Theor Appl Genet (2006) 113:1551–1561
DOI 10.1007/s00122-006-0402-3

ORIGINAL PAPER

Segregation distortion in *Arabidopsis* C24/Col-0 and Col-0/C24 recombinant inbred line populations is due to reduced fertility caused by epistatic interaction of two loci

Ottó Törjék · Hanna Witucka-Wall · Rhonda C. Meyer · Maria von Korff · Barbara Kusterer · Carsten Rautengarten · Thomas Altmann

Received: 14 July 2006 / Accepted: 18 August 2006 / Published online: 19 September 2006

Abstract A new large set of reciprocal recombinant inbred lines (RILs) was created between the *Arabidopsis* accessions Col-0 and C24 for quantitative trait mapping approaches, consisting of 209 Col-0 × C24 and 214 C24 × Col-0 F_7 RI lines. Genotyping was performed using 110 evenly distributed framework single nucleotide polymorphism markers, yielding a genetic map of 425.70 cM, with an average interval of 3.87 cM. Segregation distortion (SD) was observed in several genomic regions during the construction of the genetic map. Linkage disequilibrium analysis revealed an association between a distorted region at the bottom of chromosome V and a non-distorted region on chromosome IV. A detailed analysis of the RILs for these two regions showed that an SD occurred when homozygous Col-0 alleles on chromosome IV coincided with homozygous C24 alleles at the bottom of chromosome V. Using nearly isogenic lines segregating for the distorted region we confirmed that this genotypic composition leads to reduced fertility and fitness.

Communicated by C. Quiros.

Ottó Törjék and Hanna Witucka-Wall contributed equally to this work.

O. Törjék (✉) · H. Witucka-Wall · R. C. Meyer · M. von Korff · C. Rautengarten · T. Altmann
Department of Genetics, University of Potsdam, Karl-Liebknecht-Str. 24-25, 14476 Potsdam-Golm, Germany
e-mail: toerjek@mpimp-golm.mpg.de

T. Altmann
Max-Planck-Institute of Molecular Plant Physiology, Am Muehlenberg 1, 14476 Golm, Germany

B. Kusterer
Institute of Plant Breeding, Seed Science, and Population Genetics, University of Hohenheim, 70593 Stuttgart, Germany

Introduction

Recombinant inbred lines (RILs) are an important resource for investigation and genetic mapping of quantitative and qualitative traits in many plant species. Based on the high degree of homozygosity they represent a permanent so-called "immortal" segregating population allowing the inherent environmental error to be reduced by replication and providing an efficient system for quantitative trait locus (QTL) analysis (Koornneef et al. 1997; Borevitz et al. 2002). Key factors for such mapping studies are the number of identified recombination events, the marker density and the trait complexity. In this context RILs are first of all an excellent source for mapping of QTL to broad genomic regions (Peleman et al. 2005; Darvasi 1998). An important advantage of the RILs is that after genotyping, the same population can be used for the mapping of any number traits for which the parents differ (Koornneef et al. 2004), thus reducing the costs and the effort required to identify QTL positions. In *Arabidopsis* several well-characterised RIL populations exist, which allowed the mapping and identification of many QTL. For an extensive review about the currently available RILs developed for different accessions see Koornneef et al. (2004), El-Lithy et al. (2006). Once QTL have been identified, the confirmation and fine-mapping can be facilitated by using nearly-isogenic lines (NILs) (Koornneef et al. 1997; Zamir 2001).

In recent years, segregation distortion (SD) has been reported in a variety of taxa. SD leads to a deviation of the observed genotypic frequencies from the expected Mendelian ratios. Several mechanisms leading to SD have been described for *Drosophila* (Lyttle 1993), mouse (Silver 1993) and *Silene* (Taylor and Ingvarsson 2003). A summary of studies on SDs found in plants during 1986–1995 has been reported by Xu et al. (1997). Referencing only few later examples here, SD has been observed in maize (Lu et al. 2002), Populus (Yin et al. 2004), rice (Wang et al. 2005) and mungbean (Lambrides et al. 2004). The molecular background of this phenomenon is poorly understood, but it is known that SD can be caused by a variety of environmental, physiological and genetic factors. It can affect either the male or female germ line, or it is a result of post-zygotic selection (Lyttle 1991; Hormaza and Herrero 1992). SD is often attributed to pollen-pistil incompatibilities (Lord and Russell 2002), gametic competition (Lu et al. 2002), negative epistatic interactions (Li et al. 1997) or gamete abortion (Sano 1990). SD observed in RIL populations derived via single-seed descent represents an effect of both genetic and environmental factors cumulated in multiple generations. In the present work, we report on a new large set of reciprocal RILs created between the *Arabidopsis* accessions Col-0 and C24. We show that an SD region in the *Arabidopsis* genome is associated with reduced fertility and is predominantly caused by an epistatic interaction of two loci residing on chromosomes IV and V.

Materials and methods

Creation of RILs

Two reciprocal sets of RILs were developed from a cross between the two *A. thaliana* accessions C24 (originated from J.P. Hernalsteens, Vrije Universiteit Brussels) and Col-0 (from G. Rédei; Univ. Missouri-Columbia, USA). F_2 plants were propagated by self-pollination using the single-seed descent method to the F_7 generation, where genotyping and bulk amplification was performed. The final population consisted of 209 Col-0 × C24 F_7 and 214 C24 × Col-0 F_7 individuals, which we refer to as the Col-0 × C24 RILs and C24 × Col-0 RILs, respectively.

DNA-isolation and genotyping

Genomic DNA was extracted from 50 mg leaf tissue of one F_7 plant according to the modified protocol of Benito et al. (1993), Törjék et al. (2002) converted to a 96-well format. The leaf tissue collected in Collection Microtubes (Qiagen, Hilden, Germany) was frozen in liquid N_2 and powdered using a Retsch MM301 extractor (Retsch GmbH, Haan, Germany) at 20 Hz for 2 × 1 min. The fine powder was mixed with 300 µl of Extraction Buffer (100 mM Tris–HCl pH 8.0, 50 mM EDTA pH 8.0, 500 mM NaCl) and 75 µl of 10% SDS and then incubated at 60°C for 20 min. After adding 200 µl of 3 M Potassium-acetate, pH 4.8 followed by incubation at 0°C for 10 min, the samples were centrifuged (5,600–6,000 × *g*, 10 min) and the DNA was precipitated from the supernatant by adding 1/10 volume of Na-acetate (pH 5.2) and two volumes of 96% ethanol for 5 min at room-temperature. Afterwards DNA was collected (5,600–6,000 × *g*, 10 min), solved in 150 µl TE1 (10 mM Tris–HCl pH 8, 1 mM EDTA) and treated with RNAse (1 µl/sample, 10 mg/ml). After two washing steps (70% ethanol) the DNA was eluted in 150 µl TE1 buffer.

The collected DNA samples were genotyped with a set of 110 framework markers (Table 1) established for MALDI-ToF analysis (performed by GAG-Bioscience GmbH, Bremen, Germany) as described in Törjék et al. (2003) and Schmid et al. (2006). Missing MALDI-ToF genotypes (2%) were substituted using SNaPshot assays developed for the same framework marker set. More details about the markers can be obtained from the website: http://www.mpimp-golm.mpg.de/arab-diversity/.

Map construction, analysis of LD

The linkage analysis was carried out using the Joinmap software version 3.0 (Van Ooijen and Voorrips 2001), which implements a weighted least-squares method adding the markers sequentially into the map. Marker order was permutated after each marker was added. Linkage between two markers was declared significant in two-point analyses when the LOD score (log 10 of likelihood odds ratio) exceeded a threshold of 3.5. After determination of the linkage groups and linear alignment of marker loci along them, recombination frequencies between marker loci were estimated by multi-point analyses and transformed into centiMorgans (cM) by using Haldane's (1919) mapping function. The map was constructed with all 110 scored single nucleotide polymorphism (SNP) markers (Table 1) including both reciprocal RIL populations together.

Estimation of pairwise linkage disequilibrium (LD) between SNP markers was carried out using the software PowerMarker v3.25 (Liu and Muse 2005, http://www.powermarker.net/). D' values were calculated by

Table 1 Set of 110 framework SNP markers used for genotyping of RILs and for construction of genetic map

Chromosome 1			Chromosome 2			Chromosome 3			Chromosome 4			Chromosome 5		
Marker	Physical position	Genetical position	Marker	Physical position	Genetical position	Marker	Physical position	Genetical position	Marker	Physical position	Genetical position	Marker	Physical position	Genetical position
MASC03771	174605	0.00	MASC05502	132623	0.00	MASC03898	276925	0.00	MASC04123	300135	0.00	MASC00144	506801	0.00
MASC07014	1189392	4.51	MASC05477	1124046	7.48	MASC03001	971413	2.01	MASC04725	1091297	5.88	MASC04860	1193468	2.35
MASC03758	2312713	8.79	MASC06808	2025926	10.32	MASC05312	2087119	5.78	MASC05042	2187167	8.33	MASC04531	2073108	5.61
MASC09203	3487296	12.42	MASC05962	2995450	13.90	MASC03344	3463241	10.75	MASC09225	2905965	8.33	MASC09207	3284128	8.80
MASC02475	4759749	16.07	MASC05803	3531067	13.90	MASC02947	4625704	13.98	MASC04685	4436613	9.56	MASC05127	4243050	11.64
MASC03658	5855350	19.83	MASC05857	4405145	13.90	MASC04608	5551329	17.71	MASC02668	5326601	20.55	MASC03612	5209985	14.52
MASC03911	7144350	22.23	MASC05360	5328235	13.49	MASC04279	6535963	20.20	MASC09213	6810745	33.51	MASC09208	6537447	18.61
MASC05029	8168405	26.47	MASC05657	7669169	24.69	MASC02841	7829419	24.37	MASC03275	8167827	40.38	MASC09209	7773667	23.45
MASC05303	9370965	30.28	MASC02747	8577949	29.90	MASC04516	8980078	32.25	MASC03263	8959714	43.08	MASC03559	8866230	30.04
MASC09223	10895982	36.63	MASC02600	9532715	33.38	MASC04523	9943260	38.02	MASC02548	9819803	46.87	MASC04983	10136884	37.60
MASC02998	11442495	39.01	MASC02644	10577823	39.31	MASC02648	11050700	45.99	MASC04642	10992260	52.60	MASC03952	11108011	40.80
MASC09204	12534056	42.71	MASC09221	11573651	42.53	MASC04262	12490580	49.48	MASC09214	11916062	55.99	MASC01174	12145995	40.80
MASC02577	13841417	48.07	MASC06104	12668467	48.48	MASC05045	14600527	50.04	MASC09215	12793881	58.06	MASC01361	13849125	42.97
MASC04127	15508237	49.93	MASC05434	13438594	52.33	MASC04819	15734377	52.25	MASC03154	13864490	60.40	MASC04275	14990317	48.75
MASC04209	16224562	51.93	MASC09222	14524299	55.32	MASC07090	16479540	57.09	MASC04005	14992413	63.08	MASC01582	16086029	53.87
MASC03340	17446072	57.26	MASC05386	15934504	60.73	MASC01171	16479665	57.09	MASC04199	16085270	65.98	MASC03128	17185298	59.33
MASC00545	18580358	61.26	MASC06025	17035571	64.41	MASC09224	18817123	67.69	MASC09216	16977975	70.83	MASC04317	18282539	62.17
MASC03754	19663589	64.97	MASC00371	18585623	69.23	MASC03218	20180613	73.95	MASC05258	17744374	74.07	MASC04298	19161824	64.69
MASC09205	20497212	67.37	MASC02812	19820979	72.39	MASC02788	21060362	78.11				MASC02675	20198601	67.04
MASC04170	21894195	71.81				MASC09218	22134402	80.91				MASC04591	21429747	70.72
MASC03447	23645218	76.30				MASC04925	22943559	82.33				MASC04394	22403190	75.35
MASC03631	24843636	81.66				MASC09219	23722561	85.53				MASC01545	23417207	78.78
MASC03684	26054726	88.42										MASC04576	24373942	81.81
MASC03930	27455076	93.15										MASC09211	25596187	86.26
MASC03765	28492717	96.38										MASC04350	26886066	92.62
MASC09206	29956051	101.04												

Physical (base positions on chromosomes. MIPS annotation. version 171102; Schoof et al. 2002) and genetic (cM values. Joinmap 3.0; Van Ooijen and Voorrips 2001) distances are given for each marker. For additional information about the markers see http://www2.mpiz-koeln.mpg.de/masc/ and http://www.mpimp-golm.mpg.de/arab-diversity

expectation maximization algorithm, as implemented in the software package. Tests for significant pairwise LD were conducted with χ^2 statistic and sequential Bonferroni correction of P values (Sokal and Rohlf 1995).

Results

RIL populations, genetic and physical map and allele frequencies

The whole set of 209 Col-0 × C24 and 214 C24 × Col-0 RI lines developed by a single seed descent (SSD) method was genotyped. Allele and genotype frequencies were estimated for both subpopulations separately. Global statistical analysis showed no difference in the overall allele frequencies between the two subpopulations. For the Col-0 × C24 RILs we calculated 50.6% Col-0 alleles, 47.2% C24 alleles, 2.1% heterozygous, 0.1% missing genotypes, and for C24 × Col-0 RILs: 50.2% Col-0 alleles, 47.2% C24 alleles, 2.5% heterozygous, 0.1% missing genotypes.

The genetic map constructed based on the complete set of 409 RILs has a length of 425.7 cM. One hundred and five loci were separated with an average distance of 3.87 cM (Table 1, Figs. 1, 2). No recombination events were found on chromosome II between markers MASC05962, MASC05803 and MASC05957, on chromosome III between MASC07090 and MASC01171, on chromosome IV between MASC05042 and MASC09225 as well as on chromosome V between MASC03952 and MASC01174 (these groups of markers were thus assigned to single loci). The largest genetic distances were found on chromosome IV between markers MASC02668 and MASC09213 as well as between MASC04685 and MASC02668 (12.96 and 10.99 cM, respectively). Similarly large distances occurred on chromosome II between three non-separated markers (MASC05962, MASC05803 and MASC05857) and MASC05657 (10.79 cM) as well as on chromosome III between two non-separated markers (MASC07090 and MASC01171) and MASC09224 (10.60 cM).

For all linkage groups all markers could be mapped according to their physical order except marker MASC05360 on chromosome II (physical position at 5.33 Mb), which was genetically mapped to position 13.49 cM on top of MASC05962, MASC05803, MASC05857 genetically mapped (without separation) to position 13.9 cM but which occupy more upstream physical positions: 2.99, 3.53 and 4.40 Mb, respectively.

Evidence of segregation distortion

The expected segregation ratio in RILs is 1:1 for each parental allele at each marker locus. Significant deviations (at $P < 0.05$) were observed for several genomic regions (Fig. 3) leading to a SD of up to 2.5:1 (Col-0:C24) at the marker loci MASC07090 and MASC01171 on chromosome III (present in both reciprocal RIL populations). Substantial SD was observed for chromosome IV for the markers MASC05042 and MASC09225, 2.34:1 (C24:Col-0), only in the Col-0 × C24 lines. Further markers with significant SD were: Chr. I—MASC03684, MASC03930, MASC03765, MASC09206 with the Col-0 allele overrepresented in Col-0 × C24; Chr. III—MASC03898, MASC03001 with the C24 allele overrepresented C24 × Col-0; MASC02648, MASC04262, MASC05045, MASC04819, MASC 09224, MASC03218, MASC02788, MASC09218, MASC04925, MASC09219 with the Col-0 allele overrepresented in Col-0 × C24; MASC02648, MASC 04262, MASC05045, MASC04819, MASC09224 with the Col-0 allele overrepresented in C24 × Col-0; Chr. IV—MASC04123, MASC04725, MASC04685, MASC02668 with the C24 allele overrepresented in Col-0 × C24; Chr. V—MASC09211, MASC04350 with the Col-0 allele overrepresented in both Col-0 × C24 and C24 × Col-0 RILs.

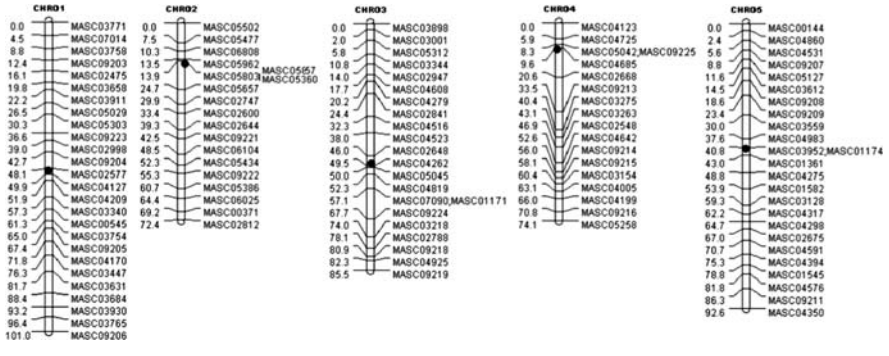

Fig. 1 Genetic map obtained from the reciprocal Col-0 × C24/C24 × Col-0 RIL population. The SNP markers are listed in Table 1

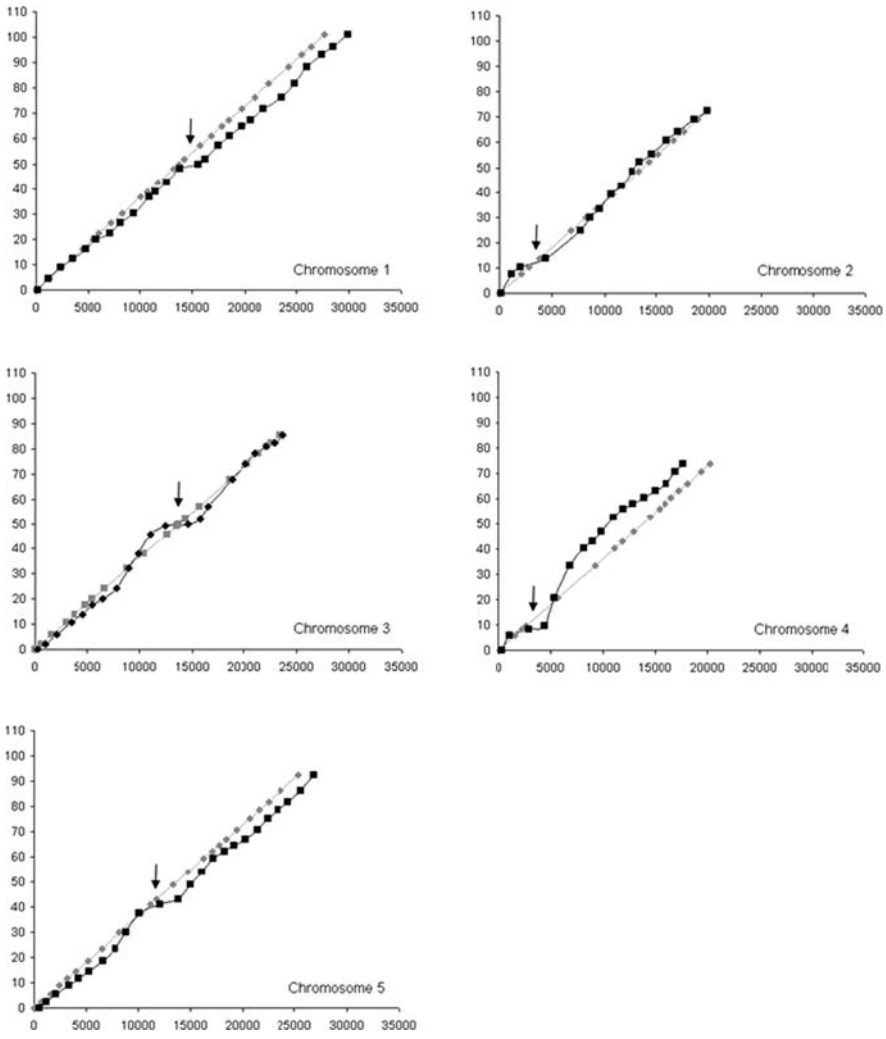

Fig. 2 Relation between the genetic and physical distances along the five *Arabidopsis* chromosomes. *Arrows* indicate the centromere positions. The *black line* represents detected local values, the *grey line* the calculated whole-genome mean ratio (274.23 kb). *X*-axis: physical distance (kb); *y*-axis: genetic distance (cM)

To test if pair wise interactions (associations) between some of the distorted regions exist, we performed LD analyses with the obtained genotype data using the PowerMarker program. As shown in Fig. 4 high LD values ($D' > 0.2$, $P < 0.05$) exist for all linked neighbouring marker positions. These high values are due to the close proximity of the markers (rarely separated by recombination events in RILs). High LD values have also been observed between unlinked genomic regions. The strongest was detected between the markers MASC09215 (chromosome IV) and MASC04350 (chromosome V) ($D = 0.60$, $P = 0.01$). Detailed analysis of these two markers showed that lines carrying the Col-0 allele at marker MASC09215 had for the unlinked marker position MASC04350 a highly significant deviation from the expected (49.22% homozygous Col-0:1.56% heterozygous:49.22% homozygous C24) segregation ratio: 93 Col-0 × C24 RILs segregated 77 Col-0 (82.8%):6 heterozygous (6.45%): 10 C24 (10.75%) and 102 C24 × Col-0 RILs segregated 90 Col-0 (88.24%):6 heterozygous (5.88%):6 C24 (5.88%) (Fig. 5a, b). This extreme MASC04350 segregation deviance was not detected in the subset of RILs bearing C24 genotypes at the MASC09215 locus. On the other hand, the deviance for the MASC09215 marker alleles was less drastic as RIL subsets having either Col-0 and C24 MASC04350 alleles have been formed (data not shown). Due to this observation, we suppose that the SD is caused by the Col-0 allele at the locus MASC09215 and that plants homozygous Col-0 at MASC09215/IV and homozygous C24 at MASC04350/IV have low fitness (Fig. 5c).

To confirm this hypothesis, we created segregating plants by selfing a nearly isogenic line (NIL, N70-BC_3F_1) with Col-0 homozygous background, which was heterozygous only for the bottom of the chromosome V encompassing also the position MASC04350. Analysing 218 descendants returned 49 plants with strongly reduced fertility and a phenotypic segregation of 1 (reduced fertility):3.45 (normal fertility) supporting the hypothesis above. Marker analysis using the markers MASC01519 and K9I9/43-44 (Table 2) confirmed that the reduced

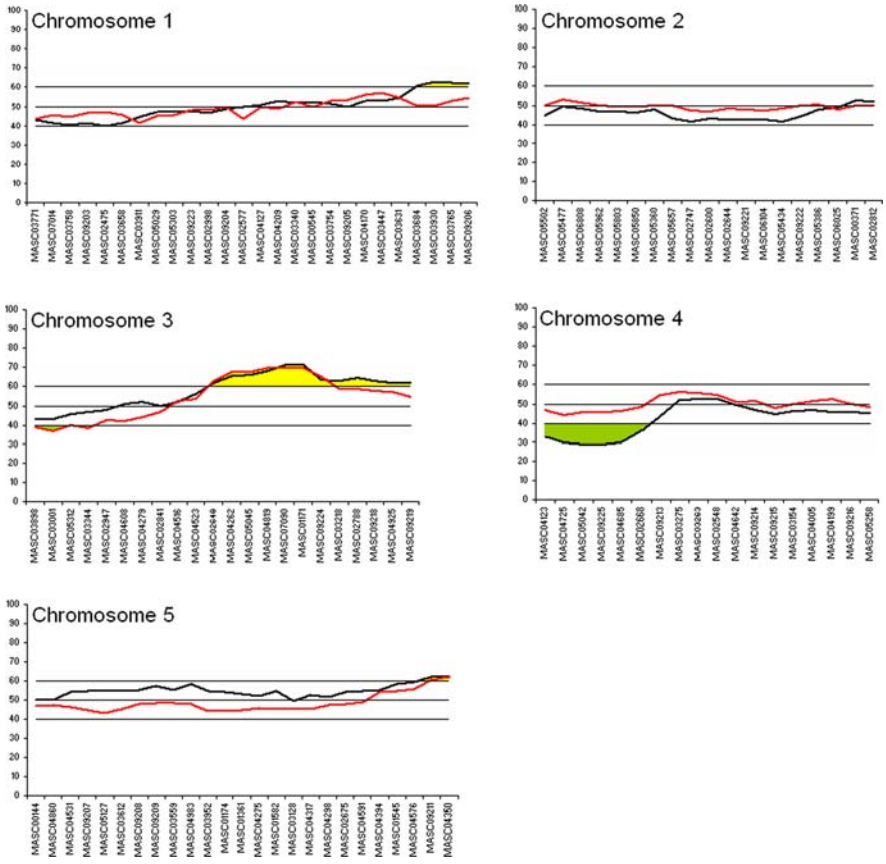

Fig. 3 Segregation between Col-0 and C24 alleles along the five chromosomes in Col-0 × C24 (*black line*) and C24 × Col (*red line*) RIL subpopulations. The percentage of the Col-0 allele is represented along the *y*-axis. Distorted regions deviating from the expected 1:1 ratio ($P < 0.05$) are indicated with *yellow* (if Col-0 alleles are overrepresented) or *green* (if C24 alleles are overrepresented) shading

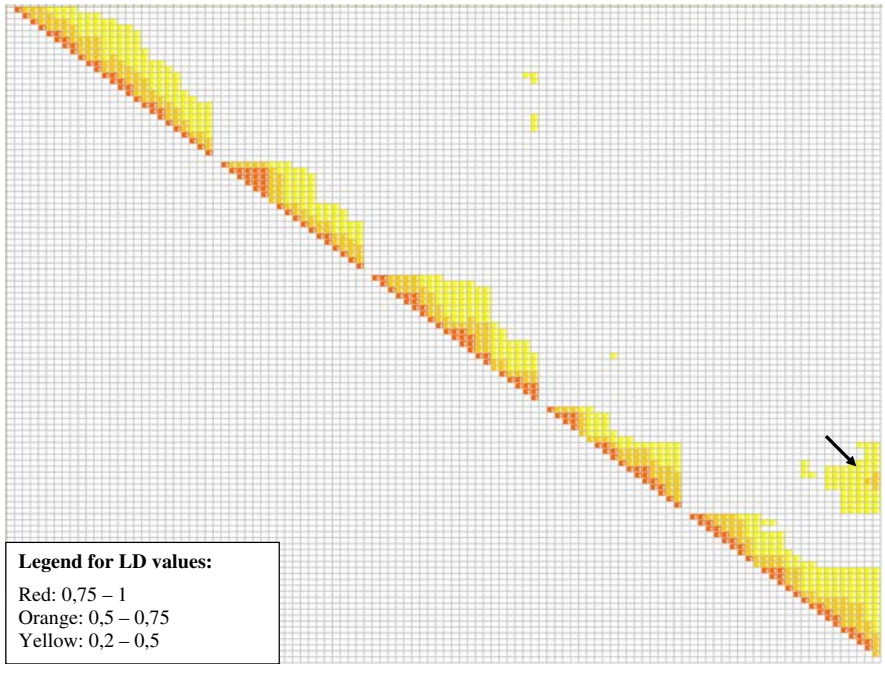

Fig. 4 Intergenomic interactions in RILs along the five chromosomes. The *arrow* indicates the highest non-linked intergenomic D' value found between the markers MASC09215 (chromosome 4) and MASC04350 (chromosome 5) $D' = 0.60$

fertility phenotype occurs in plants with homozygous C24 genotypes at the bottom of the chromosome V. This subset of plants produced smaller and shorter siliques with fewer or no seeds (Fig. 6b, d) compared to the plants, which were either heterozygous or homozygous Col-0 for the same region (Fig. 6a, c).

To investigate if the reduced fertility affected male or female function, flowers collected from plants with

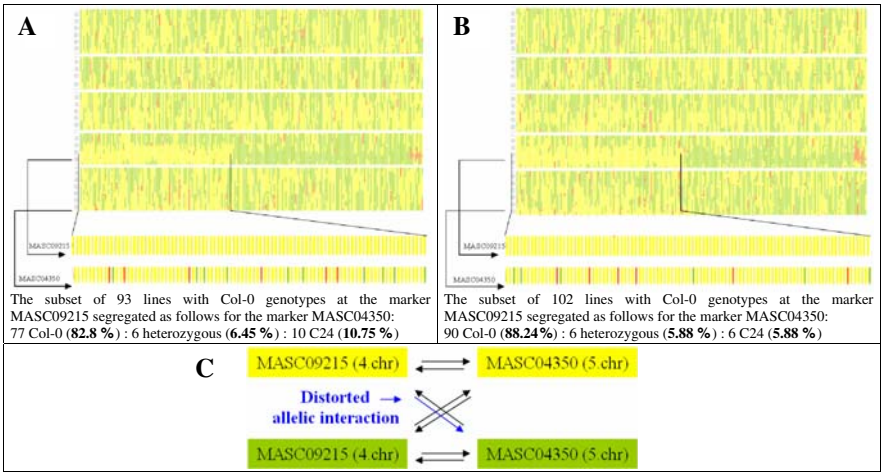

Fig. 5 Graphical genotypes of 209 Col-0 × C24 F7 (**a**), 214 C24 × Col-0 F7 (**b**) RI lines sorted based on the genotype at marker MASC09215. Segregation distortion in RILs is caused by Col-0 allele at the marker MASC09215 (**c**). Legend for genotypes: *yellow* Col-0, *green* C24, *red* heterozygous

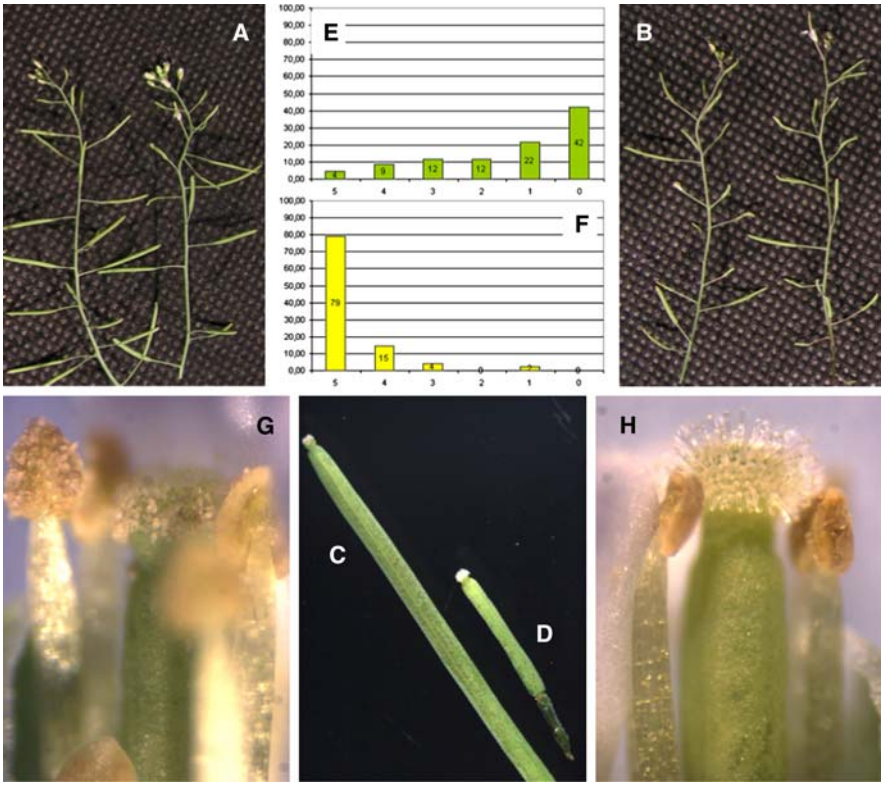

Fig. 6 Phenotype of plants with reduced fertility (distorted allelic interaction) compared to the lines with normal fertility. **a** Line with normal fertility—all siliques are filled with seeds; **b** line with reduced fertility (most of the siliques are short and empty); **c** silique of a line with normal fertility having the full seed set; **d** empty silique of a line with reduced fertility; **e** capacity of pollen production in the lines with reduced fertility (category 0 correspond to no pollen, 5 to pollen amount typical for normal flowers); **f** capacity of pollen production in plants with normal fertility; **g** Flower with normal floral organs and pollen grains; **h** Flower of a reduced fertility plant without pollen grains

normal and with reduced fertility have been examined microscopically. In contrast to the normal plants that produced abundant amounts of pollen, the flowers of the reduced fertility plants in general developed much fewer pollen (Fig. 6g, h) and 42% of them (39 of 93 flowers collected from 31 plants) completely lacked pollen. No morphological abnormalities in other floral organs were observed. To test whether the seeds in the few siliques of those plants with reduced fertility are not a result of crosspollination, the seeds from these plants (12 siliques/12 plants) were genotyped for the markers MASC01519 and K9I9/43-44 flanking the region of interest on both sides (Table 2). They all showed the homozygous C24 genotype and occasional crosspollination can therefore be excluded to be the cause of this seed formation (data not shown). Furthermore, when pollen of four independent plants with reduced fertility was used to pollinate Col-0 or C24 flowers, normal amounts of viable seeds were produced. Reciprocal crosses using plants with reduced fertility as females also produced seeds without visible differences for distinct combinations.

Twenty-one plants from the BC_3F_2 progeny (218 descendants of the NIL N70) were identified as recombinant for the region between MASC01519 and K9I9/43-44. Using five additional markers between these two

Table 2 New markers used for mapping on the bottom of the chromosome 5

Marker name	Chromosome position (base)	Primer (5′–3′)	Polymorphism (C24/Col)
MASC01619	25889628	F-AGAGATGAGCGCGACGGAGC R-CATGAACGTCGCGGACTTCG S-TTTTTTTTTTTGACGGAAACTACGCTCTCACCTG	C/T
MASC01519	26096352	F-ATCCGATAGAAGACGCATGAGC R-TTGTCACAGACCCGAAAAGACG S-TTTTTTTTTTTTTTTTGCGGCGGAGACGGCGAAGCAGAG	T/G
MASC06738	26193193	F-TTGTGTTTGGTCAGTGTCATCA R-CGAGGAAAAGAGACTGCCATC S-TTTTTTTTTTTTTTTTTTAGAAAG GAAACA	G/T
MASC07101	26381972	F-CGTCGTTGGGCAAGTACAAT R-TTCCGAGATTTCAAGGTTCA S-TTTTTTTTTTTTTTTAACACCGTCTTTAACCGCCG	T/A
MASC01090	26469329	F-GTGCTCTGTATGCAGGTGCACG R-TTCTGATCAACCATGGCTGCTG S-TTTTTTTTTTTTTTTTTTGGGGATA CAGAGGAGATTTTGGTG	G/A
K1F13/52-53	Between 26580794 26581516	F-TGCAGTCCACATGGGCTTCC R-AGCTGATCTTCACCATGCTTGG	717/722
MDD7	26895915	F-GATTAAAGGTAAGTGGAGCAGA GC R-CACGAGAGAATGCGTGCAAGG	155/152
K9I9/43-44	26980554	F-CAGACTGTCACCCAGTTCTG R- GTACCGGACCTACTGATTGTA S-TTTTTTTTTTACTATAGTTTATCCAGT	A/T

MASC01619. MASC01519. MASC06738. MASC07101. MASC01090. K9I9/43-44 have been detected using SNaPshot reactions. For detection of length-polymorphisms (K1F13/52-53 and MDD7) the amplicons were labelled with Cy5.5 dCTP (Amerham Biosciences) and subsequently analysed on LI-COR 4300 DNA Analyser

F forward primer, *R* reverse primer, *S* extension primer

positions the locus causing the reduced fertility could be mapped to a 112-kb region (data not shown) flanked by the markers MASC01090 and K1F13/52-53.

Discussion

In this study, we present two new large reciprocal subsets of RILs that were developed for the purpose of QTL mapping. The genotyping of the developed 209 Col-0 × C24 and 214 C24 × Col-0 F_7 RI lines was performed in 110 evenly distributed framework positions (average physical distance 1.15 Mb) applying the Maldi-ToF and SNaPshot methodology, which have an error rate of less than 1% (previously shown by Törjék et al. 2003; Schmid et al. 2006). Due to the similar overall recombination rates and allele-frequencies in both subpopulations, a single integrated genetic map was created yielding a total length of 425.7 cM. All markers could be mapped according to their physical order except for the marker MASC05360 and the marker group MASC05962/MASC05803/MASC05857, which were placed in reverse order. A possible explanation for this are genomic rearrangements, which are typical for the pericentromeric part of the chromosome II (Lin et al. 1999). The authors reported mitochondrial DNA-insertion in this genomic region (the four markers above cover the chromosome positions from 2995450 to 5328235 bp) showing sequence polymorphisms between the accessions Col-0 and C24. In addition, the pericentric region is also characterised by the presence of a large number of transposable elements (Lin et al. 1999).

The constructed genetic map contains 105 genetic positions in contrast to the physical map with 110 sites. This is due to the missing recombination events among a group of three adjacent markers and among further three pairs of adjacent markers. The average genetic distance between two adjacent markers was 3.87 cM with a maximum value of 12.96 cM for the MASC02668–MASC09213 marker pair on chromosome IV. Except for the regions with high and low recombination rates the markers (Table 1, Fig. 1) are evenly distributed throughout the genetic map. Variation of recombination frequencies across the genome has also been noted in other RIL populations. The recombination frequency deduced from the Col/Ler RILs (Mayer et al. 1999) varied between 50 and 200 kb cM^{-1} on the chromosome arms and decreased to 1,000 kb cM^{-1} through the centromeric region of

chromosome IV. In our RILs we have an average recombination frequency of 274.23 kb cM^{-1} (Fig. 2) and the regions of low recombination frequency coincided with the centromers as expected. Taking the top of the chromosome IV, we observed for example a high recombination frequency of 134.55 kb cM^{-1} between the first two markers (MASC04123/300135 bp and MASC04725/1091297 bp) while no recombination events were detected in the next region between the third and fourth markers adjacent to the centromere (MASC05042/2187167 bp and MASC09225/2905965 bp). The comparison between our genetic map and other maps is not easy since different marker sets were used in different studies. Nevertheless the profiles of genetic/physical distances along the five chromosomes have similarities in different RIL populations, such as in Bay-0 × Shahdara (38 markers) (Loudet et al. 2002) or in the Ler × An-1 (65 markers) Ler × Kas-2 (78 markers) marker and Ler × Kond (75 markers) RILs (El-Lithy et al. 2006). Using a higher marker density the contrasts between the genomic regions with low and high recombination frequencies become more visible such as in our case.

In the two subsets of F_7 RILs a total number of 1910 (Col-0 × C24) and 2017 (C24 × Col-0) detected recombinations have been fixed, which corresponds to an average of 9.28 recombinations (1.86/chromosome) in individual RI lines with an average segment size of 32.06 cM. Counting each breakpoint as one recombination, an average of 37.76 fixed recombinations per interval can be observed over the whole set of RILs. Taking this relatively high number of recombinations represented in the RILs and the comparatively high number of evenly distributed genetic markers with a maximal marker distance of 12.96 cM, we think that our RIL population provides an excellent resource for QTL mapping. In addition, the construction of two reciprocal subpopulations genotyped with the same marker set is a versatile feature of this RIL population which opens the opportunity to investigate cytoplasmatic effects on QTL. The complete set of RILs and their genotype data will be available upon publication.

SD has been frequently found during the construction of genetic linkage maps in many different plant species. In the family *Brassicaceae* it was observed in different inter- (Boivin et al. 2004) and intraspecific (Bechsgaard et al. 2004; Kuittinen et al. 2004) crosses. SD has commonly been detected in different *Arabidopsis* RIL populations, where the expected segregation would be 1:1 for each parental allele. Alonso-Blanco et al. (1998) described SD in Ler × Cvi F_8 RILs on chromosomes I, III and V with a largest magnitude of distortion (1:2.3—an excess of Ler-alleles) on the top of the chromosome V. In Ler × Col F_8 RILs (Lister and Dean 1993) the bottom of the chromosome I and the top of the chromosome V showed aberrant segregation. The Bay-0 × Shahdara F_6 RILs developed by Loudet et al. (2002) contained regions with significant overrepresentation of one or the other parental allele on the lower part of chromosome II, on the top and the lower part of chromosome IV and on the bottom of chromosome V. El-Lithy et al. (2006) found markers clustered in certain genomic regions with SD ratios ranging from 1.4:1 to 2.1:1 in the three different F_9 RIL populations Ler × An-1, Ler × Kas-2 and Ler × Kond. In the population of 96 Nd × Col RILs (Werner et al. 2005) SD was detected at the bottom of chromosome I, on the top of chromosome II and IV and at the lower part of chromosome V. In the sets of F_7 RILs described in this paper we found aberrant segregation ratios at the bottom of chromosomes I and V, at the top of chromosomes III and IV and in the lower region of chromosome III. A complete and accurate comparison of the SD regions identified in different RILs is not easy since the marker sets used for genotyping differed, but it seems that some of the SD regions overlap such as for example the SD at the bottom of the chromosome I (Ler × Col, Nd × Col, Ler × Cvi and Col × C24 RILs), at the top of chromosome IV (Bay-0 × Shahdara, Nd × Col, Ler × An-1 and Col × C24) or at the bottom of chromosome V (Bay-0 × Shahdara, Ler × Kas-2, Ler × Kond, Col × C24 and C24 × Col). But there are also genomic regions with SD unique to specific RIL populations.

A significantly high LD of 0.6 ($P < 0.01$) was found between the distorted region at the bottom of chromosome V and the middle of chromosome IV, where the marker-allele frequencies were not skewed (Fig. 4). By means of RILs focusing on the two involved loci at the markers MASC09215 and MASC04350, we showed that the segregation distortion occurs when a homozygous Col-0 allele on chromosome IV coincided with homozygous C24 alleles at the bottom of chromosome V (Fig. 5). Forming subsets of RILs containing different alleles at both loci, we found that the highest distortion could be observed if the subset is homozygous Col-0 at the marker locus MASC09215.

The effects of the unfavourable allele interactions were tested using a BC_3F_1 NIL, the progeny of which segregated for the bottom of the chromosome V and was homozygous Col-0 for the remaining part of the genome. Among the offspring the plants having homozygous C24 genotypes showed a highly reduced fertility, which causes the SD observed in the RILs. The phenotype has a recessive character causing a reduced amount of pollen or no pollen production by flowers

having the unfavourable allele-combination. The penetrance of the phenotype is not complete. When pollen of low fertility plants was used to pollinate plants with different allelic compositions (Col-0 accession, C24 accession, Col-0 background with heterozygous segment on the bottom of chromosome V, Col-0 background with C24 segment on the bottom of chromosome V) all combinations produced seeds. Similarly pollination of the low fertility plants with pollen collected from different genotypes (Col-0 accession, C24 accession, Col-0 background with heterozygous segment on the bottom of chromosome V, Col-0 background with C24 segment on the bottom of chromosome V) leaded to well-developed seeds. According to these observations it is most likely that the major effect causing the low fertility is on male gametogenesis. The exact developmental stage that is affected, however, remains to be determined.

The region harbouring the responsible gene(s) on chomosome 5 has been confined to a 112 kb region flanked with the markers MASC01090 and K1F13/52-53. We are now poised to begin fine-scale mapping and subsequently map-based cloning of the gene(s) causing reduced fertility at the bottom of chromosome V. This work is facilitated by the existence of a NIL, which segregates only for the bottom of chromosome V.

The mode of interaction between the alleles of the gene(s) in the region at the bottom of chromosome V and the responsible gene(s) on chromosome IV remains to be further investigated and will require the isolation of the latter. A C24 NIL heterozygous only for the respective chromosome IV region is available and will be used to fine-map this locus. Isolation of the affected genes and elucidation of the underlying mechanism in *Arabidopsis* could yield important information for the causes of similar phenomena in crops and may be useful to engineer male sterility systems applicable to hybrid breeding.

Acknowledgments We thank Maik Zehnsdorf, Melanie Lück and Monique Zeh for excellent technical assistance, and Katrin Seehaus and Steffi Zimmermann for plant care. This work was supported by a grants of the Bundesministerium für Bildung und Forschung (BMBF) (#0312275A, #0313061 as part of the GABI program) to T.A., grants of the Deutsche Forschungsgesellschaft (DFG) to T.A. and R.C.M. (AL 387/6-1, 6-2) and to B.K. (ME 931/4-1, 4-2), a grant of the European Community to T.A. (QLG2-CT-2001-01097), and by the Max-Planck-Society.

References

Alonso-Blanco C, Peeters AJ, Koornneef M, Lister C, Dean C, van den Bosch N, Pot J, Kuiper MT (1998) Development of an AFLP based linkage map of Ler, Col and Cvi *Arabidopsis thaliana* ecotypes and construction of a Ler/Cvi recombinant inbred line population. Plant J 14:259–271

Bechsgaard J, Bataillon T, Schierup MH (2004) Uneven segregation of sporophytic self-incompatibility alleles in *Arabidopsis lyrata*. J Evol Biol 17:554–561

Benito C, Figueras AM, Zaragoza C, Gallego FJ, de la Pena A (1993) Rapid identification of *Triticeae* genotypes from single seeds using the polymerase chain reaction. Plant Mol Biol 21:181–183

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

Borevitz J, Maloof J, Lutes J, Tsegaye D, Redfern JL, Trainer GT, Werner JD, Asami T, Berry CC, Weigel D, Chory J (2002) Quantitative trait loci controlling light and hormone response in two accessions of *Arabidopsis thaliana*. Genetics 160:683–696

Darvasi A (1998) Experimental strategies for the genetic dissection of complex traits in animal models. Nat Genet 18:19–24

El-Lithy ME, Bentsink L, Hanhart CJ, Ruys GJ, Rovito D, Broekhof JLM, van der Poel HJA, van Eijk MJT, Vreugdenhil D, Koornneef M (2006) New *Arabidopsis* recombinant inbred line populations genotyped using SNPWave and their use for mapping flowering-time quantitative trait loci. Genetics 172:1867–1876

Haldane JBS (1919) The combination of linkage values, and the calculation of distances between the loci of linked factors. J Genet 1919:299–309

Hormaza JI, Herrero M (1992) Pollen selection. Theor Appl Genet 83:663–672

Koornneef M, Alonso-Blanco C, Peeters AJM (1997) Genetic approaches in plant physiology. New Phytol 137:1–8

Koornneef M, Alonso-Blanco C, Vreugdenhil D (2004) Naturally occurring genetic variation in *Arabidopsis thaliana*. Ann Rev Plant Biol 55:141–172

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

Lambrides CJ, Godwin ID, Lawn RJ, Imrie BC (2004) Segregation distortion for seed testa color in mungbean (*Vigna radiata* L. Wilcek). J Hered 95:532–535

Li Z, Pinson SRM, Paterson AH, Park WD, Stansel JW (1997) Genetics of hybrid sterility and hybrid breakdown in an intersubspecific rice (*Oryza sativa* L.) population. Genetics 145:1139–1148

Lin X, Kaul S, Rounsley S, Shea TP, Benito MI, Town CD, Fujii CY, Mason T, Bowman CL, Barnstead M et al (1999) Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*. Nature 402:761–768

Lister C, Dean C (1993) Recombinant inbred lines for mapping RFLP and phenotypic markers in *Arabidopsis thaliana*. Plant J 4:745–750

Liu K, Muse SV (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics 21:2128–2129

Lord EM, Russell SD (2002) Mechanisms of pollination and fertilization in plants. Annu Rev Cell Dev Biol 18:81–105

Loudet O, Chaillou S, Camilleri C, Bouchez D, Daniel-Vedele F (2002) Bay-0 × Shahdara recombinant inbred line population: a powerful tool for the genetic dissection of complex traits in *Arabidopsis*. TAG 104:1173–1184

Lu H, Romero-Severson J, Bernardo R (2002) Chromosomal regions associated with segregation distortion in maize. Theor Appl Genet 105:622–628

Lyttle TW (1991) Segregation distorters. Annu Rev Genet 25:511–557

Lyttle TW (1993) Cheaters sometimes prosper distortion of Mendelian segregation by meiotic drive. Trends Genet 9:205–210

Mayer K, Schuller C, Wambutt R, Murphy G, Volckaert G, Pohl T, Dusterhoft A, Stiekema W, Entian KD, Terryn N et al (1999) Sequence and analysis of chromosome 4 of the plant *Arabidopsis thaliana*. Nature 402:769–777

Peleman JD, Wye C, Zethof J, Sørensen AP, Verbakel H, van Oeveren J, Gerats T, van der Voort JR (2005) Quantitative trait locus (QTL) isogenic recombinant analysis: a method for high-resolution mapping of QTL within a single population. Genetics 171:1341–1352

Sano Y (1990) The genic nature of gamete eliminator in rice. Genetics 125:183–191

Schmid KJ, Törjék O, Meyer R, Schmuths H, Hoffmann MH, Altmann T (2006) Evidence for pleistocene refugia and postglacial expansion of *Arabiodpsis thaliana* from genome-wide SNP markers. TAG 112:1104–1114

Schoof H, Zaccaria P, Gundlach H, Lemcke K, Rudd S, Kolesov G, Arnold R, Mewes HW, Mayer KF (2002) MIPS Arabidopsis thaliana Database (MAtDB): an integrated biological knowledge resource based on the first complete plant genome. Nucleic Acids Res 30:91–93

Silver LM (1993) The peculiar journey of a selfish chromosome: mouse t haplotypes and meiotic drive. Trends Genet 9:250–254

Sokal RR, Rohlf FJ (1995) Biometry: the principles and practice of statistics in biological research. W. H. Freeman and Company, New York

Taylor DR, Ingvarsson PK (2003) Common features of segregation distortion in plants and animals. Genetica 117:27–35

Törjék O, Bucherna N, Kiss E, Homoki H, Finta-Korpelova Z, Bócsa I, Nagy I, Heszky LE (2002) Novel male-specific molecular markers (MADC5, MADC6) in hemp. Euphytica 127:209–218

Törjék O, Berger D, Meyer RC, Müssig C, Schmid KJ, Sörensen TR, Weisshaar B, Mitchell-Olds T, Altmann T (2003) Establishment of a high-efficiency SNP-based framework marker set for *Arabidopsis*. Plant J 36:122–140

Van Ooijen JW, Voorrips RE (2001) JoinMap® version 3.0: software for the calculation of genetic linkage maps. Plant Research International, Wageningen

Wang C, Zhu C, Zhai H, Wan J (2005) Mapping segregation distortion loci and quantitative trait loci for spikelet sterility in rice (*Oryza sativa* L.). Genet Res 86:97–106

Werner JD, Borevitz J, Warthmann N, Trainer GT, Ecker JR et al (2005) Quantitative trait locus mapping and DNA array hybridization identify an FLM deletion as a cause for natural flowering-time variation. Proc Natl Acad Sci USA 102:2460–2465

Xu Y, Zhu L, Xiao J, Huang N, McCouch SR (1997) Chromosomal regions associated with segregation distortion of molecular markers in F2, backcross, doubled haploid and recombinant inbred populations in rice (*Oryza sativa* L.). Mol Gen Genet 253:535–545

Yin TM, DiFazio SP, Gunter LE, Riemenschneider D, Tuskan GA (2004) Large-scale heterospecific segregation distortion in Populus revealed by dense genetic map. Theor Appl Genet 109:451–463

Zamir D (2001) Improving plant breeding with exotic genetic libraries. Nat Rev Genet 2:983–989