
Review

Comparative genomics of Brassicaceae crops

Ashutosh Sharma*^{1,2)}, Xiaonan Li³⁾ and Yong Pyo Lim³⁾

¹⁾ Graduate School of Agricultural Science, Tohoku University, Aoba, Sendai, Miyagi 981-8555, Japan

²⁾ Present address: Institute of Molecular, Cell, and Systems Biology, College of Medical, Veterinary, and Life Sciences, University of Glasgow, Glasgow G12 8QQ, United Kingdom

³⁾ Molecular Genetics and Genomics Laboratory, Department of Horticulture, Chungnam National University, Daejeon 305-764, Republic of Korea

The family Brassicaceae is one of the major groups of the plant kingdom and comprises diverse species of great economic, agronomic and scientific importance, including the model plant *Arabidopsis*. The sequencing of the *Arabidopsis* genome has revolutionized our knowledge in the field of plant biology and provides a foundation in genomics and comparative biology. Genomic resources have been utilized in *Brassica* for diversity analyses, construction of genetic maps and identification of agronomic traits. In Brassicaceae, comparative sequence analysis across the species has been utilized to understand genome structure, evolution and the detection of conserved genomic segments. In this review, we focus on the progress made in genetic resource development, genome sequencing and comparative mapping in *Brassica* and related species. The utilization of genomic resources and next-generation sequencing approaches in improvement of *Brassica* crops is also discussed.

Key Words: DNA markers, functional genomics, linkage map, Brassicaceae.

Introduction

Comparative genomics is the study of similarities and differences at the genomic level to make inferences about the functions and evolution of various biological processes. This is an important field to study genome evolution, sequence collinearity and transfer of information from extensively studied model organisms to species of commercial interest. Genome sequencing of *Arabidopsis*, a member of the Brassicaceae family, has revolutionized our knowledge in every field of plant biology and laid a foundation for genomics and comparative biology.

The family Brassicaceae is one of the major groups of the plant kingdom, comprising of 340–360 genera and over 3,700 species distributed worldwide (Warwick *et al.* 2006). Many species within the family are of great economic, agronomic and scientific importance. Some examples of these include the following: *Brassica napus* and *B. juncea* (oil-seed crops); *B. rapa* (turnip, leaf vegetable); *B. oleracea* (cabbage, cauliflower, Kale, broccoli); *Raphanus sativus* (vegetable) and *Arabidopsis thaliana* (model plant). The six most cultivated species of the genus *Brassica* comprises the three diploid genomes of *B. rapa* (AA, $2n = 20$), *B. nigra*

(BB, $2n = 16$) and *B. oleracea* (CC, $2n = 18$) together with three amphidiploid species, *B. juncea* (AABB, $2n = 36$), *B. napus* (AACC, $2n = 38$) and *B. carinata* (BBCC, $2n = 34$). Cytogenetics and hybridization studies have demonstrated that amphidiploid species are natural hybrids of diploids and the six *Brassica* species are interlinked (Fig. 1, U 1935). Genome evolution and comparative sequence analysis of Brassicaceae have also confirmed the interrelationship of the six *Brassica* species at the molecular level (Schmidt and Bancroft 2011). In Brassicaceae, genomic studies are mainly focused on cultivated *Brassica* and their diploid progenitor species, which are compared with the *Arabidopsis* genome. With the inception of the Multinational Brassica Genome Project (MBGP) in 2002, the international *Brassica* community agreed to develop more resources for *Brassica* crops and genome sequencing. In the last decade, significant advances have been made in generation of genomics resources and translational research, which will aid *Brassica* crop improvement (Augustine *et al.* 2013, Schmidt and Bancroft 2011). Recently, sequencing of *B. rapa* and ancestral diploids of *Brassica* has expanded comparative genomic studies, providing resources for the identification of candidate genes of agronomic traits. Comparative mapping of *Brassica* species with the *Arabidopsis* genome helps in understanding conserved genetic architecture and genome evolution and the identification and functional analysis of genes for important agronomic traits. Genome-wide synteny

Communicated by T. Nishio

Received November 6, 2013. Accepted February 16, 2014.

*Corresponding author (e-mail: ashutosh@glasgow.ac.uk)

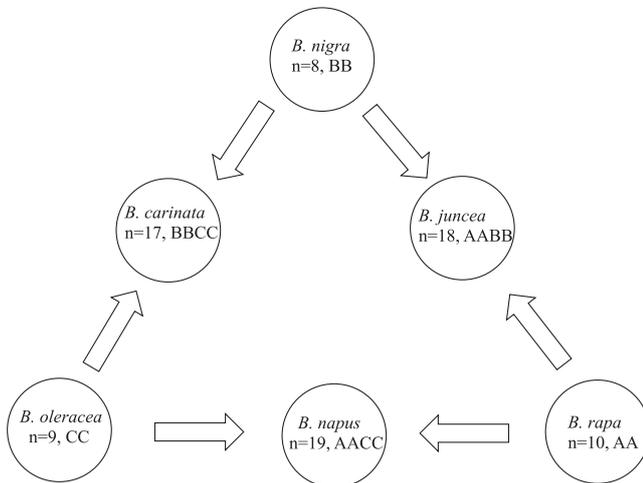


Fig. 1. Genomic relationships among six cultivated *Brassica* species represented by ‘Triangle of U’. Adapted from UN, 1935.

analyses between the *Arabidopsis* and *Brassica* A, B, and C genomes have identified conserved chromosomal blocks and elucidated genome rearrangements and karyotype diversification.

Next-generation sequencing (NGS) techniques have been utilized to develop cost-effective and efficient methods for single nucleotide polymorphism (SNP) discovery, genotyping and gene expression studies. In some *Brassica* species, these techniques have been used for the identification of SNP markers and the construction of linkage maps. Transcriptome analysis has also been used to find different gene expression profiles in response to abiotic and biotic stress and in understanding gene regulatory mechanisms. In this review, we emphasize the advancement of resource development in Brassicaceae, comparative mapping and the recent progress made in sequencing *Brassica* and related species. We focus on the genomics and genetic improvements made in six cultivated crops of *Brassica* and *Raphanus*.

Linkage maps are a prerequisite of comparative mapping

Linkage maps are valuable sources for identification and map-based cloning of important genes, analyses of QTLs for agronomic traits, and comparative mapping. Linkage maps have been developed in almost all the major crops with the advancement of DNA markers, such as RFLP (Restriction Fragment Length Polymorphism), AFLP (Amplified Fragment Length Polymorphism) and sequence-based markers like SSR (Simple Sequence Repeats) and SNPs. Comparative studies of linkage maps between species are useful in predicting diversity, genome evolution and organization. A number of genetic linkage maps have been generated in *Brassica*, utilizing different sets of markers and mapping populations. Linkage maps provide a basis for genetic architecture analysis of the genome and sequence-

based marker data facilitate comparative mapping and the study of genomic relations among species.

In *Brassica rapa*, during the last decade, more than ten linkage maps have been developed mainly based on molecular markers (RFLP, RAPD, AFLP, and SSR) using different mapping populations (Kim *et al.* 2006, Lou *et al.* 2008, Sakamoto *et al.* 2008, Song *et al.* 1991, Suwabe *et al.* 2004, 2006, Wang *et al.* 2004), which has made comparative analysis with each other difficult without a common reference map. Choi *et al.* (2007) constructed the first reference genetic map for *B. rapa* using doubled haploid lines derived from a cross between two diverse Chinese cabbage (*B. rapa* ssp. *pekinensis*) inbred lines, ‘‘Chiifu-401-42’’ and ‘‘Kenshin-402-43’’. The reference linkage map was updated with the addition of 156 BAC-end SSR markers (Kim *et al.* 2009) and subsequently was used for high-density integrated map construction (Li *et al.* 2010). Recently, the genome of the *B. rapa* inbred line Chiifu-401-42 has been completely sequenced under the *Brassica rapa* Genome Consortium (Wang, X. *et al.* 2011) and the reference linkage map has facilitated assignment of sequence scaffolds to the chromosomes.

Brassica oleracea, representing the C genome of *Brassica*, comprises various vegetables, one of which, cabbage (*B. oleracea* var. *capitata*), has been considered for the genome sequencing project. In *B. oleracea*, more than ten linkage maps have been developed using RFLP, AFLP or SSR markers in different mapping populations (Iniguez-Luy *et al.* 2009, Okazaki *et al.* 2007, Schmidt and Bancroft 2011). Integrated maps in *B. oleracea* have also been constructed with RFLP and AFLP markers by Kianian and Quiros (1992) and Sebastian *et al.* (2002). Available expressed sequence tag (EST) sequences of *Arabidopsis* and *Raphanus* have also been explored to construct several other maps and have allowed comparison of the *B. oleracea* genome with the *Arabidopsis* genome (Ashutosh *et al.* 2012, Babula *et al.* 2003, Kifuji *et al.* 2013, Kowalski *et al.* 1994, Lan *et al.* 2000). A high-density linkage map using Sequence-Related Amplified Polymorphism (SRAP) markers was developed in *B. oleracea* and identified QTLs of curd formation in cauliflower (Gao *et al.* 2007). In *Brassica*, 56,465 non-redundant SSR markers identified from *B. oleracea* whole-genome shotgun sequences were preferentially located on the C genome, and of these 752 markers showed polymorphism among six *B. napus* varieties (Li, H. *et al.* 2011). As the *B. oleracea* genome sequencing project was launched, a high-density reference map was drafted including 602 SSRs and 625 SNP markers generated from whole-genome shotgun sequences by NGS, covering 1197.9 cM (Wang, W. *et al.* 2012). This is also the first map that has allowed the assembled scaffold to be anchored to pseudochromosomes, which has significantly contributed to *Brassica* genome studies.

Brassica nigra (BB), one of the diploid *Brassica*, has not been studied extensively at the genomic level relative to other *Brassica* species despite a rich source of agronomically

important genes in terms of disease resistance, drought tolerance and seed oil quality (Chevre *et al.* 1996, Sjödin and Glimelius 1989, Struss *et al.* 1996). In *B. nigra*, the first linkage map was developed by Truco and Quiros (1994) using isozymes, RFLP and RAPD markers. In total, 124 markers were assigned to 11 linkage groups covering a total distance of 677 cM. A comprehensive linkage map of *B. nigra* was constructed with 160 DNA probes from *Arabidopsis* and identified 284 homologous loci covering a 750 cM region (Lagercrantz 1998).

Brassica napus (AACC), a major oilseed crop, is a highly accessed *Brassica* in terms of genetics and genomics. More than 30 linkage maps have been developed using various types of mapping populations and different molecular markers for different agronomic traits. The linkage maps in *B. napus* were developed using AFLP (Mei *et al.* 2009, Qiu *et al.* 2006, Radoev *et al.* 2008), RFLP (Parkin *et al.* 1995, Uzunova *et al.* 1995), SRAP (Sun *et al.* 2007), and SSR (Piquemal *et al.* 2005, Wang, J. *et al.* 2011) markers for developmental traits, seed quality and disease resistance. These maps have provided valuable information for rapeseed improvement and also in genome structure analysis. Recently, a high-density SNP linkage map, consisting of 5,764 SNP and 1603 PCR markers, was developed by integrating four DH populations to detect polymorphism level and linkage disequilibrium across different collections (Delourme *et al.* 2013).

Brassica juncea (AABB), one of the six cultivated *Brassica*, is the major oilseed crop of India. Relative to *B. napus*, genetic and genomic studies in *B. juncea* have been done less intensively, but in recent years the international community has given more attention to *B. juncea* because of its resistance to salinity and seed shattering. Early linkage maps in *B. juncea* were developed with RFLP and AFLP markers to investigate various traits (Axelsson *et al.* 2000, Christianson *et al.* 2006, Pradhan *et al.* 2003). A high-density linkage map in *B. juncea* was developed using AFLP, RFLP, SSR and gene-based markers with a total of 1,148 loci covering 1,840 cM of 18 linkage groups (Ramchiary *et al.* 2007). Although these linkage maps were useful for breeding and tagging of important traits, they provided limited information for comparative mapping.

Recent work on *B. carinata* (BBCC), one of the six cultivated amphidiploids, suggests it has better adaptability and productivity in semi-arid and temperate areas compared to oilseed rape. Being resistant to various diseases and biotic stress, *B. carinata* is suitable to cultivate in temperate environments (Getinet *et al.* 1996) and is also a potential crop in biofuel production (Cardone *et al.* 2003). Although genetic diversity analysis of this species was carried out, limited work has been done on genomic studies. Recently, a linkage map of *B. carinata* has been constructed and 212 loci were assigned to seventeen linkage groups covering a region of 1703 cM (Guo *et al.* 2012).

Raphanus sativus (radish), a member of Brassicaceae, is used all over the world as a vegetable crop with an edible

taproot. Although *Raphanus* is an economically important crop, genetic and genomic research has not progressed as in *B. rapa* and *B. napus*. A number of genetic maps have been developed in *R. sativus* using RFLP, AFLP, SSR and EST-SNP to analyze QTLs for disease resistance, root shape, flowering time, and pigmentation (Bett and Lydiat 2003, Budahn *et al.* 2009, Hashida *et al.* 2013, Kamei *et al.* 2010, Tsuru *et al.* 2005, Yu, X. *et al.* 2013, Zou, Z. *et al.* 2013). EST-based SNP and SSR were utilized to construct dense linkage maps and alignment of marker sequences to known *Brassica* sequences identified extensive chromosome homology among Brassicaceae (Li, F. *et al.* 2011, Shirasawa *et al.* 2011). The *Brassica* SSR and BAC-end sequence markers have also been explored in *R. sativus* in identification of QTLs for *Fusarium* wilt resistance trait (Yu, X. *et al.* 2013).

Comparative mapping for identification of conserved genomic segments

Arabidopsis, a member of Brassicaceae, is closely related to *Brassica* at the genomic sequence level, and shows around 85–90% identity in the exonic regions (Schmidt 2002). The fact that molecular markers of a *Brassica* species are transferrable to other *Brassica* species helps comparative mapping studies between cultivated *Brassica* species and with *A. thaliana*, as well as other Brassicaceae crops. On the basis of comparative mapping studies between *A. thaliana* and the ancestral karyotypes, 24 crucifer genomic blocks (A–X) have been proposed by Schranz *et al.* (2006), which are now widely accepted by scientific communities. By comparative genetic mapping between *Arabidopsis* and *Brassica* species, the presence of segmental duplications and genome rearrangements of *Brassica* A, B and C genomes was proposed and confirmed at the micro or macro level (Navabi *et al.* 2013, O'Neill and Bancroft 2000, Parkin *et al.* 2005). Lukens *et al.* (2003) attempted a comparison of mapped RFLP probe sequences of *B. oleracea* with the *Arabidopsis* genome sequence and identified 34 genomic collinear regions. In *B. oleracea*, through cDNA or BAC sequence comparison with *Arabidopsis* and *B. rapa*, they identified conserved collinearity for gene order and content of specific chromosomal segments (Li *et al.* 2003, Qiu *et al.* 2009). In *B. napus*, by sequencing of mapped RFLP probes and comparing these with the *Arabidopsis* genome, Parkin *et al.* (2005) identified 21 genomic blocks linked to the A and C genomes. Most of these conserved segments were found in six copies, which confirm the proposed hexaploid ancestor for the diploid *Brassica* progenitors. These genomic segments could be duplicated and rearranged in the present-day *B. napus* genome. Panjabi *et al.* (2008) extended comparative work in *B. juncea* and used *Arabidopsis*-based polymorphic intron PCR markers to identify conserved chromosomal regions and evolutionary relationships of the A, B and C genomes of *Brassica*. BAC- and SSR-based linkage maps of *B. rapa* (Choi *et al.* 2007, Kim *et al.* 2009) were

adopted as references to anchor sequence contigs of the international *B. rapa* genome sequence project and facilitated identification of conserved genomic blocks between *Arabidopsis* and *B. rapa*. The genome sequence of *B. rapa* provides an important resource for comparative mapping (BrGSP, Wang, X. *et al.* 2011). A conserved genomic restructuring in *B. napus* was confirmed by comparative mapping of dense linkage maps based on SSR and SNP markers with *Arabidopsis* and *B. rapa* (Bancroft *et al.* 2011, Wang, J. *et al.* 2011). A linkage map was developed in *B. carinata* and comparative mapping with the *Arabidopsis* sequence identified conserved ancestral building blocks (Guo *et al.* 2012). Recently, *B. nigra* BAC libraries have been sequenced and compared with *Arabidopsis* chromosome 4 and homologous *Brassica* A and C genomes, identifying conserved collinearity for gene content and order (Navabi *et al.* 2013). Extensive chromosome sequence homology was also revealed in *Raphanus* by comparing an EST-SNP-based linkage map (Li, F. *et al.* 2011) and an EST-SSR map with sequences of *Arabidopsis* and *B. rapa* genomes (Shirasawa *et al.* 2011). Very recently, comparing the whole-genome sequence of *B. rapa* with genome sequences or genetic maps of other crucifer species, the conserved genomic block boundaries were re-defined for seven ancestral karyotype blocks, deciphering the diploid ancestral genome of mesohexaploid *B. rapa* (Cheng *et al.* 2013).

Functional genomic regions and candidate genes have also been identified by a comparative mapping approach in *Brassicaceae* (Schmidt and Bancroft 2011). Mapping and identification of candidate genes in *B. rapa* by comparative genomic study have been reported: for example, cloning of flowering time *FLC* genes (Schranz *et al.* 2002), clubroot resistance genes syntenic to the *Arabidopsis* chromosome (Suwabe *et al.* 2006), and mapping QTLs for flowering time (Li *et al.* 2009). QTLs for clubroot resistance were identified in *B. oleracea* and comparative analysis of resistance genes was performed between *B. rapa* and *B. oleracea* (Nagaoka *et al.* 2010). In *B. juncea*, comparative mapping of QTL regions for aliphatic glucosinolate with the corresponding *Arabidopsis* sequence identified candidate genes regulating the aliphatic glucosinolate biosynthetic pathway (Bisht *et al.* 2009). In *B. oleracea*, candidate genes for male fertility were identified by comparing sequence-tagged markers with genome sequences of *Arabidopsis* and *B. rapa* (Ashutosh *et al.* 2012). As the *B. napus* genome sequence is not available, sequences of diploid progenitors *B. rapa* and *B. oleracea* were utilized in comparative mapping with *Arabidopsis* to identify candidate genes of QTLs for seed weight in *B. napus* (Cai *et al.* 2012). Li *et al.* (2013) have identified five major functional conserved genomic regions containing QTLs for morphological and yield traits between A, B, C subgenomes of *B. rapa*, *B. juncea* and *B. napus*. The knowledge gained from comparative analysis has revealed high-level sequence collinearity across *Brassicaceae* and helps in understanding genome evolution and polyploidization. Comparative genomic studies give confidence

in identifying orthologous candidate genes for important agronomic traits in *Brassica* crops and help in generating an integrated linkage map of species.

Next-generation genotyping techniques in *Brassica*

Recent advances of NGS technology have facilitated the discovery of various approaches of simultaneous sequence variant analysis and genotyping. Selected genomic regions or targeted restriction fragments of pooled individuals can be sequenced in a single reaction of a massive parallel sequencing platform. The sequences are aligned to the reference genome to compare assembled individual sequences and to identify variant sites to discover SNPs. These approaches are cost-effective and highly efficient in generating large amounts of informative data. Different protocols are available, such as complexity reduction of polymorphic sequences (CRoPS, Van Orsouw *et al.* 2007), restriction-associated DNA sequencing (RADseq, Baird *et al.* 2008), genotyping by sequencing (GBS, Elshire *et al.* 2011), and diversity arrays technology (DArT, Jaccoud *et al.* 2001). Each of the above protocols has its advantages and limitations but is reliable in SNP discovery and genotyping. Comparisons of these protocols have been explained in various reviews (Davey *et al.* 2011, Nielsen *et al.* 2011). These genotyping methods have been explored only in a few *Brassica* species, although transcriptome sequence techniques have been used for SNP discovery in *B. napus* and *B. rapa* (Hu *et al.* 2012, Trick *et al.* 2009a, 2009b).

The RADseq technique used in *B. napus* identified more than 20,000 SNPs and simultaneously genotyped eight different inbred lines (Bus *et al.* 2012). This method is simple, cost-effective, efficient and an alternative to transcriptome sequencing in SNP genotyping. DArT markers developed in *Brassica* and related species have been used in molecular diversity analysis of 89 different accessions of *B. napus*, *B. rapa*, *B. juncea* and *B. carinata* (Raman *et al.* 2012). Recently, a consensus linkage map based on DArT markers has been developed in *B. napus*, consisting of 1,359 markers spanning all 19 chromosomes covering a total of 1,987.2 cM with an average map density of one marker per 1.46 cM (Raman *et al.* 2013). Most of the DArT markers sequenced and aligned with *B. rapa* and *B. oleracea* genomes are useful in comparative mapping and genome evolution studies. Wells *et al.* (2013) developed a methodology based on pooled PCR product sequencing that incorporates bar-coded amplification tags (BATs) into PCR products. Using this method, targeted gene sequences were screened in a *B. napus* population and the resulting allele scoring mapped 24 markers on the expected position of the *B. napus* linkage map. In summary, next-generation high-throughput genotyping techniques are capable of providing increased marker density for genome selection or genome-wide association studies. Furthermore, next-generation genotyping methods need to be explored in *Brassica* and *Raphanus* to generate high-density linkage maps.

EST sequences and transcriptomes

In Brassicaceae, a large amount of data of EST sequences have been generated from all major *Brassica* crops and related species. In the NCBI database, approximately 1,500,000 EST sequences are available from various tissues exposed to different stress/growth conditions of ten different species of Brassicaceae (excluding *Arabidopsis*). Love *et al.* (2005) initiated development of a *Brassica* microarray by assembling EST sequences. *Brassica* 95K EST microarray was developed through clustering and assembling 810,254 *Brassica* raw ESTs, available in 2007, in non-redundant unigenes (Trick *et al.* 2009a). These unigenes have been loaded on a web portal (<http://brassica.bbsrc.ac.uk>) and are a valuable source for comparative mapping and genome analysis. Unigene sets are useable as pseudo-reference sequences for re-sequencing projects by next-generation techniques and assembling transcriptomes (Trick *et al.* 2009b) and SNP chips (Hayward *et al.* 2012).

In radish, rich EST sequences are available in the public domain and are being used for generation of linkage maps and genetic studies. In total, 3,800 EST-SSR markers were developed from 26,606 ESTs derived from different tissues of *R. sativus* and a linkage map was constructed using 630 EST-SSR and 213 reported marker loci covering a 1,129.3 cM region (Shirasawa *et al.* 2011). Available radish ESTs in databases (Radish DB; <http://radish.plantbiology.msu.edu>) were explored to discover SNPs and to construct an *R. sativus* linkage map consisting of 726 markers, and 72 syntenic regions to *Arabidopsis* were identified (Li, F. *et al.* 2011). RNA-seq-based transcriptome profiling of *Raphanus* root was utilized for identification of genes in response to metal Pb stress and 22 genes have been validated by quantitative real-time PCR (Wang *et al.* 2013). In the latest updated information, a total of 311,799 high-quality EST sequences were generated from raw EST data and further assembled in 85,083 unigenes (RadishBase, bioinfo.bti.cornell.edu/radish) (Shen *et al.* 2013).

In recent years, with the advancement of NGS technology, it has become possible to economically re-sequence whole genomes or generate large amount of transcriptome data in a short time. These sequences have been utilized for variant analysis to develop genic and functional markers. NGS techniques have been utilized to generate transcriptome sequences in polyploid *B. napus* and to discover single nucleotide polymorphism (Hu *et al.* 2012, Trick *et al.* 2009b). Furthermore, the *B. napus* genome was dissected by transcriptome sequences of parental and mapping population leaf samples and an SNP linkage map of about 23,000 markers was constructed (Bancroft *et al.* 2011). Sequence comparison of the *B. napus* genome with its progenitors *B. rapa* and *B. oleracea* revealed genome re-arrangements and detected a track of genomic segment inheritance. Transcriptome profiling based on deep EST sequencing in *B. napus* and three other oilseed species revealed both conserved and distinct species-specific expression patterns for

genes involved in the synthesis of glycerolipids and their precursors (Troncoso-Ponce *et al.* 2011). Higgins *et al.* (2012) employed a next-generation-based RNA-seq technique to discriminate A and C genome transcriptomes in amphidiploid *B. napus* and measured the contribution of gene expression by each genome. The associative transcriptomics approach has been explored in *B. napus* to identify genomic deletions in QTL regions of glucosinolate content of seeds (Harper *et al.* 2012). A different gene expression pattern in response to water logging has also been identified in *B. napus* roots at the seedling stage (Zou, X. *et al.* 2013).

In *Brassica rapa*, abiotic stress transcriptome studies identified 56 transcription factors and 60 genes commonly expressed under various stresses (Lee *et al.* 2008). The gene expression pattern in different tissues of *B. rapa* analysed by RNA-seq revealed transcriptome complexity (Tong *et al.* 2013). Recently, in *B. juncea* (Tumourous stem mustard), transcription level analysis was performed to detect gene expression patterns at various stem development stages (Sun *et al.* 2012). In brief, the advancement of transcriptomic studies helps to understand the complexity of gene expression and regulation networks at various developmental stages and the response to biotic/abiotic stresses; in addition, high-resolution genome dissection provides resources for comparative and functional genomics.

Genome sequencing of the Brassicaceae family

Recent advances in high-throughput sequencing technology have immensely benefitted whole-genome sequencing projects in non-model organisms and have opened a new era in comparative genomics. In the Brassicaceae family, to date, the genomes of ten species have been partially or completely sequenced, including the model plant *A. thaliana* and cultivated *Brassica* species, e.g., *B. rapa* and *B. oleracea*, summarized in Table 1. The annotated *Arabidopsis* genome sequence provides a valuable reference, and genome sequences have also been utilized to develop DNA markers and a number of informative linkage maps in cultivated *Brassica* species, which have been used to identify candidate genes. Most of the ancestral progenitor sequences have been used for genome evolution studies and identification of conserved ancestral genomic segments. The sequencing project of 1,001 accessions of *A. thaliana* will enable the study of genome-wide association in this species (<http://1001genomes.org>). These sequences will provide a link of phenotypic diversity with genome variation and generate large resources for the plant community.

Sequencing of *Brassica* genomes are required for finding important genes, understanding of genome evolution and improvement of crops. Considering this, and the importance of *Brassica* crops, sequencing of the *Brassica* genomes was initiated in 2002 by the Multinational Brassica Genome Project (MBGP). The *B. rapa* Chinese cabbage (cv. Chiifu-401) was the first genome selected for sequencing because of its small genome size (529 Mb) and low frequency of

Table 1. Summary of genome sequences completed in Brassicaceae species

Species	Genome size ^a (Mb)	Number of predicted genes	% of genes ortholo- gous to <i>A. thaliana</i>	References
<i>Arabidopsis thaliana</i>	135–157	28,710	100	Arabidopsis Genome Initiative 2000
<i>Arabidopsis lyrata</i>	230–245	27,379	92	Hu <i>et al.</i> 2011
<i>Schrenkiella parvula</i>	140	28,901	80.2	Dassanayake <i>et al.</i> 2011
<i>Brassica rapa</i>	529	41,174	78.2	Wang, X. <i>et al.</i> 2011
<i>Capsella rubella</i>	210–216	26,521	88	Slotte <i>et al.</i> 2013
<i>Eutrema salsugineum</i>	314	26,521	82.7	Yang <i>et al.</i> 2013
<i>Leavenworthia alabamica</i>	316	30,343	67.7	Haudry <i>et al.</i> 2013
<i>Sisymbrium irio</i>	262	28,917	82.9	Haudry <i>et al.</i> 2013
<i>Aethionema arabicum</i>	240	23,167	72.4	Haudry <i>et al.</i> 2013
<i>Brassica oleracea</i>	696	45,758	–	Yu <i>et al.</i> 2013

^a Genome size reference from Johnston *et al.* (2005) or adapted from Haudry *et al.* 2013.

repetitive sequences. The draft genome sequence of *B. rapa* (A genome) has been published by the *Brassica rapa* Genome Sequencing Project (BrGSP) (Wang, X. *et al.* 2011). A total of 41,174 protein-encoding genes were modeled on the *B. rapa* genome by assembling 1,427 markers, and 10 pseudochromosomes have been produced. The genome sequence of another important vegetable crop, *B. oleracea* (C genome), has recently been completed using a whole-genome shotgun (WGS) sequencing strategy. A 630 Mb assembled draft genome sequence was obtained, with a scaffold N50 size of 1.457 Mb and contig size of 26.828 kb, and assigned to nine pseudochromosomes containing 45,758 predicted genes (Yu, J. *et al.* 2013, <http://www.ocri-genomics.org/bolbase/index.html>).

Recently, Haudry *et al.* (2013) sequenced three genomes of Brassicaceae species, i.e., *Leavenworthia alabamica*, *Sisymbrium irio* and *Aethionema arabicum*. Comparative analysis with the previously sequenced genomes identified 90,000 conserved noncoding sequences (CNS) in Brassicaceae that show evidence of transcriptional and post-transcriptional regulation. Currently, work is in progress to complete the genome sequences of other *Brassica* and *Raphanus* species in the near future.

Genomic resources

Sequencing technology advancement has produced vast genomic information and sequence data in major crops. In the past decade, various *Brassica* databases were integrated on a common platform to facilitate efficient utilization by diverse researchers. An open access integrated database provides annotated genome information, genetic and physical maps, molecular markers, reference maps and gene expression data. The UK *Brassica* community put initiative in this direction in 1996 by compiling *Brassica* sequences and genetic maps to create the BrassicaDB database. A major advance in knowledge sharing realized the initiation of the Multinational Brassica Genome Project (MBGP) in 2002. A number of open access databases are available in Brassicaceae with systematics information on linkage maps, QTL maps, details of mapping populations, BAC libraries, marker data, EST repositories and genome sequences. The anno-

tated *B. rapa* genome sequence is available on BRAD Brassica database (IVF-CAAS, China) and Brass ensemble (Rothensted Research, UK) web resources. Recently, the *B. oleracea* genome sequence has become available for comparative analysis on the Bolbase data source (<http://www.ocri-genomics.org/bolbase/index.html>), although the complete genome for download is yet to be released. Radish Base, a database of genetics and genomics of radish, was recently developed by Cornell University, USA, and consists of SSR, EST, and SNP marker information, linkage maps and organelle genome sequences. Currently, many genetic and genomic resources in Brassicaceae are available and are summarized in Table 2. The integrated knowledge available in the public domain will provide a platform to exchange information and a basis for crop *Brassica* enhancement.

Conclusion and perspectives

Since the accomplishment of genomic sequencing of the model plant *A. thaliana*, and later *B. rapa* and *B. oleracea*, comparative mapping between these species and important Brassicaceae crops has been possible. The presence of duplicated and repetitive DNAs complicates the proper alignment and identification of actual causal genes out of many paralogs. Genome sequence information on the other four cultivated *Brassica* genomes is still not available. Since the sequences of *Brassica* species are highly conserved, molecular markers and genomic information obtained for extensively studied *B. rapa*, *B. oleracea* and *B. napus* could be transferred to other commercial *Brassica* crops. Construction of high-density consensus genetic maps, common marker systems, and genomic sequence information is of great significance for accelerating breeding progress, as it allows comparative QTL mapping analysis, marker-assisted selection and cloning of economically important genes for desired traits.

Although many genomic resources have been established, and genomic sequencing of several Brassicaceae crops has been finished, most comparative studies are on a structural genomics level, and only a handful of genes governing important traits have been identified and functionally

Table 2. List of web resources and databases providing bioinformatics analysis and genomic resources for the Brassicaceae

Names	URL	Key contents
ACPFPG Applied Bioinformatics group	http://www.appliedbioinformatics.com.au/index.php/Main_Page http://www.brassicagenome.net/	<i>B. rapa</i> genome browser, EST-SNP data base, BrassicaDB, CMap to compare genome and genetic map
Bolbase	http://www.ocri-genomics.org/bolbase/	Genomic data of <i>B. oleracea</i> , analysis of genome structure as well as syntenic regions, browse, search and download genome of <i>B. rapa</i> and <i>A. thaliana</i>
BRAD	http://brassicadb.org/brad/	Compilation of sequence datasets including the complete sequence of <i>B. rapa</i> . Annotations of genes orthologous to those in <i>A. thaliana</i> , and genetic markers and genetic maps, BLAST server
BrassEnsembl	http://www.brassica.info/BrassEnsembl/index.html	<i>B. rapa</i> genome sequence, consensus integrated genetic maps of the <i>Brassica</i> A and C genomes
Brassica Genome Gateway	http://brassica.nbi.ac.uk	Brassica genome sequencing database, Brassica 95K uni-gene set, the <i>Brassica</i> IGF Project, BrassicaDB
Brassica.info	http://www.brassica.info/	Web-based open source to exchange information relating to <i>Brassica</i> genomics and genetics, registries of reference datasets, nomenclature standards, a compilation of ongoing public domain genome sequencing
BrassicaDB	http://brassica.nbi.ac.uk/BrassicaDB/	Comprehensive sequence data set, genetic maps and markers in <i>Brassica</i> species, BLAST server, physical maps
CropStoreDB	http://www.cropstoredb.org	A collection of datasets related to plant and crop genetics, <i>Brassica</i> data implemented
Radish Base	http://bioinfo.bti.cornell.edu/cgi-bin/radish/index.cgi	Assembled and annotated ESTs, predicted metabolic pathways, EST-SSR, SNP markers, and genetic maps
Radish database	http://radish.plantbiology.msu.edu/index.php/Main_Page	EST sequences, linkage maps, SNP and SSR markers, radish genome sequence updates

characterized. Thus, the genomic and comparative genomic resources that are being established are only a starting point for exploring the variation within Brassicaceae. In the near future, functional genomics should increasingly be used to identify desired genes for directed gene-assisted selection of economically important traits, and to detect genetic variation within the species, by combining various techniques, such as transcriptomic analysis and high-throughput genotyping and phenotypic characterization, to study the expression of duplicated genes under different environmental conditions.

Acknowledgments

Ashutosh Sharma was a recipient of a research fellowship from the Japan Society for the Promotion of Science for Foreign Scientists and his work was supported in part by the Program for Promotion of Basic and Applied Research for Innovations in Bio-oriented Industry (BRAIN), Japan. Work of YPL and XNL was supported by a grant from the Next-Generation BioGreen 21 Program (Plant Molecular Breeding Center No. PJ007992), Rural Development Administration, Republic of Korea. We are grateful to Prof Gareth Jenkins and Stuart Sullivan, University of Glasgow, for their valuable comments on this article.

Literature Cited

- Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408: 796–815.
- Ashutosh, B. Sharma, T. Shinada, Y. Kifuji, H. Kitashiba and T. Nishio (2012) Molecular mapping of a male fertility restorer locus of *Brassica oleracea* using EST-based SNP markers and analysis of a syntenic region in *Arabidopsis thaliana* for identification of genes encoding PPR proteins. *Mol. Breed.* 30: 1781–1792.
- Augustine, R., G.C. Arya, D.M. Nambiar, R. Kumar and N.C. Bisht (2013) Translational genomics in Brassica crops: challenges, progress, and future prospects. *Plant Biotechnol. Rep.* 8: 65–81.
- Axelsson, T., C.M. Bowman, A.G. Sharpe, D.J. Lydiate and U. Lagercrantz (2000) Amphidiploid *Brassica juncea* contains conserved progenitor genomes. *Genome* 43: 679–688.
- Babula, D., M. Kaczmarck, A. Barakat, M. Delseny, C.F. Quiros and J. Sadowski (2003) Chromosomal mapping of *Brassica oleracea* based on Arabidopsis ESTs: complexity of comparative mapping. *Mol. Gen. Genet.* 268: 656–665.
- Baird, N.A., P.D. Etter, T.S. Atwood, M.C. Currey, A.L. Shiver, Z.A. Lewis, E.U. Selker, W.A. Cresko and E.A. Johnson (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One* 3: e3376.
- Bancroft, I., C. Morgan, F. Fraser, J. Higgins, R. Wells, L. Clissold, D. Baker, Y. Long, J. Meng, X. Wang *et al.* (2011) Dissecting the

- genome of the polyploid crop oilseed rape by transcriptome sequencing. *Nat. Biotechnol.* 29: 762–766.
- Bett, K.E. and D.J. Lydiate (2003) Genetic analysis and genome mapping in *Raphanus*. *Genome* 46: 423–430.
- Bisht, N.C., V. Gupta, N. Ramchiary, Y.S. Sodhi, A. Mukhopadhyay, N. Arumugam, D. Pental and A.K. Pradhan (2009) Fine mapping of loci involved with glucosinolate biosynthesis in oilseed mustard (*Brassica juncea*) using genomic information from allied species. *Theor. Appl. Genet.* 118: 413–421.
- Budahn, H., H. Peterka, M.A. Mousa, Y. Ding, S. Zhang and J. Li (2009) Molecular mapping in oil radish (*Raphanus sativus* L.) and QTL analysis of resistance against beet cyst nematode (*Heterodera schachtii*). *Theor. Appl. Genet.* 118: 775–782.
- Bus, A., J. Hecht, B. Huettel, R. Reinhardt and B. Stich (2012) High-throughput polymorphism detection and genotyping in *Brassica napus* using next-generation RAD sequencing. *BMC Genomics* 13: 281.
- Cai, G., Q. Yang, Q. Yang, Z. Zhao, H. Chen, J. Wu, C. Fan and Y. Zhou (2012) Identification of candidate genes of QTLs for seed weight in *Brassica napus* through comparative mapping among *Arabidopsis* and *Brassica* species. *BMC Genet.* 13: 105.
- Cardone, M., M. Mazzoncini, S. Menini, V. Rocco, A. Senatore, M. Seggiani and S. Vitolo (2003) *Brassica carinata* as an alternative oil crop for the production of biodiesel in Italy: agronomic evaluation, fuel production by transesterification and characterization. *Biomass and Bioenergy* 25: 623–636.
- Cheng, F., T. Mandaková, J. Wu, Q. Xie, M.A. Lysak and X. Wang (2013) Deciphering the diploid ancestral genome of the mesohexaploid *Brassica rapa*. *Plant Cell* 25: 1541–1554.
- Chevre, A.M., F. Eber, P. This, P. Barret, X. Tanguy, H. Brun, M. Delseny and M. Renard (1996) Characterization of *Brassica nigra* chromosomes and of blackleg resistance in *B. napus*-*B. nigra* addition lines. *Plant Breed.* 115: 113–118.
- Choi, S.R., G.R. Teakle, P. Plaha, J.H. Kim, C.J. Allender, E. Beynon, Z.Y. Piao, P. Soengas, T.H. Han, G.J. King *et al.* (2007) The reference genetic linkage map for the multinational *Brassica rapa* genome sequencing project. *Theor. Appl. Genet.* 115: 777–792.
- Christianson, J.A., S.R. Rimmer, A.G. Good and D.J. Lydiate (2006) Mapping genes for resistance to *Leptosphaeria maculans* in *Brassica juncea*. *Genome* 49: 30–41.
- Dassanayake, M., D.H. Oh, J.S. Haas, A. Hernandez, H. Hong, S. Ali, D.J. Yun, R.A. Bressan, J.K. Zhu, H.J. Bohnert *et al.* (2011) The genome of the extremophile crucifer *Thellungiella parvula*. *Nat. Genet.* 43: 913–918.
- Davey, J.W., P.A. Hohenlohe, P.D. Etter, J.Q. Boone, J.M. Catchen and M.L. Blaxter (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nat. Rev. Genet.* 12: 499–510.
- Delourme, R., C. Falentin, B. Fomeju, M. Boillot, G. Lassalle, I. André, J. Duarte, V. Gauthier, N. Lucante, A. Marty *et al.* (2013) High-density SNP-based genetic map development and linkage disequilibrium assessment in *Brassica napus*. *BMC Genomics* 14: 120.
- Elshire, R.J., J.C. Glaubitz, Q. Sun, J.A. Poland, K. Kawamoto, E.S. Buckler and S.E. Mitchell (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* 6: e19379.
- Gao, M., G. Li, B. Yang, D. Qiu, M. Farnham and C. Quiros (2007) High-density *Brassica oleracea* linkage map: identification of useful new linkages. *Theor. Appl. Genet.* 115: 277–287.
- Getinet, A., G. Rakow and R.K. Downey (1996) Agronomic performance and seed quality of Ethiopian mustard in Saskatchewan. *Can. J. Plant Sci.* 76: 387–392.
- Guo, S., J. Zou, R. Li, Y. Long, S. Chen and J. Meng (2012) A genetic linkage map of *Brassica carinata* constructed with a double haploid population. *Theor. Appl. Genet.* 125: 1113–1124.
- Harper, A.L., M. Trick, J. Higgins, F. Fraser, L. Clissold, R. Wells, C. Hattori, P. Werner and I. Bancroft (2012) Associative transcriptomics of traits in the polyploid crop species *Brassica napus*. *Nat. Biotechnol.* 30: 798–802.
- Hashida, T., R. Nakatsuji, O. Budahn, O. Schrader, H. Peterka, T. Fujimura, N. Kubo and M. Hirai (2013) Construction of a chromosome-assigned, sequence-tagged linkage map for the radish, *Raphanus sativus* L. and QTL analysis of morphological traits. *Breed. Sci.* 63: 218–226.
- Haudry, A., A.E. Platts, E. Vello, D.R. Hoen, M. Leclercq, R.J. Williamson, E. Forczek, Z. Joly-Lopez, J.G. Steffen, K.M. Hazzouri *et al.* (2013) An atlas of over 90,000 conserved noncoding sequences provides insight into crucifer regulatory regions. *Nat. Genet.* 45: 891–898.
- Hayward, A., A.S. Mason, J. Dalton-Morgan, M. Zander, D. Edwards and J. Batley (2012) SNP discovery and applications in *Brassica napus*. *J. Plant Biotechnol.* 39: 1–12.
- Higgins, J., A. Magusin, M. Trick, F. Fraser and I. Bancroft (2012) Use of mRNA-Seq to discriminate contributions to the transcriptome from the constituent genomes of the polyploid crop species *Brassica napus*. *BMC Genomics* 13: 247.
- Hu, T.T., P. Pattyn, E.G. Bakker, J. Cao, J.F. Cheng, R.M. Clark, N. Fahlgren, J.A. Fawcett, J. Grimwood, H. Gundlach *et al.* (2011) The *Arabidopsis lyrata* genome sequence and the basis of rapid genome size change. *Nat. Genet.* 43: 476–481.
- Hu, Z., S. Huang, M. Sun, H. Wang and W. Hua (2012) Development and application of single nucleotide polymorphism markers in the polyploid *Brassica napus* by 454 sequencing of expressed sequence tags. *Plant Breed.* 131: 293–299.
- Iniguez-Luy, F., L. Lukens, M.W. Farnham, R.M. Amasino and T.C. Osborn (2009) Development of public immortal mapping populations, molecular markers and linkage maps for rapid cycling *Brassica rapa* and *Brassica oleracea*. *Theor. Appl. Genet.* 120: 31–43.
- Jaccoud, D., K. Peng, D. Feinstein and A. Kilian (2001) Diversity arrays: a solid state technology for sequence information independent genotyping. *Nucleic Acids Res.* 29: E25.
- Johnston, J.S., A.E. Pepper, A.E. Hall, Z.J. Chen, G. Hodnett, J. Drabek, R. Lopez and H.J. Price (2005) Evolution of genome size in Brassicaceae. *Ann. Bot.* 95: 229–235.
- Kamei, A., M. Tsuru, N. Kubo, T. Hayashi, N. Wang, T. Fujimura and M. Hirai (2010) QTL mapping of clubroot resistance in radish (*Raphanus sativus* L.). *Theor. Appl. Genet.* 120: 1021–1027.
- Kianian, S.F. and C.F. Quiros (1992) Trait inheritance, fertility and genomic relationships of some n = 9 *Brassica* species. *Genet. Resour. Crop Evol.* 39: 165–175.
- Kifuji, Y., H. Hanzawa, Y. Terasawa, Ashutosh and T. Nishio (2013) QTL analysis of black rot resistance in cabbage using newly developed EST-SNP markers. *Euphytica* 190: 289–295.
- Kim, H.R., S.R. Choi, J. Bae, C.P. Hong, S.Y. Lee, M.D. Hossain, D.V. Nguyen, M. Jin, B.S. Park, J.W. Bang *et al.* (2009) Sequenced BAC anchored reference genetic map that reconciles the ten individual chromosomes of *Brassica rapa*. *BMC Genomics* 10: 432.
- Kim, J.S., T.Y. Chung, G.J. King, M. Jin, T.J. Yang, Y.M. Jin, H. Kim and B.S. Park (2006) A sequence tagged linkage map of *Brassica rapa*. *Genetics* 174: 29–39.
- Kowalski, S.P., T.H. Lan, K.A. Feldmann and A.H. Paterson (1994)

- Comparative mapping of *Arabidopsis thaliana* and *Brassica oleracea* chromosomes reveals islands of conserved organization. *Genetics* 138: 499–510.
- Lagercrantz, U. (1998) Comparative mapping between *Arabidopsis thaliana* and *Brassica nigra* indicates that *Brassica* genomes have evolved through extensive genome replication accompanied by chromosome fusions and frequent rearrangements. *Genetics* 150: 1217–1228.
- Lan, T.H., T.A. DelMonte, K.P. Reischmann, J. Hyman, S.P. Kowalski, J. McFerson, S. Kresovich and A.H. Paterson (2000) An EST-enriched comparative map of *Brassica oleracea* and *Arabidopsis thaliana*. *Genome Res.* 10: 776–788.
- Lee, S.C., M.H. Lim, J.A. Kim, S.I. Lee, J.S. Kim, M. Jin, S.J. Kwon, J.H. Mun, Y.K. Kim, H.U. Kim *et al.* (2008) Transcriptome analysis in *Brassica rapa* under the abiotic stresses using *Brassica* 24 K oligo microarray. *Mol. Cells* 26: 595–605.
- Li, F., H. Kitashiba, K. Inaba and T. Nishio (2009) A *Brassica rapa* linkage map of EST-based SNP markers for identification of candidate genes controlling flowering time and leaf morphological traits. *DNA Res.* 16: 311–323.
- Li, F., Y. Hasegawa, M. Saito, S. Shirasawa, A. Fukushima, T. Ito, H. Fujii, S. Kishitani, H. Kitashiba and T. Nishio (2011) Extensive chromosome homoeology among Brassicaceae species were revealed by comparative genetic mapping with high-density EST-based SNP markers in radish (*Raphanus sativus* L.). *DNA Res.* 18: 401–411.
- Li, G., M. Gao, B. Yang and C.F. Quiros (2003) Gene to gene alignment between the *Brassica* and *Arabidopsis* genomes by transcriptional mapping. *Theor. Appl. Genet.* 107: 168–180.
- Li, H., X. Chen, Y. Yang, J. Xu, J. Gu, J. Fu, X. Qian, S. Zhang, J. Wu and K. Liu (2011) Development and genetic mapping of microsatellite markers from whole genome shotgun sequences in *Brassica oleracea*. *Mol. Breed.* 28: 585–596.
- Li, X., N. Ramchiary, S.R. Choi, D. Van Nguyen, M.J. Hossain, H.K. Yang and Y.P. Lim (2010) Development of a high density integrated reference genetic linkage map for the multinational *Brassica rapa* Genome Sequencing Project. *Genome* 53: 939–947.
- Li, X., N. Ramchiary, V. Dhandapani, S.R. Choi, Y. Hur, I.S. Nou, M.K. Yoon and Y.P. Lim (2013) Quantitative trait loci mapping in *Brassica rapa* revealed the structural and functional conservation of genetic loci governing morphological and yield component traits in the A, B, and C subgenomes of *Brassica* species. *DNA Res.* 20: 1–16.
- Lou, P., J. Zhao, H. He, C. Hanhart, D.P. DelCarpio, R. Verkerk, J. Custers, M. Koornneef and G. Bonnema (2008) Quantitative trait loci for glucosinolate accumulation in *Brassica rapa* leaves. *New Phytol.* 179: 1017–1032.
- Love, C.G., A.J. Robinson, G.A. Lim, C.J. Hopkins, J. Batley, G. Baker, G.C. Spangenberg and D. Edwards (2005) *Brassica* ASTRA: an integrated database for *Brassica* genomic research. *Nucleic Acids Res.* 33: 656–659.
- Lukens, L., F. Zou, D. Lydiate, I. Parkin and T. Osborn (2003) Comparison of a *Brassica oleracea* genetic map with the genome of *Arabidopsis thaliana*. *Genetics* 164: 359–372.
- Mei, D.S., H.Z. Wang, Q. Hu, Y.D. Li, Y.S. Xu and Y.C. Li (2009) QTL analysis on plant height and flowering time in *Brassica napus*. *Plant Breed.* 128: 458–465.
- Nagaoka, T., M.A.U. Doullah, S. Matsumoto, S. Kawasaki, T. Ishikawa, H. Hori and K. Okazaki (2010) Identification of QTLs that control clubroot resistance in *Brassica oleracea* and comparative analysis of clubroot resistance genes between *B. rapa* and *B. oleracea*. *Theor. Appl. Genet.* 120: 1335–1346.
- Navabi, Z.K., T. Huebert, A.G. Sharpe, I. Bancroft, I.A. Parkin and C.M.O. Neill (2013) Conserved microstructure of the *Brassica* B Genome of *Brassica nigra* in relation to homologous regions of *Arabidopsis thaliana*, *B. rapa* and *B. oleracea*. *BMC Genomics* 14: 250.
- Nielsen, R., J.S. Paul, A. Albrechtsen and Y.S. Song (2011) Genotype and SNP calling from next-generation sequencing data. *Nat. Rev. Genet.* 12: 443–451.
- O'Neill, C.M. and I. Bancroft (2000) Comparative physical mapping of segments of the genome of *Brassica oleracea* var. *alboglabra* that are homoeologous to sequenced regions of chromosomes 4 and 5 of *Arabidopsis thaliana*. *Plant J.* 23: 233–243.
- Okazaki, K., K. Sakamoto, R. Kikuchi, A. Saito, E. Togashi, Y. Kuginuki, S. Matsumoto and M. Hirai (2007) Mapping and characterization of FLC homologs and QTL analysis of flowering time in *Brassica oleracea*. *Theor. Appl. Genet.* 114: 595–608.
- Panjabi, P., A. Jagannath, N.C. Bisht, K.L. Padmaja, S. Sharma, V. Gupta, A.K. Pradhan and D. Pental (2008) Comparative mapping of *Brassica juncea* and *Arabidopsis thaliana* using intron polymorphism (IP) markers: homeologous relationships, diversification and evolution of the A, B and C *Brassica* genomes. *BMC Genomics* 9: 113.
- Parkin, I.A.P., A.G. Sharpe, D.J. Keith and D.J. Lydiate (1995) Identification of the A and C genomes of amphidiploid *Brassica napus* (oilseed rape). *Genome* 38: 1122–1131.
- Parkin, I.A.P., S.M. Gulden, A.G. Sharpe, L. Lukens, M. Trick, T.C. Osborn and D.J. Lydiate (2005) Segmental structure of the *Brassica napus* genome based on comparative analysis with *Arabidopsis thaliana*. *Genetics* 171: 765–781.
- Piquemal, J., E. Cinquin, F. Couton, C. Rondeau, E. Seignoret, I. Doucet, D. Perret, M.J. Villeger, P. Vincourt and P. Blanchard (2005) Construction of an oilseed rape (*Brassica napus* L.) genetic map with SSR markers. *Theor. Appl. Genet.* 111: 1514–1523.
- Pradhan, A.K., V. Gupta, A. Mukhopadhyay, N. Arumugam, Y.S. Sodhi and D. Pental (2003) A high density linkage map in *Brassica juncea* (Indian mustard) using AFLP and RFLP markers. *Theor. Appl. Genet.* 106: 607–614.
- Qiu, D., C. Morgan, J. Shi, Y. Long, J. Liu, R. Li, X. Zhuang, Y. Wang, X. Tan, E. Dietrich *et al.* (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.
- Qiu, D., M. Gao, G. Li and C. Quiros (2009) Comparative sequence analysis for *Brassica oleracea* with similar sequences in *B. rapa* and *Arabidopsis thaliana*. *Plant Cell Rep.* 28: 649–661.
- Radoev, M., H. Becker and W. Ecke (2008) Genetic analysis of heterosis for yield and yield components in rapeseed (*Brassica napus* L.) by quantitative trait locus mapping. *Genetics* 179: 1547–1558.
- Raman, H., R. Raman, M.N. Nelson, M.N. Aslam, R. Rajasekaran, N. Wratten, W.A. Cowling, A. Kilian, A.G. Sharpe and J. Schondelmaier (2012) Diversity Array Technology markers: Genetic diversity analyses and linkage map construction in rapeseed (*Brassica napus* L.). *DNA Res.* 19: 51–65.
- Raman, H., R. Raman, A. Kilian, F. Detering, Y. Long, D. Edwards, I.A.P. Parkin, A.G. Sharpe, M.N. Nelson, N. Larkan *et al.* (2013) A consensus map of rapeseed (*Brassica napus* L.) based on diversity array technology markers: applications in genetic dissection of qualitative and quantitative traits. *BMC Genomics* 14: 277.
- Ramchiary, N., K.L. Padmaja, S. Sharma, V. Gupta, Y.S. Sodhi, A. Mukhopadhyay, N. Arumugam, D. Pental and A.K. Pradhan (2007) Mapping of yield influencing QTL in *Brassica juncea*: implications

- for breeding of a major oilseed crop of dryland areas. *Theor. Appl. Genet.* 115: 807–817.
- Sakamoto, K., A. Saito, N. Hayashida, G. Taguchi and E. Matsumoto (2008) Mapping of isolate-specific QTL for clubroot resistance in Chinese cabbage (*Brassica rapa* L. ssp. *pekinensis*). *Theor. Appl. Genet.* 117: 759–767.
- Schmidt, R. (2002) Plant genome evolution: lessons from comparative genomics at the DNA level. *Plant Mol. Biol.* 48: 21–37.
- Schmidt, R. and I. Bancroft (2011) Genetics and genomics of the Brassicaceae. *Plant genetics and genomics: crops and models*, vol 9. Springer, New York.
- Schranz, M.E., P. Quijada, S.B. Sung, L. Lukens, R.M. Amasino and T.C. Osborn (2002) Characterization and effects of the replicated time gene *FLC* in *Brassica rapa*. *Genetics* 162: 1457–1468.
- Schranz, M.E., M.A. Lysak and T. Mitchell-Olds (2006) The ABC's of comparative genomics in the Brassicaceae: building blocks of crucifer genomes. *Trends Plant Sci.* 11: 535–542.
- Sebastian, R.L., E.C. Howell, G.J. King, D.F. Marshall and M.J. Kearsey (2002) An integrated AFLP and RFLP *Brassica oleracea* linkage map from two morphologically distinct doubled-haploid mapping populations. *Theor. Appl. Genet.* 100: 75–81.
- Shen, D., H. Sun, M. Huang, Y. Zheng, X. Li and Z. Fei (2013) Radish-Base: a database for genomics and genetics of radish. *Plant Cell Physiol.* 54: e3.
- Shirasawa, K., M. Oyama, H. Hiraikawa, S. Sato, S. Tabata, T. Fujioka, C. Kimizuka-Takagi, S. Sasamoto, A. Watanabe, M. Kato *et al.* (2011) An EST-SSR linkage map of *Raphanus sativus* and comparative genomics of the Brassicaceae. *DNA Res.* 18: 221–232.
- Sjödín, C. and K. Glimelius (1989) Transfer of resistance against *Phoma lingam* to *Brassica napus* by asymmetric somatic hybridization combined with toxin selection. *Theor. Appl. Genet.* 78: 513–520.
- Slotte, T., K.M. Hazzouri, J.A. Agren, D. Koenig, F. Maumus, Y.L. Guo, K. Steige, A.E. Platts, J.S. Escobar, L.K. Newman *et al.* (2013) The *Capsella rubella* genome and the genomic consequences of rapid mating system evolution. *Nat. Genet.* 45: 831–835.
- Song, K.M., J.Y. Suzuki, M.K. Slocum, P.H. Williams and T.C. Osborn (1991) A linkage map of *Brassica rapa* base on restriction fragment length polymorphism loci. *Theor. Appl. Genet.* 82: 296–304.
- Struss, D., C.F. Quiros, J. Plieske and G. Röbbelen (1996) Construction of *Brassica* B genome synteny groups based on chromosomes extracted from three different sources by phenotypic, isozyme and molecular markers. *Theor. Appl. Genet.* 93: 1026–1032.
- Sun, Q., G.F. Zhou, Y.F. Cai, Y.H. Fan, X.Y. Zhu, Y. Liu, X. He, J. Shen, H. Jiang, D. Hu *et al.* (2012) Transcriptome analysis of stem development in the tumorous stem mustard *Brassica juncea* var. *tumida* Tsen et Lee by RNA sequencing. *BMC Plant Biol.* 12: 53.
- Sun, Z., Z. Wang, J. Tu, J. Zhang, F. Yu, P.B.E. McVetty and G. Li (2007) An ultradense genetic recombination map for *Brassica napus*, consisting of 13551 SRAP markers. *Theor. Appl. Genet.* 114: 1305–1317.
- Suwabe, K., H. Iketani, T. Nunome, A. Ohyama, M. Hirai and H. Fukuoka (2004) Characteristics of microsatellites in *Brassica rapa* genome and their potentials of utilization for comparative genomics in cruciferae. *Breed. Sci.* 54: 85–90.
- Suwabe, K., H. Tsukazaki, H. Iketani, K. Hatakeyama, M. Kondo, M. Fujimura, T. Nunome, H. Fukuoka, M. Hirai and S. Matsumoto (2006) Simple sequence repeat-based comparative genomics between *Brassica rapa* and *Arabidopsis thaliana*: the genetic origin of clubroot resistance. *Genetics* 173: 309–319.
- Tong, C., X. Wang, J. Yu, J. Wu, W. Li, J. Huang, C. Dong, W. Hua and S. Liu (2013) Comprehensive analysis of RNA-seq data reveals the complexity of the transcriptome in *Brassica rapa*. *BMC Genomics* 14: 689.
- Trick, M., F. Cheung, N. Drou, F. Fraser, E.K. Lobenhofer, P. Hurban, A. Magusin, C.D. Town and I. Bancroft (2009a) A newly-developed community microarray resource for transcriptome profiling in *Brassica* species enables the confirmation of *Brassica*-specific expressed sequences. *BMC Plant Biol.* 9: 50.
- Trick, M., Y. Long, J. Meng and I. Bancroft (2009b) Single nucleotide polymorphism (SNP) discovery in the polyploid *Brassica napus* using Solexa transcriptome sequencing. *Plant Biotechnol. J.* 7: 334–346.
- Troncoso-Ponce, M.A., A. Kilaru, X. Cao, T.P. Durrett, J. Fan, J.K. Jensen, N.A. Thrower, M. Pauly, C. Wilkerson and J.B. Ohlrogge (2011) Comparative deep transcriptional profiling of four developing oilseeds. *Plant J.* 68: 1014–1027.
- Truco, M.J. and C.F. Quiros (1994) Structure and organization of the B genome based on a linkage map in *Brassica nigra*. *Theor. Appl. Genet.* 89: 590–598.
- Tsuro, M., K. Suwabe, N. Kubo, S. Matsumoto and M. Hirai (2005) Construction of a molecular linkage map of radish (*Raphanus sativus* L.), based on AFLP and *Brassica*-SSR markers. *Breed. Sci.* 55: 107–111.
- U, N. (1935) Genome analysis in *Brassica* with special reference to the experimental formation of *B. napus* and peculiar mode of fertilization. *Jpn. J. Bot.* 7: 389–452.
- Uzunova, M., W. Ecke, K. Weissleder and G. Röbbelen (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 Orsouw, N.J., R.C.J. Hogers, A. Janssen, F. Yalcin, S. Snoeijers, E. Verstege, H. Schneiders, H. van der Poel, J. van Oeveren, H. Verstegen *et al.* (2007) Complexity reduction of polymorphic sequences (CRoPS™): A novel approach for large-scale polymorphism discovery in complex genomes. *PLoS One* 2: e1172.
- Wang, F., L. Li, H. Li, L. Liu, Y. Zhang, J. Gao and X. Wang (2012) Transcriptome analysis of rosette and folding leaves in Chinese cabbage using high-throughput RNA sequencing. *Genomics* 99: 299–307.
- Wang, J., D.J. Lydiate, I.A.P. Parkin, C. Falentin, R. Delourme, P.W.C. Carionl and G.J. King (2011) Integration of linkage maps for the amphidiploid *Brassica napus* and comparative mapping with *Arabidopsis* and *Brassica rapa*. *BMC Genomics* 12: 101.
- Wang, M., F.L. Zhang, X.D. Meng, X.C. Liu, X.Y. Zhao and Z.C. Fan (2004) A linkage map construction for Chinese cabbage based on AFLP markers using DH population. *Acta Horti. Sinica* 19: 28–33.
- Wang, W., S. Huang, Y. Liu, Z. Fang, L. Yang, W. Hua, S. Yuan, S. Liu, J. Sun, M. Zhuang *et al.* (2012) Construction and analysis of a high-density genetic linkage map in cabbage (*Brassica oleracea* L. var. *capitata*). *BMC Genomics* 13: 523.
- Wang, X., H. Wang, J. Wang, R. Sun, J. Wu, S. Liu, Y. Bai, J.H. Mun, I. Bancroft, F. Cheng *et al.* (2011) The genome of the mesopolyploid crop species *Brassica rapa*. *Nat. Genet.* 43: 1035–1039.
- Wang, Y., L. Xu, Y. Chen, H. Shen, Y. Gong, C. Limeria and L. Liu (2013) Transcriptome profiling of radish (*Raphanus sativus* L.) root and identification of genes involved in response to lead (Pb) stress with next generation sequencing. *PLoS One* 8: e66539.
- Warwick, S.I., A. Francis and I.A. Al-Shehbaz (2006) Brassicaceae: species checklist and database on CD-ROM. *Plant Syst. Evol.* 259: 249–258.
- Wells, R., M. Trick, F. Fraser, E. Soumpourou, L. Clissold, C. Morgan,

- J. Pauquet and I. Bancroft (2013) Sequencing-based variant detection in the polyploid crop oilseed rape. *BMC Plant Biol.* 13: 111.
- Yang, R., D.E. Jarvis, H. Chen, M.A. Beilstein, J. Grimwood, J. Jenkins, S. Shu, S. Prochnik, M. Xin, C. Ma *et al.* (2013) The reference genome of the halophytic plant *Eutrema salsugineum*. *Front. Plant Sci.* 4: 46.
- Yu, J., M. Zhao, X. Wang, C. Tong, S. Huang, S. Tehrim, Y. Liu, W. Hua and S. Liu (2013) Bolbase: a comprehensive genomics database for *Brassica oleracea*. *BMC Genomics* 14: 664.
- Yu, X., S.R. Choi, N. Ramchiary, X. Miao, S.H. Lee, H.J. Sun, S. Kim, C.H. Ahn and Y.P. Lim (2013) Comparative mapping of *Raphanus sativus* genome using *Brassica* markers and quantitative trait loci analysis for the Fusarium wilt resistance trait. *Theor. Appl. Genet.* 126: 2553–2562.
- Zou, X., X. Tan, C. Hu, L. Zeng, G. Lu, G. Fu, Y. Cheng and X. Zhang (2013) The transcriptome of *Brassica napus* L. roots under waterlogging at the seedling stage. *Int. J. Mol. Sci.* 14: 2637–2651.
- Zou, Z., M. Ishida, F. Li, T. Kakizaki, S. Suzuki, H. Kitashiba and T. Nishio (2013) QTL analysis using SNP markers developed by next-generation sequencing for identification of candidate genes controlling 4-methylthio-3-butenyl glucosinolate contents in roots of radish, *Raphanus sativus* L. *PLoS One* 8: e53541.