Theor Appl Genet (2007) 115:777–792
DOI 10.1007/s00122-007-0608-z

ORIGINAL PAPER

The reference genetic linkage map for the multinational *Brassica rapa* genome sequencing project

Su Ryun Choi · Graham R. Teakle · Prikshit Plaha · Jeong Hee Kim · Charlotte J. Allender · Elena Beynon · Zhong Yun Piao · Pilar Soengas · Tae Ho Han · Graham J. King · Guy C. Barker · Paul Hand · Derek J. Lydiate · Jacqueline Batley · David Edwards · Dal Hoe Koo · Jae Wook Bang · Beom-Seok Park · Yong Pyo Lim

Received: 17 March 2007 / Accepted: 2 July 2007 / Published online: 24 July 2007

Abstract We describe the construction of a reference genetic linkage map for the *Brassica* A genome, which will form the backbone for anchoring sequence contigs for the Multinational *Brassica rapa* Genome Sequencing Project. Seventy-eight doubled haploid lines derived from anther culture of the F_1 of a cross between two diverse Chinese cabbage (*B. rapa* ssp. *pekinensis*) inbred lines, 'Chiifu-401-42' (C) and 'Kenshin-402-43' (K) were used to construct the map. The map comprises a total of 556 markers, including 278 AFLP, 235 SSR, 25 RAPD and 18 ESTP, STS and CAPS markers. Ten linkage groups were identified and designated as R1–R10 through alignment and orientation using SSR markers in common with existing *B. napus* reference linkage maps. The total length of the linkage map was 1,182 cM with an average interval of 2.83 cM between adjacent loci. The length of linkage groups ranged from 81 to 161 cM for R04 and R06, respectively. The use of 235 SSR markers allowed us to align the A-genome chromosomes of *B. napus* with those of *B. rapa* ssp. *pekinensis*. The development of this map is vital to the integration of genome sequence and genetic information and will enable the international research community to share resources

Communicated by H. C. Becker.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-007-0608-z) contains supplementary material, which is available to authorized users.

S. R. Choi · P. Plaha · J. H. Kim · Y. P. Lim (✉)
Department of Horticulture, Genome Research Center,
Chungnam National University, Kung-Dong 220,
Yusong-Gu, Daejeon 305-764, South Korea
e-mail: yplim@cnu.ac.kr

Z. Y. Piao
College of Horticulture, Shenyang Agricultural University,
120 Dongling Road, Shenyang, Liaoning, China

G. R. Teakle · C. J. Allender · G. C. Barker · P. Hand
Warwick HRI, University of Warwick,
Wellesbourne, Warwick CV35 9EF, UK

G. J. King
Centre for Crop Genetic Improvement, Rothamsted Research,
Harpenden, Hertfordshire AL5 2JQ, UK

E. Beynon · D. J. Lydiate
Saskatoon Research Centre, Agriculture and Agri-Food Canada,
107 Science Place, Saskatoon, SK, Canada S7N 3V3

T. H. Han
Department of Applied Plant Science,
Chonnam National University,
Gwangju 500-757, South Korea

J. Batley · D. Edwards
Plant Biotechnology Centre,
Primary Industries Research Victoria,
Bundoora, VIC 3086, Australia

D. H. Koo · J. W. Bang
Department of Biology,
Chungnam National University,
Daejeon 305-764, South Korea

B.-S. Park
Brassica Genomics Team,
National Institute of Agricultural Biotechnology,
Suwon 441-707, South Korea

P. Soengas
Misión Biológica de Galicia (CSIC), Apartado 28,
36080 Pontevedra, Spain

P. Plaha
Advanced Centre of Hill Bioresources and Biotechnology,
HP Agricultural University, Palampur 176062, India

and data for the improvement of *B. rapa* and other cultivated *Brassica* species.

Introduction

The genus *Brassica* is one of the core genera in the subtribe *Brassicinae* and includes a number of crops with a wide spectrum of adaptation for cultivation under varied agro climatic conditions. Of the six cultivated species of *Brassica*, *B. rapa* (AA, $2n = 20$), *B. nigra* (BB, $2n = 16$) and *B. oleracea* (CC, $2n = 18$), are monogenomic diploids which have contributed by hybridization to the allopolyploids (amphidiploid) *B. juncea* (AABB, $2n = 36$), *B. napus* (AACC, $2n = 38$) and *B. carinata* (BBCC, $2n = 34$). Besides providing oil, vegetables, fodder and condiment crops, *Brassicas* are the important sources of dietary fibre, vitamin C and other nutritionally beneficial factors such as anticancer compounds (Fahey and Talalay 1995). The *Brassica* A genome is of major economic and calorific importance, contributing ~12% of the world-wide edible vegetable oil supplies via the crops *B. napus* (oilseed rape, Canola) and *B. juncea* (mustard oil). *B. rapa* comprises a number of morphologically diverse crops, including Chinese cabbage, Pakchoi, turnip and broccoletto, as well as oilseeds that include yellow and brown Sarsons (Gomez-Campo and Prakash, 1999). Together, these crops contribute significantly towards world food and fodder production.

Assigning molecular markers to the linkage groups and constructing genetic maps is an important step towards analyzing the genomes of crop species. Such maps provide a better insight into genome organization, evolution of the crop species, and synteny with related species. In addition, the maps are useful for tagging and cloning genes of economically important traits, marker assisted breeding and gene pyramiding.

A number of genetic linkage maps based on a range of marker types, including Restriction Fragment Length Polymorphisms (RFLPs), Random Amplified Polymorphic DNA (RAPD), Simple sequence Repeats (SSRs) and Amplified Fragment Length Polymorphisms (AFLPs), have been produced for *B. rapa* (Song et al. 1991; Chyi et al. 1992; Teutonico and Osborn, 1994; Nozaki et al. 1997; Kim et al. 2006; Suwabe et al. 2006; Soengas et al. 2007). PCR-based markers have been widely used in the construction of genetic linkage maps for *B. oleracea* (Cheung et al. 1997; Sebastian et al. 2000), *B. nigra* (Truco and Quiros 1994), *B. juncea* (Pradhan et al. 2003) and *B. napus* (Foisset et al. 1996; Lombard and Delourme 2001; Piquemal et al. 2005). However, there is only limited published data on sequence-tagged PCR-based markers mapped in *B. rapa* (Suwabe et al. 2006), especially markers that may provide anchors to the *B. rapa* genome and are readily transferable to other populations.

The nomenclature and orientation of many earlier *Brassica* linkage groups, including those of different linkage maps for the same species, has to some extent been arbitrary and hence published maps are often difficult to compare. In the context of whole genome sequencing, it is essential to align and orientate the linkage maps that relate to the same genome, and to adopt a common linkage group nomenclature, allowing establishment of reference linkage maps. The A and C genome linkage group nomenclature described by Parkin et al. (1995) for *B. napus* has been adopted for a number of other maps of the C-genome of *B. oleracea*, including Sebastian et al. (2000) and Howell et al. (2002), and for the A-genome of *B. rapa* (Suwabe et al. 2006; Kim et al. 2006). There is a requirement for unambiguous assignment of marker loci associated with sequence-tagged marker assays and the ability to identify data provenance. In order to achieve this, it is important to identify common sequence-tagged anchor loci. SSR markers are readily transferable since they exhibit a high degree of polymorphism and locus-specificity. For *B. rapa* 228 SSRs have recently been described, of which 90% successfully amplified in other *Brassica* species, and 40% amplified in Arabidopsis (Suwabe et al. 2002, 2004). SSRs have also been described for other *Brassica* species and Arabidopsis and a summary of these can be found at the *Brassica* microsatellite information exchange (http://www.brassica.info/ssr/SSRinfo.htm).

Chinese cabbage (*B. rapa* ssp. pekinensis) is considered a typical representative of the *Brassica* A genome (Lim et al. 2006). Due to the small genome size relative to the other *Brassica* species together with the extensive genetics and genomics resources available, it has been adopted as the subject for the first whole genome sequence arising from the Multinational *Brassica* Genome Project (MBGP). The Chinese cabbage inbred line Chiifu-401-42 has been selected as the sequencing template. In this study, we report the construction of a reference molecular genetic linkage map of *B. rapa* ssp. *pekinensis* which is aligned to the A genome map of *B. napus* (Parkin et al. 1995) and describe its relationship to the Arabidopsis genome. It has been agreed that in the MBGP (see www.brassica.info), the sequencing contigs will be aligned using this map. The map is constructed using Simple Sequence Repeats (SSR), Amplified Fragment Length Polymorphism (AFLP), Random Amplified Polymorphic DNA (RAPD), Sequence Tagged Sites (STS), Expressed Sequence Tag Polymorphism (ESTP), and Cleaved Amplified Polymorphism Sequence (CAPS) markers segregating in a doubled haploid (DH) population derived from the F_1 involving inbred lines, 'Chiifu-401-42' and 'Kenshin-402-43'.

Materials and methods

Plant material and DNA isolation

Two diverse inbred lines of Chinese cabbage (*Brassica rapa* ssp. *pekinensis*), 'Chiifu-401-42' and 'Kenshin-402-43', were used as the parents for developing a mapping population. Seventy-eight doubled haploid (DH) lines, derived from anther culture of five F_1 plants ('Chiifu' × 'Kenshin') were used as the mapping population, named CKDH. DNA was isolated from fresh leaf material of the parental and 78 DH lines following the procedure of Guillemaut and Laurence (1992) or by using the Qiagen 96-DNeasy 96 Plate Kit.

Analysis of molecular markers

AFLP

The original AFLP procedure as described by Vos et al. (1995) was followed with minor modifications. Genomic DNA was digested with two different enzyme combinations, *Eco*RI/*Mse*I and *Pst*I/*Mse*I, in a total volume of 25 µl at 37°C overnight. After inactivation of restriction enzymes at 75°C, adapters were ligated in a total volume of 50 µl at 37°C for 3 h. A 5 µl aliquot of tenfold dilution of the restriction/ligation mixture was pre-amplified with primers complementary to the adaptor sequence and including one additional selective nucleotide at the 3′ end of an *Eco*RI or *Pst*I directed primer (E + A or P + G) and an *Mse*I directed primer (M + C). Pre-amplification reaction contained 2.5 mM dNTPs, 1XPCR buffer (1.5 mM $MgCl_2$), 30 ng of each primer, and 1 unit *Taq* DNA polymerase. The amplicon was diluted 50-fold and used as template DNA in selective amplification reactions, which were carried out using with one or two primers additional selective nucleotides at the 3′ end for all the enzyme digests. The number following E, P corresponds to the selective nucleotide(s) at the 3′ end of *Eco*RI, *Pst*I or *Mse*I primers. Selective PCR was performed with 1XPCR buffer (1.5 mM $MgCl_2$), 2.5 mM dNTPs, 30 ng *Mse*I + CNN or *Mse*I + CN, 15 ng *Pst*I + GNN or *Eco*RI + ANN, and 0.4 unit *Taq* DNA polymerase. Fifty-nine selective primer combinations, 40 of *Eco*RI/*Mse*I and 19 of *Pst*I/*Mse*I, were used to generate AFLP fragments from pre-amplified fragment templates.

To generate AFLP marker data, 40 *Eco*RI (36 at CNU and four at WHRI) and three *Pst*I primers were end labeled with ^{33}P-ATP. AFLP fragments amplified with the remaining 16 *Pst*I/*Mse*I primer combinations were detected by silver staining. All amplifications were performed in a Perkin-Elmer 9700 Thermal-cycler with the PCR conditions described by Vos et al. (1995). The adapter and primer sequences used are presented below:

Enzyme	Adapter sequence	Pre-amplification primers
*Eco*RI	Oligo 1 5'-ctcgtagactgcgtacc-3'	5'-gactgcgtaccaattca-3'
	Oligo 2 3'-catctgacgcatggttaa-5'	
*Pst*I	Oligo 1 5'-ctcgtagactgcgtacatgca-3'	5'-gactgcgtacatgcag-3'
	Oligo 2 3'-catctgacgcatgt-5'	
*Mse*I	Oligo 1 5'-gacgatgagtcctgag-3'	5'-gatgagtcctgagtaac-3'
	Oligo 2 3'-tactcaggactcat-5'	

Silver staining was done based on the conditions described by Piao et al. (2004).

AFLP markers were named according to the nomenclature method of Keygene (http://www.keygene.com) with the modification that the suffix was given a number relating to the relative position on the gel rather than the actual band size resulting from primer combinations employed. Multiple loci, resulting from the same primer combination, were given identical names with the addition of a number suffix (see locus nomenclature below).

SSR

SSR repeats were identified from the BAC-end sequences of the KBrH library of the inbred line 'Chiifu-401-42' (Park et al. 2005). A total of 2,902 end-sequences were analyzed using 1,451 BAC clones of this library. Forty one sequences were found to contain repeat motifs, and primer pairs were designed from the flanking sequences using Primer 3 software (Rozen and Skaletsky 2000). The criteria used for designing the primers were that the size of amplified DNA fragments were in the range of 100–400 bp, the difference of T_m between the two primers within a pairs was less than 3°C, the primer T_m ranged from 55 to 63°C, and the GC contents were greater than 35%. DNA amplification was carried out in volumes of 20 µl, containing 1 unit of *Taq* polymerase, 0.5 µM of primers, 200 µM of dNTPs, 1.5–2.0 mM $MgCl_2$, 1XPCR buffer and 40 ng of genomic DNA as templates. The PCR profile was: initial 5 min at 95°C, and 30–35 cycles, each with 30 s DNA denaturation at 94°C, 45 s at appropriate annealing temperature and 60 s extension at 72°C, and final extension of 7 min at 72°C. PCR was carried out in a Perkin-Elmer Thermal-cycler and the PCR products separated on 2% NuSieve 3:1 Agarose (FMC) or 6% polyacrylamide gel. The marker assay names for these SSRs are taken from the BAC clone in which they were identified and the details of the primers for loci mapped here are given in Table 1.

For the PBC SSRs, primers were synthesized for 470 SSR repeats, identified from *Brassica napus* EST and genomic

Table 1 Primers for new SSR, STS, ESTP and CAPS marker assays

Marker assay name	Assay type	Forward primer	Reverse primer
KBRH138G23	SSR	TTTGACATCGTGCAATGCTA	TTGGGCTGGTCCTGAAGATA
KBRH139B23	SSR	ATCTCATGGTTGGTTCACCG	ATTTCCAAAACACACACGCA
KBRH143D22	SSR	GATGTGATACTTTGGCGACGG	TGAAGGATAATATGGTCTTGGCC
KBRH143F19	SSR	GCATGCAAGCTTGGAACTGAT	CAGTCACGCTTTCTGACGAAAA
KBRH143H15	SSR	TCTGCATCAAAATGCTAAAATGA	TGATCTTTTAGAAACAAAGATCGAG
KBRH143K20	SSR	CAAATGTCTCAAGACACATAAACCA	CTAAAGCAGCAATTGGGTGTTC
KBRH040J24	STS	CTTGCGCACGACGTAAACAG	TGGCTCTCCAGAGGCTAACG
KBRH043E02	STS	ATGCAAGCTTCATGGTGTCA	CATCAGCAAAATTTCATTTGTGT
KBRH048O11	STS	GCCTCTACCTGGCTTCAGCA	TCATTTGGCGCATACTTCCA
PW161	STS	TTCCTCTCGCTCTCCTCG	TCGAAGGAGAAGCCCACT
PW258	STS	CTCCTTCTCAGCATTCCCTTTGG	TCGGATTCTTCTCACTCCACATACG
PW243	STS	CTCCTTCTCAGCATTCCCTTTGG	ACAACAAACTGGGACTGGGACAG
PW186	STS	GACGGTGACGACCAATCAGAGCA	GCGTTCGCCATAGACGAGTCCAA
PX129	STS	AGCGCCAAGAAGTCATTT	AAGGAGAGAGATGACACTGATG
E018	ESTP	AACGAAACAGCCACTAGAAACA	CCCTCTCACAGCCCTCAG
E039	ESTP	CTTGAGTGCTCAGGTCAAAGC	GAACCCTTACCCCCAAGACTAC
E060	ESTP	GGGTTTACATCCTCATCGGA	CGCAAAGAGATAAGCTTTGGA
E101	ESTP	ATTGAGATTCTGGCCAAGGAG	AGGGAATATCAAGATCCACCAC
E120	ESTP	ATCATAACCCTCAGGTTTGACATC	ACATCAAGCTCCTCTCTGGGTA
E129	ESTP	AGATGGTAAAAGAGCACAAGCC	TTCAAGCTACCGATCCAACTG
E138	ESTP	TGCTATCACAGTAGGGATTGCTT	CACTCCCACTCCTCCTAGTCC
30AFAI	CAPS	GCTTCACATGTGTCTTAAACGG	TTCAGGAGTGTGATCCCAAGT
91STYl and 91XBAI	CAPS	CTTCCTCCTCGTTTCGGTCT	GTTGGGTAAATGAACCTTCGTG
GOL1	SSR	CAATATGTCCACCACCACCA	GGAGAAGCAGAAGGAGCTGA
GOL2	SSR	AGACATCCCACATCGGCTAC	GACCCAAGACCCAAGACTCA
GOL3	SSR	ACTCACTTTTGTTGGGCGTC	GGAGCCGCTTTCTCTACCTT
EJU1	SSR	GGTGAAAGAGGAAGATTGGT	AGGAGATACAGTTGAAGGGTC
EJU2	SSR	TTCACATCTTCTTCATCTTCC	TTGCTATTCGTTCTCAGTCTC
EJU3	SSR	CCTCTTTTAATTCAAACAAGAAATCA	TTCGGACAATGGCAGTGATA
EJU4	SSR	CACCTTATCATCTCTCTATCCC	CCTCTGTTTCTCTCCTTGTG
EJU5	SSR	GGCACGTACATGGAGGATTC	TGTTGGTCGAGCTGTTTCAG
EJU6	SSR	TCTCTCACCTGCCTTGTCT	ACTCCTCGGTAATGCCTC
ENA1	SSR	ACTTCCACAAACAGAGCAAC	TATGAAGACACCAAAGAGCC
ENA10	SSR	ATCGTCTCCTCTCATCTCAA	ATTACATCCTCCACCTTCTTC
ENA13	SSR	CTGCGTTCCAGTACCCTCTC	TCATCCTTAATGGTCCTGCC
ENA14	SSR	CTTACGGTGGAAATGCTG	TCGCTGGTGCTAAACTTG
ENA15	SSR	TAAACGGGAACTACCTCTATG	CTCTGCTCTTTCTTCTGACTG
ENA16	SSR	GAAGATAGAGAGAGGCGTGAG	GAAGATAGAGAGAGGCGTGAG
ENA17	SSR	CAGTTATTTCGCCTCGTCT	TATTTGTGGTCTGTTATTGGA
ENA18	SSR	TTAAAATGAAACCCACCCGA	TGTTGGGCAACATCCATTTA
ENA19	SSR	AAGTTACCAAGGAGAGGACAG	AAAGGGACGCTACAAGTCA
ENA2	SSR	GATGGTGATGGTGATAGGTC	GAAGAGAAGGAGTCAGAGATG
ENA20	SSR	GATGGAGGAAGAAGACAAGAC	TCTGAACTACCAAAGCCAAC
ENA21	SSR	GAGTGTTTGGAGCAGATGA	GGAGACTTTGCCTTTGTGT
ENA22	SSR	TTTGTAGACGAACAGCCACG	AGAATCGCATTTGATGGAGG
ENA23	SSR	GCTGTGCCAGTTCCTCTTTC	TCATTCCAAATGGCCTTACC
ENA25	SSR	ACACCCTCCTTCTCCTCTC	GCTTTGTTGAGTATCTTCGTC
ENA26	SSR	AAGAACGTGATCTCCATCGC	CGAGCTCCAAAGCAGATACC

Table 1 continued

Marker assay name	Assay type	Forward primer	Reverse primer
ENA27	SSR	AAAGGACAAAGAGGAAGGGC	TTGAAATCAAATGAGAGTGACG
ENA28	SSR	GGAGTCCGAGCGTTATGAAT	CTTCATCGACCCACCTTGTT
ENA3	SSR	ATCCCTTCTCACAGGTTTACT	GTCAAGTTTCTCTCCACACC
ENA4	SSR	ACTTCTCTTTATTCACTTCCCA	GAGGGTGGTTGGTTCATT
ENA5	SSR	CCCCCTTTCTATCTCCTCCA	GAATTCGGGTCTCGATTTGA
ENA6	SSR	CTCGTCTTCTTCACCTACAAC	CTGACATCTTTCTCACCCAC
ENA7	SSR	GCTCACACCTACAAGACAGAC	ACTAATAAAGCAACCAAGCAA
ENA8	SSR	ACTGAGAGCAACAACAACAAC	GTAGAGACGGAACCCTGA
ENA9	SSR	CCTGAGACCAACCTACTCCT	ATCTTCAACACGCATACCAC

sequences using SSR Primer (Robinson et al. 2004). The forward primers were 5'-labelled with one of three different ABI fluorescent dyes (FAM, HEX or NED). Amplifications were carried out in a 25 µl reaction volume containing 25 ng DNA, 2.5 µl 10× PCR reaction buffer (Qiagen, CA, USA), 15 pmol forward and reverse primers, 200 µM of each dNTP and 2 unit HotStart Taq polymerase (Qiagen). After an initial hot start at 95°C for 15 min, the following cycling parameters were employed: denaturation at 94°C for 1 min, annealing at 55°C for 1 min and extension at 72°C for 1 min. After 35 rounds of amplification, a final extension step was performed at 72°C for 10 min. All PCR reactions were performed in a PE9700 DNA thermal cycler (Perkin-Elmer). Following amplification, for multiplex loading, PCR products labeled with the three different dyes were pooled, ethanol precipitated and mixed with deionised formamide and GeneScan™-500 LIZ™ (Applied Biosystems, CA, USA) size standard. The SSRs were separated by capillary electrophoresis on an ABI3730xl (Applied Biosystems) and results were analyzed using Genotyper® Software (Applied Biosystems).

In addition to these, a further 1,107 SSRs from a number of sources (Lowe et al. 2004; Suwabe et al. 2002, 2006; AAFC Consortium primers—http://brassica.agr.gc.ca; see also Microsatellite Information Exchange) were also used. Fourteen of the AAFC Consortium primer sequences used in this study are available upon request from A. Sharpe, Agriculture and Agri-Food Canada for research purposes (sN11641, sN11707, sN13039, sN1919, sN2025, sN3523R, sN8474, sNRD71, sNRE30, sR3688, sR4047, sR6068, sR7178, and sS1949). In order to confirm the laboratory-to-laboratory reproducibility, five markers were screened in duplicate in two different laboratories using the same lines to confirm the genotype.

RAPD

About 10- and 12-base random oligonucleotide sequences from the Operon Technologies (USA) and WA (Japan), respectively, were used as RAPD primers. These were designated as OP or WA with an additional letter indicating the primer.

PCR amplification was carried out in volumes of 25 µl, containing 1 unit of *Taq* polymerase, 200 µM of dNTPs, 5 pM primer, 1.5 mM $MgCl_2$, and 50 ng of template genomic DNA. Samples were amplified for 50 cycles using the following thermal profile: 60 s at 94°C, 60 s at 35°C and 60 s at 72°C. Samples were pre-denaturated for 2 min at 94°C and final elongation was performed for 3 min at 72°C to ensure complete DNA extension. All amplifications were repeated to ensure the fidelity of the DNA profiles generated. Amplified fragments were separated on 1.5% agarose gel, stained with ethidium bromide and scored visually.

The primer details for the STS, ESTP and CAPS markers mapped in this study are given in Table 1.

STS

Thirty primer pairs were designed to KBrH BAC-end sequences. PCR amplification was performed in volumes of 20 µl, containing 1 unit of *Taq* polymerase, 0.5 µM of primers, 200 µM of dNTPs, 1.5–2.0 mM $MgCl_2$, 1× PCR buffer and 40 ng of genomic DNA as template.

PCR profile included initial 5 min at 95°C, and 30–35 cycles, each with 30 s DNA denaturation at 94°C, 45 s annealing at the appropriate temperature, and 60 s extension at 72°C, and final extension of 7 min at 72°C. The PCR products were separated on 2% Nusieve 3:1 Agarose (FMC) or 6% polyacrylamide gel.

Another set of STS markers were designed from the sequences of the RFLP clones used in the construction of a *B. rapa* linkage map (Teutonico and Osborn 1994). STS marker assay names are designated by the name of BAC or RFLP clone.

ESTP

One hundred and fifty ESTP primer pairs were designed using cDNA sequences of 'Chiifu-401-42'. PCR conditions

for reaction mixture and amplification profiles were the same as used for STS. The cDNA-based markers were designated as E followed by the number of primer.

CAPS

Monomorphic products of STS and ESTP assays were further tested for polymorphism in a CAPS assay against a set of 20 restriction endonucleases and the products resolved on 2% Nusieve 3:1 Agarose (FMC) or 6% polyacrylamide gels. These markers are designated by the number referring to the primer pair and the restriction endonuclease used.

Linkage analysis and map construction

Markers which were reproducibly polymorphic between the parent lines were scored in the DH population. Linkage analysis and map construction were performed using JoinMap Version 3.0 (Stam 1993; Van Ooijen and Voorrips 2001). Linkage groups were identified in the LOD grouping threshold range of 3.2–5.0, and linkage groups were assigned as R1–R10, corresponding to the A-genome linkage groups (N1–N10) of *B. napus*, using SSR markers previously mapped in the species.

Maps were generated for each linkage group using a recombination frequency below 0.45 and LOD score above 0.5 for all the marker pairs within each linkage group. A "ripple" procedure was performed after the addition of each marker and the "jump" thresholds were set to 5. Recombination frequencies were converted to centiMorgans (cM) with Kosambi's method for map-distance calculation (Kosambi 1944).

Locus nomenclature

A new convention is used here to name loci unambiguously, and to associate provenance in relation to marker assays. The locus name is constructed from different domains and is described in Bakus-Naur Form (Marcotty and Ledgard, 1986) as: <3-letter institute code>_<single letter marker assay type designator>< MARKER ASSAY NAME><single letter locus counter | 2-digit leading zero number locus counter | 3-digit leading zero allele size>, where different domains within the name as indicated above are bound by angle brackets and are case-sensitive. The institute codes refer to the laboratory institution where each assay was performed. Those used here are: aaf, Agriculture and Agri-Food Canada; cnu, Chungnam National University; hri, Warwick HRI; pbc, Plant Biotechnology Centre. The single letter marker assay type designators follow the recommendations of De Vicente et al. (2004), and are: a, AFLP; e, ESTP; f, RFLP; m, SSR; p, CAPS; r, RAPD; t, STS. Marker assay names are as described in the sections above or as previously published. Where previously published marker assay names are used, non-alphanumeric characters have been removed and any lower case letters converted to upper case in order to distinguish the marker assay name from the other domains of the locus name. The default letter counter for a locus is "a" and additional loci are designated by sequential letters. Where loci have previously been assigned, the product size and map position obtained here were compared with the original data in order to assign the correct locus. For some marker types, such as AFLPs and RAPDs, many loci are detected and where possible are referred to here by their band size in nucleotides. For three common markers, (hri_aE11M50_3>330, hri_aE12M47_1>330, hri_aE12M47_2>330) used in the Soengas et al. (2007), just followed the former name in the part of size by a digit number which corresponds to the number of the polymorphic band starting with the highest molecular weight band scored. A summary of this convention will also be available at http://www.brassica.info.

Comparison with the Arabidopsis genome sequence

The sequences of each sequence-tagged SSR, STS, ESTP and CAPS locus in the CKDH map were aligned with Arabidopsis genome sequences using BLASTN. The results of detected homologies ($E < 10^{-20}$) are reported with respect to the Arabidopsis sequence tiling path BACs encompassing the homologous sequence (http://atensembl.arabidopsis.info/Arabidopsis_thaliana_TAIR/index.html). The Arabidopsis chromosome and TIGR 5.0 bp co-ordinates of the end of the BAC orientated to the top of the chromosome are summarized in supplemental Table 4.

Results

Generation of markers and polymorphism survey

Data from the polymorphism survey using primers for assaying different kinds of markers are presented in supplemental Table 1. Of the 1,205 SSR primers from a range of different sources used, 389 (32.3%) showed polymorphic banding pattern between 'Chiifu-401-42' and 'Kenshin-402-43', and 235 were placed on the map. Of these, seven SSR marker assays (ENA23, GOL1, NA12A01, Na12E02, SJ4322A, SN2713 and SORD79) detected more than one segregating locus. Four of the 30 STS primers, designed from the BAC-end sequences, showed polymorphism between the parents. Sequencing of RFLP clones generated 137 STS primer pairs. Of these, 17 revealed polymorphism between the parents in 2% agarose or 6% polyacrylamide gels, and another four after subsequent restriction enzyme digestion (i.e. conversion to CAPS assays). One hundred and fifty primer

pairs generated from the cDNA sequences and 34 primers showed polymorphic ESTP loci directly on agarose gels without additional manipulation of PCR products, whereas the remaining 61 revealed monomorphic banding pattern. Of the latter primers, 13 revealed polymorphism between the parents after digestion with restriction enzymes. To generate RAPD fragments, 28 primers were used, which amplified one to four polymorphic bands per primer in the parents, corresponding to an average of 1.4 markers per primer.

The extent of AFLP polymorphism generated with the two enzyme combinations is given in Supplemental Table 2 and a summary of the polymorphism detected for each primer combination in Supplemental Table 3. Forty *Eco*RI/*Mse*I and nineteen *Pst*I/*Mse*I selective primer combinations were used to generate a total of 3,647 bands, of which 24% were found to be polymorphic and 87% of these were unambiguously scored.

A new nomenclature format

The accurate exchange of genetic information is becoming increasingly important as researchers aim to maximize the value of their genotype data by integrating it with existing mapping and genomics information. To achieve this reliably, specified data standards and formats are required. De Vicente et al. (2004) initiated this by developing a set of descriptors for genetic markers. Here, we adopt their locus nomenclature proposals to create a flexible format that will accommodate loci detected both by new and previously published marker assays. This should contribute to increased quality assurance and tracking of data provenance. The locus name is divided into clearly defined domains, with the marker assay name being discriminated by being written completely in upper case. Due to the potential ambiguities associated with establishing whether duplicate marker assays performed in different laboratories on different samples are detecting the same loci, the inclusion of the institute code provides a means for discriminating between these results.

Linkage groups assembled and distribution of markers

A total of 556 loci were assigned to 10 linkage groups with LOD values of 3.2–5.0. The genetic map had a total length

Table 2 Salient features of molecular linkage map of *B. rapa*

Linkage group of *B. rapa*	No. of markers	Density (markers/cM)	No. of intervals[a]	No. of gaps[b]	Average interval (cM)	Length (cM)
R01	46	2.96	40	0	3.40	136.3
R02	40	3.01	31	1	3.89	120.6
R03	79	1.95	63	1	2.45	154.2
R04	27	3.02	23	0	3.54	81.5
R05	72	1.82	49	0	2.67	130.9
R06	79	2.03	62	0	2.59	160.7
R07	74	1.23	55	0	1.66	91.3
R08	36	2.78	30	0	3.34	100.2
R09	59	1.95	49	0	2.35	115.1
R10	44	2.08	39	0	2.35	91.6
Total/average	556	2.13	418	2	2.83	1182.3

[a] Adjacent markers > 1 cM

[b] Distance between adjacent markers ≥ 15 cM

Table 3 Summary of SSRs used to identify and orientate the linkage groups of *B. rapa* corresponding of the A-genomes of *B. napus* and *B. juncea*

Assigned linkage group of *B. rapa*	Known linkage group of *B. napus*	No. of SSR loci	Known linkage group of *B. juncea*	No. of SSR loci
R01	N01	13	J01	4
R02	N02	15	J02	
R03	N03	16	J03	6
R04	N04	5	J04	1
R05	N05	10	J05	5
R06	N06	13	J06	8
R07	N07	12	J07	3
R08	N08	6	J08	
R09	N09	17	J09	
R10	N10	9	J10	2
Total		116		29

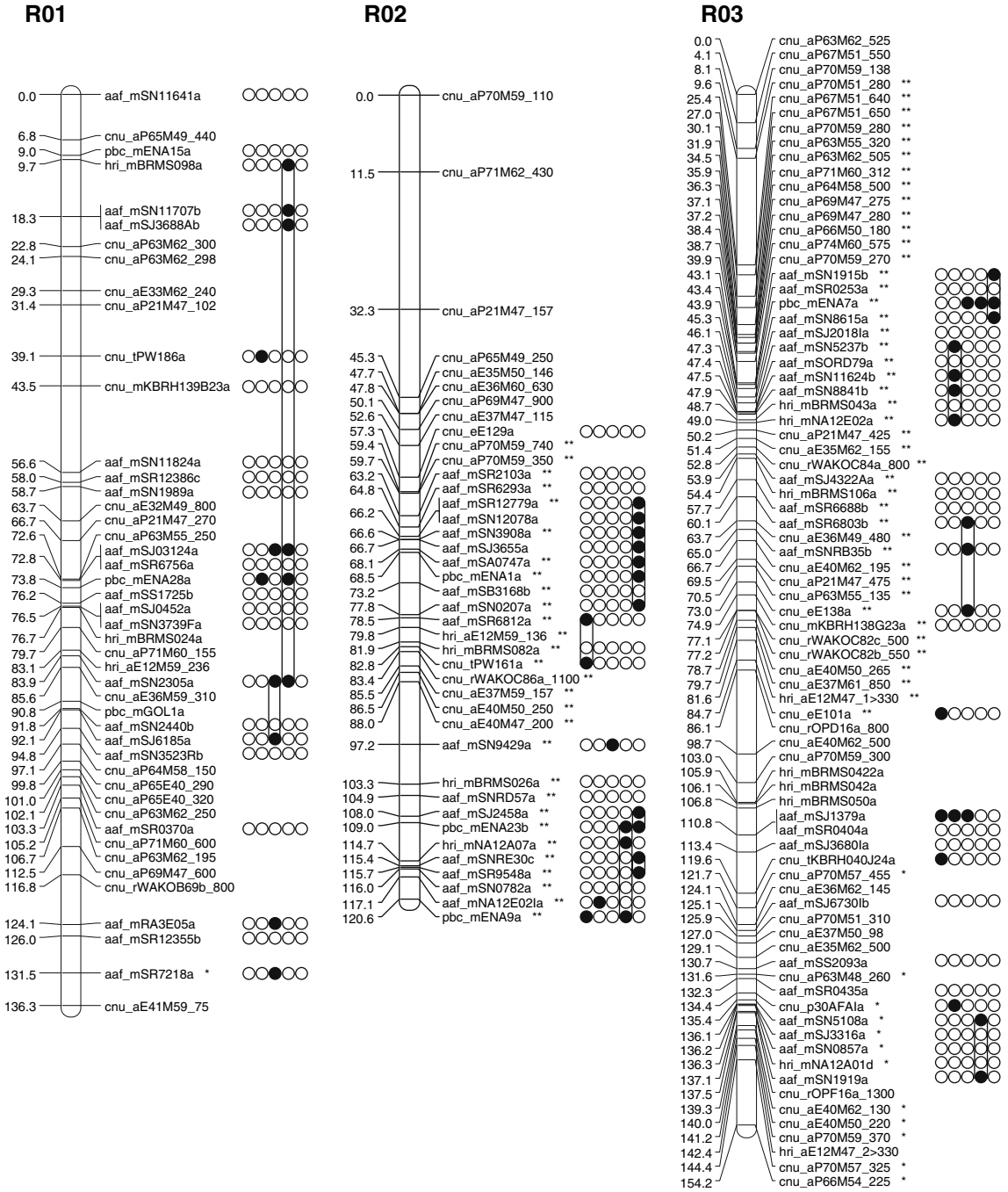

Fig. 1 Linkage map of *B. rapa* L. ssp. *pekinensis*. Cumulative recombination distances are shown on the left side and marker loci on the right of the linkage groups. The map shows the distribution of 556 loci along 10 linkage groups (R01–R10). Total map length is 1,182 cM. * and ** represent distorted marker segregation to 'Kenshin-402-43' and 'Chiifu-401-42', respectively. Loci tested for homology with Arabidopsis using BLASTN are shown with five circles to the right following the convention of Lan et al. (2000). These circles from left to right represent Arabidopsis chromosomes 1–5, respectively. Filled circles indicate the detection of homologous sequences with $E < 10^{-20}$. Groups of two or more markers showing homology to a neighboring region of an Arabidopsis chromosome are boxed

of 1,182 cM, with an average distance of 2.1 cM between two loci (Table 2). Using SSR markers common to the reference A genome of *B. napus* (Parkin et al. 1995) linkage groups R1–R10 were assembled (Fig. 1). The length of the linkage groups ranged from 81 to 161 cM for R04 and R06, respectively. The number of markers in the 10 linkage groups ranged from 27 to 79. The number of intervals, as determined on the basis of mapped points in the linkage groups, varied from 23 (R04) to 63 (R03) with an average interval of 2.8 cM. In the final linkage map 97.2% of the genome was within 5 cM of a marker; the two largest gaps observed were 21 and 16 cM on R02 and R03, respectively.

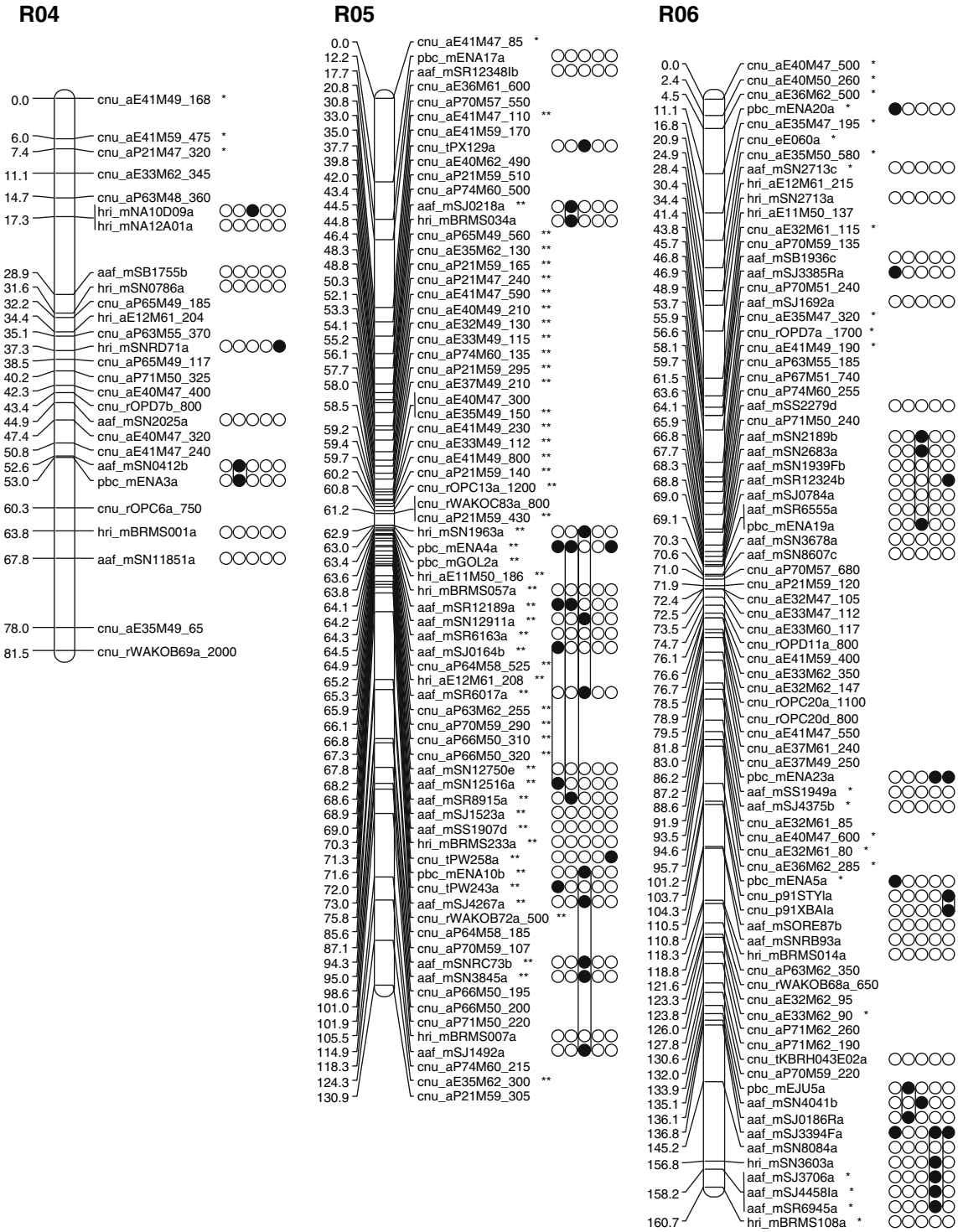

Fig. 1 continued

In all, 278 AFLP, 235 SSR, 25 RAPD, 8 STS, 7 ESTP and 3 CAPS markers were mapped on the linkage groups. The distribution of the individual marker types varied along the linkage groups (Fig. 1). The distribution of AFLP and SSR markers along the linkage groups ranged from 14 to 45 and 10 to 33, respectively. RAPD markers located across the 9 linkage groups ranged from 1 to 5.

Skewed segregation of markers

Skewed segregation of markers is a common feature in most of the *Brassica* linkage maps, and particularly ones based on DH lines due to preferential selection of genotypes responsive to microspore or anther culture (Chuong et al. 1988; Ferreira et al. 1994; Cloutier et al 1995; Uzunova

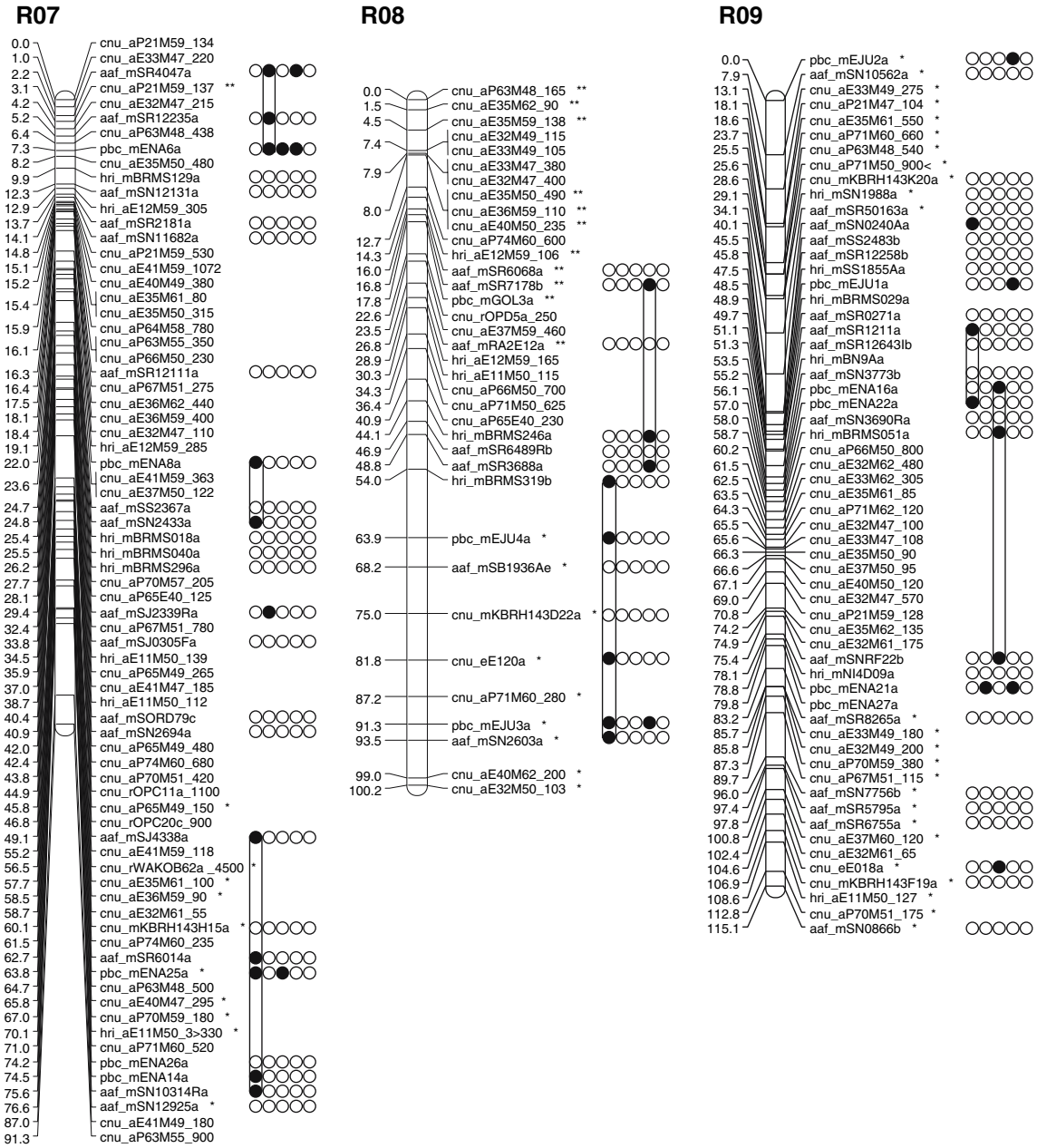

Fig. 1 continued

et al. 1995; Foisset et al. 1996; Takahate and Keller 1991). Of the 556 markers assigned to the linkage groups, 213 (38.3%) showed deviation from the expected segregation ratio of 1:1 ($P \leq 0.05$) (Fig. 1). This was a general property of the markers and not restricted to a particular marker type. In general, the loci showing distorted segregation were more skewed towards the male 'Kenshin-402-43' alleles irrespective of the type of marker used.

The markers showing biased segregation were distributed along all the linkage groups, except R01. In five linkage groups, R04, R06, R07, R09 and R10, the cluster of markers were distorted towards 'Kenshin-402-43'. However, for R02 and R05, the clusters of segregating markers were skewed towards 'Chiifu-401-42' alleles. The markers on R03 and R08 exhibiting segregation distortion were skewed towards both parents. The clusters of distorted markers were generally randomly distributed across the linkage groups. However, in R08, the only cluster of markers showing skewed segregation was located on the distal end of the chromosome.

Aligning the linkage groups to the A-genome chromosomes of *B. napus* and *B. juncea*

In tetraploid *B. napus* linkage groups N01 to N10 represent the A-genome, whereas linkage groups N11–N19 represent

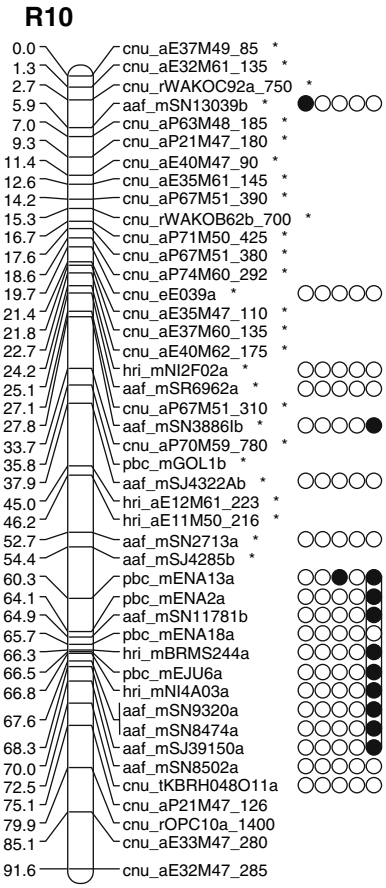

Fig. 1 continued

the C genome (Parkin et al. 1995; Sharpe et al. 1995). Of the 235 SSR markers used in the construction of the CKDH map, 116 could be assigned to the linkage groups of the A genome of *B. napus* (Microsatellite Information Exchange; AAFC Consortium, Lowe et al. 2004; Table 3). *B. juncea* is another allotetraploid *Brassica* species comprising the A and B genomes. The *B. juncea* A genome linkage groups are referred to as J01–J10 and equate to N01–N10 (Axelsson et al. 2000). The CKDH linkage group assignments established with SSRs common to the *B. napus* maps were further supported by an additional 29 SSRs that mapped to the corresponding A genome linkage groups of *B. juncea* (Table 3).

A comparison of the detailed marker order between the *B. napus* and *B. rapa* maps reveals general conservation of marker order for most linkage groups, with several minor rearrangements. A small number of minor rearrangements in marker order are not surprising as genetic linkage maps are the statistically most likely order generated by a particular software package, and do not necessarily represent the correct order. In addition, although all care was taken to minimize genotype scoring errors, even small numbers of errors may have a significant impact on relative map position. Further, two more significant rearrangements were observed. Firstly, a small inversion covering 10 cM on R03 was suggested by the altered order of loci in the region of aaf_mSN5237b and aaf_mSR6688b relative to the order in *B. napus*. A second possible rearrangement relative to *B. napus* covering 50 cM on R07 is suggested by the altered order of loci between aaf_mSN12925a and aaf_mSS2367a. There were insufficient markers relating to *B. juncea* to compare conservation of marker order.

Comparison of marker order with the Arabidopsis genome sequence

In total 232 sequence-tagged markers were used in the construction of the CKDH map and these detected 240 loci (Fig. 1). BLASTN was used to test homology between the originating *Brassica* sequences and the Arabidopsis genome, with a significance threshold of $E < 10^{-20}$ (reported to $E < 10^{-5}$ in Supplemental Table 4). About 120 of the 232 sequence-tagged markers detected 153 homologous regions, located on 131 different Arabidopsis BAC clones (Table 4, Supplemental Table 4).

The five chromosomes of Arabidopsis were represented by homologous sequences distributed throughout the ten *B. rapa* linkage groups. Groups of two or more markers showing homology to a neighboring region of an Arabidopsis chromosome are considered here as evidence for synteny (Fig. 1). On this basis 29 groups of markers could be recognized, ranging from two to nine markers in a group. In combination these groups covered 344 cM, representing 29% of the CKDH map that could be aligned to the Arabidopsis genome. The longest syntenic region is on R01 and is supported by six markers that span 8–17 Mb on At4. Particularly notable regions of synteny are on linkage group R02 and Arabidopsis chromosome 5 (At5) and also R10 and At5. The three syntenic regions on R02 together span 13 cM and correspond to three regions on At5: 21.1–23.5 Mb, around 24.5 Mb and around 8.5 Mb. Within the 21.6–24.1 Mb region there is also evidence of reshuffling, although some of this could also possibly be explained by imprecision in marker order. The region on R10 spans 8 cM on the map and again correspond to a region on At5; around 2.1–4.3 Mb. This region also appears to show evidence of reshuffling.

Discussion

Linkage map of *B. rapa*

The reference linkage map comprises a total of 556 markers and grouped in ten linkage groups with an average interval of 2.83 cM between adjacent loci.

One of the purposes of this map is to support the genome sequencing project. The nomenclature and orientation of

Table 4 Number of Arabidopsis genome sequences detected with *B. rapa* sequence-tagged markers

Brassica LG	No. markers[a]	Markers with hits[b]	Total hits[c]	At1	At2	At3	At4	At5	Mean loci/marker
R01	24	10	13	0	2	5	6	0	0.54
R02	25	17	20	3	1	1	3	12	0.8
R03	32	16	21	3	7	5	3	3	0.66
R04	10	4	4	0	2	1	0	1	0.4
R05	27	18	22	6	5	9	0	2	0.81
R06	36	17	26	5	2	5	10	4	0.72
R07	26	11	16	9	4	2	2	0	0.61
R08	13	8	9	5	0	0	4	0	0.69
R09	28	10	11	3	1	4	3	0	0.39
R10	19	11	13	1	0	1	0	11	0.68
Total	240 (232[e])	122 (120[f])	155 (153[g])	35	24	33	31	33	
Expected[d]				40	27	34	24	31	

[a] Sequenced markers on the *B. rapa* genetic map

[b] Markers highly similar to an Arabidopsis BAC (BLAST $E = <10^{-20}$)

[c] The number of BACs in the Arabidopsis genome containing sequences that are similar to the markers

[d] Expected number of total hits if *B. rapa* markers were distributed equally on separate Arabidopsis chromosomes in proportion to their physical length

[e] The total number of unique markers after taking into account duplicated loci

[f] Total number of markers with homology to at least one Arabidopsis BAC clone

[g] Total number of unique homologies after taking into account duplicate loci

many earlier *Brassica* linkage groups, including those of different linkage maps for the same species, has to some extent been arbitrary and hence published maps are often difficult to compare. The format used here eliminates the use of non-alphanumeric characters that are frequently incorporated into locus names and which are a potential source of errors when parsing data between databases. In addition, the default use of a locus counter suffix prevents the confusion that can result when new loci are identified. Although this format creates yet another name, there is already a pressing need to track marker/locus synonyms. Those presented here will be referred to as the reference locus name, against which synonyms will be tracked in future publications and databases.

A number of different types of markers, including 278 AFLP, 235 SSR, 25 RAPD and 18 ESTP, STS and CAPS markers, were mapped on the linkage groups. Although the CKDH population is derived from an intrasubspecific Chinese cabbage cross, the overall level of AFLP polymorphism detected is higher than that reported in barley (Becker et al. 1995) and rice (Maheswaran et al. 1997). The number of polymorphic bands obtained by different primer combinations ranged from 4 to 26 with an average of 15. This compares with average values of 22 and 17 polymorphic bands per primer combination in the linkage maps reported for *B. napus* (Lombard and Delourme 2001) and *B. juncea* (Pradhan et al. 2003), respectively. Although there was no notable skew in the distribution of AFLP markers in the present map, *Eco*RI/*Mse*I AFLP markers showed some tendency to cluster. Several markers grouped as clusters within the central regions of the linkage groups, which might represent the centromeric/pericentromeric regions of linkage groups (e.g. in R1, R4 and R6) due to the low frequency of recombination characteristic of these regions. However, the distribution of *Pst*I/*Mse*I combination based AFLP markers were more or less random across the linkage groups. Such clustering of *Eco*RI/MseI markers in specific regions is reported in the AFLP-based linkage maps of potato (Van Eck et al. 1995), barley (Qi et al. 1998), soybean (Keim et al. 1997), Arabidopsis (Carlos et al. 1998), maize (Vuylsteke et al. 1999), Guinea yam (Mignouna 2002), *B. oleracea* (Sebastian et al. 2000) and *B. juncea* (Pradhan et al. 2003). The markers generated by the *Eco*RI/*Mse*I enzyme combination were clustered mainly near the centromeric regions, perhaps because of the abundance of AT-rich segments which are not sensitive to GC methylation, and more target sequences of the restriction enzymes (Castiglioni et al. 1999; Vuylsteke et al. 1999). Consequently, there is a higher probability of identifying *Eco*RI/*Mse*I AFLP markers than *Pst*I/*Mse*I AFLP markers and RFLP in the highly repetitive regions of the genome (Castiglioni et al. 1999).

We mapped 556 loci on 10 linkage groups which covered a total length of 1,182 cM. In comparison to this, the previous maps of *B. rapa* reported by Song et al. (1991), Chyi et al. (1992), Teutonico and Osborn (1994), Suwabe

et al. (2006), Kim et al. (2006), and Soengas et al. (2007) reported total map lengths of 1,850, 1,876, 1,785, 1,006, 1,287 and 663 cM, respectively with 280, 360, 139, 262, 545 and 246 as the corresponding number of mapped loci. The differences in map lengths in different studies is usually mostly attributable to scoring errors, but has been attributed to the type of marker, number of individuals, number of markers, recombination frequency and LOD values, and the software employed (Gosselin et al. 2002). Qi et al. (1996) reported that the length and the observed genome coverage in barley was greater with MAPMAKER than with JOINMAP. Such a difference was attributed to how each program calculates map distance when the actual interference differs from that assumed (Qi et al. 1996: Gosselin et al. 2002). Pradhan et al. (2003) also observed reduction in the total genetic length although they mapped more markers in comparison to other maps in *B. juncea*.

The diploid *Brassica* genomes contain large replicated blocks of collinear segments both within and between linkage groups. These are thought to have been derived from a polyploid ancestor, although the exact mechanism by which this occurred is still under debate (discussed in Parkin et al. 2005). In the linkage map generated here only seven SSR markers detected multiple loci (see above) and no evidence of conserved blocks of synteny can be deduced from this. This low level of detection of the replication within the genome is a result of the marker types used to generate this map. The AFLP and RAPD techniques are valuable as they are able to detect many polymorphic loci simultaneously. However, the products are anonymous and so homology between them is not readily detected without further characterization. SSR markers are designed to discrete loci and are usually located in non-coding sequences which are less well conserved between the replicated blocks than coding regions. Previous studies have included other marker types, such as RFLPs, which usually encompass some coding sequence, and these have proved efficient at detecting polymorphisms in replicated loci (Parkin et al. 2005).

The tendency of markers to be skewed towards 'Kenshin-403-43' is in agreement with studies in which the inbred Chinese cabbage line 'Kenshin-402-43' has been found to be particularly responsive to regeneration (Sato et al. 1989; Kuginuki et al. 1997) and a separate study in which Kenshin 402-43 was found to be more responsive to regeneration than Chiifu 401-42 (Yang et al., unpublished). The preferential transmission of chromosomal segments, probably carrying allele(s) for increased anther culture responsiveness of 'Kenshin-402-43' in the DHs, might have resulted in more markers exhibiting segregation bias towards this parent. Previously, Foisset et al. (1996) reported the possibility that anther/microspore culture resulted in clusters of markers biased towards the responsive parental line. This phenomenon has often been reported in rapeseed (Chuong et al. 1988; Ferreira et al. 1994; Uzunova et al. 1995; Cloutier et al. 1995) and *B. oleracea* (Takahate and Keller 1991).

The clusters of distorted markers were generally randomly distributed across the linkage groups. In *B. napus*, markers with distorted segregation ratios have been reported to cluster in specific regions of linkage groups, and each cluster comprises loci biased towards the alleles of the same parental line (Foisset 1996). A higher percentage of markers (38%) showed skewed segregation in this DH population than previously reported for F_2 *B. rapa* populations by Song et al. (1991) (3%), Nozaki et al. (1997) (17%), Chyi et al. (1992) (24%), Suwabe et al. (2006) (16%) and Soengas et al. (2007) (16%). In addition to biased selection of parental genotypes during DH line production, other reported explanations for the frequent occurrence of segregation distortions include loss of chromosomes (Kasha and Kao 1970), genetic isolation mechanisms (Zamir and Tadmor 1986), the presence of viability genes (Beavis and Grant 1991) and the occurrence of gene-conversion events (Nicolas and Rossignol 1983; Nag et al. 1989). Segregation distortions can also be generated by non-homologous and/or homoeologous recombination (Armstrong and Keller 1982).

The use of 235 SSR markers enabled us to align and orientate all 10 linkage groups of *B. rapa* ssp. *pekinensis* with the corresponding A-genome reference linkage groups of *B. napus*. In addition to a number of minor differences in marker order, which may represent real differences in genomic arrangement or artefacts of the mapping process, two larger rearrangements were also observed. It is likely that such rearrangements may have occurred subsequent to speciation. This study further demonstrates the value of SSRs for the alignment of different linkage maps. Those used here will provide a valuable additional set of mapped loci for this purpose. To this end, a number of these SSRs have recently been used to align two other *B. rapa* linkage maps to this map, one segregating for clubroot resistance (Suwabe et al. 2006) and the R-o-18 x B162 map which segregates for *Xanthomonas campestris* resistance (Soengas et al. 2007).

Recently, Koo et al. (2004) reported the construction of a high resolution karyotype of *B. rapa* ssp. *pekinensis* by chromosome *in situ* hybridization. The integration of the cytogenetic and genetic map will be an important goal for future work. Howell et al. (2002) assigned all the 9 linkage groups of a *B. oleracea* genetic map to the corresponding 9 chromosomes through FISH using BAC clones as probes. The BAC clones were definitively assigned to the linkage map positions through development of locus-specific PCR assays. We are using linkage group specific ESTs to identify the corresponding BAC clones in a KBrH library, which will be involved in the FISH assay to integrate the genetic map and physical map of *B. rapa* ssp. *pekinensis*.

Comparison with the Arabidopsis genome sequence

With the criteria used here for considering synteny, nearly a quarter of the *B. rapa* genetic map could be aligned to the Arabidopsis genome. The markers used in the construction of the CKDH map provide little information on the replicated nature of the *B. rapa* genome. However, the finding that regions on both R2 and R10 correspond to At5 suggests that these may in fact be duplicate regions. This is confirmed by the maps of Parkin et al. (2005) and Kim et al. (2006) which have 16 and eight markers in common between the two linkage groups, respectively.

Seventy-eight percent of the markers showing homology to Arabidopsis genome sequences fall into the syntenic groups. For the remaining singleton markers there is either insufficient marker density to establish the size of the homologous segment or they are the result of extensive chromosomal reshufflings such as inversions, duplications and deletions which have occurred since the divergence of these two species (Fig. 1). In addition, many of the syntenic regions presented here overlap, which provides further evidence for genomic rearrangement.

The CKDH map adds considerably to the accumulating evidence relating to the organization of the *Brassica* A genome, and provides a key resource for comparison to Arabidopsis and the anchoring of the *Brassica* genome sequence. Other maps that also contribute to this are the *B. rapa* maps of Suwabe et al. (2006) and Kim et al. (2006) and the *B. napus* maps of Parkin et al. (2005) and Qui et al. (2006). Comparison of the relationship between the CKDH map and Arabidopsis with the relationship found in these additional maps demonstrates a high degree of consensus. Significant regions of correspondence are with At3 and At4 on R01, At5 on R02, At2, At4 and At5 on R03, At2 on R04, At1, At2 and At3 on R05, At5 on R06, At1 and At2 on R07, At1 on R08 and At5 on R10. In combination, these maps provide increased detail of the comparative relationship between the Arabidopsis and *B. rapa* genomes. With this increased detail, however, there seems to be an emerging picture of ever increasing complexity.

The Multinational *Brassica rapa* Genome Sequencing Project (http://www.brassica.info/mbgp/mbgp1.htm) is now underway. The aim of this project is to generate the reference genome sequence for the A genome of *Brassica* based upon the same *B. rapa* ssp. *pekinensis* inbred line 'Chiifu-401-42' for which alleles segregate in the map presented here. A BAC library has already been constructed (Park et al. 2005) in this inbred line, which will be used for genome sequencing. To date, at least 90,000 BACs (representing ~20% of the genome) have end-sequence tags. More than 530 seed BACs have been sequenced (a further 15% of the gene space), and BAC by BAC sequencing has already commenced on 5 chromosomes by international consortia. The linkage map reported here is therefore a key resource in undertaking future structural and functional genomic studies in *B. rapa*.

Acknowledgements This study was supported by grants in Korea from Bio-Green 21 Program, Rural Development Administration from Ministry of Agriculture and Forestry; and Korean Science and Engineering Foundation (R21-2004-000-10010-0). In the UK the work was funded by the Biotechnology and Biological Sciences Research Council. We wish to thank Neale Grant, Carol Evered, Volkan Cevik, Jason Pole, and James Lynn at the Warwick HRI in UK, K Suwabe, M Matsumoto and M Hirai of the Kyoto Prefectural University in Japan, and Hayley Mountford, Clare Hopkins of the Plant Biotechnology Centre in Australia, Akihito Wakatsuki, Sakata Seed Corporation in Japan, Thomas C. Osborn at Seminis Vegetable Seeds Inc. in USA, and Andrew Sharpe, Clarke Wayne at the AAFC Saskatoon Research Centre in Canada for the contribution of the BLAST information.

References

Armstrong KC, Keller WA (1982) Chromosome pairing in haploids of *Brassica oleracea*. Can J Genet Cytol 24:735–739

Axelsson T, Bowman CM, Sharpe AG, Lydiate DJ, lagercrantz U (2000) Amphidiploid Brassica juncea contains conserved progenitor genomes. Genome 43:679–688

Beavis WD, Grant D (1991) A linkage map based on information from 4 F_2 populations of maize (*Zea mays* L.). Theor Appl Genet 82:636–644

Becker J, Vos P, Kuiper M, Salamini F, Heun M (1995) Combined mapping of AFLP and RFLP markers in barley. Mol Gen Genet 249:65–73

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

Castiglioni P, Ajmone-Marsan P, van Wijk R, Motto M (1999) AFLP markers in a molecular linkage map of maize: codominant scoring and linkage group distribution. Theor Appl Genet 99:425–431

Cheung WY, Champagne G, Hubert N, Landry BS (1997) Comparison of the genetic maps of *Brassica napus* and *Brassica oleracea*. Theor Appl Genet 94:569–582

Chuong PV, Deslauriesrs C, Kott LS, Beversdorf WD (1988) Effects of donor genotype and bud sampling on microspore culture of *Brassica napus*. Can J Bot 66:1653–1637

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

Cloutier S, Cappadocia M, Landry BS (1995) Study of microspore-culture responsiveness in oilseed rape (*Brassica napus* L.) by comparative mapping of a F_2 population and two microspore-derived populations. Theor Appl Genet 91:841–847

De Vicente C, Thomas M, Adriana A (2004) IPGRI report, Descripors for genetic markers technologies

Fahey JW, Talalay P (1995) The role of crucifers in cancer chemoprotection. In: Gustin DL, Flores HE (eds) Phytochemicals and Health. American Society of Plant Physiologists, Rockville, pp 87–93

Ferreira ME, Williams PH, Osborn TC (1994) RFLP mapping of *Brassica napus* using doubled haploid lines. Theor Appl Genet 89:615–621

Foisset N, Delourme R, Barret P, Hubert N, Landry BS, Renard M (1996) Molecular-mapping analysis in *Brassica napus* using isozyme, RAPD and RFLP markers on a doubled haploid progeny. Theor Appl Genet 93:1017–1025

Gomez-Campo C, Prakash S (1999) Origin and domestication. In: Gomez-Campo C (eds) Biology of Brassica Coenospecies. Elsevier, Amsterdam, pp 33–58

Gosselin I, Zhou Y, Bousquet J, Isabel N (2002) Megagametophte-derivered linkage maps of white spruce (*Picea glaucea*) based on RAPD, SCAR and ESTP markers. Theor Appl Genet 104:987–997

Guillemaut P, Laurence MD (1992) Isolation of plant DNA: a fast, inexpensive and reliable method. Plant Mol Biol Rep 10:60–65

Howell EC, Baker GC, Jones GH, Kearsey MJ, King GJ, Kop EP, Ryder CD, Teakle GR, Vicente JG, Armstrong J (2002) Integration of cytogenetic and genetic linkage maps of *Brassica oleracea*. Genetics 161:1225–1234

Kasha KJ, Kao KN (1970) High frequency haploid production in barley (*Hordeum vulgare* L.). Nature 225:874–876

Keim P, Schupp JM, Travis SE, Clayton K, Zhu T, Shi L, Ferreira A, Webb DM (1997) A high-density soybean genetic map based on AFLP markers. Crop Sci 37:537–543

Kim JS, Chung TY, King GJ, Jin M, Yang T-J, Jin Y-M, Kim H-I, Park B-S (2006) A sequence-tagged linkage map of *Brassica rapa*. Genetics 174:29–39

Koo DH, Plaha P, Lim YP, Hurr Y-K, Bang J-W (2004) A high resolution karyotype of *Brassica rapa* ssp. *pekinensis* revealed by pachytene analysis and multicolor fluorescence in situ hybridization. Theor Appl Genet 109:1346–1352

Kosambi DD (1944) The estimation of map distance from recombination values. Ann Eugen 12:172–175

Kuginuki Y, Nakamura K, Hida K, Yosikawa H (1997) Varietal differences in embryogenic and regenerative ability in microspore culture of Chinese cabbage (*Brassica rapa* L. ssp. *pekinensis*). Breed Sci 47:341–346

Lim YP, Plaha P, Choi SR, Uhm T, Hong CP, Bang JW, Hur YK (2006) Toward unraveling the structure of *Brassica rapa* genome. Physiol Plant 126:585–591

Lombard V, Delourme R (2001) A consensus linkage map for rapeseed (*Brassica napus* L.): construction and integration of three individual maps from DH populations. Theor Appl Genet 103:491–507

Lowe AJ, Moule C, Trick M, Edwards KJ (2004) Efficient large-scale development of microsatellites for marker and mapping applications in *Brassica* crop species. Theor Appl Genet 108:1103–1112

Maheswaran M, Subudhi PK, Nandi S, Xu JC, Parco A, Yang DC, Huang N (1997) Polymorphism distribution and segregation of AFLP markers in a doubled-haploid rice population. Theor Appl Genet 94:39–45

Marcotty M, Ledgard H (1986) The world of programming languages. Springer-Verlag, Berlin, p 41

Mignouna HD, Mank RA, Eills THN, Van den Bosch N, Asoedu R, Ng SYC, Peleman J (2002) A genetic linkage map of Guinea yam (*Dioscorea rotundata Poir*) based on AFLP markers. Theor Appl Genet 105:716–725

Nag DK, White MA, Petes TD (1989) Palindromic sequences in heteroduplex DNA inhibit mismatch repair in yeast. Nature 340:318–320

Nicolas A, Rossignol JL (1983) Gene conversion: point mutation heterozygosities lower heteroduplex formation. EMBO J 2:2265–2270

Nozaki T, Kumazaki A, Koba T, Ishikawa K, Ikehashi H (1997) Linkage analysis among loci for RAPDs, isozymes and some agronomic traits in *Brassica campestris* L. Euphytica 95:115–123

Park JY, Koo DH, Hong CP, Lee SJ, Jeon JW, Lee SH, Yun PY, Park BS, Kim HR, Bang JW, Plaha P, Bancroft I, Lim YP (2005) Physical mapping and microsynteny of *Brassica rapa* ssp. *pekinensis* genome corresponding to a 222 kb gene-rich region of Arabidopsis chromosome 4 and partially duplicated on chromosome 5. Mol Gen Genomics 274(6):579–588

Parkin IA, Sharpe AG, Keith DJ, Lydiate DJ (1995) Identification of the A and C genomes of amphiploid *Brassica napus* (oilseed rape). Genome 38:1122–1133

Parkin IA, Gulden SM, Sharpe AG, Lukens L, Trick M, Osborn TC, Lydiate DJ (2005) Segmental structure of the *Brassica napus* genome based on comparative analysis with *Arabidopsis thaliana*. Genetics 171(2):765–781

Piao ZY, Deng YQ, Choi SR, Park YJ, Lim YP (2004) SCAR and CAPS mapping of *CRb*, a gene conferring resistance to *Plasmodiophora brassicae* in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*). Theor Appl Genet 108:1458–1465

Piquemal J, Cinquin E, Couton F, Rondeau C, Seignoret E, Doucet I, Perret D, Villeger MJ, Vincourt P, Blanchard P (2005) Construction of an oilseed rape (*Brassica napus* L.) genetic map with SSR markers. Theor Appl Genet 111:1514–1523

Pradhan AK, Gupta V, Mukhopadhyay A, Arumugam N, Sodhi YS, Pental D (2003) A high-density linkage map in *Brassica juncea* (Indian mustard) using AFLP and RFLP markers. Theor Appl Genet 106:607–614

Qi X, Stam P, Lindhout P (1996) Comparison and integration of four barley genetic maps. Genome 39:379–394

Qi X, Stam P, Lindhout P (1998) Use of locus specific AFLP markers to construct a high-density molecular map in barley. Theor Appl Genet 96:376–384

Qui D, Morgan C, Shi J, Long Y, Liu J, Zhuang X, Wang Y, Tan X, Diedricht E, Weihmann T, Everett C, Vanstraelen S, Beckett P, Fraser F, Trick M, Barnes S, Wilmer J, Schmidt R, Li J, Meng J, Bancroft I (2006) A comparative linkage map of oilseed rape and its use for QTL analysis of seed oil and erucic acid content. Theor Appl Genet 114:67–80

Robinson AJ, Love CG, Batley J, Baker G, Edwards D (2004) Simple sequence repeat marker loci discovery using SSR Primer. Bioinformatics 20:1475–1476

Rozen S, Skaletsky HJ (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) Bioinformatics methods and protocols: methods in molecular biology. Humana Press, Totowa, NJ, pp 365–386

Sato T, Nishio T, Hirai M (1989) Plant regeneration from isolated microspore culture of Chinese cabbage (*Brassica campestiris* spp. *pekinensis*). Plant Cell Rep 8:486–488

Sebastian RL, Howell EC, King GJ, Marshall DF, Kearsey MJ (2000) An integrated AFLP and RFLP *Brassica oleracea* linkage map from two morphologically distinct doubled-haploid mapping populations. Theor Appl Genet 100:75–81

Sharpe AG, Parkin IAP, Keith DJ, Lydiate DJ (1995) Frequent nonreciprocal translocations in the amphidiploid genome of oilseed rape (*Brassica napus*). Genome 38:1112–1121

Soengas P, Hand P, Vicente JG, Pole JM, Pink DAC (2007) Identification of quantitative trait loci for resistance to *Xanthomonas campestris* pv. *campestris* in *Brassica rapa*. Theor Appl Genet 114:637–645

Song KM, Suzuki JY, Slocum MK, Williams PH, Osborn TC (1991) A linkage map of *Brassica rapa* (syn. *campestris*) based on restriction fragment length polymorphism loci. Theor Appl Genet 82:296–304

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

Suwabe K, Iketani H, Nunome T, Kage T, Hirai M (2002) Isolation and characterization of microsatellites in *Brassica rapa* L. Theor Appl Genet 104:1092–1098

Suwabe K, Iketani H, Nunome T, Ohyama A, Hirai M, Fukuoka H (2004) Characteristics of microsatellites in *Brassica rapa* genome and their potential utilisation for comparative genomics in cruciferae. Breed Res 54:85–90

Suwabe K, Tsukazaki H, Iketani H, Hatakeyama K, Kondo M, Fujimura M, Nunome T, Fukuoka H, Hirai M, Matsumoto M (2006) SSR-based comparative genomics between *Brassica rapa* and *Arabidopsis thaliana*: the genetic origin of clubroot resistance. Genetics 173:309–319

Takahate Y, Keller WA (1991) High frequency embryogenesis and plant regeneration in isolated microspore culture of *Brassica oleracea* L. Plant Sci 74:235–242

Teutonico RA, Osborn TC (1994) Mapping of RFLP and qualitative trait loci in *Brassica rapa* and comparison to the linkage maps of *B. napus*, *B. oleracea*, and *Arabidopsis thaliana*. Theor Appl Genet 89:885–894

Truco MJ, Quiros CF (1994) Structure and organization of the B genome based on a linkage map in *Brassica nigra*. Theor Appl Genet 89:590–598

Uzunova M, Ecke W, Weissleder K, Röbbelen G (1995) Mapping the genome of rapeseed (*Brassica napus* L.) I. Construction of an RFLP linkage map and localization of QTLs for seed glucosinolate content. Theor Appl Genet 90:194–204

Van Eck HJ, Van der Voort JR, Draaistra J, Van Zandvoort P, Van Enckevort E, Segers B, Peleman J, Jacobsen E, Helder J, Bakker J (1995) The inheritance and chromosomal localization of AFLP markers in a non-inbred potato offspring. Mol Breed 1:397–410

Van Ooijen JW, Voorrips RE (2001) JoinMap ® Version 3.0: software for the calculation of genetic linkage maps. CPRO-DLO, Wageningen

Vos P, Hogers R, Bleeker M, Reijans M, van der Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucl Acids Res 23:4407–4414

Vuylsteke M, Mank R, Antonise R, Bastiaans E, Senior ML, Stuber CW, Melchinger AE, Lübberstedt T, Xia XC, Stam P, Zabeau M, Kuiper M (1999) Two high-density AFLP® linkage maps of *Zea mays* L.: analysis of distribution of AFLP markers. Theor Appl Genet 99:921–935

Zamir D, Tadmor Y (1986) Unequal segregation of nuclear genes in plants. Bot Gaz 147:355–358