

Comparative gene mapping in *Arabidopsis lyrata* chromosomes 6 and 7 and *A. thaliana* chromosome IV: evolutionary history, rearrangements and local recombination rates

AKIRA KAWABE¹, BENGT HANSSON¹†, ALAN FORREST¹,
JENNY HAGENBLAD¹‡ AND DEBORAH CHARLESWORTH¹*

¹Institute of Evolutionary Biology, University of Edinburgh, Ashworth Laboratories King's Buildings, West Mains Road, Edinburgh EH9 3JT, UK

(Received 10 February 2006 and in revised form 8 May 2006)

Summary

We have increased the density of genetic markers on the *Arabidopsis lyrata* chromosomes AL6 and AL7 corresponding to the *A. thaliana* chromosome IV, in order to determine chromosome rearrangements between these two species, and to compare recombination fractions across the same intervals. We confirm the two rearrangements previously inferred (a reciprocal translocation and a large inversion, which we infer to be pericentric). By including markers around the centromere regions of *A. thaliana* chromosomes IV and V, we localize the AL6 centromere, and can localize the breakpoints of these chromosome rearrangements more precisely than previously. One translocation breakpoint was close to the centromere, and the other coincided with one end of the inversion, suggesting that a single event caused both rearrangements. At the resolution of our mapping, apart from these rearrangements, all other markers are in the same order in *A. lyrata* and *A. thaliana*. We could thus compare recombination rates in the two species. We found slightly higher values in *A. thaliana*, and a minimum estimate for regions not close to a centromere in *A. lyrata* is 4–5 centimorgans per megabase. The mapped region of AL7 includes the self-incompatibility loci (*S*-loci), and this region has been predicted to have lower recombination than elsewhere in the genome. We mapped 17 markers in a region of 1.23 Mb surrounding these loci, and compared the approximately 600 kb closest to the *S*-loci with the surrounding region of approximately the same size. There were significantly fewer recombination events in the closer than the more distant region, supporting the above prediction, but showing that the low recombination region is very limited in size.

1. Introduction

The chromosomes of *A. thaliana* differ by several rearrangements from those of closely related plants (reviewed by Koch & Kiefer, 2005; Yogeewaran *et al.*, 2005). Chromosome structure is highly conserved in other *Arabidopsis* relatives (Lysak *et al.*, 2003), even between *Capsella rubella* and *A. lyrata*, which diverged much longer ago than *A. lyrata* and

A. thaliana, and comparative genetic mapping confirms this (Kuittinen *et al.*, 2004; Boivin *et al.*, 2004), implying that most of these changes evolved within the *A. thaliana* lineage, rather than in the other species (reviewed by Koch & Kiefer, 2005; Yogeewaran *et al.*, 2005).

One purpose of this study was to compare map distances across the same intervals in two plant species, *A. thaliana* and *A. lyrata*. If large rearrangements are not too numerous, and there are few complex rearrangements, it should be possible to compare map distances across the same intervals in different species, and make progress towards comparing recombination rates per base pair in *A. thaliana* (an inbreeder, with a mean of about 200 kb per centimorgan; see Copenhagen *et al.*, 1998) and *A. lyrata*

* Corresponding author. Institute of Evolutionary Biology, University of Edinburgh, Ashworth Laboratories King's Buildings, West Mains Road, Edinburgh EH9 3JT, UK. e-mail: deborah.charlesworth@ed.ac.uk

† Present address: Department of Animal Ecology, Lund University, Ecology Building, 223 62 Lund, Sweden.

‡ Department of Physics, Chemistry and Biology, Linköping University, 581 83 Linköping, Sweden.

(a self-incompatible species whose total DNA content is about 60% higher). Such data are of interest because it is predicted that inbreeders should have higher recombination rates than outcrossers (Charlesworth *et al.*, 1977, 1979; Hedrick *et al.*, 1978). The evidence to date is purely from chiasma frequencies per chromosome arm in a number of plant genera (Charlesworth *et al.*, 1977), and no genetic tests have so far been published, but the genus *Arabidopsis* should be suitable for such a test. A comparison of the two chromosomes corresponding to *A. thaliana* chromosome I shows that one *A. lyrata* chromosome, AL1, has very similar map intervals to those in *A. thaliana*, while intervals in AL2, which has undergone an inversion in the *A. thaliana* lineage (inversion 1 of Koch & Kiefer, 2005), are less similar (Hansson *et al.*, 2006). Overall, the results are consistent with the prediction that the more inbreeding *A. thaliana* should have more crossovers per base pair than *A. lyrata*.

As discussed in our previous paper (Hansson *et al.*, 2006), the genome sizes of *A. thaliana* and *A. lyrata* differ considerably. Their map distances may nevertheless be similar if recombination occurs largely within coding sequences, and if gene content is similar in homologous intervals in both species. Restriction of crossovers to the genes has previously been inferred from the similarity of genetic map lengths between other species with differing DNA amounts (Thuriaux, 1977), and in maize most recombination indeed seems to occur within genes (Schnable *et al.*, 1998; Xu *et al.*, 1995; Dooner & Martinez-Férez, 1997).

In studying these questions in the genus *Arabidopsis*, it is important to compare another chromosome, to test whether the results from chromosome I are general. Chromosome IV (one of two small *A. thaliana* chromosomes) is of special interest, because in *A. thaliana* it carries a nucleolar organizer region (NOR) at the tip. *A. thaliana* chromosomes carrying nucleolar organizers have less recombination than other chromosomes (Copenhaver *et al.*, 1998), although the region closest to the NOR appears to be a recombination hotspot (Drouaud *et al.*, 2005). *A. thaliana* chromosome IV also carries a 5S rDNA locus (Murata *et al.*, 1997; Fransz *et al.*, 1998), and its short arm also has a large heterochromatic knob with a large transposable element content, located in the region between the knob and the pericentromeric heterochromatin (Fransz *et al.*, 2000; Arabidopsis Genome Initiative, 2000).

Before one can compare genetic maps between two species, it is, however, essential to understand the arrangement of their genes. It should also be possible, using a combination of genetic mapping and molecular cytogenetics, to understand in detail the rearrangements that have led to the evolution of the present *A. thaliana* chromosomes from the arrangement in its relatives. Among the interesting questions

is the location of the centromeres that were not lost in the formation of the *A. thaliana* chromosomes. Like *A. thaliana* chromosome I, chromosome IV evolved from two *A. lyrata* chromosomes, but, rather than a fusion, there was a reciprocal translocation between AL6 and AL7, with the loss of an NOR probably present on *A. lyrata* chromosome 7 (Lysak *et al.*, 2003). It seems likely that the centromere of chromosome IV is descended from the AL6 centromere, because, in the *A. lyrata* map, markers from both arms of *A. thaliana* chromosome IV are on AL6, while AL7 markers form a separate block (Kuittinen *et al.*, 2004). There was also an inversion in the AL6-derived region in the *A. thaliana* lineage (inversion 2 of Koch & Kiefer, 2005). To localize the breakpoints of these chromosome rearrangements more precisely, and establish the position of the AL6 centromere, more detailed mapping is needed. A second goal of this study was therefore to add new *A. thaliana* chromosome IV markers to map chromosomes AL6 and AL7 in the same *A. lyrata* ssp. *petraea* family used previously (Kuittinen *et al.*, 2004).

This chromosome is also interesting in relation to the evolution of recombination rates, because it carries the self-incompatibility locus (*S*-locus) region. This is located in the AL7-derived region in the *A. thaliana* genetic map (Yogeeswaran *et al.*, 2005). It has been suggested that low recombination may have evolved in the *S*-locus region in order to keep incompatible combinations of alleles at the pistil and pollen loci (*SRK* and *SCR*) together in functional haplotypes (e.g. Casselman *et al.*, 2000). This hypothesis predicts that recombination rates per base pair will be lower in and around the *S*-loci than in other genome regions. Estimates of recombination frequencies for RFLP markers in genes flanking the *B. oleracea* *S*-locus, covering an estimated 24 kb, detected 4 recombinants in a family of 509 plants segregating for the highly diverged alleles BoS8 and BoS1f2, and estimated the physical length per centimorgan (cM) to be 410 kb (Casselman *et al.*, 2000). This estimate is not markedly different from the estimates for nearby regions, and is about half the rate estimated for *A. thaliana* non-centromeric regions (Copenhaver *et al.*, 1998), i.e. it does not suggest strongly suppressed recombination. However, results from other haplotypes might differ (Casselman *et al.*, 2000), since different *Brassica* *S*-haplotypes differ in the physical distances between genes (Fobis-Loisy *et al.*, 2004).

In *A. lyrata*, indirect estimates, based on fitting a model of self-incompatibility in which many alleles are maintained (Takebayashi *et al.*, 2004), suggest low recombination in this region, consistent with findings of unusually high diversity at nearby loci, and linkage disequilibrium between *SRK* alleles and alleles at the flanking loci (Kamau & Charlesworth, 2005). To test this further, genetic map data are needed, in

combination with information on physical distances between loci. Physical distances in the immediate *S*-locus region are, however, currently known for only two *A. lyrata* haplotypes, whose *SRK*–*SCR* distances differ considerably (Kusaba *et al.*, 2001), and distances in other haplotypes may also be different. The distances to the closest flanking genes are also known for these haplotypes. These distances are larger than those for the homologous *A. thaliana* region, and, unlike the distances between *SRK* and *SCR*, they differ little between the two haplotypes studied (Kusaba *et al.*, 2001). Despite the difficulty of detecting recombination across small physical distances, comparing recombination frequencies for the region with those for other parts of the chromosome, along with their physical sizes, may help test whether recombination is or is not reduced near the *S*-loci.

2. Materials and methods

(i) Plants

The genetic map used the same segregating ‘mapping population’ of *A. lyrata* ssp. *petraea* as was used by Kuittinen *et al.* (2004). Following that paper, we refer to this as the mapping family. It included 99 plants from an F₂ made by a cross between two populations, from Karhumäki (Russia) and Mjällom (Sweden), forming an F₁ sibship, followed by reciprocal crossing between two of the F₁ plants.

(ii) Markers studied

In total, 40 genes, in three categories, were analysed in this study (Supplemental Table 1). Eighteen genes are from the two *A. thaliana* chromosome IV arms, 17 of them at intervals of about 1 Mbp, plus one gene from the region studied by Hagenblad & Nordborg (2002); two of these, At4g20410 and At4g22720, are located near the *S*-locus. To determine chromosome rearrangements near the centromere region on the short arm, we added genes close to the centromere, two on each arm, plus three genes in the region between the heterochromatin knob and the centromere, at intervals of about 300–600 kb. Finally, to determine crossing over rates around the *S*-locus, 11 further genes expected to be located close to the *S*-locus were sampled at intervals of about 80–150 kb. These were mapped, along with the *SRK* gene itself, and three other genes previously shown by physical mapping of two *S*-haplotypes to be located very close to the *S*-locus (Kusaba *et al.*, 2001).

(iii) Genotyping and data analysis

Primers were designed from coding regions of *A. thaliana* genomic sequences. PCR products from one parent F₁ plant and between three and seven F₂

progeny plants were directly sequenced using the DYEnamic sequencing kit (Amersham). Except for the At4g00030 gene, at least one polymorphic site was found in the mapping family in every gene. To obtain a marker for At4g00030, primers were designed for the intergenic region between At4g00030 and At4g00040, and polymorphic sites were then obtained. We used four different genotyping methods (Hansson & Kawabe, 2005; Hansson *et al.*, 2006). Some variants were scored using large indel variants distinguishable by agarose gel electrophoresis, and some using PCR-RFLP. Small indel variants were detected using labelled primers and running the fragments on an ABI3730 sequencer. Finally, some were scored by allele-specific primer-induced fragment length polymorphism (Hansson & Kawabe, 2005) using an ABI3730 sequencer.

Recombination rates were estimated as averages for male and female meioses, because only 99 progeny were available in our mapping family, from both reciprocal crosses (see above). Mapping was done using JoinMap software 3.0 (van Ooijen & Voorrips, 2001). The map presented in Fig. 1 is the most parsimonious map constructed with the JoinMap procedure and, as in all mapping studies, the centimorgan distances are estimates with some degree of error. However, our analyses are directed towards comparing the maps in two species, over numerous intervals, mostly with large genetic map distances, and our inferences should be little affected by these minor errors in individual distances. To compare map intervals between the two species, we need to map the same intervals in both. We used the same loci as far as possible, but most of the loci we mapped have not been genetically mapped in *A. thaliana*; in such cases, we followed the approach of Hansson *et al.* (2006) and used data on the physically closest mapped *A. thaliana* loci to homologues of our genetically mapped loci. Only 32 genes could be compared between the two species (see Table 1). *A. thaliana* recombinant inbred mapping data and physical positions of markers were obtained from The Arabidopsis Information Resource (TAIR) database (<http://www.arabidopsis.org/>, version published on 1 September 2005).

3. Results

(i) Linkage of markers, and chromosome rearrangements

Fig. 1 shows the genetic map and karyotype changes in *A. thaliana* IV relative to *A. lyrata* chromosomes AL6 and AL7, which, as explained above, are probably similar in arrangement to the ancestral chromosomes. With a few exceptions, the markers are in the same order in *A. lyrata* and *A. thaliana*; the eight markers with locations differing slightly from those

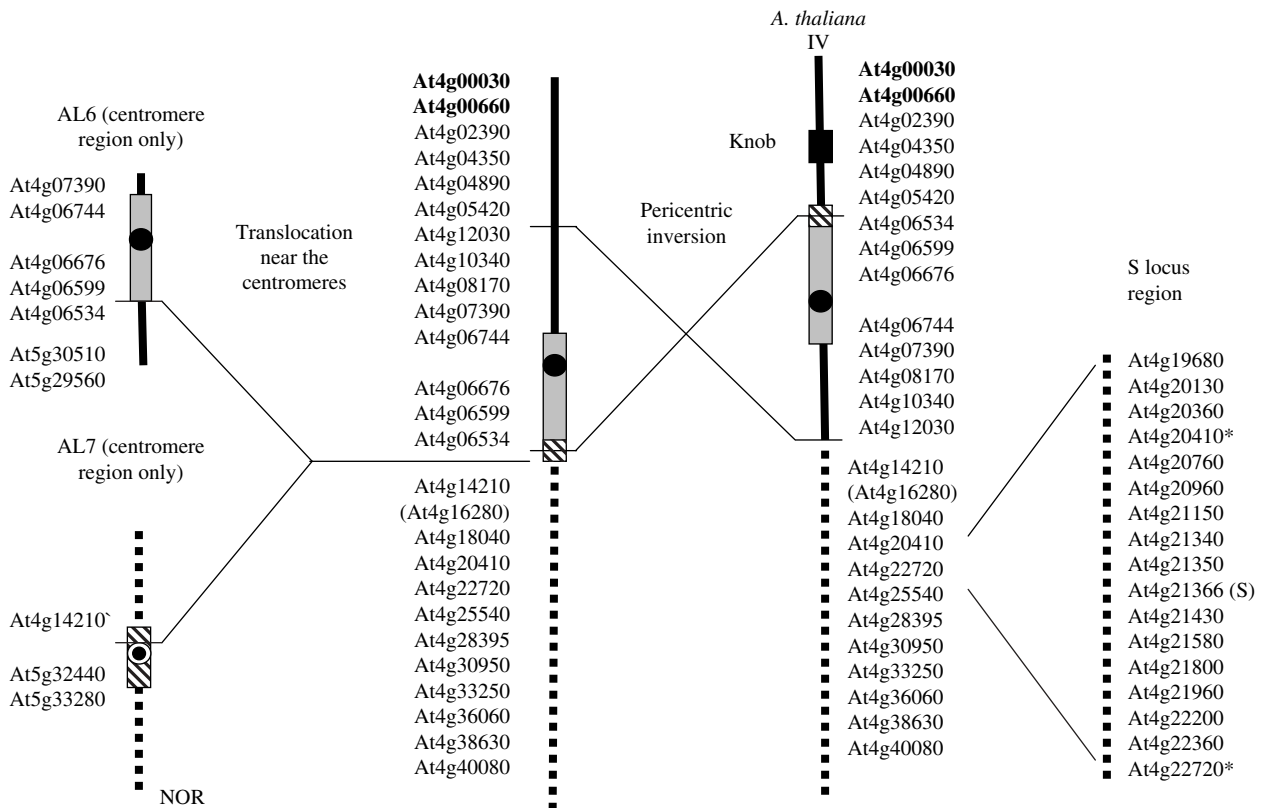


Fig. 1. Map of AL6 and AL7. Regions derived from the *A. thaliana* chromosome IV are shown as black lines, and the centromere is shown as a black disc (unbroken lines represent AL6 loci and broken lines, AL7 loci), and chromosome V regions are grey lines and its centromere is a white disc with a black dot in the centre (unbroken and broken lines again indicate AL6 and AL7). The centromeric regions are indicated by boxes. For markers near the *S*-locus, due to lack of space, the figure indicates the region and lists the genes; details of these genes are given in Supplemental Table 1.

expected are indicated with dotted lines in Fig. 2. Two rearrangements are known to have occurred in the formation of *A. thaliana* chromosome IV: a reciprocal translocation that exchanged parts of the two linkage groups corresponding to *A. lyrata* AL6 and AL7, and a large inversion in the part of *A. thaliana* chromosome IV corresponding to *A. lyrata* AL6 (see Koch & Kiefer, 2005).

Previous comparative maps for the same two species had low marker density in the region where the two rearrangements occurred. Our additional markers, especially around the centromere regions of *A. thaliana* chromosomes IV and V, localize them more precisely. In *A. thaliana* chromosome V, the break can be inferred to be close to the centromere of the ancestral chromosome, between the AL6 centromere markers At5g30510 (Fig. 1) and At5g32440 (now located in *A. lyrata* AL7), on the opposite sides of the *A. thaliana* centromere. The junction with the *A. thaliana* chromosome IV markers regions is located between At4g12030 (at the bottom of the AL6-derived region in Fig. 1) and At4g14210 (at the top of the AL7-derived region in Fig. 1). The second rearrangement detected in the map is a large pericentric inversion in the part of *A. thaliana* chromosome IV corresponding to

A. lyrata AL6. One end of this inversion is between At4g12030 and At4g14210, the same region as the junction between the AL6-derived and AL7-derived regions in *A. thaliana* chromosome IV, while the other end lies between At4g06534 and At4g05420, a region of high transposable element density in *A. thaliana* close to the CEN4 region (Arabidopsis Genome Initiative, 2000). Because one end of the inversion shares the same breakpoint with the translocation, both rearrangements may have occurred in one complex event, rather than as two independent events (discussed in detail in Kawabe *et al.*, 2006); thus the order of the events in Fig. 1 does not indicate a temporal succession, but is merely intended to make the breakpoints of the rearrangements clear.

(ii) Transposition of the *FRI* gene

We also investigated more closely the rearrangement involved in the unexpected map position of the *FRI* locus (At4g00650) in *A. lyrata*. Kuitinen *et al.* (2004) found that this *A. thaliana* chromosome IV locus mapped, unexpectedly, to AL8 (as far as is known, the rest of this linkage group carries exclusively genes found on *A. thaliana* chromosome V). In *A. thaliana*,

Table 1. Map distances of markers on *A. thaliana* chromosomes I and IV and the corresponding regions of AL1 and AL2, and of AL6 and AL7. The physical sizes are calculated from *A. thaliana*, excluding the core centromere region. Significance levels: * : $p < 0.05$, ** : $p < 0.01$

Chromosome regions	Physical and genetic map distances				Regression parameters						
	<i>A. thaliana</i>		<i>A. lyrata</i>		<i>A. thaliana</i>		<i>A. lyrata</i>				
	Number of markers	Mb	cM	cM/Mb	Mb versus cM	Slope	r value	Mb versus cM	Slope	r value	
<i>A. thaliana</i> I											
AL1	31	20.89	81.22	3.89	3.97	4.02	0.93**	3.99	0.991**	0.985	0.991**
Short (top) arm	26	11.90	55.64	4.68	4.83	4.35	0.985**	4.30	0.982**	0.974	0.982**
Long arm	5	4.77	22.70	4.76	3.84	4.43	0.963**	3.98	0.949*	0.759	0.833
AL2	9	5.33	37.41	7.02	8.90	5.29	0.892**	8.05	0.982**	1.311	0.949**
AL2 (inverted region)	3	1.88	5.98	3.19	5.48	3.13	0.880	-5.522	0.982	-1.229	0.775
<i>A. thaliana</i> IV											
AL6-derived (non-inverted)	6	2.11	22.38	10.58	6.60	10.08	0.971**	6.73	0.943**	0.584	0.907*
AL6 (inverted region)	5	2.51	7.00	2.79	2.78	2.08	0.729	-2.24	0.773	-0.967	0.950*
All AL7-derived	21	10.39	59.56	5.73	4.86	5.14	0.976**	4.07	0.973**	0.758	0.953**
AL7, excluding S-locus region	15		as for AL7 above			5.19	0.975**	4.01	0.972**	0.737	0.950**
S-locus region only	6	1.23	3.96	3.21	1.19	2.47	0.764	0.957	0.524	0.286	0.505

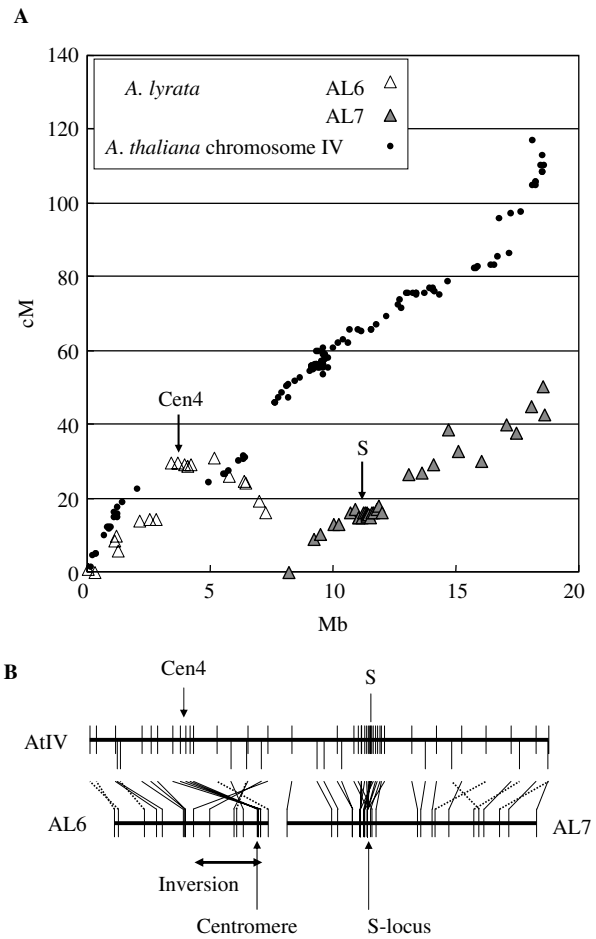


Fig. 2. (A) Comparison between physical distances (Mb) and genetic map distances (in cM) of *A. thaliana* chromosome IV, and those for the homologues in *A. lyrata* AL6 and AL7. Black dots indicate positions of markers mapped in *A. thaliana*, and triangles indicate *A. lyrata* map positions (open triangles for AL6, and filled triangles for AL7 markers). The positions of the AL6 centromere and of the S locus region are indicated. (B) The relationship between the *A. thaliana* physical map and the *A. lyrata* genetic map. Markers used by Kuittinen *et al.* (2004) are indicated as vertical bars under the horizontal bar representing the *A. thaliana* chromosome, and markers mapped in our study are shown as vertical bars crossing the bar representing the *A. thaliana* chromosome. Single markers with unexpected map positions (apart from the inversion) are indicated by dotted lines.

FRI maps at the end of the euchromatic region of the chromosome IV short arm, and should thus map to AL6 (see Fig. 1). The difference between the two species suggests a transposition in the history of *A. thaliana* of the tip of the chromosome corresponding to AL6 (a large NOR array, also on the short arm of *A. thaliana* chromosome IV, is also not found on AL6, and must have come from another location via a separate transposition event). To test the transposition hypothesis for *FRI*, we included in our map an adjacent gene in *A. thaliana* (Hagenblad & Nordborg, 2002), and an even more distal gene (At4g00030),

both located near the tip of the *A. thaliana* chromosome IV (these two genes are indicated in bold at the top of chromosome IV in Fig. 1). Both are strongly linked to AL6 genes. Thus, either *FRI* transposed to AL8 without transposition of its flanking regions, or else it was present as duplicate copies in the ancestor, and one copy was lost in each of the *A. thaliana* and *A. lyrata* lineages. Silent site divergence for *FRI* between the two species is about average (LeCorre *et al.*, 2002), which makes the second possibility unlikely.

(iii) Duplications and losses of genes

We here describe our findings about copy number differences for the mapped genes. These have importance for two reasons. First, comparative mapping requires that the same genes should be in the same order, in both species, as do our inferences about recombination rates per megabase for the large non-rearranged chromosome arm regions (see below). Second, given the expectation that the fusions were involved in the evolution of the five *A. thaliana* chromosomes from the ancestral number of eight chromosomes, genes near the rearrangement breakpoints may have been lost in the *A. thaliana* lineage. These genes will be absent from the *A. thaliana* genome, which is the sole source of markers for our mapping. Thus, to assess the possibility of gene loss in the fusion events, we cannot use mapping of *A. thaliana* genes (though any missing genes might be detected as gaps by *in situ* hybridization experiments), but might be detected as different gene copy numbers.

We found no examples of gene losses that could be attributed to the fusions, so the details are presented in Supplementary Fig. 1 and are here outlined briefly. Including our new results presented here, four genes show copy number differences, in all cases having two or more copies in *A. lyrata*, but only one in *A. thaliana* (and also in *A. glabra*; however, we cannot exclude the possibility of further copies in *A. glabra*, with other copies failing to amplify with our primers, due to divergence).

Among the non-centromeric region genes studied, 25 on *A. thaliana* chromosome I (Hansson *et al.*, 2006) and the 10 AL6 and 26 AL7 loci mapped in the present study, we found two copy number differences. In both cases, single-copy genes in *A. thaliana* are present as two unlinked copies in *A. lyrata*. The gene tree based on At1g01370 (or *HTR12*) sequences, including an outgroup, *A. glabra*, is consistent with the species' relationships (Supplementary Fig. 1), suggesting duplication in *A. lyrata*. For the other non-centromeric gene copy number difference, At4g21150, *A. lyrata*, one copy is linked to the SRK gene, like the single *A. thaliana* locus, while the other copy maps at the tip of chromosome AL7, in a region homologous with *A. thaliana* chromosome V; in *A. lyrata*, this

region has an NOR (Lysak *et al.*, 2003). The sequences suggest a duplication before the divergence of *A. glabra* from the two *Arabidopsis* species (Supplementary Fig. 1), suggesting that the locus at the tip of AL7 is not a recent duplication in *A. lyrata* but may be a case of gene loss in *A. thaliana* (Supplementary Fig. 1).

Of the 24 genes studied from regions near *A. thaliana* centromeres, we again found two copy number differences. Both are cases of closely linked gene duplicates in *A. lyrata* that are single-copy genes in *A. thaliana*. We previously found multiple sequences in *A. lyrata* of one gene, At1g43980, from the *A. thaliana* centromere region of chromosome I (Hansson *et al.*, 2006). Both copies are completely linked in *A. lyrata*, indicating a probable tandem duplication within the region of infrequent recombination near the centromere, and their sequences are considerably more similar to each other than to the *A. thaliana* sequence. The gene tree, including *A. glabra*, is consistent with the species' relationships (Supplementary Fig. 1), thus a recent duplication is likely in the *A. lyrata* lineage.

The marker amplified with primers for the centromere III region gene At3g42050 is apparently triplicated in *A. lyrata*, and gave puzzling results. From the sequence analyses, the copy labelled B in Supplementary Fig. 1 is closest to those of *A. thaliana* and *A. glabra*, suggesting that this is orthologous to the *A. thaliana* copy. The two other copies, 'A1' and 'A2' in the figure, appear to be paralogues pre-dating the species' divergence; they are probably tandem duplicated, but not very recently, as their divergence is high ($K_{\text{silent}} = 0.111$). Supplementary Fig. 2 gives the expectations for divergence (see Supplementary Fig. 2 for the simplest expectation, and see Notebaer *et al.*, 2005). Unexpectedly, the A copies, and not B, are in the location homologous with that in *A. thaliana*, and we observed no recombinants in *A. lyrata*, consistent with a centromere location. In contrast, the B copy maps near markers from the long arm of *A. thaliana* chromosome IV, in a region where gene order is otherwise conserved between *A. thaliana* and *A. lyrata*. At3g42050 is also absent from the pericentromeric location where it is found in *A. thaliana*, in three related species (Hall *et al.*, 2006). Several gene duplications and losses seem necessary to explain these results.

(iv) Comparison of *A. thaliana* and *A. lyrata* map distances

Despite the chromosome rearrangements detailed above, most of the AL7 markers (in the interval between At4g14210 and At4g40080) show synteny between *A. thaliana* and *A. lyrata* in our map, suggesting that there are no other large rearrangements in this chromosome arm, consistent with the coarse maps of Kuittinen *et al.* (2004) and Yogeewaran *et al.*

(2005). This allows us to compare map distances directly between *A. thaliana* and *A. lyrata*. The regression slopes of the linear regressions of *A. lyrata* map positions versus those in *A. thaliana* are highly significantly different from zero (Table 1). The maps are thus similar in the two species. For the AL7-derived portion of chromosome IV, the slope is significantly less than 1 (the 95% confidence interval is 0.650 to 0.866). The AL6-derived part includes a few markers in each of the two segments (the segment that is inverted in *A. thaliana*, and the non-inverted segment; see Table 1), so the confidence intervals are wide. Nevertheless, the non-inverted region has a slope significantly less than 1 (the 95% confidence interval is 0.317 to 0.851), whereas the inverted region shows no significant difference between the *A. thaliana* and *A. lyrata* maps (slope = -0.967 , with a 95% confidence interval of -0.608 to -1.326). Map intervals are thus somewhat lower in *A. lyrata*, which is surprising if the species have the same recombination rate per physical distance, given that larger physical distances are expected in *A. lyrata*.

Physical distances are not currently known in *A. lyrata*. However, we can use the *A. thaliana* physical distances to obtain minimum values of the recombination rate per megabase, assuming that physical distances are not lower than in *A. thaliana*. For the AL7 markers, the overall values are then at most 4.86 in *A. lyrata*, versus 5.73 cM/Mb in *A. thaliana* (Table 1). The inverted region yields the same values in both species (2.79 cM/Mb in *A. thaliana*, and 2.78 in *A. lyrata*), whereas the rest of the AL6-derived region again yields a lower value in *A. lyrata* (6.60 cM/Mb versus the *A. thaliana* value of 10.58; see Table 1). The overall results thus suggest greater recombination per megabase in *A. thaliana*, in agreement with the prediction outline in Section 1.

However, paired comparisons of individual intervals, using Wilcoxon signed rank tests (for all loci, or separately for AL6 and AL7 markers), are not statistically significant, as intervals differ in opposite directions between the species. Across all 29 intervals (five in the non-inverted AL6-derived region, four in the inverted region and 20 corresponding to AL7 markers), excluding the interval that cannot be compared because of the inversion, about half the intervals have higher map distances in *A. thaliana* than in *A. lyrata*, and half differ in the opposite direction.

(v) *The self-incompatibility locus region*

In the mapping family, the *S*-locus was mapped to AL7, as was previously found using *A. lyrata* ssp. *lyrata* (Yogeeswaran *et al.*, 2005). No crossovers were detected among the markers closest to the *S*-loci (Table 2). To compare recombination in the region immediately surrounding the *S*-locus with that in the

neighbouring regions, we considered meioses of both parents separately. No progeny derived from recombination in the KM 6-5 parent were found in the region between At4g20360 and At4g21580 (486 kb in *A. thaliana*). In the KM 4-2 parent no recombination events were detected between At4g20360 and At4g21960 (*A. thaliana* distance 656 kb).

On both sides of this region, however, we did detect recombination events in both parents. One event was in the KM 6-5 parent, between At4g19680 and At4g20360 (in the 287 kb region before the region surrounding to the *S*-locus), and this parent also yielded two events between At4g21580 and At4g22360 (in the 332 kb on the other side of the *S*-locus); the total from this parent is thus three events in 619 kb (versus none in the 486 kb surrounding the *S*-locus). In the KM 4-2 parent, we detected three events between At4g19680 and At4g20360 (in 287 kb on one side of the *S*-locus) and one between At4g21960 and At4g22720 (in 291 kb on the other side of the *S*-locus); this totals four events in 578 kb, versus none in the 656 kb surrounding the *S*-locus. From the average genetic map distance per megabase for AL7, i.e. 4.86 cM/Mb (based on the summed physical distances in *A. thaliana* of the map locations against physical distances; see Table 1), and assuming this value for both parents, the 486 kb region surrounding the *S*-locus in which the KM 6-5 parent yielded no recombinants should correspond to 2.36 cM. For the KM 4-2 parent, the expected genetic distance is 3.19 cM (based on 656 kb). Pooling both parents, a total of 5.5 recombination events are thus expected in this region, whereas none was observed. In many intervals, 98 meioses were scored from both the parents (Table 2), and, overall, no crossover events were detected in a region of 1142 kb closest to the *S*-locus, but seven events were detected in the more distant flanking regions, which total an only slightly larger size (1197 kb in *A. thaliana*). The numbers of recombination events differ significantly between these two regions (7/1197 versus 0/1142; $P=0.016$ by a two-tailed Fisher's exact test).

4. Discussion

(i) *Recombination rates in the two species*

Our map confirms that no large chromosome rearrangements have occurred apart from those already known (Koch & Kiefer, 2005), which we have now mapped to definite small intervals. This allowed us to compare the map distances directly between *A. thaliana* and *A. lyrata* for the many intervals not involved in rearrangements, for which it is reasonable to infer that they largely carry the same genes in the same order (though with the caveat that some copy number differences were detected among the genes we

Table 2. Numbers of recombinants between markers around the S-locus region (total length 1234 kb). The first three columns show the marker names and positions and (in bold) the informative markers in each parent of the mapping family. The other columns show the numbers of recombinants between all adjacent pairs of informative markers. Brackets indicate intervals where recombination was detected (98 meioses were scored for each informative marker in each parent). Intervals where recombination events cannot be localised to a particular interval, or where one parent was not informative, are indicated by brackets

Genes and genotypes			Interval physical sizes and genetic results		
Gene	Position in <i>A. thaliana</i> (Mb)	Phase of markers in parents KM4-2 and KM6-5	Physical size of the interval in <i>A. thaliana</i> (kb)	Numbers of recombinants in the interval	
				In parent KM4-2	In parent KM6-5
At4g19680	10.70	ab × ab	175.54	2	} 1
At4g20130	10.88	ab × aa	111.05	1	
At4g20360	10.99	ab × ab	24.00	0	0
At4g20410	11.01	ab × ab	115.07	0	0
At4g20760	11.13	ab × ab	82.93	0	} 0
At4g20960	11.21	ab × bb	66.48	0	
At4g21150	11.28	ab × aa	74.56	0	} 0
At4g21340	11.35	ab × bb	3.16	} 0	
At4g21350	11.36	bb × ab	27.40		0
At4g21366	11.38	ab × ac	24.29	} 0	0
At4g21430	11.41	bb × ab	67.89		0
At4g21580	11.48	ab × ab	97.07	0	} 1
At4g21800	11.57	ab × bb	73.38	0	
At4g21960	11.65	ab × aa	100.34	} 1	1
At4g22200	11.75	bb × ab	61.15		
At4g22360	11.81	aa × ab	129.35		
At4g22720	11.94	ab × aa			

mapped; see Section 3). Our overall results for these regions suggest a greater recombination frequency per megabase in *A. thaliana*. This is consistent with the prediction outlined above: that inbreeding species tend to have greater numbers of chiasmata per chromosome arm than related outbreeders. However, different regions differ between the species in different ways, and many individual intervals differ in recombination between the two species (not always statistically significantly). For the large non-rearranged arm corresponding to AL7, our estimate suggests that, across the same intervals, the *A. thaliana* recombination rate is about 18% higher than that in *A. lyrata*.

This is a much greater difference than estimated for *A. lyrata* chromosome 1, for which recombination rates are almost identical in the two species. The inverted region of AL6 yields very similar genetic map intervals for the two species (assuming the same physical sizes, 2.79 cM/Mb in *A. thaliana*, and 2.78 in *A. lyrata*). The other region suggests much higher recombination in *A. thaliana* (10.58 cM/Mb versus 6.60 in *A. lyrata*). In this region, it is possible that physical distances could be smaller in *A. lyrata*, because *A. thaliana* has a heterochromatic knob (whose presence is not yet clear in the *A. lyrata* karyotype). If this knob evolved in *A. thaliana*, after the split from

A. lyrata, *A. lyrata* could have smaller inter-marker distances.

As mentioned above, the physical distances corresponding to a centimorgan are not yet known for any region in *A. lyrata*. It is very difficult to determine such distances, other than by complete genome sequencing (which does not provide information about differences between different genotypes) or methods such as relating distances between genetically mapped loci to genetic distances inferred from density of recombination nodules in regions whose physical distances can be estimated (Anderson *et al.*, 2005). Population genetic estimates, based on linkage disequilibrium (LD) between polymorphisms in gene sequences, can also be made. So far, the only such estimate for *A. lyrata* suggested that the distance corresponding to 1 cM may be considerably larger than the average of about 200 kb estimated for non-centromeric regions in *A. thaliana* (Copenhaver *et al.*, 1997; Drouaud *et al.*, 2005), perhaps as large as one to two megabases (Wright *et al.*, 2003); this is, however, an overestimate, because the analysis was based on alleles sampled from a single population, which will overestimate LD (Wakeley & Lessard, 2003). It is also not likely to be accurate, as it is based on data from only a few loci. Here, our direct estimate using one arm of AL7, which is not rearranged, and therefore offers good comparisons across multiple intervals, indeed yields a considerably higher recombination rate: 4.86 cM/Mb (or 205 kb/cM), assuming the same physical size as in *A. thaliana*. This is very similar to the *A. thaliana* average, but does not take into account the probable larger DNA amount in this genome region of *A. lyrata*.

This conclusion is quite consistent with information from our previous genetic mapping results for *A. lyrata* chromosomes 1 and 2, which yielded map lengths of 83 and 73 cM, respectively (Hansson *et al.*, 2006). These chromosomes were fused to form the *A. thaliana* chromosome I, which is about 30 Mb, and is also not rearranged. Again assuming the same physical size as in *A. thaliana*, the length per centimorgan for *A. lyrata* chromosome 1 is about 250 kb, based on 32 non-centromere region markers with a mean distance apart of 0.72 Mb, very similar to the *A. thaliana* value.

For chromosome AL2 (15 markers at a mean distance apart of 0.47 Mb, but with an inversion and possibly other rearrangements), we obtain the very low estimate per centimorgan of about 100 kb (Hansson *et al.*, 2006). In *A. lyrata*, this chromosome has two arms, both shorter than the rearranged homologous arm of the *A. thaliana* chromosome I (Lysak *et al.*, 2006). This may account for the high recombination per megabase in *A. lyrata*, since plant species generally have at least one crossover per arm (Nilsson *et al.*, 1993); thus in *A. lyrata* these genes'

location across two short chromosome arms corresponds to a situation which is often associated with high recombination per physical distance (Perry *et al.*, 2001), while the same genes are on just one longer arm in *A. thaliana*. The non-inverted region of chromosome AL6 may be affected similarly. The pericentric inversion has shortened the arm of *A. thaliana* chromosome IV, compared with the arm carrying the corresponding genes in *A. lyrata*, and the estimated cM/Mb value is higher.

Even if we assume that the DNA content of each of these chromosome arms is 50% larger than the *A. thaliana* value, i.e. the same ratio as estimated for the entire genome, these values would increase to only about 400 and 150 kb/cM, respectively, and the AL7 value would become about 300. These values are probably too high, because part of the reason for the genome size difference between *A. thaliana* and *A. lyrata* could be loss of a few large genome regions during the chromosome rearrangements that formed the *A. thaliana* chromosomes. Even this probably inflated value of the physical distance corresponding to a centimorgan still suggests rather more recombination than the estimate based on LD.

(ii) *The self-incompatibility locus region*

The *S*-locus region itself is larger in the two *A. lyrata* haplotypes that have been studied, compared with the orthologous *A. thaliana* region (Kusaba *et al.*, 2001). Our comparison of genetic distance per kilobase between the *S*-locus and other regions, using *A. thaliana* physical distances is therefore conservative (i.e. biased against finding lower recombination in the *A. lyrata* *S*-locus region than elsewhere in the genome). Even so, we detect a significantly lower recombination rate in this region. In the future, other marker pairs with similar physical distances (in *A. thaliana*, or, ideally, in *A. lyrata* itself) should be mapped, to test whether they consistently have higher map distances than the genes in the *S*-locus region.

The conclusion of low recombination in this region is of interest in relation to whether *S*-haplotype-specific genetic load is expected. If the region recombines very rarely, loci within the region affected may experience deleterious mutations that will not be effectively removed by natural selection, both because only their heterozygous effects will be expressed (since the *S*-locus region will generally be heterozygous), and also because the efficacy of natural selection is weakened in non-recombining regions, due to the occurrence of hitch-hiking processes (spread of advantageous alleles, causing selective sweeps, and reduction of effective population size due to selection against linked deleterious mutations). These processes have been well studied in the context of sex chromosomes (Charlesworth & Charlesworth, 2000), and

may also act in plant *S*-locus regions (Kurian & Richards, 1997; Stone, 2004). However, the low recombination in *A. lyrata* is limited to only about 600 kb immediately surrounding the *S*-loci, which suggests that this region is unlikely to accumulate haplotype-specific genetic load (the heterostyled distylous systems may be different in this respect, as it is possible that they may have larger non-recombining regions than homomorphic incompatibility systems (Kurian & Richards, 1997). The magnitude of any such effect depends on the deleterious mutation rate for the non-recombining region as a whole, which depends on the number of genes in the region. From the *A. thaliana* genome sequence, we estimate the total number of genes in the putatively non-recombining *S*-locus region to be about 121, plus about 41 genes in the adjacent region where recombination may have occurred but cannot be definitively localized. These counts include some possible duplicate loci, but exclude the two self-incompatibility loci, both of which are pseudogenes, plus one other pseudogene. The total mutation rate for this small region should thus be about 200 times smaller than that of the entire *A. thaliana* genome.

The whole-genome deleterious mutation rate for *A. thaliana* has been estimated to be as high as 0.5 per generation, based on inbreeding depression values (Charlesworth *et al.*, 1990), but much lower values have been estimated in mutation accumulation experiments, with the plausible range lying between 0.1 and 0.6 (Schultz *et al.*, 1999; Shaw *et al.*, 2000). With a high value of 0.5, the deleterious mutation rate, μ , for the *S*-locus region might thus be as high as 0.0025. To ask whether it is likely that haplotypes for this region will commonly carry mutations with effects large enough to be detected from departures from Mendelian segregation in families, one can use this value to predict the expected equilibrium frequency of haplotypes carrying deleterious mutations (from the standard equation for mutation–selection balance for mutations with selection coefficient s , and dominance coefficients h above zero: $q = \mu/hs$). This theory, for mutations with effects in the diploid stage, predicts values of only a few per cent, for $hs > 0.1$, making it unlikely that any effect would be detected. Mutations affecting the haploid stages would be predicted to have even lower frequencies, particularly if they strongly affect performance of male or female gametophytes. It is therefore surprising that Bechsgaard *et al.* (2004) detected effects in gametophytes. Possibly, recombination suppression has lasted for enough time, and effective sizes have been small enough, that haplotypes have fixed highly deleterious mutations. More theoretical work should be done to investigate this possibility.

Another region in which recombination is infrequent is found in the part of *A. lyrata* AL6 that

is not inverted in *A. thaliana*. In *A. thaliana*, this region (between genes At4g04350, At4g04890 and At4g05420) lies between the heterochromatic knob and the centromere. There is no evident reason why these genes should not recombine in *A. lyrata*, and more markers in this region should probably be tested, and its physical size estimated in *A. lyrata*, in order to understand what the reason may be. Possibly this region has evolved a high transposon density in *A. thaliana*, since in this species (due to the inversion), but not in *A. lyrata*, it is close to the centromere region (Arabidopsis Genome Initiative, 2000), which may have led to accumulation of transposon and other non-coding junk DNA, increasing the DNA content. However, this might not lead to increased recombination, because, in other plant species, presence of transposable elements is associated with low recombination (Dooner & Martinez-Férez, 1997).

We thank O. Savolainen (University of Oulu, Finland) for DNA samples from the *A. lyrata* mapping family, H. Kuittinen for help with the linkage analyses and E. Kamau for primers. M. Lysak kindly let us see his unpublished results. The support of a post-doctoral grant from the Swedish Research Council and STINT (to B.H.) and a small research grant from the National Environmental Research Council of the UK (to D.C.) is gratefully acknowledged.

References

- Anderson, L. K., Salameh, N., Bass, H. W., Harper, L. C., Cande, W. Z., Weber, G. & Stack, S. M. (2005). Integrating genetic linkage maps with pachytene chromosome structure in maize. *Genetics* **166**, 1923–1933.
- Arabidopsis Genome Initiative (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**, 796–815.
- Bechsgaard, J., Bataillon, T. & Schierup, M. H. (2004). Uneven segregation of sporophytic self-incompatibility alleles in *Arabidopsis lyrata*. *Journal of Evolutionary Biology* **17**, 554–561.
- Boivin, K., Acarkan, A., Mbulu, R.-S., 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 Physiology* **134**, 735–744.
- Casselman, A. L., Vrebalov, J., Conner, J. A., Singhal, A., Giovanni, J., Nasrallah, M. E. & Nasrallah, J. B. (2000). Determining the physical limits of the *Brassica* *S*-locus by recombinational analysis. *The Plant Cell* **12**, 23–24.
- Charlesworth, B. & Charlesworth, D. (2000). The degeneration of Y chromosomes. *Proceedings of the Royal Society of London, Series B* **355**, 1563–1572.
- Charlesworth, B., Charlesworth, D. & Morgan, M. T. (1990). Genetic loads and estimates of mutation rates in very inbred plant populations. *Nature* **347**, 380–382.
- Charlesworth, D., Charlesworth, B. & Strobeck, C. (1977). Effects of selfing on selection for recombination. *Genetics* **68**, 213–226.
- Charlesworth, D., Charlesworth, B. & Strobeck, C. (1979). Selection for recombination in self-fertilising species. *Genetics* **93**, 237–244.

- Copenhaver, G. P., Browne, W. E. & Preuss, D. (1997). Measuring genome-wide recombination and centromere functions in *Arabidopsis meiosis*. *Molecular Biology of the Cell* **8**, 1557.
- Copenhaver, G. P., Browne, W. E. & Preuss, D. (1998). Assaying genome-wide recombination and centromere functions with *Arabidopsis tetrads*. *Proceedings of the National Academy of Sciences of the USA* **95**, 247–252.
- Dooner, H. K. & Martinez-Férez, I. M. (1997). Recombination occurs uniformly in the bronze gene, a recombination hotspot in the maize genome. *The Plant Cell* **9**, 1633–1646.
- Drouaud, J., Camilleri, C., Bourguignon, P.-Y., Canaguier, A., Bérard, A., Vezon, D., Giancola, S., Brunel, D., Colot, V., Prum, B., Quesneville, H. & Mézard, C. (2005). Variation in crossing-over rates across chromosome 4 of *Arabidopsis thaliana* reveals the presence of meiotic recombination ‘hot spots’. *Genome Research* **16**, 106–114.
- Fobis-Loisy, I., Miege, C. & Gaude, T. (2004). Molecular evolution of the S locus controlling mating in the Brassicaceae. *Plant Biology* **6**, 109–118.
- Fransz, P. F., Armstrong, S. J., Alonso-Blanco, C., Fischer, T. C., Torres-Ruiz, R. A. & Jones, G. (1998). Cytogenetics for the model system *Arabidopsis thaliana*. *Plant Journal* **13**, 867–876.
- Fransz, P. F., Armstrong, S. J., Jong, H., de, Parnell, L. D., Drunen, C. van, Dean, C., Zabel, P., Bisseling, T. & Jones, G. H. (2000). Integrated cytogenetic map of chromosome arm 4S of *A. thaliana* L.: structural organization of heterochromatic knob and centromere region. *Cell* **100**, 367–376.
- Hagenblad, J. & Nordborg, M. (2002). Sequence variation and haplotype structure surrounding the flowering time locus *FRI* in *Arabidopsis thaliana*. *Genetics* **161**, 289–298.
- Hall, A. E., Kettler, G. C. & Preuss, D. (2006). Dynamic evolution at pericentromeres. *Genome Research* **16**, 355–364.
- Hansson, B. & Kawabe, A. (2005). A simple method to score single nucleotide polymorphisms based on allele-specific PCR and primer-induced fragment-length variation. *Molecular Ecology Notes* **5**, 692–696.
- Hansson, B., Kawabe, A., Preuss, S., Kuittinen, H. & Charlesworth, D. (2006). Comparative gene mapping in *Arabidopsis lyrata* chromosomes 1 and 2 and the corresponding *A. thaliana* chromosome 1: recombination rates, rearrangements and centromere location. *Genetical Research* in press.
- Hedrick, P. W., Jain, S. K. & Holden, L. (1978). Multi-locus systems in evolution. *Evolutionary Biology* **2**, 104–184.
- Kamau, E. & Charlesworth, D. (2005). Balancing selection and low recombination affect diversity near the self-incompatibility loci of the plant *Arabidopsis lyrata*. *Current Biology* **15**, 1773–1778.
- Kawabe, A., Hansson, B., Hagenblad, J., Forrest, A. & Charlesworth, D. (2006). Centromere locations and associated chromosome rearrangements in *A. lyrata* and *A. thaliana*. *Genetics* in press.
- Koch, M. & Kiefer, M. (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*. *American Journal of Botany* **92**, 761–767.
- Kuittinen, H., Haan, A. A. de, Vogl, C., Oikarinen, S., Leppälä, 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.
- Kurian, V. & Richards, A. J. (1997). A new recombinant in the heteromorphy ‘S’ supergene in *Primula*. *Heredity* **78**, 383–390.
- Kusaba, M., Dwyer, K., Hendershot, J., Vrebalov, J., Nasrallah, J. B. & Nasrallah, M. E. (2001). Self-incompatibility in the genus *Arabidopsis*: characterization of the S locus in the outcrossing *A. lyrata* and its autogamous relative, *A. thaliana*. *The Plant Cell* **13**, 627–643.
- LeCorre, V., Roux, F. & Reboud, X. (2002). DNA polymorphism at the FRIGIDA gene in *Arabidopsis thaliana*: extensive nonsynonymous variation is consistent with local selection for flowering time. *Molecular Biology and Evolution* **19**, 1261–1271.
- Lysak, M. A., Pecinka, A. & Schubert, I. (2003). Recent progress in chromosome painting of *Arabidopsis* and related species. *Chromosome Research* **11**, 195–204.
- Lysak, M. A., Berr, A., Pecinka, A., Schmidt, R., McBreen, K. & Schubert, I. (2006). Mechanisms of chromosome number reduction in *Arabidopsis thaliana* and related Brassicaceae species. *Proceedings of the National Academy of Sciences of the USA* **103**, 5224–5229.
- Murata, M., Heslop-Harrison, J. S. & Motoyoshi, F. (1997). Physical mapping of the 5S ribosomal RNA genes in *Arabidopsis thaliana* by multi-color fluorescence *in situ* hybridization with cosmid clones. *The Plant Journal* **12**, 31–37.
- Nilsson, N. O., Sall, T. & Bengtsson, B. O. (1993). Chiasma and recombination data in plants: are they compatible? *Trends in Genetics* **9**, 344–348.
- Notebaar, R. A., Huynen, M. A., Teusink, B., Siezen, R. J. & Snel, B. (2005). Correlation between sequence conservation and the genomic context after gene duplication. *Nucleic Acids Research* **33**, 6164–6171.
- Perry, J., Palmer, S., Gabriel, A. & Ashworth, A. (2001). A short pseudoautosomal region in laboratory mice. *Genome Research* **11**, 1826–1832.
- Schnable, P. S., Hsia, A. P. & Nicolau, B. J. (1998). Genetic recombination in plants. *Current Opinion in Plant Biology* **1**, 123–129.
- Schultz, S., Lynch, M. & Willis, J. H. (1999). Spontaneous deleterious mutation rate in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the USA* **96**, 11393–11398.
- Shaw, R. G., Byers, D. L. & Darmo, E. (2000). Spontaneous mutational effects on reproductive traits of *Arabidopsis thaliana*. *Genetics* **155**, 369–378.
- Stone, J. L. (2004). Sheltered load associated with S-alleles in *Solanum carolinense*. *Heredity* **92**, 335–342.
- Takebayashi, N., Newbigin, E. & Uyenoyama, M. K. (2004). Maximum-likelihood estimation of rates of recombination within mating-type regions. *Genetics* **167**, 2097–2109.
- Thuriaux, P. (1977). Is recombination confined to structural genes on the eukaryotic genome? *Nature* **268**, 460–462.
- van Ooijen, J. W. & Voorrips, R. E. (2001). *JoinMap 3.0. Software for the Calculation of Genetic Linkage Maps*. Wageningen, The Netherlands: Plant Research International.
- Wakeley, J. & Lessard, S. (2003). Theory of the effects of population structure and sampling on patterns of linkage disequilibrium applied to genomic data from humans. *Genetics* **164**, 1043–1053.
- Wright, S. I., Lauga, B. & Charlesworth, D. (2003). Subdivision and haplotype structure in natural

- populations of *Arabidopsis lyrata*. *Molecular Ecology* **12**, 1247–1263.
- Xu, X., Hsia, A. P., Zhang, L., Nikolau, B. J. & Schnable, P. S. (1995). Meiotic recombination break points resolve at high rates at the 5' end of a maize coding sequence. *The Plant Cell* **7**, 2151–2161.
- Yogeeswaran, K., Frary, A., York, T. L., Amenta, A., Lesser, A. H., Nasrallah, J. B., Tanksley, S. D. & Nasrallah, M. E. (2005). Comparative genome analyses of *Arabidopsis* spp.: inferring chromosomal rearrangement events in the evolutionary history of *A. thaliana*. *Genome Research* **15**, 505–515.