Database Update

**Brassica database (BRAD) version 2.0: integrating and mining Brassicaceae species genomic resources**

Xiaobo Wang, Jian Wu, Jianli Liang, Feng Cheng* and Xiaowu Wang*

Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences, Beijing 100081, China

*Correspondence may also be addressed to Feng Cheng. Tel: +86 010-82105971; Fax: +86 010-62174123; Email: chengfeng@caas.cn

*Corresponding author: Tel: +86 010-82105971; Fax: +86 010-62174123; Email: wangxiaowu@caas.cn

Citation details: Wang, X., Wu, J., Liang, J. Brassica database (BRAD) version 2.0: integrating and mining Brassicaceae species genomic resources. *Database* (2015) Vol. 2015: article ID bav093; doi:10.1093/database/bav093

Received 4 March 2015; Revised 24 June 2015; Accepted 31 August 2015

**Abstract**

The Brassica database (BRAD) was built initially to assist users apply *Brassica rapa* and *Arabidopsis thaliana* genomic data efficiently to their research. However, many Brassicaceae genomes have been sequenced and released after its construction. These genomes are rich resources for comparative genomics, gene annotation and functional evolutionary studies of *Brassica* crops. Therefore, we have updated BRAD to version 2.0 (V2.0). In BRAD V2.0, 11 more Brassicaceae genomes have been integrated into the database, namely those of *Arabidopsis lyrata*, *Aethionema arabicum*, *Brassica oleracea*, *Brassica napus*, *Camelina sativa*, *Capsella rubella*, *Leavenworthia alabamica*, *Sisymbrium irio* and three extremophiles *Schrenkiella parvula*, *Thellungiella halophila* and *Thellungiella salsuginea*. BRAD V2.0 provides plots of syntenic genomic fragments between pairs of Brassicaceae species, from the level of chromosomes to genomic blocks. The Generic Synteny Browser (GBrowse_syn), a module of the Genome Browser (GBrowse), is used to show syntenic relationships between multiple genomes. Search functions for retrieving syntenic and non-syntenic orthologs, as well as their annotation and sequences are also provided. Furthermore, genome and annotation information have been imported into GBrowse so that all functional elements can be visualized in one frame. We plan to continuously update BRAD by integrating more Brassicaceae genomes into the database.

**Database URL:** http://brassicadb.org/brad/

**Introduction**

Brassicaceae is a large eudicot family that includes the model plant *Arabidopsis thaliana*. The Brassicaceae family has a remarkable diversity of species, genetics and morphotypes, as well as scientific and economic importance. Brassicaceae species have become model systems for
studies of polyploidy and evolution (1). The important
genus *Brassica* of Brassicaceae contains many vegetable,
condiment and oil species that account for about 12% of
the world’s edible vegetable oil production (http://faostat.
fao.org/). U’s triangle theory (2) has been applied to de-
scribe the relationships among six widely cultivated
*Brassica* species, the diploids *Brassica rapa* (AA), *B. nigra*
(BB) and *B. oleracea* (CC) and their allotetraploids *B. jun-
cea* (AABB), *B. napus* (AACC) and *B. carinata* (BBCC). Of
these, the *B. rapa* genome was the first to be sequenced in
2011 (3) and the original *Brassica* database was built based
on it (4).

BRAD version 1.0 (V1.0) provides *B. rapa* genome se-
dquences and gene models, as well as all the syntetic and
non-syntetic homologous gene pairs between *B. rapa* and
*A. thaliana*. On all its pages, BRAD V1.0 incorporates a
useful navigation dialog-window that provides links to
every *B. rapa* and *A. thaliana* gene ID. The small naviga-
tion window directs users by integrating relevant resource
links of the target gene. With the rapid development of
next-generation sequencing technology and the dramatic
decrease in cost, many Brassicaceae species have been
sequenced or were planned to be sequenced after BRAD
V1.0 was constructed. Recently, the genomes of *B. rapa*
sister species, *B. oleracea* and *B. napus*, have been
sequenced (5, 6) and nine other Brassicaceae species have
also been sequenced (7–13). These 13 Brassicaceae genome
datasets are a valuable resource for genome and gene stud-
ies among the closely related Brassicaceae species.

To help researchers and breeders use these recently
released Brassicaceae genome sequences efficiently in scien-
tific investigations and breeding applications, we have
updated BRAD to version 2.0 (V2.0). BRAD V2.0 contains
updated datasets and functions that include all syntetic
gene pairs between *A. thaliana* and the other Brassicaceae
species, more genome and gene sequences and gene anno-
tations, as well as syntetic figures and genome visualization
of all the incorporated Brassicaceae species in the
Genome Browser (GBrowse) (14). BRAD V2.0 provides a
comprehensive framework for comparative genomic anal-
ysis and studies of the evolution of gene function across
Brassicaceae species, especially for the *Brassica* crops.

**BRAD V2.0: feature updates**

**Overview of BRAD V2.0**

In BRAD V1.0, datasets of genome and gene sequences,
gene annotations, non-coding RNAs, transposable elements,
genetic markers and linkage maps of *B. rapa* were provided
(15, 16). A navigation dialog-window for every gene of *B.
rapa* and *A. thaliana* was provided to help users obtain all
related information. Furthermore, BLAST and GBrowse
tools (16) were embedded in BRAD for sequence align-
ment and for visualizing genomic elements, respectively.

BRAD V1.0 has now been updated to V2.0 to include
Brassicaceae genome sequences that have been released re-
cently. In BRAD V2.0, a new section has been incorporated
that shows genomic synteny and micro-fragmental synteny
among any two Brassicaceae species. An alternative pair-
wise synteny plotting tool, the Generic Synteny Browser
(GBrowse_syn) module (17) of GBrowse, has been included
to visualize local synteny relationships among multiple gen-
omes. Moreover, genome and gene sequences, gene annota-
tions and syntenic and non-syntenic orthologs between *A.
thaliana* and other Brassicaceae species have been integrated
into different sections of BRAD V2.0.

**Technical details**

All genomic data were processed using the tool SynOrths
tool (15) to generate genome and gene level synteny data-
sets. Then, syntenic figures were generated based on these
synteny datasets and stored in a MySQL (18) database.

Genome sequences, gene models and the processed
datasets, including all syntetic genes, gene annotation in-
formation and specific gene families were all imported into
MySQL, which enables multifaceted browsing and search-
ing in BRAD. Furthermore, a standalone BLAST (19) ser-
vice implemented in BRAD allows sequence searches
against Brassicaceae genomes, protein-coding gene se-
dquences and protein sequences. The GBrowse package,
which is commonly used to visualize genomic datasets, re-
mains in BRAD V2.0 to view bulk genomic elements of the
Brassicaceae species. Furthermore, the syntenic datasets
are provided not only as tabular results and pairwise-gen-
ome synteny images in the keyword search section, but
also are visualized as a multiple genome synteny compar-
ison in the GBrowse module GBrowse_syn.

**BRAD stocks: Brassicaceae genomes**

Statistics of the Brassicaceae genomic data, including gen-
ome sequences, predicted gene models, protein-coding
gene sequences and protein sequences are shown in
Table 1. In total, about 4 Gb of data have been collected in
BRAD V2.0. In addition to the original genome sequences
and gene models, seven types of annotation for the pre-
dicted genes have been generated. The annotations have
been sourced from the Swiss-Prot, TrEMBL (20), KEGG
(Kyoto Encyclopedia of Genes and Genomes) (21),
InterPro (22) and Gene Ontology (GO) (23) databases and
syntenic genes and BLASTX alignments (best hit, e-value
1E-05) of Brassicaceae genes to the *A. thaliana* genome
also have been included. The numbers of annotation records in these datasets for these species (excluding *A. thaliana*) are shown in Table 2. We used InterProScan (V48.0) ([24](#)), which includes 28,175 GO terms, to generate the InterPro domain and GO annotations. When InterProScan is updated, the GO annotations also will be updated in BRAD.

Updated feature: genome synteny analysis

Genome synteny analysis provides information for studies into the evolution of genome and gene function among species. BRAD V1.0 provided syntenic gene pairs between *B. rapa* and *A. thaliana* so that the gene information of the well-studied model plant *A. thaliana* could be used to annotate *B. rapa* genes. In BRAD V2.0, whole-genome synteny relationships between *A. thaliana* genes and the genes of other Brassicaceae species have been generated and integrated. We obtained syntenic gene pairs that ranged from 17,800 between *A. thaliana* and *Aethionema arabicum* to 59,191 between *A. thaliana* and *Camelina sativa* (Table 3 and Supplementary Tables S1 and S2). The number of tandem gene arrays is shown in Table 4; most had syntenic counterparts in the *A. thaliana* genome. These datasets can be used to investigate genomic rearrangement history, share gene annotation information and investigate functional differentiation of orthologous genes among Brassicaceae species.

Brassica crops experienced a common and relatively recent (9–15 million years ago) whole-genome triplication event after three rounds of polyploidization (γ, β and α whole-genome duplication) in Brassicaceae ([3](#), [5](#), [6](#), [8](#), [25](#)).

### Table 1. Overview of the 13 Brassicaceae genomes in BRAD V2.0

| Species        | Genome size (Mb) | No. of chromosomes | No. of genes | Status          | Source                                      |
|----------------|------------------|---------------------|--------------|-----------------|---------------------------------------------|
| *A. thaliana*   | 120              | 5                   | 27,416       | Chromosome      | TAIR ([https://www.arabidopsis.org/](https://www.arabidopsis.org/)) |
| *A. lyrata*     | 207              | 8                   | 32,670       | Chromosome      | [http://www.phytozome.net/](http://www.phytozome.net/) |
| *A. arabicum*   | 203              | 11                  | 37,839       | Scaffold        | [http://mustang.biol.mcgill.ca:8885/](http://mustang.biol.mcgill.ca:8885/) |
| *B. rapa*       | 284              | 10                  | 41,174       | Chromosome      | BRAD ([http://brassicadb.org/brad/](http://brassicadb.org/brad/)) |
| *B. oleracea*   | 540              | 9                   | 45,758       | Chromosome      | BollBase ([http://ocri-genomics.org/bollbase/index.html](http://ocri-genomics.org/bollbase/index.html)) |
| *B. napus*      | 840              | 19                  | 101,040      | Chromosome      | CoGe ([https://genomemutation.org/CoGe/](https://genomemutation.org/CoGe/)) |
| *C. sativa*     | 641              | 20                  | 94,495       | Chromosome      | [http://www.camelinadb.ca](http://www.camelinadb.ca) |
| *C. rubella*    | 135              | 8                   | 28,447       | Chromosome      | [http://www.phytozome.net/](http://www.phytozome.net/) |
| *L. alabamica*  | 174              | 11                  | 38,676       | Scaffold        | [http://mustang.biol.mcgill.ca:8885/](http://mustang.biol.mcgill.ca:8885/) |
| *S. irio*       | 259              | 7                   | 49,956       | Scaffold        | [http://mustang.biol.mcgill.ca:8885/](http://mustang.biol.mcgill.ca:8885/) |
| *S. parvula*    | 114              | 7                   | 28,901       | Chromosome      | GenBank ([http://www.ncbi.nlm.nih.gov/genbank](http://www.ncbi.nlm.nih.gov/genbank)) |
| *T. halophila*  | 243              | 7                   | 29,284       | Scaffold        | [http://www.phytozome.net/](http://www.phytozome.net/) |
| *T. salsuginea* | 234              | 7                   | 28,457       | Scaffold        | GenBank ([http://www.ncbi.nlm.nih.gov/genbank](http://www.ncbi.nlm.nih.gov/genbank)) |

*aCollaboration with project investigator for genome analysis.*

### Table 2. Numbers of annotation records in BRAD V2.0 from the Gene Ontology (GO), InterPro, KEGG, Swiss-Prot and TrEMBL databases, as well as syntenic genes (Orthologs) and BLASTX alignments (best hits) to the *A. thaliana* genome sequence

| Species        | GO          | InterPro    | KEGG        | Swiss-Prot | TrEMBL    | Orthologs | BLASTX |
|----------------|-------------|-------------|-------------|------------|-----------|-----------|--------|
| *A. arabicum*  | 32,609      | 56,964      | 28,773      | 21,689     | 20,342    | 15,754    | 26,910 |
| *A. lyrata*    | 53,457      | 64,268      | 29,723      | 22,396     | 22,232    | 17,613    | 33,896 |
| *B. rapa*      | 62,875      | 62,852      | 28,501      | 28,501     | 37,220    | 26,194    | 34,648 |
| *B. oleracea*  | 76,109      | 85,261      | 21,071      | 30,504     | 40,498    | 31,794    | 45,603 |
| *B. napus*     | 96,202      | 136,419     | 88,173      | 68,590     | 44,436    | 59,707    | 89,257 |
| *C. sativa*    | 38,109      | 59,533      | 27,927      | 22,123     | 26,195    | 20,952    | 27,973 |
| *C. rubella*   | 112,461     | 168,977     | 90,572      | 71,615     | 35,161    | 64,433    | 91,080 |
| *L. alabamica* | 42,588      | 66,021      | 33,062      | 25,932     | 22,838    | 24,259    | 32,277 |
| *S. irio*      | 38,116      | 67,961      | 35,969      | 26,150     | 24,992    | 18,224    | 33,522 |
| *S. parvula*   | 36,542      | 53,883      | 25,986      | 20,179     | 21,721    | 20,846    | 28,827 |
| *T. halophila* | 39,795      | 61,350      | 28,634      | 22,732     | 23,398    | 19,819    | 28,594 |
| *T. salsuginea*| 33,820      | 52,397      | 25,730      | 20,158     | 23,104    | 19,328    | 25,903 |
| **Total**      | **662,683** | **935,886** | **456,083** | **380,569** | **351,685** | **343,862** | **503,416** |

Database, Vol. 2015, Article ID bav093
They have three subgenomes in their genomes compared with other Brassicaceae species. *B. napus* is the allotetraploid of *B. rapa* and *B. oleracea*, thus its genome is composed of six subgenomes. Additionally, *C. sativa* experienced an independent and more recent whole-genome triplication event than the event in *S. irio*, *T. halophila*, or *T. salsuginea*. For example, by selecting *B. oleracea* as the species, inputting Bra019255 as the gene ID, setting the number of flanking genes to 10 (the default) and clicking the ‘GO’ button, the results are output in a table that appears below the search panel as shown in Figure 1D. The solid circles indicate genes. Information about a gene can be obtained by placing the cursor over a circle. Clicking on the solid circle opens a pop-up dialog-window in which navigation information for the target gene is displayed (Figure 1E). Clicking on a tandem symbol (two small dots following a gene symbol) displays the corresponding tandem gene array information at the bottom of the search page (Figure 1F).

### Table 3. Numbers of syntenic genes in 24 genomic blocks of the three subgenomes in *B. rapa* and *B. oleracea*. The three subgenomes were partitioned as described previously (3), and named as least fractionated subgenome (LF), more fractionated subgenome 1 (MF1) and more fractionated subgenome 2 (MF2)

| Genomic blocks | *B. rapa* |  |  | *B. oleracea* |  |  |
|---------------|----------|---|---|----------------|---|---|
|   | LF | MF1 | MF2 | LF | MF1 | MF2 |
| A  | 1230 | 653 | 785 | 1163 | 645 | 756 |
| B  | 109 | 110 | 109 | 88 | 110 | 109 |
| C  | 112 | 111 | 112 | 102 | 110 | 109 |
| D  | 113 | 112 | 113 | 102 | 110 | 109 |
| E  | 114 | 112 | 114 | 102 | 110 | 109 |
| F  | 115 | 112 | 115 | 102 | 110 | 109 |
| G  | 116 | 112 | 116 | 102 | 110 | 109 |
| H  | 117 | 112 | 117 | 102 | 110 | 109 |
| I  | 118 | 112 | 118 | 102 | 110 | 109 |
| J  | 119 | 112 | 119 | 102 | 110 | 109 |
| K  | 120 | 112 | 120 | 102 | 110 | 109 |
| L  | 121 | 112 | 121 | 102 | 110 | 109 |
| M  | 122 | 112 | 122 | 102 | 110 | 109 |
| N  | 123 | 112 | 123 | 102 | 110 | 109 |
| O  | 124 | 112 | 124 | 102 | 110 | 109 |
| P  | 125 | 112 | 125 | 102 | 110 | 109 |
| Q  | 126 | 112 | 126 | 102 | 110 | 109 |
| R  | 127 | 112 | 127 | 102 | 110 | 109 |
| S  | 128 | 112 | 128 | 102 | 110 | 109 |
| T  | 129 | 112 | 129 | 102 | 110 | 109 |
| U  | 130 | 112 | 130 | 102 | 110 | 109 |
| V  | 131 | 112 | 131 | 102 | 110 | 109 |
| Wa | 132 | 112 | 132 | 102 | 110 | 109 |
| Wb | 133 | 112 | 133 | 102 | 110 | 109 |
| X  | 134 | 112 | 134 | 102 | 110 | 109 |
| Total | 13 893 | 9586 | 8284 | 13 314 | 9166 | 7944 |

### Table 4. Tandem array statistics for each Brassicaceae species in BRAD V2.0

| Species | Tandem (arrays/genes) | Syntenic tandem (arrays/genes) | Ratio (syntenic tandem/tandem) (%) |
|---------|------------------------|-------------------------------|----------------------------------|
| *B. rapa* | 2041|4896 | 1570|3796 | 76.9 |
| *B. oleracea* | 1823|4223 | 1290|2960 | 70.8 |
| *S. parvula* | 1139|2700 | 1022|2545 | 89.7 |
| *A. lyrata* | 1751|4388 | 1441|3743 | 82.3 |
| *L. alabamica* | 789|1769 | 454|1026 | 57.5 |
| *C. rubella* | 1619|4377 | 1397|3691 | 86.3 |
| *S. irio* | 1760|4221 | 1080|2710 | 61.4 |
| *A. arabicum* | 1355|3557 | 880|2377 | 64.9 |
| *T. halophila* | 1414|3642 | 990|2491 | 70.0 |
| *T. salsuginea* | 1401|3378 | 975|2337 | 69.6 |
| *B. napus* | 4406|10 228 | 2317|5355 | 52.6 |
| *C. sativa* | 5713|13 961 | 1121|2722 | 19.6 |
Users can also input their own nucleotide sequences instead of gene IDs using the BLAST services (Blastn, Blastp, Blastx tBlastn and tBlastx) provided under the Tools section in BRAD V2.0. The BLAST search page allows users to search against bulk data from different Brassicaceae sequence databases such as genomes, BACs, protein-coding genes, proteins and ESTs (Figure 1G). Users will obtain related gene IDs based on the BLAST alignments as output (Figure 1H). The obtained gene IDs can be used as input for the search syntenic genes analysis described above (Figure 1B–F). Furthermore, if a user’s nucleotide sequences are not derived from gene regions, the user may still be able to obtain the location of their sequences in the genomes of certain species. This information can be used to retrieve the flanking sequences and elements, which can be visualized or downloaded from GBrowse under the Tools section in BRAD V2.0.

Visualization of synteny analysis

A new ‘Syntenic figure’ function is available under the Search section in BRAD V2.0, which can be used to better illustrate the genomic synteny relationship between two Brassicaceae species. This function can be used to plot genomic synteny relationships as two-dimensional figures. One of the four ancestral species (A. thaliana, A. lyrata, C.
rubella and S. parvula) can be selected for display on the y axis and one of eight other Brassicaceae species can be selected for display on the x axis by clicking the corresponding checkboxes. A total of 28 such figures are available (ignoring self-to-self plots). For example, if ‘Ath’ is chosen for the y axis and ‘Aly’ is chosen for the x axis, then by clicking the ‘View’ button (Figure 2A), users will obtain the image shown in Figure 2B. The lines formed by the red dots show the genomic synteny relationships between the two genome sequences. Clicking on any of the GB regions (shown in color-coded bars), such as GB ‘A’, opens a figure that shows detailed synteny information (Figure 2C). Clicking a dot, which represents a particular gene, on the GB figure will open the GBrowse_syn Web page (17) and show the 100-Kb genomic region flanking the clicked dot.

Syntenic blocks analysis for multiple genome resources

The GBrowse_syn (16) tool for visualizing synteny or collinear genomic regions among multiple genomes can be accessed from the Tools section of BRAD 2.0. GBrowse_syn uses species name and genomic position consisting of the chromosome label, and start and stop positions as input. For example, if ‘c01:601,285..801,285’ is input in the Landmark search box and A. lyrata is selected as the target species from the Genome to Search drop-down list, then the genomic region from 601 285 bp to 801 285 bp on chromosome 1 of A. lyrata will be searched (Figure 3A). By checking the boxes of A. thaliana and S. parvula (Figure 3A) and clicking the ‘Search’ button next to the Landmark search box, a visualization of syntenic blocks for the multiple genomes is obtained (Figure 3B). The sequence of the target species (in this case A. lyrata) is shown in the middle of the graph as the reference genome, and the genomes being compared with the reference are displayed above and below it. Clicking on the track of a compared species changes it into the reference species and all others become the compared genomes. Furthermore, a link to the ‘Syntenic gene’ search section is provided for each gene icon shown on the graph of multiple genome syntenies.

Discussion and conclusions

Many Brassica databases have been built to better understand and use the genomic datasets from Brassica species. These databases include the Brassica Database BRAD (http://brassicadb.org), Brassica.info (http://www.brassica.info/), BrassEnsembl (http://www.brassica.info/BrassEnsembl/index.html), BrassicaDB (http://brassica.nbi.ac.uk/BrassicaDB/), CropStoreDB (http://www.cropstoredb.org/) and BolBase (http://www.ocri-genomics.org/bolbase/index.html). These databases all have different emphasis. Brassica.info integrates information about genomic resources and releases news of projects or activities on Brassica studies. It also provides downloading services for some genome data. BrassEnsembl visualizes different sets of Brassica genomic data under a single frame. CropStoreDB provides a practical approach to managing...
crop genetic data, whereas BolBase (B. oleracea Genome Database) is focused on genomic structure comparisons of the B. oleracea genome. Unlike these other databases, BRAD uses information from genomic studies and gene function studies in the model species A. thaliana to annotate the newly sequenced genomes of Brassica species.

BRAD V2.0 is a substantially improved version of BRAD V1.0. In BRAD V2.9, more Brassicaceae genomes have been integrated, and comprehensive functional annotations of all the Brassicaceae gene models, genome and gene-level syntenic datasets and visualization tools have been provided. In addition, we have included a new application ‘Syntenic figure’ in the search section to allow users to view pairwise syntenic relationships between the Brassicaceae genomes in BRAD V2.0. We used the GBrowse_syn module to visualize multiple genome syntenies. The inclusion of bulk Brassicaceae genome datasets and new applications make BRAD V2.0 a user friendly platform from which to conveniently retrieve genomic information from the genome to gene levels. The updated BRAD V2.0 will be a valuable resource for research into comparative genomics, plant evolution and molecular biology, as well as for breeders of Brassicaceae crops.

Funding
973 program (2012CB113900 and no. 2013CB127000), the 863 program (2012AA100101), the National Natural Science Foundation of China (grant no. 31301771) and the Science and Technology Innovation Program of the Chinese Academy of Agricultural Sciences. Research was carried out in the Key Laboratory of Biology and Genetic Improvement of Horticultural Crops, Ministry of Agriculture, China.

Conflict of interest. None declared.

Supplementary Data
Supplementary data are available at Database Online.

References
1. Kagale,S., Robinson,S.J., Nixon,J. et al. (2014) Polyploid evolution of the Brassicaceae during the Cenozoic era. Plant Cell, 26, 2777–2791.
2. Nagaharu,U. (1935) Genome analysis in Brassica with special reference to the experimental formation of B. napus and peculiar mode of fertilization. Jap J Bot., 7, 389–452.
3. Wang,X., Wang,H., Wang,J. et al. (2011) The genome of the mesopolyploid crop species Brassica rapa. Nat Genet., 43, 1035–1039.
4. Cheng, F., Liu, S., Wu, J. et al. (2011) BRAD, the genetics and genomics database for Brassica plants. BMC Plant Biol., 11, e136.

5. Liu, S., Liu, Y., Yang, X. et al. (2014) The Brassica oleracea genome reveals the asymmetrical evolution of polyploid genomes. Nat Commun., 5, e3930.

6. Chalhoub, B., Denoeud, F., Liu, S. et al. (2014) Early allopolyploid evolution in the post-Neolithic Brassica napus oilseed genome. Science, 345, 950–953.

7. Hu, T.T., Pattyn, P., Bakker, E.G. et al. (2011) The Arabidopsis lyrata genome sequence and the basis of rapid genome size change. Nat Genet., 43, 476–481.

8. Haudry, A., Platts, A.E., Vello, E. et al. (2013) An atlas of over 90,000 conserved noncoding sequences provides insight into crucifer regulatory regions. Nat Genet., 45, 891–898.

9. Slotte, T., Hazzouri, K.M., Ågren, J.A. et al. (2013) The Capsella rubella genome and the genomic consequences of rapid mating system evolution. Nat Genet., 45, 913–918.

10. Dassanayake, M., Oh, D-H., Haas, J.S. et al. (2011) The genome of the extremophile crucifer Thellungiella parvula. Nat Genet., 43, 913–918.

11. Yang, R., Jarvis, D.E., Chen, H. et al. (2013) The reference genome of the halophytic plant Eutrema salsugineum. Front Plant Sci., 4, e46.

12. Wu, H.J., Zhang, Z., Wang, J.Y. et al. (2012) Insights into salt tolerance from the genome of Thellungiella salsuginea. Proc Natl Acad Sci USA, 109, 12219–12224.

13. Kagale, S., Koh, C., Nixon, J. et al. (2014) The emerging biofuel crop Camelina sativa retains a highly undifferentiated hexaploid genome structure. Nat Commun., 5, 3706–3706.

14. Donlin, M.J. (2009) Using the Generic Genome Browser (GBrowse). Curr Protoc Bioinformatics, Chapter 9, Unit 9.9.

15. Cheng, F., Wu, J., Fang, L. et al. (2012) Syntenic gene analysis between Brassica rapa and other Brassicaceae species. Front Plant Sci., 3, e198.

16. Tang, H., Bowers, J.E., Wang, X. et al. (2008) Synteny and collinearity in plant genomes. Science, 320, 486–488.

17. McKay, S.J., Vergara, I.A. and Stajich, J.E. (2010). Using the generic syntenic browser (GBrowse_syn). Curr Protoc Bioinformatics.

18. MySQL, A. (1995) MySQL: the world’s most popular open source database. MySQL AB.

19. Altschul, S.F., Madden, T.L., Schäffer, A.A. et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res., 25, 3389–3402.

20. Boeckmann, B., Bairoch, A., Apweiler, R. et al. (2003) The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. Nucleic Acids Res., 31, 365–370.

21. Kanehisa, M. and Goto, S. (2000) KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res., 28, 27–30.

22. Hunter, S., Apweiler, R., Attwood, T.K. et al. (2009) InterPro: the integrative protein signature database. Nucleic Acids Res., 37, D211–D215.

23. Consortium, G.O. (2004) The Gene Ontology (GO) database and informatics resource. Nucleic Acids Res., 32, D258–D261.

24. Zdobnov, E.M. and Apweiler, R. (2001) InterProScan—an integration platform for the signature-recognition methods in InterPro. Bioinformatics, 17, 847–848.

25. Franzke, A., Lysak, M.A., Al-Shehbaz, I.A. et al. (2011) Cabbage family affairs: the evolutionary history of Brassicaceae. Trends Plant Sci., 16, 108–116.

26. Cheng, F., Mandáková, T., Wu, J. et al. (2013) Deciphering the diploid ancestral genome of the mesohexaploid Brassica rapa. Plant Cell, 25, 1541–1554.

27. Schranz, M.E., Lysak, M.A. and Mitchell, O.T. (2006) The ABC’s of comparative genomics in the Brassicaceae: building blocks of crucifer genomes. Trends Plant Sci., 11, 353–354.

28. Schranz, M.E., Song, B.H., Windsor, A. J. et al. (2007) Comparative genomics in the Brassicaceae: a family-wide perspective. Curr. Opin. Plant Biol., 10, 168–175.