Collinearity between a 30-centimorgan segment of *Arabidopsis thaliana* chromosome 4 and duplicated regions within the *Brassica napus* genome

A.C. Cavell, D.J. Lydiate, I.A.P. Parkin, C. Dean, and M. Trick

Abstract: *Arabidopsis thaliana* (the model dicotyledonous plant) is closely related to *Brassica* crop species. Genome collinearity, or conservation of marker order, between *Brassica napus* (oilseed rape) and *A. thaliana* was assessed over a 7.5-Mbp region of the long arm of *A. thaliana* chromosome 4, equivalent to 30 cM. Estimates of copy number indicated that sequences present in a single copy in the haploid genome of *A. thaliana* ($n = 5$) were present in 2–8 copies in the haploid genome of *B. napus* ($n = 19$), while sequences present in multiple copies in *A. thaliana* were present in over 10 copies in *B. napus*. Genetic mapping in *B. napus* of DNA markers derived from a segment of *A. thaliana* chromosome 4 revealed duplicated homologous segments in the *B. napus* genome. Physical mapping in *A. thaliana* of homologues of *Brassica* clones derived from these regions confirmed the identity of six duplicated segments with substantial homology to the 7.5-Mbp region of chromosome 4 in *A. thaliana*. These six duplicated *Brassica* regions (on average 22cM in length) are collinear, except that two of the six copies contain the same large internal inversion. These results have encouraging implications for the feasibility of shuttling between the physical map of *A. thaliana* and genetic maps of *Brassica* species, for identifying candidate genes and for map based gene cloning in *Brassica* crops.

Keywords: *Arabidopsis thaliana*, *Brassica napus*, comparative mapping, collinearity.

Résumé : *Arabidopsis thaliana* (la plante modèle dicotylédone) est très apparentée aux espèces cultivées du genre *Brassica*. La colinéarité des génomes du *Brassica napus* (colza) et de l'*A. thaliana*, c'est-à-dire la conservation de l'ordre des marqueurs, a été scrutée dans un segment du bras long du chromosome 4 de l'*A. thaliana* mesurant 7,5 Mb et équivalent à 30 cM. Il a été estimé que des séquences qui sont présentes en une seule copie dans le génome haploïde de l'*A. thaliana* ($n = 5$) sont présentes en 2–8 copies dans le génome haploïde du *B. napus* ($n = 19$). Par ailleurs, il a été évalué que les séquences présentes en plusieurs copies chez l'*A. thaliana* sont présentes en plus de 10 copies chez le *B. napus*. La cartographie génétique de certains marqueurs provenant d'un segment du chromosome 4 de l'*A. thaliana* a révélé des segments homologues dupliqués chez le *B. napus*. La cartographie physique réalisée chez l'*A. thaliana* au moyen des séquences homologues aux clones du *Brassica* situés dans ces régions ont confirmé l'identité de six segments dupliqués montrant une homologie substantielle à la région de 7,5 Mb du chromosome 4 chez l'*A. thaliana*. Ces six régions dupliquées dans le génome du *Brassica* (mesurant en moyenne 22 cM) sont colinéaires à l'exception de deux d'entre elles qui possèdent une même inversion de grande taille. Ces résultats sont encourageants en vue de l'exploitation parallèle de la carte physique de l'*A. thaliana* et les cartes génétiques des espèces du genre *Brassica* pour l'identification de gènes et pour le clonage positionnel chez les espèces cultivées du genre *Brassica*.

Mots clés : *Arabidopsis thaliana*, *Brassica napus*, cartographie comparée, colinéarité.

[Traduit par la Rédaction]

Corresponding Editor: G. Jenkins.

Received June 3, 1997. Accepted September 18, 1997.

A.C. Cavell.[1] Department of Brassica and Oilseeds Research, John Innes Centre, Norwich Research Park, Colney, Norwich NR4 7UH, U.K. and Department of Cell Biology, Institute of Arable Crop Research (IACR), Long Ashton Research Station, Department of Agricultural Sciences, University of Bristol, Long Ashton, Bristol BS18 9AF, U.K.
D.J. Lydiate, I.A.P. Parkin, and M. Trick. Department of Brassica and Oilseeds Research, John Innes Centre, Norwich Research Park, Colney, Norwich NR4 7UH, U.K.
C. Dean. Department of Molecular Genetics, John Innes Centre, Norwich Research Park, Colney, Norwich NR4 7UH, U.K.

[1] Author to whom all correspondence should be sent at the following address: Department of Cell Biology, Institute of Arable Crop Research (IACR), Long Ashton Research Station, Department of Agricultural Sciences, University of Bristol, Long Ashton, Bristol BS18 9AF, U.K. (e-mail: alison.cavell@bbsrc.ac.uk).

Introduction

The *Brassica* species cultivated as important oilseed, vegetable, and fodder crops are closely related to the small cruciferous model plant *Arabidopsis thaliana* (Hedge 1976; Meyerowitz and Pruitt 1985). However, the genomes of diploid *Brassica* species are larger than that of *A. thaliana*, 468–662 Mbp compared with 145 Mbp, respectively (Arumuganathan and Earle 1991), and there is evidence for extensive internal duplications within the genomes of *Brassica oleracea* (Slocum et l. 1990) and *Brassica rapa* (Song et al. 1991; Chyi et al. 1992). Several *Brassica* crop species, including *B. napus* (oliseed rape or canola), are amphidiploid, apparently as a result of interspecific hybridization. For example, *B. napus* ($n = 19$) appears to have arisen from the hybridization of *B. rapa* (the *Brassica* A genome, $n = 10$) with *B. oleracea* (the *Brassica* C genome, $n = 9$) (U 1935; Parkin et al. 1995; Bohuon et al. 1996).

Arabidopsis thaliana is ideally suited to map-based gene cloning, because of its small genome size and the relative paucity of duplicated genes and repetitive elements (Meyerowitz and Pruitt 1985). Efforts are underway to generate ordered libraries of overlapping YAC (yeast artificial chromosome) and BAC (bacterial artificial chromosome) clones covering the whole *A. thaliana* genome, and an international programme to determine the complete DNA sequence has been initiated (Schmidt and Dean 1992; Schmidt et al. 1995). Comparative genetic mapping of several plant species, including members of the Gramineae (Kurata et al. 1994; Pereira et al. 1993) and the Solanaceae (Bonierbale et al. 1988; Prince et al. 1992), has demonstrated that conservation of marker order (collinearity) exists between genomes over considerable genetic intervals and across considerable taxonomic distances.

It has already been shown that short-range collinearity exists between the *A. thaliana* and *Brassica nigra* genomes over a region surrounding the *CO* locus on *A. thaliana* chromosome 5 (Lagercrantz et al. 1996). A general description of collinearity between the *Brassica* and *A. thaliana* genomes would make it possible to exploit ordered *A. thaliana* clone libraries and DNA sequence data as major resources for *Brassica* genome analysis. Comparative mapping between the two genera would also provide an insight into the mechanisms of genomic and karyotypic evolution, because disruptions in collinearity indicate the breakpoints of chromosomal rearrangements that have occurred during the divergence of species. The duplicated nature of the *B. napus* genome complicates collinearity studies and makes simple one-to-one relationships between the *B. napus* and *A. thaliana* genomes highly unlikely.

In this study, we assessed genome collinearity between *A. thaliana* and *B. napus* over a 7.5-Mbp region of the long arm of *A. thaliana* chromosome 4 (Schmidt et al. 1995), equivalent to a recombinational interval of 30 cM.

Materials and methods

Comparison of DNA sequences

All those genes for which the nucleotide sequences of homologues in both *A. thaliana* and *B. napus* were available were retrieved from the GenBank databases using common keywords or common E.C. (enzyme classification) numbers as the initial search criteria. The protein coding regions were assembled and aligned using the GCG BESTFIT program (Genetics Computer Group 1994) and, where multiple gene sequences existed for gene families in either species, the intergeneric pair with the greatest homology was selected.

Southern hybridization

Procedures for isolating genomic DNA, restriction enzyme digestion, gel electrophoresis, and Southern hybridization were as described by Sharpe et al. (1995), except that 15 μg of genomic DNA was used and final filter washes were carried out at a lower stringency (2× SSC (1× SSC: 0.15 M NaCl plus 0.015 M sodium citrate) plus 0.1% SDS at 65°C). These washing conditions approximated $T_m - 27$°C, allowing retention of hybrids with 27% sequence mismatch (Bonner et al. 1973; McConaughy et al. 1969). Hybridization probes were prepared by PCR or gel isolation of inserts (Sambrook et al. 1989) and were labelled with [^{32}P]dCTP (Feinberg and Vogelstein 1983).

Arabidopsis thaliana and *Brassica* clones used as probes

A set of 11 cloned fragments of the *A. thaliana* genome were used as RFLP (restriction fragment length polymophism) probes. Ten of these clones together spanned a 2.2-Mbp region flanking the *FCA* locus on *A. thaliana* chromosome 4: YAC end probes EW10D7LE, ABI6D3LE, and yUP3f7RE (Schmidt et al. 1995); cosmid subclones CC15D15, CC16N19, and g4539 (Schmidt et al. 1995); the *FCA* cDNA (C. Dean, personal communication); and random *Pst*I genomic fragments mi30, mi112, and mi330 (Ohio Stock Centre). The mi431 clone (Ohio Stock Centre) maps on *A. thaliana* chromosome 4 approximately 26 cM south of *FCA*. The *Brassica* clones were a subset of the RFLP probes described in Sharpe et al. (1995).

Estimating the copy number of sequences homologous to *A. thaliana* probes

The *A. thaliana* clones were used to probe multi-enzyme filters of *A. thaliana* and *B. napus*: that is, Southern hybridization filters carrying genomic DNA from two *A. thaliana* ecotypes (Landsberg erecta and Columbia) digested separately with 16 restriction enzymes, and also, filters carrying genomic DNA from the parents of the N-fo-61-9 mapping population of *B. napus* (Parkin et al. 1995) digested separately with *Eco*RI, *Bam*HI, *Eco*RV, *Hin*dIII, and *Xba*I.

Genetic mapping in *B. napus*

The *A. thaliana* clones were used to probe filters carrying genomic DNA from 57 doubled haploid lines of the highly polymorphic N-fo-61-9 mapping population of *B. napus* (Parkin et al. 1995; Parkin and Lydiate 1997). Polymorphic RFLP loci detected by the *A. thaliana* probes were placed on the genetic map of *B. napus*, relative to the loci mapped previously, using simple linkage analysis followed by manual fine adjustment based on minimising the number of double crossovers flanking the new loci.

Hybridization of *Brassica* probes to YACs of *A. thaliana* chromosome 4

A minimal set of 24 overlapping YAC clones, covering a 7.5-Mbp segment of *A. thaliana* chromosome 4, was kindly selected and made available by Dr. R. Schmidt (John Innes Centre, Norwich, U.K). Total yeast DNA was isolated using a colony scrape procedure. Yeast colonies were grown on solid YEPD medium (10 g/L yeast extract, 20 g/L peptone, 20 g/L glucose, 100 mg/L adenine, plus 2% agar) at 30°C, until large colonies had formed. Colonies were scraped from the plates and resuspended in 300 μL TE–SDS (0.4% SDS, 75 mM Tris–HCl (pH 7.5), plus 40 mM EDTA (pH 8.0)).To this, 400 μL phenol – chloroform – isoamyl alcohol (25:24:1 (v/v)) was added, and the samples were vortexed and then incubated at 65°C for 30 min. After centrifugation at 14 000 rpm for 10 min, the supernatant was recovered and ethanol added (10 min at room temperature) to precipitate the DNA. The DNA was recovered by centrifugation for 30 min at 14 000 rpm and resuspended in 100 μL of water. Samples were

Table 1. Computed sequence conservation at nucleotide and anino acid levels of homologous genes in *A. thaliana* and *B. napus*.

Gene product	GenBank accession numbers	% nucleotide identity	% amino acid identity
Acetyl CoA carboxylase	L27074; X77374	91	90
Acetolactate synthase	X51514; X16708	78	80
Acyl carrier protein	X13708; X16114	84	80
Cyclophilin	Z17785; M55018	83	88
Fructose 1,6-biphosphatase	Z27269; U20179	82	77
Glutamine synthetase	S69727; X72751	90	93
Isocitrate lyase	Z18772; L08482	91	96
Methionine sulphoxide reductase	X97326; X94225	85	84
Myrosinase	X79194; Z21978	80	72
Nitrate reductase	X13434; D38220	86	90
Omega-6 desaturase (plastidial)	U09503; L29214	91	93
PEP carboxykinase	Z26739; U21745	94	95
Rubisco small subunit	X13610; X55937	91	93

Table. 2. Estimated copy numbers for the sequences in the *A. thaliana* and *B. napus* genomes that are homologous to probes derived from the 2.2-Mbp region of *A. thaliana* chromosome 4 (Fig. 1).

A. thaliana clone	Copy number		No. of polymorphic loci[c]	*B. napus* linkage groups [d]
	A. thaliana [a]	*B. napus*[b]		
mi30	1	4	2	N8, N18
EW10D7LE	2	13	3	N1, N14, N5
yUP3F7RE	1	6	3	N1, N3, N17
FCA	1	2	2	N1, N11
ABI6D3LE	1	8	3	N3, N17, N11
CC16N19	1	6	2	N1, N11
CC15D15	1	6	2	N1, N11
g4539	1	4	1	N18
mi112	1	6	4	N1, N11, N8, N18
mi330	3	10	4	N1, N11, N5, N8

[a]Estimates based on hybridization to genomic *A. thaliana* DNA digested separately with 16 different restriction enzymes.

[b]Estimates based on hybridization to genomic *B. napus* DNA digested separately with the restriction enzymes *Bam*HI, *Eco*RI, *Eco*RV, *Hin*dIII, and *Xba*I.

[c]The number of loci polymorphic in the N-fo-61-9 population of *B. napus*, using genomic DNA digested with *Bam*HI, *Eco*RI, *Eco*RV, *Hin*dIII, or *Xba*I.

[d]Groups on which the polymorphic loci were mapped; N1–N19 (Fig. 2).

re-extracted with phenol – chloroform – isoamyl alcohol, precipitated with ethanol, and finally resuspended in TE (10 mM Tris–HCl plus 1 mM EDTA). Yeast DNA was digested with *Bam*HI and *Eco*RI, separated using conventional gel electrophoresis, and alkaline blotted onto Hybond N^+ membranes (Amersham). *Brassica* RFLP probes were labelled with [^{32}P]dCTP and hybridized to digested immobilised yeast DNAs. Washes were carried out at the same stringency as for the genomic Southern blots.

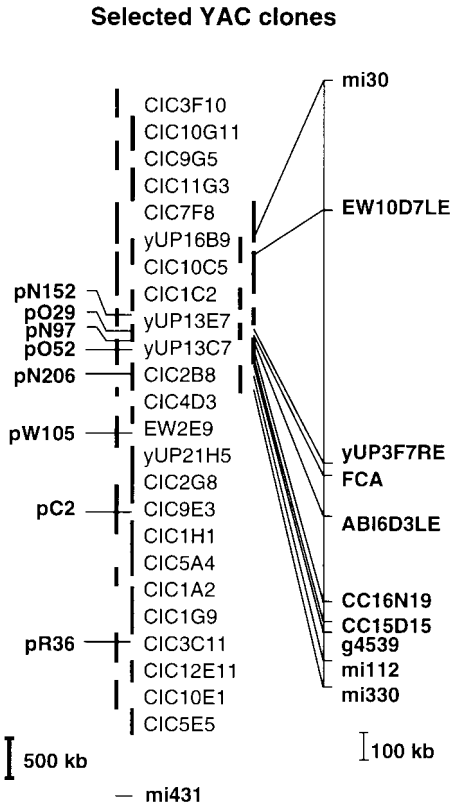
Results

Homology between *A. thaliana* and *B. napus*

Database comparisons of homologous gene sequences from *A. thaliana* and *B. napus* showed an average 87% identity over their exons (Table 1). This indicated that the degree of intergeneric sequence conservation in coding regions was likely to be universally high and allowed the selection of Southern hybridization conditions that would be expected to yield a signal from sequences exhibiting only half the degree of conservation normal for exons.

Ten *A. thaliana* clones, all derived from a small segment of chromosome 4 (Fig. 1), were used to probe multi-enzyme filters of *A. thaliana* and *B. napus*. All 10 probes gave good Southern hybridization signals against *Brassica* DNA, despite the fact that only one clone (*FCA*) was known to represent a coding sequence. The genomic copy number in both *A. thaliana* and *B. napus* of sequences homologous to the 10 *A. thaliana* clones was estimated from their hybridization patterns on the multi-enzyme filters, and the results are presented in Table 2. While most of the clones were apparently single copy in *A. thaliana*, all clones except *FCA* were present at considerably higher copy numbers in *Brassica*.

Fig. 1. Physical map positions of RFLP probes derived from *A. thaliana* and *Brassica* on the 7.5 Mbp YAC contig of the long arm of *A. thaliana* chromosome 4. Thick vertical lines denote the positions and relative lengths of the 24 selected YAC clones (CIC3F10–CIC5E5) that constitute the minimum tiling set for the 7.5-Mbp (30cM) contig on the long arm of *A. thaliana* chromosome 4 (contig III; Schmidt et al. 1995). The thin vertical line to the right of the YAC contig represents the fine-scale map positions of the 10 *A. thaliana* probes (mi30–mi330) used for RFLP mapping in *B. napus* (Fig. 2). The physical map positions on *A. thaliana* chromosome 4 of sequences homologous to eight *Brassica* probes (pN152–pR36) are indicated to the left of the YAC contig.



Genetic mapping in *B. napus* of loci homologous to probes from a 2.2-Mbp region of *A. thaliana* chromosome 4

To investigate the distribution of *B. napus* loci homologous to probes from a physically well-defined segment of *A. thaliana* chromosome 4, 10 *A. thaliana* clones from the 2.2-Mbp (9 cM) region immediately flanking *FCA* (Fig. 1) were used to probe the N-fo-61-9 mapping population. The *A. thaliana* probes detected both monomorphic and polymorphic loci in *B. napus* (Table 2). The RFLP-defined loci detected by the *A. thaliana* probes were positioned on the genetic map of the *B. napus* genome and were obviously arranged in coincident or tightly linked clusters on linkage groups N1, N3, N5, and N8 of the *Brassica* A genome and linkage groups N11, N14, N17, and N18 of the *Brassica* C genome (Fig. 2). Linkage groups N1 and N11, N3 and N17, N5 and N14, and N8 and N18 are known to be partial homoeologues (Lydiate et al. 1995). The only loci that failed to follow the general pattern of tight clustering were the loci on N1 and N11 that were homologous to mi330 (Fig. 2).

Long range collinearity between *A. thaliana* chromosome 4 and its *B. napus* homologues

To investigate further the apparently anomalous map positions of the *Brassica* loci on N1 and N11 that were homologous to mi330, and to better gauge the extent of the collinearity between *A. thaliana* chromosome 4 and its *B. napus* homologues, *Brassica* RFLP probes that detected loci on N1 in the interval flanked by loci homologous to *FCA* and mi330 were used to probe an overlapping set of *A. thaliana* YACs that span an extended portion of chromosome 4. The selected *Brassica* probes were pN97, pN152, pN206, pO29, pO52, pR36, pC2, and pW105 (Fig. 2; Sharpe et al. 1995), and all eight probes detected strongly homologous sequences within the chromosome 4 YAC contig (Figs. 1 and 3). The eight *Brassica* probes detected *A. thaliana* homologues distributed over almost the entire length of the extended contig, 7.5 Mbp or 30 cM in *A. thaliana*. Comparative mapping between *A. thaliana* chromosome 4 and the homoeologous pair of *B. napus* linkage groups, N1 and N11, revealed a large inversion in one with respect to the other (Fig. 3). In contrast, large scale comparative mapping between *A. thaliana* chromosome 4 and another pair of homoeologous chromosome segments of *B. napus* (on linkage groups N3 and N17) indicated perfect genome collinearity across the region assayed. The existence of an inversion on N1 and N11 with respect to both *A. thaliana* chromosome 4 and N3 and N17, suggests that the inversion took place in a *Brassica* chromosome that was the common ancestor of N1 and N11. Two *Brassica* probes (pW145 and pW225; Fig. 2) that detected loci mapping below the *FCA* locus on N1 (using the *A. thaliana* chromosome orientation) were used to probe the extended chromosome 4 contig in an attempt to define the bottom border of the inversion. Neither pW145 nor pW225 detected homologous sequences within the chromosome 4 YAC contig. An *A. thaliana* clone from immediately below the pR36 locus on *A. thaliana* chromosome 4 (mi431) was then used as an RFLP probe on the N-fo-61-9 mapping population of *B. napus*. The probe detected loci on N1 and N11 and thus defined the bottom of the inversion.

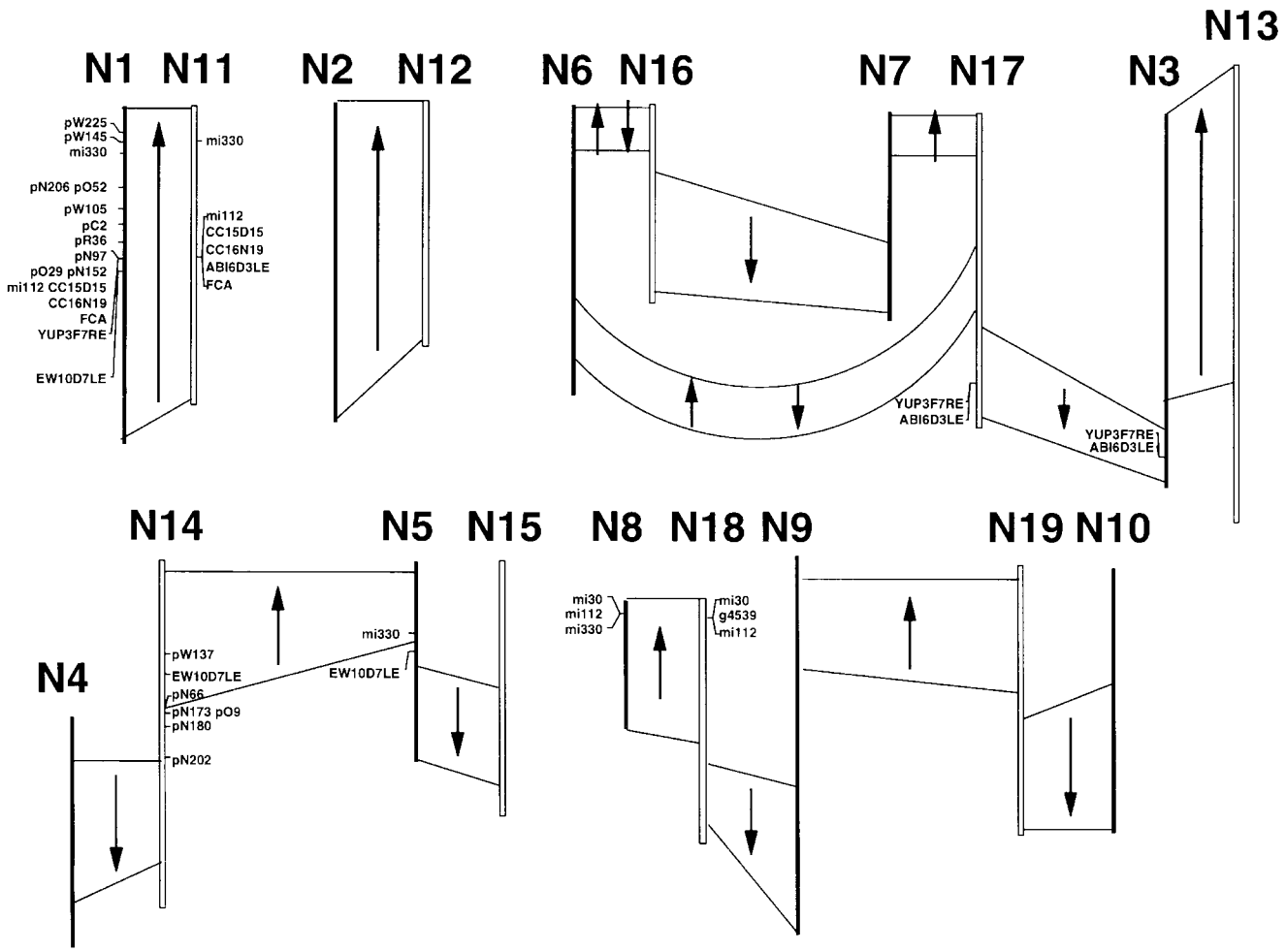
Six *B. napus* probes (namely, pN66, pN173, pN180, pN202, pO9, and pW137; Fig. 2) that detected loci on N14 flanking the locus homologous to EW10D7LE were used to probe the extended chromosome 4 contig, but none of the probes detected homologous sequences.

Discussion

DNA sequence homology between *A. thaliana* and *Brassica*

The high level of sequence conservation in protein coding regions between *A. thaliana* and *Brassica* (87% on average; Table 1) makes reciprocal Southern hybridization with cDNA probes, and hence comparative genetic mapping, between *A. thaliana* and *Brassica* feasible. Indications are that even anonymous genomic fragments cross hybridize between *A. thaliana* and *Brassica* to produce RFLP markers that are informative in comparative mapping (Lagercrantz et al. 1996; Figs. 1 and 2).

Fig. 2. Schematic map of the *B. napus* genome (redrawn from Lydiate et al. 1995) showing the map positions of *B. napus* loci homologous to the 10 *A. thaliana* probes from the 2.2 Mbp *FCA* region of chromosome 4 (mi30–mi330; Fig. 1). Solid vertical bars represent the 10 A-genome linkage groups of *B. napus* (N1–N10); open vertical bars represent the nine C-genome linkage groups of *B. napus* (N11–N19); arrows indicate the homoeologous relationships between chromosome segments from the A and C genomes within the amphidiploid *B. napus* genome (Parkin et al. 1995; Lydiate et al. 1995). The map positions on N1 of loci detected by 10 *Brassica* clones (pW225–pN152) used to probe contig III of *A. thaliana* chromosome 4, and the map positions on N14 of loci detected by a further 6 *Brassica* clones (pW137–pN202) also used to probe the *A. thaliana* contig, are also indicated.

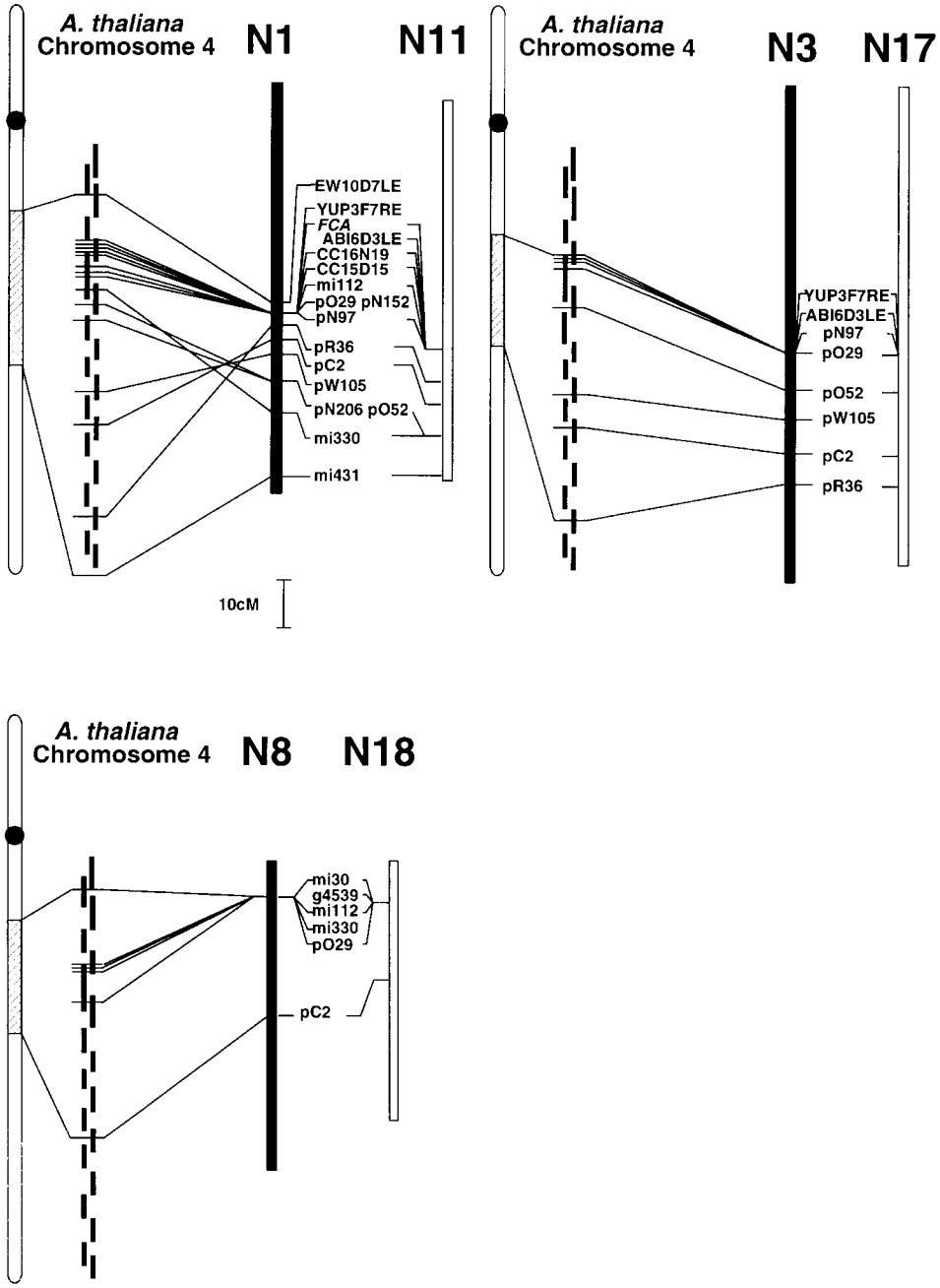


Duplicate regions of genome collinearity in *B. napus*

Loci homologous to particular DNA probes have been placed accurately on the genetic map of *B. napus,* and the same probes have been anchored to individual YACs that are both physically mapped in *A. thaliana* and integrated into the genetic map. The results demonstrate that there are six duplicate and extensively collinear regions of the *B. napus* genome, each equivalent to a 30-cM (7.5-Mbp) region of *A. thaliana* chromosome 4, and each of approximately 20–30 cM in *B. napus* (Fig. 3). *Brassica napus* is an amphidiploid–allotetraploid species formed from the interspecific hybridization of two diploid *Brassica* species, *B. rapa* (the A genome) and *B. oleracea* (the C genome) (U 1935; Parkin et al. 1995). The six collinear regions of the *B. napus* genome can be assigned to three homoeologous pairs of chromosomal segments, each with one copy in the A genome and the other copy in the C genome (Lydiate et al. 1995; Figs. 2 and 3). There is evidence that a 30-cM segment of the upper arm of *A. thaliana* chromosome 3 also corresponds to six duplicate collinear segments of the amphidiploid *B. napus* genome (Scheffler et al. 1997) and that a 7.5-cM (1.5-Mbp) segment of *A. thaliana* chromosome 5 corresponds to three duplicate collinear segments of the diploid *B. nigra* genome (Lagercrantz et al. 1996). Collinearity between a region of the top arm of *A. thaliana* chromosome 4 and its *B. napus* homologues was found to be more fragmented (Osborn et al. 1997), indicating that more rearrangement may have occured in some regions.

The regions on linkage groups N1 and N11 of *B. napus* that are related to the lower arm of *A. thaliana* chromosome 4 have a simple internal inversion (spanning approximately 20 cM) with respect to the equivalent regions in *A. thaliana* and in linkage groups N3, N17, N8, and N18 (Fig. 3). The collinear structures conserved between *Brassica* and *A. thaliana* are likely to predate the novel *Brassica* structure, suggesting that the inversion occurred after the divergence of *Brassica* and *A. thaliana* from a common ancestor, but before the divergence of *B. rapa* and *B. oleracea* from their closest common ancestor. An inversion that distinguishes one of three *B. nigra*

Fig. 3. Comparative maps of the 7.5-Mbp region of *A. thaliana* chromosome 4 and its six duplicate substantially collinear homologous segments in the *B. napus* genome. The *B. napus* linkage groups are arranged in intergenomic homoeologous pairings (Fig. 2). The 19 *A. thaliana* YAC clones (short vertical bars) are drawn as depicted in Fig. 1, except that the top 4 YACs and the bottom YAC (of the 24) have been omitted.



regions collinear with a segment of *A. thaliana* chromosome 5 has also been reported (Lagercrantz et al. 1996).

The *FCA* cDNA hybridized to only two restriction fragments in *B. napus,* in contrast to all the other probes derived from the *A. thaliana* chromosome 4 contig that hybridized to four or more restriction fragments in *B. napus* (Table 2). The loci homologous to *FCA* were mapped to homoeologous locations in linkage groups N1 and N11, while probes derived from chromosome fragments closely flanking *FCA* in *A. thaliana* (namely, yUP3F7RE and ABI6D3LE) detected coincident loci in linkage groups N1, N11, N3, and N17 (Figs. 1, 2, and 3), in addition to a number of monomorphic *B. napus* DNA fragments: three in the case of yUP3F7RE and five in the case of ABI6D3LE (Table 2). The fact that only two *FCA* homologues exist in *B. napus* is probably the result either of one or more small ancestral deletions or of accelerated sequence divergence involving the additional copies of *FCA* that would otherwise be anticipated from the general pattern of genome duplication.

Copy number, polymorphism, and genome duplication

The above investigation, together with those of Lagercrantz

et al. (1996) and Scheffler et al. (1997), suggest a general pattern of *Brassica* genome organisation where diploid *Brassica* genomes contain three representations of a basic genome (roughly equivalent in size to the genome of *A. thaliana*), with each representation being extensively collinear with the genome of *A. thaliana*, that is, exhibiting collinear stretches of chromosome 20 cM in length. The work described by Kowalski et al. (1994) is superficially at variance with this general pattern, but this apparent anomaly is probably the result of limited polymorphism and low marker densities in this early investigation, leading to a systematic underestimation of genome duplication in *Brassica*.

Using the highly polymorphic N-fo-61-9 population and *B. napus* DNA digested with *Eco*RI, *Bam*HI, *Eco*RV, *Hin*dIII, or *Xba*I, the eight single copy *A. thaliana* probes detected only 19 polymorphic loci (Table 2; Fig. 2), even though a total of 42 homologous regions of the *B. napus* genome were predicted (Table 2). This probably reflects a situation where only 45% of loci were polymorphic. As explained by Lagercrantz and Lydiate (1996), limited polymorphism reduces the appreciation of genome duplication (in a cross where 70% of loci are polymorphic, all three representations of a triplicated locus will only be mapped in 34% of cases). However, since 23 of a total of 26 polymorphic loci mapped to six discrete regions of the *B. napus* genome (Table 2), that together comprise less than 7% of the entire genetic map (Fig. 2), it is likely that almost all the monomorphic, and therefore unmapped, loci actually reside in the same six regions.

Only the two *A. thaliana* probes that were apparently homologous to duplicated sequences in *A. thaliana* (EW10D7LE and mi330; Table 2) detected *B. napus* loci in linkage groups N5 and N14, outside the six prominent regions of genome collinearity (Figs. 2 and 3). The fact that *Brassica* probes that detected loci flanking those detected by EW10D7LE and mi330 in *B. napus* linkage groups N5 and N14 failed to detect any loci in the prominent regions of genome collinearity on linkage groups N1, N11, N3, N17, N8, and N18, or within the *A. thaliana* chromosome 4 YAC contig, suggests that the loci homologous to EW10D7LE and mi330 in *B. napus* linkage groups N5 and N14 are likely to be related to duplicated loci within the *A. thaliana* genome that map to regions outside the chromosome 4 contig.

Practical implications of *Brassica – A. thaliana* genome collinearity

The established collinearity between large portions of the *Brassica* and *A. thaliana* genomes, often over genetic distances in excess of 10 cM, will promote the efficient utilisation of physically mapped clones and sequence data from *A. thaliana* to assist in genetic mapping and gene isolation in *Brassica* crops. Once the comparative genetic mapping of the *Brassica* and *A. thaliana* genomes has been completed, and assuming that the level of collinearity observed in this and other investigations (Lagercrantz et al. 1996; Scheffler et al. 1997) exists over the whole genome, it will be possible to determine which part of the *A. thaliana* genome is homologous to any specific interval of the *B. napus* genome. In the regions of the *B. napus* genome that are collinear with the *A. thaliana* chromosome 4 contig, 1 cM of the *B. napus* genetic map is equivalent to approximately 250 kbp of the *A. thaliana* physical map, a distance that can be covered by 1–2 YAC clones. The pattern and frequency of recombination is highly conserved between *B. napus* crosses (Parkin and Lydiate 1997), but the precise physical relationships between the *A. thaliana* and *B. napus* genomes, and in particular the relative gene spacings, remain to be established and may well vary from region to region.

The duplicated nature of *Brassica* genomes has several ramifications, including the possibility that the majority of traits in *B. napus* are likely to be controlled by interactions between homoeologous sets of duplicated genes (Scheffler et al. 1997). This will make it imperative that major loci controlling agronomic traits are precisely defined in *Brassica* before using information from *A. thaliana* for the cloning of candidate genes.

Acknowledgements

This work was supported by a Biotechnology and Biological Sciences Research Council (BBSRC) Competitive Strategic Grant to the John Innes Centre and by the BBSRC Plant and Animal Genome Analysis Programme (PAG01526). AC was supported by a John Innes Foundation postgraduate studentship. We thank Dr. Renate Schmidt and Melanie Stammers for assistance with the *A. thaliana* and YAC methodologies.

References

Arumuganathan, K., and Earle, E.D. 1991. Nuclear DNA content of some important plant species. Plant Mol. Biol. Rep. **9**: 208–218.

Bohuon, E.J.R., Keith, D.J., Parkin, I.A.P., Sharpe, A.G., and Lydiate, D.J. 1996. Alignment of the conserved C genomes of *Brassica oleracea* and *Brassica napus*. Theor. Appl. Genet. **93**: 833–839.

Bonierbale, M.W., Plaisted, R.L., and Tanksley, S.D. 1988. RFLP maps based on a common set of clones reveal modes of chromosomal evolution in potato and tomato. Genetics, **120**: 1095–1103.

Bonner, T.I., Brenner, D.J., Neufeld, B.R., and Britten, R.J. 1973. Reduction in the rate of DNA reassociation by sequence divergence. J. Mol. Biol. **81**: 123–135.

Chyi, Y.S., Hoenecke, M.E., and Sernyk, J.L. 1992. A genetic linkage map of restriction fragment length polymorphism loci for *Brassica rapa* (syn. *campestris*). Genome, **35**: 746–757.

Feinberg, A.P., and Vogelstein, B. 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132**: 6–13.

Genetics Computer Group 1994. Program manual for the Wisconsin Package, version 8. Genetics Computer Group, Madison, Wis.

Hedge, I.C. 1976. A systematic and geographical survey of the old world Cruciferae. *In* The biology and chemistry of the Cruciferae. *Edited by* A.J. Macleod and B.M.J. Jones. Academic Press, London. pp. 1–46.

Kowalski, S.P., Lan, T.-H., Feldmann, K.A., and Paterson, A.H. 1994. Comparative mapping of *Arabidopsis thaliana* and *Brassica oleracea* chromosomes reveals islands of conserved organization. Genetics, **138**: 1–12.

Kurata, N., Moore, G., Nagamura, Y., Foote, T., Yano, M., Minobe, Y., and Gale, M.D. 1994. Conservation of genome structure between rice and wheat. Bio/Technology (New York), **12**: 276–278.

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

Lagercrantz, U., Puterill, J., Coupland, G., and Lydiate, D.J. 1996. Comparative mapping in *Arabidopsis* and *Brassica*, fine scale genome collinearity and congruence of genes controlling flowering time. Plant J. **9**: 13–20.

Lydiate, D., Dale, P., Lagercrantz, U., Parkin, I., and Howell, P. 1995.

Selecting the optimum genetic background for transgenic varieties, with examples from *Brassica*. Euphytica, **85**: 351–358.
McConaughy, B.L., Laird, C.D., and McCarthy, B.J. 1969. Nucleic acid reassociation in formamide. Biochemistry, **8**: 3289–3295.
Meyerowitz, E.M., and Pruitt, R.E. 1985. *Arabidopsis thaliana* and plant molecular genetics. Science (Washington, D.C.), **229**: 1214–1218.
Osborn, T.C., Kole, C., Parkin, I.A.P., Sharpe, A.G., Kuiper, M., Lydiate, D.J., and Trick, M. 1997. Comparison of vernalization-responsive flowering time genes in *Brassica rapa*, *B. napus* and *Arabidopsis thaliana*. Genetics, **146**: 1123–1130.
Parkin, I.A.P., Sharpe, A.G., Keith, D.J., and Lydiate, D.J. 1995. Identification of the A and C genomes of amphidiploid *Brassica napus* (oilseed rape). Genome, **38**: 1122–1131.
Parkin, I.A.P., and Lydiate, D.J. 1997. Conserved patterns of chromosome pairing and recombination in *Brassica napus* crosses. Genome, **40**: 496-504.
Pereira, M.G., Lee, M., Bramel-Cox, P., Woodman, W., Doebley, J., and Whitkus, R. 1993. Construction of an RFLP map in sorghum and comparative mapping in maize. Genome, **37**: 236–243.
Prince, J.P., Pochard, E., and Tanksley, S.D. 1992. Construction of a molecular linkage map of pepper and a comparison of synteny with tomato. Genome, **36**: 404–417.
Sambrook, J., Fritsch, E.F., and Maniatis, T. 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
Scheffler, J.A., Sharpe, A.G., Schmidt, H., Sperling, P., Parkin, I.A.P., Lühs, W., Lydiate, D.J., and Heinz, E. 1997. Desaturase multigene families of *Brassica napus* arose through genome duplication. Theor. Appl. Genet. **94**: 583–591.
Schmidt, R., and Dean, C. 1992. Physical mapping of the *Arabidopsis thaliana* genome. *In* Genome analysis. Vol. 4. Strategies for physical mapping. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York. pp. 71–98.
Schmidt, R., West, J., Love, K., Lenehan, Z., Lister, C., Thompson, H., Bouchez, D., and Dean, C. 1995. Physical map and organization of *Arabidopsis thaliana* chromosome 4. Science (Washington, D.C.), **270**: 480–483.
Sharpe, A.G., Parkin, I.A.P., Keith, D.J., and Lydiate, D.J. 1995. Frequent nonreciprocal translocations in the amphidiploid genome of oilseed rape (*Brassica napus*). Genome, **38**: 1112–1121.
Slocum, M.K., Figdore, S.S., Kennard, W.C., Suzuki, J.Y., and Osborn, T.C. 1990. Linkage arrangement of restriction fragment length polymorphism loci in *Brassica oleracea*. Theor. Appl. Genet. **80**: 57–64.
Song, K.M., Suzuki, J.Y., Slocum, M.K., Williams, P.H., and Osborn, T.C. 1991. A linkage map of *Brassica rapa* (syn. *campestris*) based on restriction fragment length polymorphism loci. Theor. Appl. Genet. **82**: 296–304.
U, N. 1935. Genomic analysis in *Brassica* with special reference to the experimental formation of *B. napus* and peculiar mode of fertilisation. Jpn. J. Bot. **7**: 389–452.