	12608	over
GO:0010075	regulation of meristem growth	P	2.90E-05	1.40E-07	49	128	1852	12582	over
GO:0009664	plant-type cell wall organization	P	2.30E-04	1.20E-06	48	136	1853	12574	over
GO:0006412	translation	P	2.80E-04	1.60E-06	129	538	1772	12172	over
GO:0009506	plasmodesma	C	4.20E-04	2.60E-06	131	555	1770	12155	over
GO:0043481	anthocyanin accumulation in tissues in response to UV light	P	1.20E-03	9.00E-06	27	60	1874	12650	over
GO:0009908	flower development	P	1.50E-03	1.20E-05	143	641	1758	12069	over
GO:0007169	transmembrane receptor protein tyrosine kinase signaling pathway	P	3.80E-03	3.40E-05	28	70	1873	12640	over
GO:0042254	ribosome biogenesis	P	5.70E-03	5.80E-05	55	196	1846	12514	over
GO:0008356	asymmetric cell division	P	9.00E-03	9.50E-05	9	8	1892	12702	over
GO:0005576	extracellular region	C	9.00E-03	9.50E-05	164	794	1737	11916	over
GO:0009965	leaf morphogenesis	P	1.00E-02	1.10E-04	39	125	1862	12585	over
GO:0019199	transmembrane receptor protein kinase activity	F	1.70E-02	2.00E-04	12	18	1889	12692	over
GO:0009955	adaxial/abaxial pattern specification	P	1.80E-02	2.20E-04	19	43	1882	12667	over
GO:0009505	plant-type cell wall	C	2.30E-02	2.90E-04	47	171	1854	12539	over
GO:0048767	root hair elongation	P	2.80E-02	3.60E-04	35	115	1866	12595	over
GO:0009825	multidimensional cell growth	P	2.80E-02	3.60E-04	25	70	1876	12640	over
GO:0010817	regulation of hormone levels	P	2.80E-02	3.70E-04	68	281	1833	12429	over
GO:0019843	rRNA binding	F	3.00E-02	4.00E-04	16	34	1885	12676	over
GO:0016878	acid-thiol ligase activity	F	3.10E-02	4.10E-04	12	20	1889	12690	over
GO:0032440	2-alkenal reductase [NAD(P)] activity	F	3.30E-02	4.50E-04	47	175	1854	12535	over
GO:0008236	serine-type peptidase activity	F	3.90E-02	5.60E-04	30	95	1871	12615	over
GO:0006805	xenobiotic metabolic process	P	4.20E-02	6.20E-04	5	2	1896	12708	over
GO:0006655	phosphatidylglycerol biosynthetic process	P	4.40E-02	6.70E-04	18	44	1883	12666	over

592 loci over-expressed in necrotic F2 (reduced to most specific terms)

GO Term	Name	Type	FDR	single test p-Value	# in test group	# in reference group	# non annot test	# non annot reference group	Over/Under
	Fisher's Exact Test								
	Test-Set: 592OEinnecrotic_ttest.txt								
	Test all Gene Ontology terms if they are enriched in a test group when compared to a reference group using Fisher's Exact Test with Multiple Testing Correction of FDR (Benjamini and Hochberg).								
GO:0006623	protein targeting to vacuole	P	4.00E-05	1.70E-08	18	94	451	14048	over
GO:0005777	peroxisome	C	1.50E-04	1.20E-07	23	177	446	13965	over
GO:0016192	vesicle-mediated transport	P	8.00E-03	7.90E-06	36	473	433	13669	over
GO:0016773	phosphotransferase activity, alcohol group as acceptor	F	1.10E-02	1.20E-05	62	1048	407	13094	over
GO:0008756	o-succinylbenzoate-CoA ligase activity	F	4.40E-02	6.60E-05	4	4	465	14138	over
GO:0016301	kinase activity	F	4.40E-02	6.70E-05	67	1233	402	12909	over
GO:0005886	plasma membrane	C	4.50E-02	7.60E-05	120	2588	349	11554	over

1709 loci under-expressed in necrotic F2 (reduced to most specific terms)

Fisher's Exact Test								
Test-Set: 1709UEinnecrotic_ttest.txt								
Test all Gene Ontology terms if they are enriched in a test group when compared to a reference group using Fisher's Exact Test with Multiple Testing Correction of FDR (Benjamini and Hochberg).								

GO Term	Name	Type	FDR	single test p-Value	# in test group	# in reference group	# non annot test	# non annot reference group	Over/Under
GO:0003735	structural constituent of ribosome	F	2.90E-34	1.60E-37	87	106	1345	13073	over
GO:0005730	nucleolus	C	3.60E-24	4.00E-27	104	238	1328	12941	over
GO:0022627	cytosolic small ribosomal subunit	C	4.90E-21	7.50E-24	32	11	1400	13168	over
GO:0022625	cytosolic large ribosomal subunit	C	2.60E-18	4.70E-21	32	17	1400	13162	over
GO:0001510	RNA methylation	P	3.40E-11	9.40E-14	40	72	1392	13107	over
GO:0010103	stomatal complex morphogenesis	P	5.60E-09	2.00E-11	43	103	1389	13076	over
GO:0010075	regulation of meristem growth	P	2.80E-07	1.40E-09	45	132	1387	13047	over
GO:0009664	plant-type cell wall organization	P	3.10E-07	1.60E-09	46	138	1386	13041	over
GO:0009506	plasmodesma	C	1.40E-05	1.10E-07	110	576	1322	12603	over
GO:0043481	anthocyanin accumulation in tissues in response to UV light	P	1.50E-05	1.30E-07	26	61	1406	13118	over
GO:0006412	translation	P	5.40E-05	5.40E-07	105	562	1327	12617	over
GO:0007169	transmembrane receptor protein tyrosine kinase signaling pathway	P	5.10E-04	5.80E-06	25	73	1407	13106	over
GO:0009825	multidimensional cell growth	P	9.00E-04	1.10E-05	24	71	1408	13108	over
GO:0019843	rRNA binding	F	1.10E-03	1.30E-05	16	34	1416	13145	over
GO:0009505	plant-type cell wall	C	1.20E-03	1.40E-05	42	176	1390	13003	over
GO:0032440	2-alkenal reductase [NAD(P)] activity	F	1.70E-03	2.30E-05	42	180	1390	12999	over
GO:0009965	leaf morphogenesis	P	3.40E-03	4.70E-05	33	131	1399	13048	over
GO:0006655	phosphatidylglycerol biosynthetic process	P	4.60E-03	6.60E-05	17	45	1415	13134	over
GO:0008356	asymmetric cell division	P	5.80E-03	9.00E-05	8	9	1424	13170	over
GO:0048767	root hair elongation	P	7.10E-03	1.10E-04	30	120	1402	13059	over
GO:0048437	floral organ development	P	8.30E-03	1.40E-04	69	384	1363	12795	over
GO:0005576	extracellular region	C	8.70E-03	1.40E-04	128	830	1304	12349	over
GO:0007020	microtubule nucleation	P	1.00E-02	1.70E-04	19	60	1413	13119	over
GO:0009220	pyrimidine ribonucleotide biosynthetic process	P	1.20E-02	2.20E-04	26	101	1406	13078	over
GO:0006913	nucleocytoplasmic transport	P	1.30E-02	2.30E-04	32	138	1400	13041	over
GO:0009955	adaxial/abaxial pattern specification	P	1.30E-02	2.30E-04	16	46	1416	13133	over
GO:0010817	regulation of hormone levels	P	1.40E-02	2.70E-04	55	294	1377	12885	over
GO:0042545	cell wall modification	P	1.60E-02	3.10E-04	29	122	1403	13057	over
GO:0048439	flower morphogenesis	P	1.80E-02	3.40E-04	16	48	1416	13131	over
GO:0009933	meristem structural organization	P	2.00E-02	3.90E-04	42	209	1390	12970	over
GO:0060771	phyllotactic patterning	P	2.10E-02	4.20E-04	4	1	1428	13178	over
GO:0000462	maturation of SSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA)	P	2.10E-02	4.20E-04	4	1	1428	13178	over
GO:0048438	floral whorl development	P	2.60E-02	5.20E-04	58	325	1374	12854	over
GO:0045036	protein targeting to chloroplast	P	2.90E-02	6.00E-04	16	51	1416	13128	over
GO:0016441	posttranscriptional gene silencing	P	3.40E-02	7.20E-04	38	189	1394	12990	over
GO:0009855	determination of bilateral symmetry	P	3.60E-02	7.80E-04	22	87	1410	13092	over
GO:0047668	amygdalin beta-glucosidase activity	F	4.10E-02	9.40E-04	3	0	1429	13179	over
GO:0008460	dTDP-glucose 4,6-dehydratase activity	F	4.10E-02	9.40E-04	3	0	1429	13179	over
GO:0031957	very long-chain fatty acid-CoA ligase activity	F	4.10E-02	9.40E-04	3	0	1429	13179	over
GO:0080082	esculin beta-glucosidase activity	F	4.10E-02	9.40E-04	3	0	1429	13179	over
GO:0080081	4-methylumbelliferyl-beta-D-glucopyranoside beta-glucosidase activity	F	4.10E-02	9.40E-04	3	0	1429	13179	over
GO:0010159	specification of organ position	P	4.10E-02	9.40E-04	3	0	1429	13179	over
GO:0009240	isopentenyl diphosphate biosynthetic process	P	4.40E-02	1.00E-03	35	173	1397	13006	over
GO:0045338	farnesyl diphosphate metabolic process	P	4.80E-02	1.20E-03	4	2	1428	13177	over
GO:0048830	adventitious root development	P	4.80E-02	1.20E-03	4	2	1428	13177	over
GO:0046658	anchored to plasma membrane	C	4.90E-02	1.20E-03	13	39	1419	13140	over