DOI: 10.1534/genetics.103.022343

Comparing the Linkage Maps of the Close Relatives *Arabidopsis lyrata* and *A. thaliana*

Helmi Kuittinen,*[,1] Anita A. de Haan,*[,2] Claus Vogl,[†] Sami Oikarinen,*[,3] Johanna Leppälä,* Marcus Koch,[‡,4] Tom Mitchell-Olds,[‡] Charles H. Langley[§] and Outi Savolainen*

**Department of Biology, University of Oulu, 90401 Oulu, Finland, [†]Institute of Animal Husbandry and Genetics, Veterinary Medicine, University of Wien, A-1210 Wien, Austria, [‡]Department of Genetics and Evolution, Max-Planck Institute of Chemical Ecology, 07745 Jena, Germany and [§]Center for Population Biology, University of California, Davis, California 95616*

Manuscript received September 18, 2003
Accepted for publication June 1, 2004

ABSTRACT

We have constructed a genetic map of *Arabidopsis lyrata*, a self-incompatible relative of the plant model species *A. thaliana*. *A. lyrata* is a diploid ($n = 8$) species that diverged from *A. thaliana* ($n = 5$) ~5 MYA. Mapping was conducted in a full-sib progeny of two unrelated F_1 hybrids between two European populations of *A. lyrata* ssp. *petraea*. We used the least-squares method of the Joinmap program for map construction. The gross chromosomal differences between the two species were most parsimoniously explained with three fusions, two reciprocal translocations, and one inversion. The total map length was 515 cM, and the distances were 12% larger than those between corresponding markers in the linkage map of *A. thaliana*. The 72 markers, consisting of microsatellites and gene-based markers, were spaced on average every 8 cM. Transmission ratio distortion was extensive, and most distortions were specific to each reciprocal cross, suggesting cytoplasmic interactions. We estimate locations and most probable genotype frequencies of transmission ratio distorting loci (TRDL) with a Bayesian method and discuss the possible reasons for the observed distortions.

ARABIDOPSIS thaliana has become the plant molecular biology model organism partly because of its very small genome size. It has also become a focus of plant comparative genomics studies (Schmidt 2002). Since the sequencing of the *A. thaliana* genome (Arabidopsis Genome Initiative 2000), the genomic organization of *A. thaliana* has been compared with that of many other species. This work has revealed a complex history of genome-wide duplication and subsequent gene loss (Lan *et al.* 2000; Vision *et al.* 2000; Simillion *et al.* 2002). Even in rather remote relatives, such as rice, tomato, and potato, syntenic stretches have been found (Devos *et al.* 1999; Ku *et al.* 2000; Gebhardt *et al.* 2003).

Comparisons with more closely related species have shown that the rate of chromosomal evolution in the Brassicaceae is quite rapid. Kowalski *et al.* (1994) examined rates of chromosomal evolution between *Brassica oleracea* and *A. thaliana* and found 26 rearrangements. In a more detailed comparative mapping effort between *B. nigra* and *A. thaliana*, the conserved segments were only ~8 cM in length (Lagercrantz 1998). As the divergence time between Brassica and *A. thaliana* is estimated to be ~20 MY, this gives rise to a high rate of chromosomal evolution. The rapid rate of chromosomal evolution was assigned to the polyploid Brassica lineage. A comparison of *A. thaliana* and the even closer relative, *Capsella rubella*, has shown extensive microsynteny (Acarkan *et al.* 2000; Rossberg *et al.* 2001). However, the closest relatives of *A. thaliana* have not been mapped comparatively at the whole-genome level.

In this article we construct a genetic map of *A. lyrata*, the closest relative of *A. thaliana* (Koch *et al.* 1999). This species is estimated to have diverged from *A. thaliana* ~5 MYA (Koch *et al.* 2000). The synonymous divergence is 12–15% (*e.g.*, Wright *et al.* 2003 and references therein). *A. lyrata*, like many other close relatives, has eight chromosomes (Jones 1963) while *A. thaliana* has five. The genome size is known to be about one and a quarter to twice as large as that of *A. thaliana* (T. Mitchell-Olds, unpublished results; E. Earle, unpublished results). The species can be crossed with *A. thaliana* (Měsíček 1967), and the hybrids produce backcross progeny, but there seems to be no recombination within chromosomes in the hybrids (Nasrallah *et al.* 2000). *A. lyrata* is self-incompatible (Jonsell *et al.* 1995) and outcrossing (Schierup 1998; Kärkkäinen *et al.* 1999) and has become a species of interest for population

[1]*Corresponding author:* Department of Biology, P.O. Box 3000, University of Oulu, 90401 Oulu, Finland. E-mail: helmi.kuittinen@oulu.fi

[2]*Present address:* De Ruiter Seeds, 2661 CZ Bergschenhoek, The Netherlands.

[3]*Present address:* Department of Medicine, University of Tampere, 33520 Tampere, Finland.

[4]*Present address:* Heidelberg Institute for Plant Science, University of Heidelberg, 69120 Heidelberg, Germany.

genetics (van Treuren *et al.* 1997; Savolainen *et al.* 2000; Mitchell-Olds 2001; Clauss *et al.* 2003; Wright *et al.* 2003).

The markers in the map are largely orthologs of *A. thaliana* flowering time genes. In addition, we have used other genes and microsatellite loci. The goal of the map is to examine patterns of chromosomal evolution at a large-scale structural level between these closely related species, *i.e.*, to answer questions such as: What kinds of changes have given rise to the difference of chromosome numbers of five and eight? Are these whole-arm translocations or chromosome fusions, or many smaller changes? Are there as many changes per million years between *A. lyrata* and *A. thaliana* as between *A. thaliana* and Brassica? Recalling that the genome size is larger than that in *A. thaliana*, what is the genetic size of the map? Where have the reductions of *A. thaliana* genome size occurred?

We used a three-generation pedigree, as we tried to avoid the possible effects of inbreeding depression in this self-incompatible species. As the mapping cross is between diverged populations of *A. lyrata*, we also examined the extent and possible causes of transmission ratio distortion in the cross. We used an extension of the Bayesian method of Vogl and Xu (2000) for analyzing transmission ratio distortion to an outbred full-sib family.

A further aim of the map is to provide a tool for *A. lyrata* genomic work. Once the patterns of the chromosomal rearrangements are known, it will be possible to examine the changes in more detail. In analyzing quantitative traits, the locations of quantitative trait loci between species can be efficiently compared. As this is an outcrossing self-incompatible species, there are no recombinant inbred lines available, but the plants genotyped for the map have been vegetatively propagated. The genotypic data and the vegetative clones are available, so the map can be used as such for QTL mapping or can be expanded for more detailed mapping.

MATERIALS AND METHODS

Cross: From each of the *A. lyrata* ssp. *petraea* populations, Mjällom (Sweden) and Karhumäki (Russia; for more information on the populations see van Treuren *et al.* 1997), two random individuals were chosen to produce two independent hybrid plants (K12 × M6: KM4-2 and M3 × K10: KM6-5). These two F_1 plants were crossed with each other to produce F_2 progeny, using both plants as pollen donor and recipient. This crossing scheme was chosen to avoid inbreeding depression in the F_2 progeny. In reciprocal 1, KM6-5 was the pollen recipient and KM4-2 was the pollen donor. In reciprocal 2, KM4-2 was the pollen recipient and KM6-5 was the pollen donor. The four grandparents, KM4-2, KM6-5, and 76-195 F_2 progeny in about equal proportions from the two reciprocal crosses were scored for marker genotypes. Most markers were scored for ~90 individuals and some loci that showed transmission ratio distortion were scored for additional individuals, leading to varying numbers of missing individuals for the markers (see supplemental Table 1 at http://www.genetics.org/supplemental/).

Markers: Altogether 72 different markers were used (Table 1). We used microsatellites (Clauss *et al.* 2002) and gene-based markers (Kuittinen *et al.* 2002) described earlier for various Brassicaceous species and some *A. thaliana* microsatellites (Bell and Ecker 1994). In addition, 47 other gene-based markers were adopted in this study. For them amplification primers were obtained from the Arabidopsis cleaved amplified polymorphic sequence marker database (http://www.arabidopsis.org/), or they were designed on the basis of the *A. thaliana* sequence in the way described in Kuittinen *et al.* (2002). The fragment sizes ranged from 0.5 to 2.4 kb. Amplification was carried out using standard conditions ($MgCl_2$ 1.5–3.3 mM, annealing temperature 48°–60°). Polymorphisms in the genes were discovered either by cutting the PCR products of the F_1 parents and some progeny with a set of 8–20 restriction enzymes or by sequencing the PCR product of the F_1 plants. Either the progeny genotypes at a segregating site were scored by cutting the PCR products with a suitable restriction enzyme and visualizing the fragments on agarose gel or, if no suitable enzyme was available, single nucleotide polymorphisms were scored with the ABI Prism SNaPshot Multiplex kit (Applied Biosystems, Foster City, CA). Some indel polymorphisms were scored as length variation on agarose gels. For some genes with indel polymorphisms a new primer with fluorescent end label was used to produce short fragments (<350 bp) that could be run on sequencing gels. Sequencing gels were run on the ABI Prism 377 sequencer and fragments were detected with the Genescan program and their sizes determined with the Genotyper program (Applied Biosystems). Polymorphism for the *ADH1* locus was detected by heteroduplex formation. PCR products were heated at 95° for 5 min, then at 65° for 1 hr, and then allowed to cool slowly to room temperature. Electrophoretic separation was then performed using a 1× MDE gel (FMC Bioproducts, Rockland, ME) run in 0.6× TBE for ~4.5 hr at 30 W at a constant temperature of 10°. Microsatellites were separated on 6% polyacrylamide or on 4% Metaphor gels and they were visualized either with silver staining or with FMBIO fluorescent gelscanner. Most microsatellites were run on sequencing gels using a fluorescent end label in one of the primers as described above for indel polymorphisms in the genes. The polymorphism could be scored in a codominant fashion at all loci except *CRY2* and *SLL2*, which were scored as dominant markers.

Segregation: In this cross a marker can segregate in only one of the parents (*aa* × *ab*, *ab* × *aa*) or in both parents (*ab* × *ab* if there are two alleles, *ab* × *ac* if there are three alleles, and *ab* × *cd* if there are four alleles). There were 15 fully informative loci; *i.e.*, either three or four alleles were segregating as described above. These included many microsatellites and some genes in which two polymorphic sites in different configurations could be scored. Most of the markers (41) segregated in the *ab* × *ab* configuration, and 16 segregated in only one of the F_1 parents. Fit of the proportions of different marker genotypes in the F_2 progeny to the expected 1:1, 1:2:1, 1:1:1:1, or 1:3 ratios was tested with χ^2 tests separately in both reciprocal crosses with the Joinmap program (Stam 1993). We used a locus-by-locus significance level of $\alpha = 0.05$.

The four genotypes in a fully informative cross should be found in equal numbers. Processes in both gametic and zygotic stages can give rise to deviations from the expected 1:1:1:1 segregation ratio. In the case of a fully informative marker or a marker that is segregating in a backcross manner, it is possible to dissect which of the alleles present in the F_1 parents is in excess or is deficient. If the bias is due to selection against some allele(s) at the gametic stage, the observed genotype

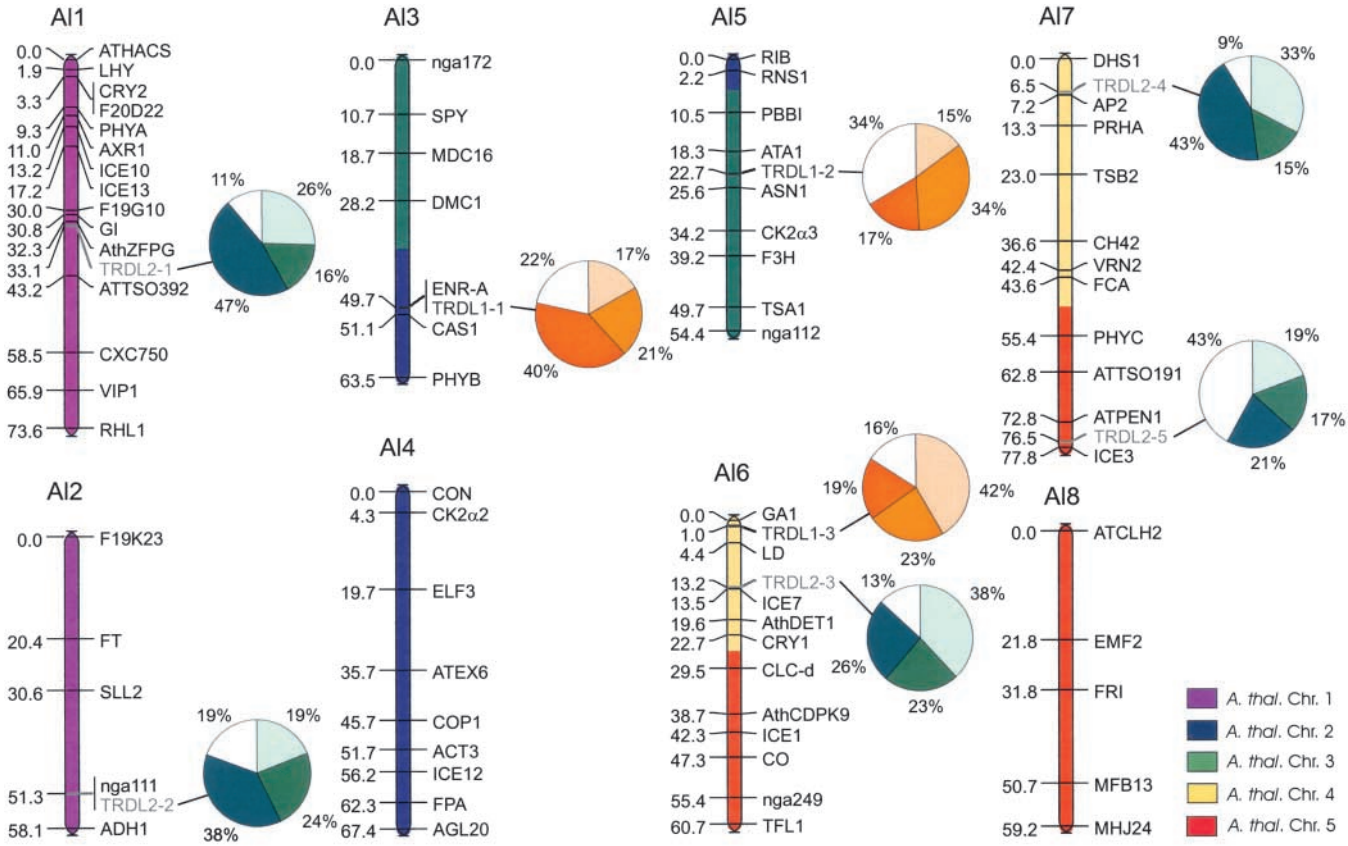

FIGURE 1.—The *A. lyrata* map constructed with the Joinmap method. The homology with the *A. thaliana* chromosomes is indicated by color. The positions and the most probable genotype frequencies at the transmission ratio distortion loci are indicated with pie diagrams (orange for reciprocal 1 and green for reciprocal 2). The order of the different genotypes, clockwise starting from the upper right, is: K_1K_2, K_1M_2, M_1K_2, and M_1M_2 from the cross between KM4-2 (K_1M_1) and KM6-5 (K_2M_2). K_1 and K_2 alleles are from Karhumäki and M_1 and M_2 alleles are from Mjällom.

frequencies should match expected frequencies that are calculated from the observed allele frequencies, assuming random union of gametes from each F_1 parent. This is also expected if there is zygotic selection with additive fitness values. If the observed genotype frequencies do not match the expected, selection on the zygote(s) is indicated. χ^2 tests were applied to test for deviations from equal allele frequencies in each F_1 parent, if possible, and to test for deviations from the random union of gametes, given the observed allele frequencies.

Map construction: We used the Joinmap program (version 3.0; STAM 1993) that uses a weighted least-squares method that adds markers sequentially in the map. The map was constructed with all scored markers, including those with significant transmission ratio distortion. The linkage groups (LGs) were determined using a LOD score threshold of 3.5. The Kosambi mapping function was used for calculation of the pairwise distances. For each LG, a map was constructed using thresholds of LOD 0.01 and a recombination frequency of 0.49 for pairs of markers, to use all linkage information. A threshold of 3 was used for an increase of the total χ^2 for the difference between observed and calculated map distances per added degrees of freedom, and a LOD threshold of 12 was used for triplet formation. Marker order was permutated after each marker was added. For most linkage groups all markers could be mapped after this. In linkage group AL1 the marker *VIP1*, in linkage group AL7 the marker *PRHA*, and in linkage group AL4 the markers *CK2a* and *CON* were mapped after relaxing the thresholds.

After map construction the location and effect of possible transmission ratio distortion loci (TRDL) in the chromosomes were estimated by applying the Bayesian method of VOGL and XU (2000). We use an extension of the method to outbreed full-sib data, such as in FISHMAN *et al.* (2001). It allows several multiplicatively interacting TRDL in a chromosome and can use partially informative marker information. As the segregation ratios appeared to differ among the two reciprocals, we performed a TRDL analysis separately in the two reciprocals. The maximum number of TRDL was set to four per chromosome. To obtain the posterior probabilities of the TRDLs, 10,000 simulations were run for each chromosome. The most likely position and the genotype frequencies were obtained.

RESULTS

Linkage groups: Markers were first assigned to linkage groups separately in the two reciprocal data sets. With a LOD 3 criterion there were no contradictory groupings. Linkage groups were, however, fragmented to different degrees in the two reciprocals, due to low sample sizes. Recombination rates between markers did not differ significantly between the reciprocals for the vast majority of comparisons as judged with the Joinmap program. In the end, the map was constructed with both recipro-

cals in one data set. The markers were divided into eight linkage groups, AL1–8 (Figure 1).

Translocations and fusions: As *A. lyrata* has a higher chromosome number than *A. thaliana*, we expected some large-scale rearrangements between the two genomes. Three breakages/fusions and two reciprocal translocations are sufficient to explain most of the differences in the maps of the two species (Figure 1). We inferred a breakage/fusion in the lower arm of chromosome 1, in the lower arm of chromosome 2, and in the lower arm of chromosome 5 of *A. thaliana*. AL3 consists of markers from the upper part of *A. thaliana* chromosome 3 and of markers from the middle part of *A. thaliana* chromosome 2. AL5 consists of markers from the upper part of *A. thaliana* chromosome 2 and the lower part of chromosome 3. This suggests a reciprocal translocation between *A. thaliana* chromosomes 2 and 3. AL6 is a combination of the upper arm of *A. thaliana* chromosome 4 and the upper arm of chromosome 5. Similarily, AL7 consists of markers in the lower part of *A. thaliana* chromosome 4, and markers from the middle part of *A. thaliana* chromosome 5. Thus a reciprocal translocation has occurred between these two chromosomes.

Marker order: The marker order of *A. lyrata*, to a great extent, was similar to that of the *A. thaliana* recombinant inbred map (Lister and Dean 1993; http://www.arabidopsis.org/). Compared with *A. thaliana*, pairs of closely linked markers in AL1 were reversed in order (*LHY-ATHACS*, *PHYA-AXR*, *GI-F19G10*), as were *ICE3* and *ATPEN1* at the lower arm of AL7. In these cases, the markers were closely situated and the actual order could well be consistent with the order in *A. thaliana*. The five markers in the upper end of AL6 (*GA1*, *LD*, *ICE7*, *AthDET1*, *CRY1*) were clearly in a different order compared to *A. thaliana* (*LD*, *GA1*, *CRY1*, *AthDET1*, *ICE7*). At least one inversion (*CRY1*, *DET1*, *ICE7*) can be inferred from the data. There is more uncertainty in the order of the closely linked markers *LD* and *GA1*.

In one case a locus was out of its *A. thaliana* context. The *FRI* locus, situated in chromosome 4 in *A. thaliana*, was mapped to AL8, which is homologous to the lower arm of *A. thaliana* chromosome 5. It was linked to the *EMF2* gene and microsatellite *MFB13*. Using a BLAST search we found in the *A. thaliana* chromosome 5 close to the *EMF2* gene a genomic region with high homology to a part of the *FRI* gene. A hypothetical protein was annotated across this sequence.

Length of the maps: The lengths of the linkage groups varied from 54.4 to 77.8 cM and summed to a total length of 515 cM. There was, on average, one marker every 8 cM. The total genome length is not likely to exceed this much because we selected markers at the ends of the *A. thaliana* linkage groups, and there are only ~13 Mb altogether between the telomeres and the outermost markers or between the markers flanking the three inferred chromosomal fusions in *A. thaliana*. The markers used in the *A. lyrata* map cover 483 cM in the five chromosomes of *A. thaliana* (TAIR database at http://www.arabidopsis.org/). Considering the colinear parts of the *A. thaliana* and *A. lyrata* maps, the Joinmap procedure resulted in genetic distances that were, on average, only 12% larger than those in the *A. thaliana* map. Since the *A. thaliana* map was also constructed with Joinmap, this comparison should be unbiased. The largest differences in the map distances between the two species are in AL2 and AL8, which correspond to the lower arms of *A. thaliana* chromosomes 1 and 5. These LGs are about twice as large in *A. lyrata* as in the corresponding areas in *A. thaliana*. AL4 was also extended to 140% compared to the homologous region in *A. thaliana*. The largest reductions in size were observed in the lower part of AL6 and in the upper part of AL7.

Transmission ratio distortion: In reciprocals 1 and 2, there were 18 (25%) and 25 (35%) distorted markers at the locus-by-locus significance level $\alpha = 0.05$ (Table 1 and supplemental Table 1 at http://www.genetics.org/supplemental/). The two reciprocals were distorted at only seven common loci. Thus, in total, there were 36 distorted markers (50%). Significant segregation distortion was found in 25 of the 72 markers tested (35%) with the χ^2 tests at $\alpha = 0.05$ when the data from both reciprocals were combined. The tests detected more distortion in separate data sets in spite of smaller sample sizes, obviously because the distortions were different in the two reciprocals. Distorted markers were usually linked to other similarily distorted markers so that certain chromosomal areas were distorted in the same direction, indicating a biological reason for the distortion instead of a genotyping error. If the genotype of a certain marker is mistyped frequently, the segregation ratio of this marker may be different from its neighboring markers. Hence, the algorithm for TRDL analysis would infer two spurious TRDLs with opposite effects to either side of a mistyped marker. We did not find evidence of closely spaced TRDLs of opposite effect. This indicates no major problems with genotyping in the data set. If only areas with at least two similarily distorted linked loci were taken into account, we obtained three distorted areas in reciprocal 1 and seven distorted areas in reciprocal 2. Only one of the areas (in AL6) was common to both reciprocals. The results from the Bayesian TRDL analysis were mostly congruent with those from the simple χ^2 tests. With a conservative cutoff value (<50% of the runs of the posterior distribution resulted in zero TRDL in a chromosome) there was support for three TRDL in reciprocal 1 and for five TRDL in reciprocal 2 (Figure 1).

Of the 11 markers in AL6, 8 were distorted in one or the other reciprocal cross (Table 1 and supplemental Table 1 at http://www.genetics.org/supplemental/). The Bayesian method strongly suggested a TRDL at the top of this linkage group in both reciprocals (TRLD1-3, TRLD2-3), as the posterior proportion resulting in one TRDL was

69% in both cases (Figure 1). The distortions were similar in the two crosses; in both cases there was an excess of Karhumäki alleles. Conditional on the number of observed gametes, the observed and expected genotype frequencies were similar. In reciprocal 1 (KM6-5 × KM4-2) there were three additional distorted areas. The Bayesian analysis detected a TRDL close to *ENR-A* in AL3 with 56% posterior probability (TRDL1-1). *ENR-A* was the only significantly distorted marker in the region detected with χ^2 tests, showing an excess of heterozygotes. In AL4 there was an excess of heterozygotes at two adjacent distorted marker loci (*ATEX6* and *COP1*; Table 1 and supplement Table 1 at http://www.genetics.org/supplemental/). The Bayesian method did not, however, support a TRDL in this linkage group. In AL5 there was an excess of Mjällom alleles at four linked marker loci; 53% of the runs supported a model with one TRDL (TRDL1-2). The Mjällom allele from KM6-5 was especially favored, while the observed genotype frequencies followed the expected. In reciprocal 2 (KM4-2 × KM6-5), the Bayesian analysis gave support for one TRDL (65.4% posterior proportion) in AL1 (TRDL2-1). There was an excess of the Karhumäki allele from KM6-5, and comparison of observed and expected genotype frequencies revealed additionally a deficiency of homozygotes at markers close to this TRDL. Some markers at the upper part of the LG were also distorted but in another direction, and the Bayesian analysis did not detect a TRDL there. In AL2 there was a deficiency of homozygotes at two linked loci, and the one TRDL model (TRDL2-2) had strong support (51% posterior proportion). In AL4 there was an excess of Karhumäki alleles at two linked loci (*CON* and *CK2*α) with no difference in observed and expected genotype frequencies. The Bayesian analysis did not support a TRDL in this linkage group. In a group of loci in AL7 there was a deficiency of Mjällom alleles and evidence for an excess of heterozygotes. In another set of linked loci in AL7, Mjällom alleles from KM4-2 were in excess and the observed genotype frequencies matched those expected, given the gametic frequencies. The Bayesian analysis supported two TRDL in these regions with a posterior proportion of 49% (TRDL2-4 and TRDL2-5).

DISCUSSION

Chromosomal evolution: Given that the basic chromosome number among most close relatives of *A. thaliana* is eight and that lower chromosome numbers in some species result from reduction of chromosome number (Jones 1963; Koch *et al.* 1999), the differences in chromosome number between *A. thaliana* and *A. lyrata* are most parsimoniously accounted for by three fusions in the *A. thaliana* lineage. Additionally, we inferred an inversion (*CRY1-DET-ICE7*) and two reciprocal translocations. One of the breakpoints of the other reciprocal translocation (AL6/AL7) is in an interval that is known to contain a centromere in *A. thaliana* (*PHYC-CLC-d*), and the other breakpoint is next to *CRY1*, which is close to the centromere. In the other reciprocal translocation (AL3/AL5) the breakpoints were also close to the centromere positions in *A. thaliana*.

In *A. thaliana* the *FRI* locus is on chromosome 4 but in *A. lyrata* it was mapped to a linkage group that is orthologous to *A. thaliana* chromosome 5. However, *A. thaliana* has a highly homologous sequence on chromosome 5, which lacks parts of the coding region, probably being a pseudogene. The nucleotide divergence between the two *A. thaliana* sequences was half of that between either *A. thaliana* sequence and the sequence of *A. lyrata*. This suggests that the duplication occurred recently in the *A. thaliana* lineage. Duplication and deletion of paralogous genes are known to have broken colinearity between different taxa and within species (*e.g.*, Ku *et al.* 2000; Vision *et al.* 2000; Fu and Dooner 2002; Li and Gill 2002).

The colinearity of the maps of these closely related species is apparently very high. High colinearity has also been found between genomes of *A. thaliana* and *C. rubella* (Brassicaceae) that have diverged 6.2–9.8 MYA (Acarkan *et al.* 2000). More markers could evidently reveal smaller-scale rearrangements, and only sequencing of long genomic contigs would reveal the details of genome evolution. Extensive microcolinearity was evident between sequences of *C. rubella* and *A. thaliana* (Rossberg *et al.* 2001), but several small-scale chromosomal rearrangements were detected between the more distant *B. oleracea* and *A. thaliana* sequences (Quiros *et al.* 2001).

The three fusions, two translocations, and one inversion that have occurred during the 5 MY after the *lyrata-thaliana* divergence (Koch *et al.* 2000) would give a rate estimate of 0.6 rearrangement/MY/genome. In the family Brassicaceae the rate of chromosomal evolution is exceptionally high. Kowalski *et al.* (1994) found 17 translocations and nine inversions between *B. oleracea* and *A. thaliana* with a map that had a resolution comparable to our map (15 cM between markers in *A. thaliana*). With a higher-resolution comparative map, Lagercrantz (1998) inferred that an average colinear segment is 8 cM and as many as 90 chromosomal rearrangements differentiate the genomes of *B. nigra* and *A. thaliana*. Lukens *et al.* (2003) obtained a somewhat larger estimate for an average colinear segment (11.6 cM) in *B. oleracea*, which corresponds to 3.8 Mbp in *A. thaliana*. With a divergence time of 14–24 MY (Yang *et al.* 1999; Koch *et al.* 2000), 26 rearrangements would result in a rate of 0.93–0.54 rearrangements/MY/genome and 90 rearrangements in 3.2–1.9 rearrangements/MY/genome. The rate of chromosomal evolution after the *lyrata-thaliana* divergence is thus lower than the estimate between *Brassica* and *A. thaliana*, if data with a comparable resolution are considered, and clearly lower than estimates from high-resolution map-

ping. The rates in comparisons among other plant taxa such as Lycopersicon-Solanum-Capsicum, Oryza-Zea-Sorghum, or *Gossypium* ssp. vary between 0.17 and 0.45 rearrangements/ MY (see Lagercrantz 1998, Table 3), and the estimate for *A. lyrata* and *A. thaliana* (0.6) is only slightly higher than that. However, the comparisons depend on the resolution of the map, on the criteria used for orthology, and on the estimates of divergence times.

It has been suggested that most of the chromosomal rearrangements between *A. thaliana* and Brassica have occurred as a result of polyploidization that took place after the divergence of Arabidopsis and Brassica (Lagercrantz and Lydiate 1996; Lagercrantz 1998; Lukens *et al.* 2003). The majority of rearrangements would thus have occurred in the Brassica lineage. The rather low number of chromosomal rearrangements in our study is in line with the hypothesis of a lower rate of genome evolution in the Arabidopsis lineage. Lagercrantz (1998) suggests that the ephemeral life styles of wild Brassica species would also favor fixation of chromosomal rearrangements that would be partially deleterious as heterozygotes. This should actually be true also for the weedy selfer *A. thaliana*. In contrast, *A. lyrata* is a perennial outcrosser that is likely to maintain more permanent populations. The location of *A. lyrata* markers in Brassica maps is currently not known. When the Brassica genome project proceeds, it will hopefully be possible to compare maps of the Arabidopsis and Brassica species in detail. This would allow defining whether the rearrangements have occurred in the lineage leading to *A. lyrata* or to *A. thaliana*.

The DNA content of *A. lyrata*, as measured by flow cytometry, is 125–185% as much as that in *A. thaliana* (T. Mitchell-Olds, unpublished results; E. Earle, unpublished results). We have regularly obtained only one sequence that corresponds to the one *A. thaliana* locus when developing PCR-based markers. On the basis of this evidence, gene duplications do not account for the higher genome size of *A. lyrata* compared to *A. thaliana*. Instead, this implies changes in noncoding DNA content. It has been suggested that intron sizes as well as intergenic regions may be larger in species with higher DNA content and that genes in large genomes may reside in gene-rich islands. Centromeric heterochromatin-associated sequences have often been shown to contribute to the variation in genome sizes. The reasons for genome-size differences are still not clear. There is some evidence for shorter introns in *A. thaliana* than in *A. lyrata* or Brassica (Wright *et al.* 2003) and shorter intergenic spacers in *A. thaliana* than in *B. oleracea* (Quiros *et al.* 2001). Despite the larger genome size, the length of the genetic map of *A. lyrata* was only slightly larger than that of *A. thaliana*.

Transmission ratio distortion: Transmission ratio distortion can result from many different processes that can act either before or after fertilization (reviewed, *e.g.*, in Korbecka *et al.* 2002 and in Fishman *et al.* 2001). Distortion of gametic frequencies can result from meiotic drive or meiotic disruptions associated with chromosomal rearrangements. Recombinant gametes may have differential fitness due to aneuploidy or epistatic interactions. Gametes may also have differential fertilization success. Selection may act after fertilization so that different zygotic genotypes have differential survival. Segregating deleterious mutations as well as male sterility and self-incompatibility alleles are known to cause segregation distortions. Much of the distortion in wide crosses is thought to be due to epistatic interactions between loci, *i.e.*, Dobzhansky-Muller incompatibilities (*e.g.*, Orr and Turelli 2001). With the cross design we aimed at excluding the influence of inbreeding depression or self-incompatibility alleles, but many other factors could have caused the observed distortions.

The extent of transmission ratio distortion was surprisingly high as the cross was within species. In each reciprocal cross, 25 and 35% of the loci were distorted at the 0.05 significance level, and taken together, 50% of the markers deviated from the Mendelian ratios. Remarkably, only one of the distorted regions was common to both reciprocal crosses and eight of the regions were present in only one of the crosses. The high levels of distortion found here are at the upper range of those normally found in crosses within species, and they resemble more those detected in crosses between species (reviewed in Jenczewski *et al.* 1997; Whitkus 1998; Fishman *et al.* 2001). This indicates that the Swedish Mjällom and the Russian Karhumäki populations are widely divergent. This is consistent with earlier information from molecular markers. Nei's standard genetic distance estimated from allozyme data was 0.39 and G_{st} based on microsatellites was 0.59 between the Karhumäki and Mjällom populations (van Treuren *et al.* 1997). Markers in a cross between two other European *A. lyrata* populations also exhibit a high level of segregation distortion (J. Leppälä *et al.*, unpublished results). Thus strong distortion is not specific to these two populations.

The mapping population was an outbred F_2 population with a maximum of four segregating alleles. We also used two reciprocal crosses. This experimental setup allowed obtaining information on different aspects of plant life cycle. With single-locus analyses segregation distortion could be assigned to the two reciprocals and to some extent to separate parents, depending on the informativeness of the loci. We could also analyze whether there was evidence for zygotic selection. The Bayesian analysis combined information from all markers and gave the most probable TRDL genotype frequencies.

In AL6, segregation distortion was similar in the two reciprocals. Mjällom alleles were selected against, and there was no indication of recessivity of the disfavored Mjällom allele (conditional on the gametic frequencies,

TABLE 1

Markers used in the map construction

Locus name	Identification/ BAC location in *A. thaliana*	Type of polymorphism	Forward primer	Reverse primer	Segregation distortion
			Linkage group AL1		
ATHACS	F22L4	Microsat	b	b	—
LHY	At1g01060	Indel	CAATAACAAAGCAGCGAGAGC	TGGCAGTTATACTTGGAGGAGAA	#
F20D22	At1g04120	Microsat	b	b	—
CRY2	At1g04400	*Dra*I	TTCCTGTCTTCATTTGGTGTCC	GGGTTTCGCATAGTTTGTTCC	#
PHYA	At1g09570	*Msp*I	CGTCATGCAAACTATCAGTGCTCAAACCC	GATTACTCAACCTCAGTGCG	—
AXR1	At1g05180	*Sty*I	GTGGTGTTGGTAGCATCA	CTCTTTTTTCTCTTCCCTGG	—
ICE10	F12F1	Microsat	b	b	—
ICE13	At1g13220	Microsat	b	b	#
F19G10	At1g23010	Microsat	b	b	##
GI	At1g22770	Snp	GCCAAAGATGATGAAGAA	TGTTGCTGGTAGACGAC	#
AthZFPG	At1g24625	Microsat	b	b	###
ATTSO392	At1g30630	Microsat	b	b	###
CXC750	At1g31580	*Mva*I	ATAATAGATAAAGAGCCCCACAC	CGCCAAAGACTACGAAATGATC	—
VIP1	At1g43700	*Acs*I	AATAGACAATCTGCGGCGAGGTC	GCCCGTTTGTACTCATCTGTTGGT	—
RHL1	At1g48380	*Sma*I	GGGACAAAAGAGGAGAATCCAG	CTTCATCAGATGTATTGCCTTCAC	—
			Linkage group AL2		
F19K23-483	At1g62050	Microsat	b	b	*
FT	At1g65480	*Ban*I	CAAGTCCTAGCAACCCTCACC	TCGCGAGTGTTGAAGTTCTGG	—
SLL2	At1g66680	*Eco*RI	CATGTACTGGGATTCAGTGTCC	CGTCCTTTGTGTGGTTACACG	—
nga111	F28P22	Microsat	c	c	##
ADH1	At1g77120	Duplex	ACCACCGGACAGATTATTCG	CCCAGAAGTAAACATCGGTGTG	##
			Linkage group AL3		
nga172	T21P5	Microsat	c	c	—
SPY	At3g11540	*Hae*III	TTGAGAGGCCTATGTATGCTGAA	AAAGCTTCTGCATATGTGGGATT	—
MDC16	MDC16	Microsat	b	b	—
DMC1	At3g22880	*Rsa*I	CTGAGGCCAAAGTTGACAA	GCCTGGAAAGCATCTGAG	—
ENR-A	At2g05990	Indel	GAAATATTGGTTGGGACTTGG	AGCTTTGGCAGAACTCATACCT	***
CAS1	At2g07050	*Rsa*I	ACCCACTGCCACCTGAG	GCGGTTCTTATAGCCTTTTCTC	—
PHYB	At2g18790	Snp	TACTCGAAGCTAAATGATCTCAA	TACAAAGGCTTCTTGATTCTTTA	—
			Linkage group AL4		
CON	At2g21320	*Taq*I	CCCTCTGTTGCTCCTGCGAC	GAAACCGCCTATGTGTTCTCTTCC	#
CK2α2	At2g23080	*Hinf*I	GAGTCTCTCGTTGTTCAATGGG	TTTGGAGTGCTCATCTCTAACAATA	##
ELF3	At2g25930	Microsat	CGGAAGGACTGATATACAAGC	TGTTGGGTGTTCTGAAGAT	—
ATEX6	At2g28950	*Hpa*II	GAATCGTCCCCGTCTCTTTCC	TTCCCATGAAAGTCTGACCAAAC	***
COP1	At2g32953	*Msp*I	a	a	*
ACT3	At2g37620	*Msp*I	ATGGCTGATGGTGAAGAC	CAGATGAGCTGGTTTTGG	—
ICE12	At2g39010	Microsat	b	b	—
FPA	At2g43410	*Dde*I	AATCTTTGGGTCGGTAGCCTAAC	AATGAAAGCTGTCTTGCGTTCTC	—
AGL20	At2g45660	Indel	ACAAGCAGACAAGTGACTTTCTC	TGAAGAACAAGGTAACCCAATG	—

(*continued*)

TABLE 1
(Continued)

Locus name	Identification/ BAC location in *A. thaliana*	Type of polymorphism	Forward primer	Reverse primer	Segregation distortion
			Linkage group AL5		
RIB	At2g01720	*Eco*RV	CAGTCCATTTGCTGTTGTTGAG	GCAAGCGAAGCCACAAATACTA	—
RNS1	At2g02990	*Taq*I Snp	TTTGATTTCTTCTACTTCGTCCAAC	GGAATTCGATCTCAGCTCCAC	*
PBB1	At3g27430	*Hae*III	CCGTAGATGTTCCGCCAAAG	CGAGTACAAGAGCAGCAGAGACA	**
ATA1	At3g42960	*Dde*I	GGCGAACTCAGACAAAAGATTATTC	TGTAACCACCATCAACAACCAAGT	*
ASN1	At3g47340	Snp	GGATGTTCCGATGATTCTCAGG	AAACTCCAAATGGAACATCAGTCA	—
CK2α3	At3g50000	*Sty*I *Sfu*I	CGATGTTAACGTGATCCGACC	TTGCAGTTAATCTGTCTTGATGAT	*
F3H	At3g51240	*Hin*dIII	a	a	—
TSA1	T14E10	*Hin*fI	TCTTGGTAGCATGATTCTCAGTC	CCTTTCCGCTTACAGATGATC	—
nga112	At3g62650	Microsat	b	b	—
			Linkage group AL6		
GA1	At4g02780	*Hin*fI	GGATCTCCTCTCAATGTCGC	ACATATGCCATTGTCGGTCC	***, ###
LD	At4g02560	*Pvu*II	CAGATGATTCGACCAAGAA	CATCATAACTCCCACCTTAC	—
ICE7	AT4g11410	Microsat	b	b	**, ##
AthDET1	At4g10180	Microsat	b	b	**, ##
CRY1	At4g08920	*Msp*I	TGTTGCTTCTCTTCTTGATGTTG	TTCCGCTGTTGAGTCTTCAGT	**, #
CLC-d	At5g26240	*Hpa*II	AAGGTCCTCTTGTGCATACAGGTG	CGCCACATAAGTTGACTTCTCCAC	#
AthCDPK9	MQM1	Microsat	b	b	**
ICE1	At5g16520	Microsat	b	b	—
CO	At5g15840	Snp	AGTTTGATGGTGTGGGAGATTG	TCACCTTGCCTAAGTTCAAAGC	*, #
nga249	At5g08550	Microsat	b	b	—
TFL1	At5g03840	*Alu*I	GGAGAATATGGGAACTAGAGTGAT	GTCTTTGCTTCTGCCTGAA	*
			Linkage group AL7		
DHS1	At4g39980	*Hae*III	GGCTGATGCTGCTGTTGGTAAAGC	GGCCTGATCTGCGGACTGCTCTGATCAGG	*, ###
AP2	At4g36920	Snp	GGACTGTGGGAAACAAGTTTAC	GCCGATGATTAATGAAATGACC	*, ##
PRHA	At4g29940	*Dde*I	GGACTAGGCGACGCCATTCGCCAACTG	CAGCTTTTGAGGGAAGCTCAGTCTCTGC	###
TSB2	At4g27070	*Hin*fI	TGTCCCCGAAACTCTAATGC	AAGTGGCTGTTCCAGAGTGG	#
CH42	At4g18480	*Hae*III	ATGGCGTCTCTTCTTGGAACATCTTCTTCTG	GCTGAAAATCTCGGCGAACTTCTCGG	—
VRN2	At4g16845	*Dde*I	AACATCTTTCACCTTCGCTCTCTA	GCTATCCCGGTCAGACATTACTT	—
FCA	At4g16280	*Rsa*I	CAATCAGATCACTCTTCCTG	TGAATATTTAACAAACCCACA	—
PHYC	At5g35840	*Taq*I Snp	a	a	##
ATTSO191	At5g37780	Microsat	b	b	##
ATPEN1	At5g42600	Indel	GAGCAATTAAGTTTCCACTG	TCATGGCCCTTGTAAAT	###
ICE3	MPO12	Microsat	b	b	##
			Linkage group AL8		
ATCLH2	At5g43860	*Dra*I	CTTCATCAAGAAACGCCTTTG	AATCCATAATCTCCGCCGTTG	—
EMF2	At5g51230	Snp	CAGGCGGTAAATGTCTCCCTCA	GCATGCCCATGGAATGTGAC	—
FRI	At4g00650	*Dde*I *Ksp*I	GGCGGCGCAATCCTCAAC	TACGAGGAGACCTGTGATTGG	—
MFB13	MFB13	Microsat	CCAATTTTGGTTTTCTACCTTAGGC	GTGATATCATGAACCAAAACCTACAAG	—
MHJ24	MHJ24	Microsat	b	b	*

The names and identification of the loci in *A. thaliana*, the type of the polymorphism scored (and if PCR-RFLP, the restriction enzyme used), and the primer sequences. Primer sequences described elsewhere are from Kuittinen *et al.* (2002) (a), Clauss *et al.* (2002) (b), or Bell and Ecker (1994) (c). Flowering-time-related loci are underlined. Segregation distortion in reciprocal 1: *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$; in reciprocal 2: #$P < 0.05$, ##$P < 0.01$, ###$P < 0.001$.

genotypic frequencies did not deviate from those expected). This suggests that the distortion is not due to segregation of a deleterious recessive allele that was fixed in the Mjällom population. The result is consistent with gametic selection or zygotic selection with additive gene effects. Because most markers were only partially informative, we know that the distortion occurred at least in the pollen parent but cannot exclude distortion in the mother. Differential pollen-tube growth has been suggested as a reason for much of the segregation distortion observed in, *e.g.*, *Mimulus guttatus* (Fishman *et al.* 2001). This kind of selection could be the reason for the segregation distortion in AL6, but would need additional data to be confirmed.

All the other distorted areas were specific for each reciprocal cross. The difference between reciprocals could result from interaction between cytoplasmic and nuclear factors, because the cytoplasms of the two F_1 plants originated from different populations. It is expected that mitochondrial and nuclear genes originating from the same population would interact positively. In several species mitochondrial genes are known to cause male sterility while nuclear genes may restore fertility (*e.g.*, Budar and Pelletier 2001). If the impairment of pollen function were induced by the mitochondrial genotype of the pollen parent, the Karhumäki allele should be favored in KM4-2 in reciprocal 1 and the Mjällom allele in KM6-5 in reciprocal 2. There was no evidence for this. Cytoplasmic interactions could also occur in other stages, as suggested by Tiffin *et al.* (2001). The two F_1 plants had different nuclear genotypes, and also polymorphic nuclear factors that influence pollen, style, or endosperm functions could potentially have caused incompatibility in the cross, resulting in different distortions in the reciprocals.

Several gametic and zygotic reproductive barriers were found in a cross between Japonica and Indica rice varieties (Harushima *et al.* 2001, 2002). In this cross, deviations were evident not only in gametic ratios but also in many cases from expected zygotic frequencies (TRDL1-1, TRDL2-1, TRDL2-2, and TRDL2-4), suggesting selection on the zygote.

Segregation distortion exhibited different patterns, so not only are many genes responsible, but also they possibly act through different mechanisms. Further crosses, as well as data on seed and pollen fertility, would be needed to distinguish between the contribution of cytoplasmic and nuclear factors and to find out in which stage the distortion occurs.

A practical concern of transmission ratio distortion is that it can interfere with mapping. Distortion of two loci in the same direction may result in spurious linkages between them or in a decrease in the distances between them, and loci distorted in different directions may inrease the genetic distance. Strong nonadditive interactions between linked markers can influence the map construction. Single transmission ratio distortion loci (Hackett and Broadfoot 2003) or several multiplicatively interacting transmission ratio distortion loci, however, should have little influence on a dense marker map when multipoint methods are used (C. Vogl and S. Xu, unpublished results). As the marker density in the map was reasonably high, it is unlikely that transmission ratio distortion has had a major effect on the linkage map. The data contained much missing information because for some markers more individuals were sampled than for others. Missing information can result in incorrect marker order especially at small intervals (Hackett and Broadfoot 2003). Potential nonadditively interacting linked viability loci and uneven numbers of genotyped markers could have an effect on the map construction. However, we have made rather conservative inferences on the marker order. The current *A. lyrata* map is likely to be quite robust, because in a cross between two other *A. lyrata* populations—with a smaller number of markers, with practically no missing individuals, and with extensive transmission ratio distortion at different locations compared to this map—the same order and similar distances between the common markers were obtained (J. Leppälä and O. Savolainen, unpublished results).

Comparative mapping as a tool: The almost complete colinearity of the *A. lyrata* map with the map of the model species *A. thaliana* implies that we can use the positional information from the Arabidopsis genome. For instance, after QTL mapping in *A. lyrata*, it will be possible to select candidate genes from the corresponding region in *A. thaliana* for further analysis, such as complementation or association studies. The current cross will be useful for QTL mapping as such. A set of genotyped F_2 plants has been micropropagated and can be used for evaluation of the desired phenotypes. The markers can also be used in other crosses. Many of the amplification primers are likely to be applicable in close species and can be used for comparative mapping. Additional markers can easily be added to the current map to increase the density of the map or to obtain information about the location of a specific gene.

We are grateful to Bryan Lazzaro for his contribution to the lab work and to Patrik Waldmann for comments on the manuscript. This work was supported by the Finnish Research Council for Biosciences and Environment.

LITERATURE CITED

Acarkan, A., M. Rossberg, M. Koch and R. Schmidt, 2000 Comparative genome analysis reveals extensive conservation of genome organization for *Arabidopsis thaliana* and *Capsella rubella*. Plant J. **23:** 55–62.

Arabidopsis Genome Initiative, 2000 Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature **408:** 796–815.

Bell, C., and J. Ecker, 1994 Assignment of 30 microsatellite loci to the linkage map of *Arabidopsis*. Genomics **19:** 137–144.

Budar, F., and G. Pelletier, 2001 Male sterility in plants: occurrence, determinism, significance and use. C. R. Acad. Sci. III **324:** 543–550.

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

Clauss, M., H. Cobban and T. Mitchell-Olds, 2003 Population genetics of tandem trypsin inhibitor genes in *Arabidopsis* species with contrasting ecology and life history. Mol. Ecol. **12:** 1287–1299.

Devos, K. M., J. Beales, Y. Nagamura and T. Sasaki, 1999 *Arabidopsis*-rice: Will colinearity allow gene prediction across the eudicot-monocot divide? Genome Res. **9:** 825–829.

Fishman, L., A. J. Kelly, E. Morgan and J. H. Willis, 2001 A genetic map in the *Mimulus guttatus* species complex reveals transmission ratio distortion due to heterospecific interactions. Genetics **159:** 1701–1716.

Fu, H., and H. Dooner, 2002 Intraspecific violation of genetic colinearity and its implications in maize. Proc. Natl. Acad. Sci. USA **99:** 9573–9578.

Gebhardt, C., B. Walkemeier, H. Henselewski, A. Barakat, M. Delseny *et al.*, 2003 Comparative mapping between potato (*Solanum tuberosum*) and *Arabidopsis thaliana* reveals structurally conserved domains and ancient duplications in the potato genome. Plant J. **34:** 529–541.

Hackett, C. A., and L. B. Broadfoot, 2003 Effects of genotyping errors, missing values and segregation distortion in molecular marker data on the construction of linkage maps. Heredity **90:** 33–38.

Harushima, Y., M. Nakagahra, M. Yano, T. Sasaki and N. Kurata, 2001 A genome-wide survey of reproductive barriers in an intraspecific hybrid. Genetics **159:** 883–892.

Harushima, Y., M. Nakagahra, M. Yano, T. Sasaki and N. Kurata, 2002 Diverse variation of reproductive barriers in three intraspecific rice crosses. Genetics **160:** 313–322.

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

Jones, B. M. G., 1963 Experimental taxonomy of the genus *Arabis*. Ph.D. Thesis, University of Leicester, Leicester, UK.

Jonsell, B., K. Kustås and I. Nordal, 1995 Genetic variation in *Arabis petraea*, a disjunct species in northern Europe. Ecography **18:** 321–332.

Kärkkäinen, K., H. Kuittinen, R. van Treuren, C. Vogl, S. Oikarinen *et al.*, 1999 Genetic basis of inbreeding depression in *Arabis petraea*. Evolution **53:** 1354–1365.

Koch, M., J. Bishop and T. Mitchell-Olds, 1999 Molecular systematics and evolution of *Arabidopsis* and *Arabis*. Plant Biol. **1:** 529–537.

Koch, M., B. Haubold and T. Mitchell-Olds, 2000 Comparative evolutionary analysis of *chalcone synthase* and *alcohol dehydrogenase* loci in *Arabidopsis*, *Arabis* and related genera (Brassicaceae). Mol. Biol. Evol. **17:** 1483–1498.

Korbecka, G., P. G. L. Klinkhamer and K. Vrieling, 2002 Selective embryo abortion hypothesis revisited—a molecular approach. Plant Biol. **4:** 298–310.

Kowalski, S. P., T.-H. Lan, K. A. Feldman and A. H. Paterson, 1994 Comparative mapping of *Arabidopsis thaliana* and *Brassica oleracea* chromosomes reveals islands of conserved organization. Genetics **138:** 499–510.

Ku, H.-M., T. Vision, J. Liu and S. D. Tanksley, 2000 Comparing sequenced segments of the tomato and *Arabidopsis* genomes: large-scale duplication followed by selective gene loss creates a network of synteny. Proc. Natl. Acad. Sci. USA **97:** 9121–9126.

Kuittinen, H., M. Aguadé, D. Charlesworth, A. De Haan, B. Lauga *et al.*, 2002 Primers for 22 candidate genes for ecological adaptations in Brassicaceae. Mol. Ecol. Notes **2:** 258–262.

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.

Lagercrantz, U., and D. Lydiate, 1996 Comparative genome mapping in Brassica. Genetics **144:** 1903–1910.

Lan, T.-H., T. A. Delmonte, K. P. Reischmann, J. Hyman, S. P. Kowalski *et al.*, 2000 An EST-enriched comparative map of *Brassica oleracea* and *Arabidopsis thaliana*. Genome Res. **10:** 776–788.

Li, W., and B. S. Gill, 2002 The colinearity of the Sh2/A1 orthologous region in rice, sorghum and maize is interrupted and accompanied by genome expansion in the Triticaceae. Genetics **160:** 1153–1162.

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

Lukens, L., F. Zou, D. Lydiate, I. Parkin and T. Osborn, 2003 Comparison of a *Brassica oleracea* genetic map with the genome of *Arabidopsis thaliana*. Genetics **164:** 359–372.

Měsíček, J., 1967 The chromosome morphology of *Arabidopsis thaliana* (L.) Heynh. and some remarks on the problem of *Hylandra suecica* (Fr.) Love. Folia Geobotanica Phytotaxonomica **2:** 433–436.

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

Nasrallah, M. E., K. Yogeeswaran, S. Snyder and J. B. Nasrallah, 2000 *Arabidopsis* species hybrids in the study of species differences and evolution of amphiploidy in plants. Plant Physiol. **124:** 1605–1614.

Orr, A., and M. Turelli, 2001 The evolution of postzygotic isolation: accumulating Dobzhansky-Muller incompatibilities. Evolution **55:** 1085–1094.

Quiros, C., F. Grellet, J. Sadowski, T. Suzuki, G. Li *et al.*, 2001 Arabidopsis and Brassica comparative genomics: sequence, structure and gene content in the *ABI1-Rps2-Ck1* chromosomal segment and related regions. Genetics **157:** 1321–1330.

Rossberg, M., K. Theres, A. Acarkan, R. Herrero, T. Schmitt *et al.*, 2001 Comparative sequence analysis reveals extensive microcolinearity in the *lateral suppressor* regions of the tomato, *Arabidopsis*, and *Capsella* genomes. Plant Cell **13:** 979–988.

Savolainen, O., C. H. Langley, B. P. Lazzaro and H. Fréville, 2000 Contrasting patterns of nucleotide polymorphism at the *alcohol dehydrogenase* locus in the outcrossing *Arabidopsis lyrata* and the selfing *Arabidopsis thaliana*. Mol. Biol. Evol. **17:** 645–655.

Schierup, M. H., 1998 The effect of enzyme heterozygosity on growth in a strictly outcrossing species, the self-incompatible *Arabis petraea* (Brassicaceae). Hereditas **128:** 21–31.

Schmidt, R., 2002 Plant genome evolution: lessons from comparative genomics at the DNA level. Plant Mol. Biol. **48:** 21–37.

Simillion, C., K. Vandepoele, M. C. Van Montagu, M. Zabeau and Y. Van De Peer, 2002 The hidden duplication past of *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. USA **99:** 13627–13632.

Stam, P., 1993 Construction of integrated genetic linkage maps by means of a new computer package: JoinMap. Plant J. **3:** 739–744.

Tiffin, P. M. S. Olson and L. C. Moyle, 2001 Asymmetrical crossing barriers in angiosperms. Proc. R. Soc. Lond. Ser. B **268:** 861–867.

van Treuren, R., H. Kuittinen, K. Kärkkäinen, E. Baena-Gonzalez and O. Savolainen, 1997 Evolution of microsatellites in *Arabis petraea* and *Arabis lyrata*, outcrossing relatives of *Arabidopsis thaliana*. Mol. Biol. Evol. **14:** 220–229.

Vision, T. J., D. G. Brown and S. D. Tanksley, 2000 The origins of genomic duplications in *Arabidopsis*. Science **290:** 2114–2117.

Vogl, C., and S. Xu, 2000 Multipoint mapping of viability and segregation distorting loci using molecular markers. Genetics **155:** 1439–1447.

Whitkus, R., 1998 Genetics of adaptive radiation in Hawaiian and Cook Islands species of Tetramolopium (Asteraceae). II. Genetic linkage map and its implications for interspecific breeding barriers. Genetics **150:** 1209–1216.

Wright, S., B. Lauga and D. Charlesworth, 2003 Rates and patterns of molecular evolution in inbred and outbred *Arabidopsis*. Mol. Biol. Evol. **19:** 1407–1420.

Yang, Y.-W., K.-N. Lai, P.-Y. Tai and W.-H. Li, 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.

Communicating editor: D. Charlesworth