

Centromere Locations and Associated Chromosome Rearrangements in *Arabidopsis lyrata* and *A. thaliana*

Akira Kawabe, Bengt Hansson,¹ Jenny Hagenblad,² Alan Forrest and Deborah Charlesworth³

Institute of Evolutionary Biology, University of Edinburgh, Edinburgh EH9 3JT, United Kingdom

Manuscript received February 15, 2006

Accepted for publication April 25, 2006

ABSTRACT

We analyzed linkage and chromosomal positions of genes in *A. lyrata* ssp. *petraea* that are located near the centromere (CEN) regions of *A. thaliana*, using at least two genes from the short and long arms of each chromosome. In our map, genes from all 10 *A. thaliana* chromosome arms are also tightly linked in *A. lyrata*. Genes from the regions on the two sides of CEN5 have distant map localizations in *A. lyrata* (genes on the *A. thaliana* short-arm genes are on linkage group AL6, and long-arm genes are on AL7), but genes from the other four *A. thaliana* centromere regions remain closely linked in *A. lyrata*. The observation of complete linkage between short- and long-arm centromere genes, but not between genes in other genome regions that are separated by similar physical distances, suggests that crossing-over frequencies near the *A. lyrata* ssp. *petraea* centromere regions are low, as in *A. thaliana*. Thus, the centromere positions appear to be conserved between *A. thaliana* and *A. lyrata*, even though three centromeres have been lost in *A. thaliana*, and the core satellite sequences in the two species are very different. We can now definitively identify the three centromeres that were eliminated in the fusions that formed the *A. thaliana* chromosomes. However, we cannot tell whether genes were lost along with these centromeres, because such genes are absent from the *A. thaliana* genome, which is the sole source of markers for our mapping.

THE chromosomes of *Arabidopsis thaliana* have undergone several rearrangements since this species' common ancestor with related species. This is evident from the fact that *A. thaliana* has only five chromosomes *vs.* eight in outgroup species such as *A. lyrata* and *Capsella rubella*. Comparative genetic mapping (BOIVIN *et al.* 2004; KUITTINEN *et al.* 2004; KOCH and KIEFER 2005; YOGESWARAN *et al.* 2005) and molecular cytological analyses (LYSAK *et al.* 2003) has revealed the major structural changes that occurred in the *A. thaliana* lineage (summarized in Figure 1).

As far as can be ascertained from the low-density genetic maps published so far, the changes during the evolution of the *A. thaliana* chromosomes are as follows (the order of events is unknown). There were two reciprocal translocations: one between the *A. lyrata* chromosomes AL3 and AL5 (corresponding to *C. rubella* chromosomes C and E, respectively), forming *A. thaliana* chromosome III and part of chromosome II, and the other between AL6 and AL7 (*C. rubella* chromosomes F and G, respectively), forming *A. thaliana* chromosome IV and part of chromosome V. In addition, three fusion events occurred: one between AL1 and

AL2 (*C. rubella* chromosomes A and B, respectively), forming *A. thaliana* chromosome I, and two others involving the chromosomes with the reciprocal translocations. One of these fused AL4 (*C. rubella* chromosome D) to a chromosome resulting from the first translocation, forming the present *A. thaliana* chromosome II, and the other fused AL8 (*C. rubella* chromosome H) to a chromosome resulting from the second one, forming *A. thaliana* chromosome V. Two large inversions are also inferred: in the present *A. thaliana* chromosome I (inversion 1 of KOCH and KIEFER 2005) and in chromosome IV (inversion 2 of KOCH and KIEFER 2005) at one end of the part translocated from AL6. For three further inversions, the lineages in which they occurred cannot currently be determined, because markers are sparse in the relevant regions in either *C. rubella* or *A. lyrata*. Inversion 3, near the breakpoint of the fusion with AL8, has been detected only in *C. rubella*, and inversions 4 and 8 have been detected only in *A. lyrata* at the end of the translocated parts of AL7 and AL6, respectively (parts of *A. thaliana* chromosomes V and IV, respectively). To resolve the arrangements in these species and determine definitively the lineages in which they occurred, more markers were needed for mapping of these regions, or, alternatively, *in situ* hybridization studies.

An especially interesting question that can be studied by more detailed mapping is how these rearrangements affected the centromeres. How did the current *A. thaliana*

¹Present address: Department of Animal Ecology, Lund University, Ecology Bldg., SE-223 62 Lund, Sweden.

²Present address: Department of Physics, Chemistry and Biology, Linköping University, SE-581 83 Linköping, Sweden.

³Corresponding author: University of Edinburgh, Ashworth Lab, King's Bldgs., W. Mains Rd., Edinburgh EH9 3JT, United Kingdom.

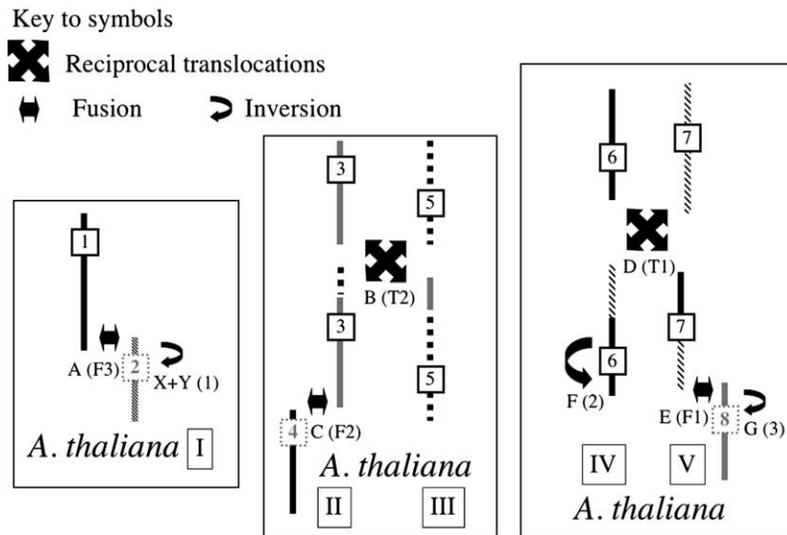


FIGURE 1.—Chromosome rearrangements in the history of *A. thaliana*. The *A. thaliana* chromosome numbers are shown at the bottom. The diagram shows the chromosomes in the ancestor, assuming that this was similar to the state in *A. lyrata* and *C. rubella*. The chromosomes are not drawn to scale. The centromeres of the eight ancestral chromosomes are shown as numbered squares, and the arms of the different ancestral chromosomes are shown as different types of lines; the three centromeres lost in the formation of the *A. thaliana* chromosomes are shaded, while the surviving centromeres are solid. Fusions and reciprocal translocations are noted, and inversions are denoted by curved arrows and numbered according to KOCH and KIEFER (2005). The rearrangement events are labeled with the letters used by YOGESWARAN *et al.* (2005) and the numbers used by KOCH and KIEFER (2005); in the latter case, T was added to the reciprocal translocation numbers, and F for fusions, to distinguish them from inversions with the same numbers.

centromeres originate, and are the centromere positions conserved? If not, this is due to neo-centromere formation, as in examples such as the primate X chromosome (VENTURA *et al.* 2001). This possibility is particularly interesting in these *Arabidopsis* species, because the centromere satellite families of *A. arenosa*, *A. halleri*, and *A. lyrata*, which are the closest relatives of *A. thaliana*, are very different from the *A. thaliana* centromere satellite sequences (KAMM *et al.* 1995; KAWABE and NASUDA 2005). It is thus important to test whether emergence of new satellite families involved formation of new centromeres. The alternative is that the new sequences emerged and replaced an existing centromere's satellites and then became homogenized between different centromere regions in these species' genomes.

The evidence from the sparse maps already published evidently suggests that centromeres 2, 4, and 8 were the three centromeres that became lost during the evolution of the *A. thaliana* genome (in the fusions forming chromosomes I, II, and V; see Figure 1). These conclusions rest on the assumption that no large-scale centromere displacement or chromosome rearrangements occurred, which has not previously been tested. For *A. thaliana* chromosomes I and II, genes from both arms map to single *A. lyrata* chromosomes, AL1 and AL3, respectively (KUITTINEN *et al.* 2004), consistent with the above assumption, but the situation for chromosome V is not evident.

To test these assumptions explicitly, and to delimit the scale of any possible centromere displacements, we have attempted to identify the *A. lyrata* centromere locations by mapping more finely spaced markers in genes flanking the *A. thaliana* centromeres of each of the five *A. thaliana* centromere regions. This should also allow us to infer the centromere positions in at least some of

the *A. lyrata* chromosomes and to develop reconstructions of the details of the rearrangement events that can be tested in future work (see Figure 1, which we will discuss at greater length below).

Because we start from known *A. thaliana* genes, we cannot map any genes lost in events such as the hypothetical loss of a centromere (see Figure 1). This means that we are unlikely to be able to discover the precise centromere locations of the three *A. lyrata* chromosomes involved in fusions. However, fusion events are expected to cause loss of a small numbers of genes, since large deletions are deleterious (KHUSH and RICK 1968). Moreover, there is no evidence in *A. thaliana* for non-duplicated genome tracts that might correspond to regions that were duplicates in the ancestral genome, but recently became single copy (BLANC *et al.* 2003). Thus the fusions often may have been between chromosomes with centromeres close to their ends. This implies either that the chromosomes in the ancestor in which the fusions occurred were telocentric or acrocentric or that inversions occurred before or during the fusion events to create this situation (see Figure 1). Because no *A. lyrata* or *C. rubella* chromosomes are telocentric (ALI *et al.* 2005), the occurrence of inversions seems likely.

MATERIALS AND METHODS

Genes analyzed: We chose two genes from each chromosome arm of *A. thaliana*, in addition to nine markers on chromosome I analyzed previously (HANSSON *et al.* 2006). From each chromosome arm, we chose one single-copy gene located as close as possible to the core centromere region, and one gene each at least 100 kb distal to the first gene (KUMEKAWA *et al.* 2000, 2001; HOSOUCHI *et al.* 2002). To ensure single-copy genes, we checked that there are no homologs with sequence identity >80% in the *A. thaliana* genome. The genes, with their

putative functions, are listed in supplemental Table 1 at <http://www.genetics.org/supplemental/>, which also gives the sequences of the primers used; these were designed on the basis of *A. thaliana* sequences.

Typing and data analyses: A mapping family of 99 F₂ progeny plants made from crosses between plants from two different subpopulations of *A. lyrata* ssp. *petraea* (KUITTINEN *et al.* 2004) was used in the study. When variants heterozygous in only one parent can be scored, we can score only 99 meioses, but when scorable variants are heterozygous in both parents, 198 meioses were scored.

To detect polymorphisms and develop PCR-based markers in the genes selected, sequences of one parent F₁ plant and three to seven F₂ progeny plants were determined for each gene (supplemental Table 1 at <http://www.genetics.org/supplemental/>). For eight loci, single-nucleotide polymorphisms (SNPs) were typed by PCR-RFLP. Large-indel variants distinguishable on agarose gels were used to genotype one locus, small-indel variants that could be scored in an ABI 3730 capillary sequencer with fluorescent-labeled primers for three loci, and six loci with SNPs were scored with allele-specific primer-induced fragment length polymorphism, also detected by an ABI 3730 capillary sequencer (HANSSON and KAWABE 2005). Including the nine markers close to the centromere of the *A. thaliana* chromosome I, we typed a total of 27 markers (from 24 genes). Details of the typing methods are given in HANSSON *et al.* (2006).

To infer the history of genes found to be duplicates in *A. lyrata*, we used an outgroup species (*Arabis glabra* = *Turritis glabra*). Sequences of the duplicated genes from the *A. lyrata* ssp. *petraea* mapping family and *T. glabra* were deposited in GenBank. The accession numbers of the new sequences reported and analyzed here are as follows: DQ487163–DQ487167 for At1g43980, DQ487168–DQ487175 for At3g42050, and DQ487176–DQ487178 for At4g21150. The genes from *A. thaliana* centromeric regions are DQ487163–DQ487175.

RESULTS

Conserved centromere positions in *A. thaliana* and *A. lyrata*:

Our markers in genes located near the centromere regions of the 10 chromosome arms of *A. thaliana* in all cases are linked in *A. lyrata* (Figure 2). These results support the previous results indicating the conservation of linkage groups representing most of the chromosome arms between these two species (KUITTINEN *et al.* 2004), but go beyond those results in that they show that even genes very closely linked to the centromere in *A. thaliana* are in the same locations in the *A. lyrata* map. They are also consistent with the conclusion that the *A. thaliana* CEN3 and CEN5 are in the same location as in *C. rubella*, *A. arenosa*, and *Olimarabidosis pumila* (HALL *et al.* 2006). Below, we discuss our evidence that these regions contain the centromeres of these *A. lyrata* chromosomes. For four *A. thaliana* chromosomes, we also find linkage between short- and long-arm genes in the *A. lyrata* map. These centromere regions (*A. thaliana* CEN1 and the putative centromere of AL1, *A. thaliana* CEN2 and the putative AL3 centromere, *A. thaliana* CEN3 and AL5, and *A. thaliana* CEN4 and AL6) are thus conserved between *A. thaliana* and *A. lyrata*, with no evidence of chromosome rearrangements in the immediately surrounding regions (see Figure 2).

CEN5 is more complex. The formation of chromosome V of *A. thaliana* involved a reciprocal translocation (see above and Figure 1), and an understanding of this event is critical for interpreting the origin of CEN5. Previous comparative molecular cytogenetic analyses of *A. thaliana* and *A. lyrata* could not precisely locate the breakpoint of the reciprocal translocation creating the *A. thaliana* chromosomes IV and V (between AL6 and AL7). Chromosome IV BAC probes were used in FISH experiments (LYSAK *et al.* 2003), but no BAC clones containing large amounts of repetitive elements were used, so that the details of event regions close to this centromere may not be resolved. It was concluded that the translocation event probably occurred near the centromere region because the transposition breakpoint in the chromosome corresponding to AL7 in our terminology is close to the pericentromere heterochromatic region (Figure 3E of LYSAK *et al.* 2003). In our *A. lyrata* genetic map, genes from the short arm of the *A. thaliana* chromosome V are completely linked to genes from the *A. thaliana* CEN4 region (Figure 2), suggesting that one breakpoint of the translocation lies between the markers closest to the core centromere in the two arms of this chromosome. Assuming that the centromere position is conserved between *A. thaliana* chromosome IV and AL6, as our mapping results indicate (Figure 2 and KAWABE *et al.* 2006), the breakpoint is probably in the pericentromeric region of the ancestral species, although its position still cannot be determined precisely (Figure 2).

We examined our results to see whether any gene copy number differences are suggested, because it has been found that the centromeric regions of *A. thaliana* have significantly fewer tandemly duplicated genes than noncentromere regions (ZHANG and GAUT 2003). We detected four events in total and estimated duplication time using the *T. glabra* sequence as an outgroup (HANSSON *et al.* 2006; KAWABE *et al.* 2006), two involving centromere region genes. In one such case, on chromosome I, At1g43980, the results suggest a recent duplication in the *A. lyrata* lineage, rather than loss in *A. thaliana* (HANSSON *et al.* 2006). The other is the apparent triplication in *A. lyrata* of At3g42050, but the events involved appear complex and cannot be inferred in detail. Given that a much smaller number of centromeric than noncentromeric genes have been studied, these results suggest *more* such events in the centromere regions, which is inconsistent with the results based on an initial annotation of the complete genome sequence (ZHANG and GAUT 2003). That study did not include comparisons with related species, so that the ancestral states are not known, leaving it unclear whether the difference is due to loss of tandem duplicates in *A. thaliana* centromeric regions or to the occurrence of more such duplications in noncentromeric than centromeric regions. However, neither of the above possibilities explains our results. Instead, our data suggest that both

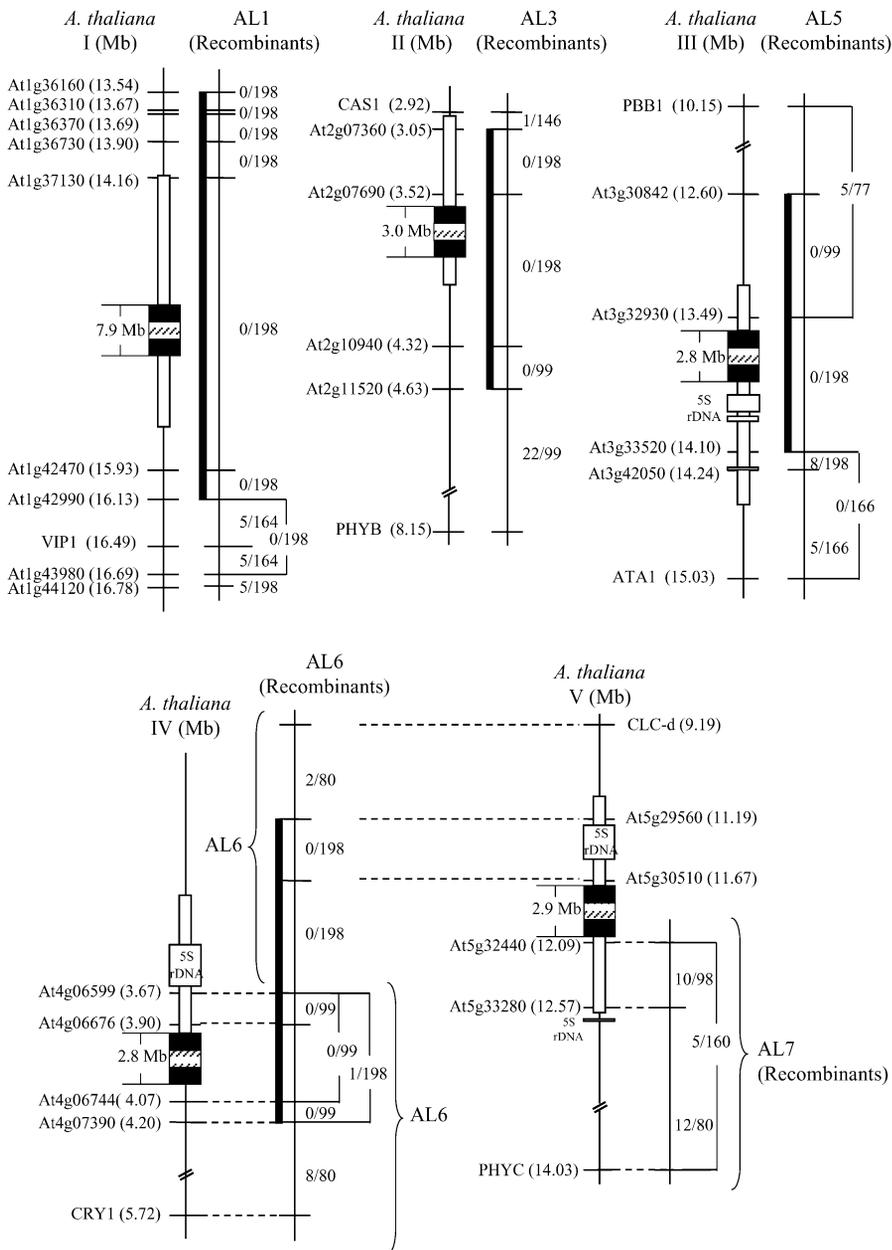


FIGURE 2.—Map positions of the centromere-located genes of each chromosome, showing the core centromere regions and their physical sizes in *A. thaliana*, the markers that we mapped in *A. lyrata*, and the distances in megabases between the genes in *A. thaliana*. To the right of each chromosome for *A. lyrata* are the numbers of recombinants observed in our mapping family. The regions where we detected no recombination are shown as solid lines at the left of AL chromosomes.

of these cases are tandem duplications in *A. lyrata* since the split from the *A. thaliana* lineage (KAWABE *et al.* 2006).

Frequencies of recombination between markers: Many organisms, including *A. thaliana*, have very low crossing-over frequencies in the regions around the “core” centromere regions. In *A. thaliana*, this is true for all five chromosomes (COPENHAVER *et al.* 1998, 1999). The markers used in this study are not always located within these regions, but are nevertheless closely linked to the *A. thaliana* centromeres. To test whether our putative *A. lyrata* centromeres are indeed centromeres, we tested whether markers from *A. thaliana* centromere region genes show evidence for restricted recombination in *A. lyrata* and then compared the estimated recombination rates with expected values on the basis of prob-

able physical distances (in *A. thaliana*) between the markers.

The physical distance corresponding to a centimorgan is not yet known in *A. lyrata*, but our genetic mapping results for *A. lyrata* chromosomes 1, 2, and 7 yield some rough estimates of the physical length per centimorgan for noncentromere regions of *A. lyrata* chromosomes. For chromosome 1, assuming the same physical size as in *A. thaliana*, we estimate a value of ~250 kb/cM (4.0 cM/Mb) and ~100 kb/cM (8–9 cM/Mb) for the *A. lyrata* chromosome 2; if we assume that the DNA content of each of these chromosome arms is 50% larger than the *A. thaliana* value, the ratio as estimated for the entire genome, these values are increased (HANSSON *et al.* 2006). Our estimate from AL7, which, unlike AL6 (which includes an inversion), is not

rearranged, yields an estimate of ~ 200 kb/cM (or 4–5 cM/Mb), assuming the same physical size as in *A. thaliana*.

If recombination is not suppressed in the region covered by our centromere region markers, we should therefore observe recombinants across the distances between these markers, which, even in *A. thaliana*, are several megabases (Figure 2). The *A. thaliana* core centromere region consists of repetitive sequences, including short satellite sequences and transposable elements, and always exceeds 2 Mb (HOSOUCHI *et al.* 2002). Given its larger genome size (JOHNSTON *et al.* 2005), at least similar amounts of DNAs may be expected at *A. lyrata* centromeres. If the core satellite regions are at the same locations in *A. thaliana* and *A. lyrata*, the physical distances between our markers should thus probably be >3 Mb and should correspond to several map units (Figure 2). With our family size, we expect to find recombinants between markers at distances of a few centimorgans, and we indeed almost always detect recombinants between markers from noncentromere regions of the *A. thaliana* genome across comparable physical distances (HANSSON *et al.* 2006). Therefore we should be able to detect the presence of nonrecombining regions in the *A. lyrata* chromosomes, if these indeed exist near the putative centromere regions identified by our markers, and if they indeed are regions of low recombination.

We scored multiple markers from each *A. thaliana* centromere and found that almost all the centromere markers are very closely linked in *A. lyrata* (Figure 2), except for the lack of linkage of genes from the two arms of the chromosome V centromere region (see above). All four markers from the *A. thaliana* chromosome II centromere region are completely linked in our family, although we observed occasional recombinants between the most distally located pairs of genes on the other chromosomes (Figure 2). For the two genes on the long arm of chromosome V, we observed nine recombinants in the 99 mapping family progeny, while the short-arm genes are completely linked to chromosome IV centromere region genes. Thus these centromeres have apparently remained in the same regions, flanked by the same set of closest loci, in both species.

DISCUSSION

Conserved centromere positions in *A. thaliana* and *A. lyrata*: The genes from each *A. thaliana* centromere region showed almost complete linkage in *A. lyrata*, suggesting both that the centromere locations have not changed in the evolution of the *A. thaliana* chromosomes and that there are regions with very low crossing-over frequencies around the *A. lyrata* centromere regions. Our results do not establish definitively that the centromere locations are completely unchanged in

A. thaliana, *i.e.*, that long- and short-arm genes are also on different sides of *A. lyrata* centromere satellite arrays, but this seems likely. Our observation that linkages of short- and long-arm genes are maintained between *A. thaliana* and *A. lyrata* narrows down the possibility for centromere movements to within the region spanned by the nonrecombining genes, whose order cannot be determined by genetic mapping alone. This conclusion is consistent with the finding that the *A. arenosa* BAC clone containing genes orthologous to those close to the *A. thaliana* CEN3 also contains *A. arenosa* centromeric satellite sequences (HALL *et al.* 2006).

A. lyrata and *A. thaliana* have different centromere satellite sequences in their core centromere regions (KAWABE and NASUDA 2005). The conclusion that the centromeres are in all cases at least close to the same map locations suggests that the loss of old satellite sequences and the increase of new ones probably occurred within established centromeres, *i.e.*, that the satellite sequences changed without centromere movement or rearrangement of the surrounding chromosome structure, even when the orientation of some short- or long-arm genes differed between the species. In turn, this implies that homogenization of centromere satellite sequences throughout the entire set of chromosomes must have involved interchromosomal transposition by some form of sequence exchange between nonhomologous chromosomes, such as gene conversion, but does not require reciprocal translocations involving the core centromere regions. In some primates, different centromere sequences have also been observed in species without changes in chromosome structure, although small sequential pericentromeric inversion events followed by emergence of neo-centromere sequences at the initial position cannot be excluded (VENTURA *et al.* 2001).

Are chromosome rearrangements associated with centromere regions? Our new map results suggest the occurrence of one reciprocal translocation between the centromere regions of the two chromosomes AL6 and 7 without loss of genes. The chromosome fusion events that reduced the chromosome number from eight to five in *A. thaliana* involved loss of three centromeres, possibly involving loss of large amounts of the genome. This could partially account for *A. thaliana*'s genome size being considerably smaller than *A. lyrata*'s (JOHNSTON *et al.* 2005). If the centromere regions in the ancestral species contained large amounts of repetitive sequences, it is possible that these regions might have contained few genes and could be lost. However, some evidence suggests that pervasive small changes have also occurred throughout the genome. Intron sizes tend to be smaller than in *A. lyrata* orthologs on the basis of a sample of 19 genes (WRIGHT *et al.* 2002), and a survey suggests consistently larger intergenic physical distances in *A. lyrata* in the FLC gene region (SANYAL and JACKSON 2005).

There are two possibilities for chromosome fusion and centromere loss without loss of many genes. A simple Robertsonian translocation could cause loss of one small chromosome arm along with the chromosome's centromere region. However, the inferred morphology in the ancestral species of the chromosomes involved in fusions does not suggest that they were telocentric. Alternatively, an inversion could occur before or together with the chromosome fusion, preserving most of the chromosome arm. Interestingly, two of the three chromosomes that are inferred to have lost their centromere regions in fusion events, AL2 and AL8, do indeed have large inversions near the locations of the fusions in the *A. thaliana* chromosomes (KOCH and KIEFER 2005). We find that inversion 1 (in AL2) is present in *A. lyrata* (HANSSON *et al.* 2006), as well as in *C. rubella* (Boivin *et al.* 2004). Thus this inversion should be attributed to the *A. thaliana* lineage. An inversion has also been detected cytologically in the relevant part of AL4, another chromosome involved in a fusion event, but whose centromere is not at the distal end of the chromosome (LYSAK *et al.* 2006; note that this is not shown in Figure 1, which reviews previously established rearrangements by genetic mapping).

Overall, it thus appears likely that of the eight major chromosome rearrangements previously inferred, seven occurred at or near the centromere regions inferred in the ancestral species. The rearrangement confirmed in this study (a translocation between AL6 and AL7 and an inversion in AL6) probably involved pericentromeric regions. The reciprocal translocation between AL3 and AL5 is thus the only rearrangement that did not involve a centromere region. The involvement of centromere regions in the rearrangements may be due to the presence of repetitive sequences in such regions, which can induce ectopic exchanges leading to chromosome breaks; thus independent events might occur in the same region at separate times. Pericentromeric regions of chromosomes contain abundant transposable elements and some of these retain their mobility (FRANK *et al.* 1997; MIURA *et al.* 2001; SINGER *et al.* 2001). Activation of transposable elements, especially the cut-and-paste type, can cause chromosome breaks and create sticky ends, which can lead to rearrangements.

Alternatively, some of these events may have occurred simultaneously as single complex events sharing the same breakpoints. The numbers of events previously proposed would then be reduced from eight to a minimum of five: fusion/breakage 1 and inversion 3 in *A. thaliana* chromosome V (events E and G in YOGESWARAN *et al.* 2005), fusion/breakage 3 and inversion 1 in *A. thaliana* chromosome I (events A, X, and Y in YOGESWARAN *et al.* 2005), and translocation 1 and inversion 2 in *A. thaliana* chromosomes IV and V (events D and F in YOGESWARAN *et al.* 2005).

In other Arabidopsis relatives (LYSAK *et al.* 2003), however, chromosome structure is highly conserved, even between *C. rubella* and *A. lyrata*. These species di-

verged about twice as long ago as *A. lyrata* and *A. thaliana* (KOCH *et al.* 2000), suggesting that chromosome rearrangement is considerably accelerated in the *A. thaliana* lineage. This is consistent with theoretical models predicting a tendency toward higher chromosome rearrangement rates in inbreeding than in outcrossing species and observations of this tendency in plant karyotypes (LANDE 1984; CHARLESWORTH 1992). The *C. rubella* map, however, provides an exception. This map was made using the self-compatible annual *C. rubella* crossed with the self-incompatible perennial, *C. grandiflora*, and shows no evidence for chromosome rearrangements having accumulated in the former. The divergence time between these taxa has not yet been estimated and could be too recent for chromosome rearrangements to have occurred.

With our probable interpretations of the rearrangements in the genetic map in *A. thaliana*, we can suggest possible locations of the *A. lyrata* centromeres that cannot be mapped, which can be tested in future work. No centromeres were apparently lost in the formation of *A. thaliana* chromosome IV by a reciprocal translocation between *A. lyrata* chromosomes 6 and 7. In the cases where centromeres were lost in fusions, it is likely that, in the ancestral species, they were located close to the inversion breakpoints (on the basis of the consideration that deletions involving many genes are unlikely). Since the *A. lyrata* map is similar to that of the outgroup species, *C. rubella*, the ancestral state for each chromosome was probably similar to that of *A. lyrata*. In the case of *A. thaliana* chromosome I, the centromere of *A. lyrata* chromosome 2 was thus lost, and its location is evidently within the large gap in the current genetic map (HANSSON *et al.* 2006). It will be interesting in the future to map *A. lyrata* chromosomes 4 and 8 more precisely, to determine whether gaps are also present on these chromosomes in the regions that were lost in forming *A. thaliana* chromosomes II and V. It may be informative to search for *A. lyrata* genes that are absent in *A. thaliana* (in EST sequences or in the forthcoming genome sequence) and to map them. If this eliminates the gaps, it would imply that many genes were lost in these fusions.

We thank O. Savolainen (University of Oulu, Finland) for DNA samples from the *A. lyrata* mapping family, H. Kuitinen for help with linkage analyses, and M. Lysak for discussions of unpublished FISH results. Support of a postdoctoral grant from the Swedish Research Council and Swedish Foundation for International Cooperation in Research and Higher Education (to B.H.) and a small research grant from the National Environmental Research Council of the United Kingdom (to D.C.) is gratefully acknowledged.

LITERATURE CITED

- ALI, H. B., M. A. LYSAK and I. SCHUBERT, 2005 Chromosome localization of rDNA in the Brassicaceae. *Genome* **48**: 341–346.
- BLANC, G., K. HOKAMP and K. H. WOLFE, 2003 A recent polyploidy superimposed on older large-scale duplications in the Arabidopsis genome. *Genome Res.* **13**: 137–144.

- BOIVIN, K., A. ACARKAN, R.-S. MBULU, O. CLARENZ and R. SCHMIDT, 2004 The Arabidopsis genome sequence as a tool for genome analysis in Brassicaceae. A comparison of the Arabidopsis and *Capsella rubella* genomes. *Plant Physiol.* **134**: 735–744.
- CHARLESWORTH, B., 1992 Evolutionary rates in partially self-fertilizing species. *Am. Nat.* **140**: 126–148.
- COPENHAVER, G. P., W. E. BROWNE and D. PREUSS, 1998 Assaying genome-wide recombination and centromere functions with Arabidopsis tetrads. *Proc. Natl. Acad. Sci. USA* **95**: 247–252.
- COPENHAVER, G. P., K. NICKEL, T. KUROMORI, M. I. BENITO, S. KAUL *et al.*, 1999 Genetic definition and sequence analysis of Arabidopsis centromeres. *Science* **286**: 2468–2474.
- FRANK, M. J., D. LIU, Y. F. TSAY, C. USTACH and N. M. CRAWFORD, 1997 Tag1 is an autonomous transposable element that shows somatic excision in both Arabidopsis and tobacco. *Plant Cell* **9**: 1745–1756.
- HALL, A. E., G. C. KETTLER and D. PREUSS, 2006 Dynamic evolution at pericentromeres. *Genome Res.* **16**: 355–364.
- HANSSON, B., and A. KAWABE, 2005 A simple method to score single nucleotide polymorphisms based on allele-specific PCR and primer-induced fragment-length variation. *Mol. Ecol. Notes* **5**: 692–696.
- HANSSON, B., A. KAWABE, S. PREUSS, H. KUITTINEN and D. CHARLESWORTH, 2006 Comparative gene mapping in *Arabidopsis lyrata* chromosomes 1 and 2 and the corresponding *A. thaliana* chromosome 1: recombination rates, rearrangements and centromere location. *Genet. Res.* **87**: 75–85.
- HOSOUCI, T., N. KUMEKAWA, H. TSURUOKA and H. KOTANI, 2002 Physical map-based sizes of the centromeric regions of *Arabidopsis thaliana* chromosomes 1, 2, and 3. *DNA Res.* **9**: 117–121.
- JOHNSTON, J. S., A. E. PEPPER, A. E. HALL, Z. J. CHEN, G. HODNETT *et al.*, 2005 Evolution of genome size in Brassicaceae. *Ann. Bot.* **95**: 229–235.
- KAMM, A., I. GALASSO, T. SCHMIDT and J. S. HESLOP-HARRISON, 1995 Analysis of a repetitive DNA family from *Arabidopsis arenosa* and relationships between Arabidopsis species. *Plant Mol. Biol.* **27**: 853–862.
- KAWABE, A., and S. NASUDA, 2005 Structure and genomic organization of centromeric repeat in Arabidopsis species. *Mol. Genet. Genomics* **272**: 593–602.
- KAWABE, A., B. HANSSON, A. FORREST, J. HAGENBLAD and D. CHARLESWORTH, 2006 Comparative gene mapping in *Arabidopsis lyrata* chromosomes 6 and 7 and *A. thaliana* chromosome IV: evolutionary history, rearrangements and local recombination rates. *Genet. Res.* (in press).
- KHUSH, G. S., and C. M. RICK, 1968 Cytogenetic analysis of the tomato genome by means of induced deficiencies. *Chromosoma* **23**: 452–484.
- KOCH, M., and M. KIEFER, 2005 Genome evolution among Cruciferous plants: a lecture from the comparison of the genetic map of three diploid species—*Capsella rubella*, *Arabidopsis lyrata* subsp. *petraea*, and *A. thaliana*. *Am. J. Bot.* **92**: 761–767.
- 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.
- KUITTINEN, H., A. A. DE HAAN, C. VOGL, S. OIKARINEN, J. LEPPÄLÄ *et al.*, 2004 Comparing the linkage maps of the close relatives *Arabidopsis lyrata* and *A. thaliana*. *Genetics* **168**: 1575–1584.
- KUMEKAWA, N., T. HOSOUCI, H. TSURUOKA and H. KOTANI, 2000 The size and sequence organization of the centromeric region of *Arabidopsis thaliana* chromosome 5. *DNA Res.* **7**: 315–321.
- KUMEKAWA, N., T. HOSOUCI, H. TSURUOKA and H. KOTANI, 2001 The size and sequence organization of the centromeric region of *Arabidopsis thaliana* chromosome 4. *DNA Res.* **8**: 285–290.
- LANDE, R., 1984 The expected fixation rate of chromosomal inversions. *Evolution* **38**: 743–752.
- LYSAK, M. A., A. PECINKA and I. SCHUBERT, 2003 Recent progress in chromosome painting of Arabidopsis and related species. *Chromosome Res.* **11**: 195–204.
- LYSAK, M. A., A. BERR, A. PECINKA, R. SCHMIDT, K. MCBREEN *et al.*, 2006 Mechanisms of chromosome number reduction in *Arabidopsis thaliana* and related Brassicaceae species. *Proc. Natl. Acad. Sci. USA* **103**: 5224–5229.
- MIURA, A., S. YONEBAYASHI, K. WATANABE, T. TOYAMA, H. SHIMADA *et al.*, 2001 Mobilization of transposons by a mutation abolishing full DNA methylation in Arabidopsis. *Nature* **411**: 212–214.
- SANYAL, A., and S. A. JACKSON, 2005 Comparative genomics reveals expansion of the *FLC* region in the genus Arabidopsis. *Mol. Genet. Genomics* **9**: 1–9.
- SINGER, T., C. YORDAN and R. A. MARTIENSSSEN, 2001 Robertson's Mutator transposons in *A. thaliana* are regulated by the chromatin-remodeling gene decrease in DNA methylation (*DDMI*). *Genes Dev.* **15**: 591–602.
- VENTURA, M., N. ARCHIDIACONO and M. ROCCHI, 2001 Centromere emergence in evolution. *Genome Res.* **11**: 595–599.
- WRIGHT, S. I., B. LAUGA and D. CHARLESWORTH, 2002 Rates and patterns of molecular evolution in inbred and outbred Arabidopsis. *Mol. Biol. Evol.* **19**: 1407–1420.
- YOGESWARAN, K., A. FRARY, T. L. YORK, A. AMENTA, A. H. LESSER *et al.*, 2005 Comparative genome analyses of Arabidopsis spp.: inferring chromosomal rearrangement events in the evolutionary history of *A. thaliana*. *Genome Res.* **15**: 505–515.
- ZHANG, L., and B. S. GAUT, 2003 Does recombination shape the distribution and evolution of tandemly arrayed genes (TAGs) in the *Arabidopsis thaliana* genome? *Genome Res.* **13**: 2533–2540.

Communicating editor: J. A. BIRCHLER