

---

# Genome Triplication Drove the Diversification of *Brassica* Plants

# 10

Feng Cheng, Jian Wu, Jianli Liang and Xiaowu Wang

---

## Abstract

*Brassica* species are significant in diversity. First, it has many close but different species, such as *Brassica rapa*, *Brassica oleracea*, *Brassica nigra*, *Brassica napus*, *Brassica juncea*, etc, many of which are important crops. Second, for each *Brassica* species, it is rich in morphotypes, they have distinctive and impressive traits, such as the heading leaves and enlarged roots in *B. rapa* (Chinese cabbage and turnip) or *B. oleracea* (cabbage and kohlrabi), and the enlarged inflorescences in *B. oleracea*, i.e. broccoli, cauliflower. All these *Brassic*as are evolved from a common hexaploidy ancestor that experienced a whole genome triplication (WGT) event. Studies show that WGT drove the diversification of *Brassica* plants in both the speciation and booming of morphotypes. Following WGT, the extensive block reshuffling and chromosome reduction of the triplicated diploid ancestor through rediploidization process as well as hybridization promoted the *Brassica* speciation. The biased gene retention and subgenome dominance effect further promoted function evolution of multi-copy genes, and finally lead to the expansion of rich morphotypes in *Brassic*as. Conclusively, the WGT event plays important role in driving the diversification of *Brassic*as by initiating the genome and gene-level evolution.

---

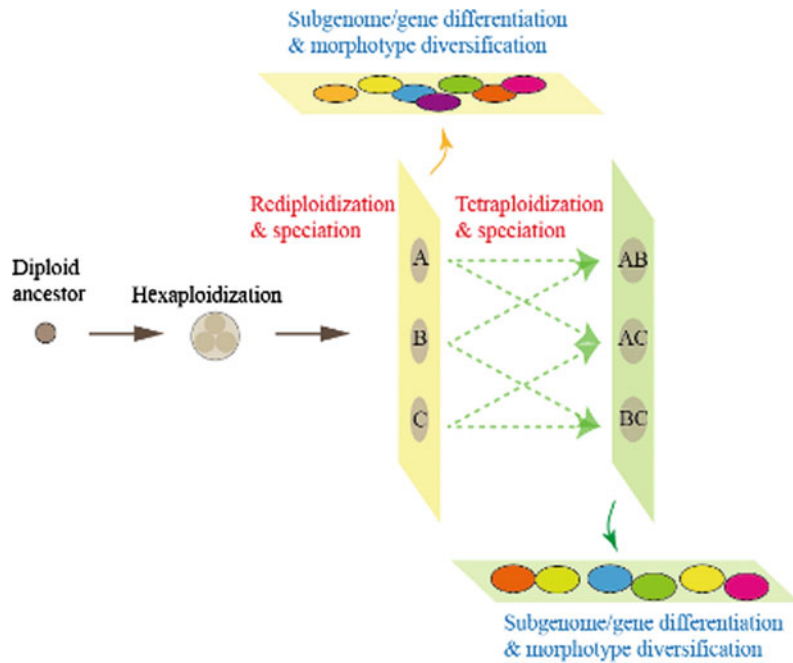
## 10.1 Diversification of *Brassica* Species

The rich diversity in phenotypes or morphotypes is the most distinct feature for *Brassica* plants. There are two levels of diversification for *Brassica* (Fig. 10.1): (1) Many *Brassica* species evolved from a common ancestor, and they are

---

F. Cheng · J. Wu · J. Liang · X. Wang (✉)  
Institute of Vegetables and Flowers, Chinese  
Academy of Agricultural Sciences, Beijing 100081,  
China  
e-mail: wangxiaowu@caas.cn

**Fig. 1** Two levels of diversification for *Brassica* species following the WGT event. Level 1 is the speciation that occurred during the process of rediploidization and tetraploidization after WGT. Level 2 is the morphotype boom after the speciation, which is likely to have been driven by the biased fractionation of subgenomes, as well as the functional differentiation of multicopy genes generated from WGT



close relatives to each other. (2) There are abundant morphotypes for each *Brassica* species. First, the six main *Brassica* crops, referred to above, whose relationships were described by U-triangle (Nagaharu 1935) represent the rich speciation of *Brassica*. Second, each of these *Brassica* species evolved into rich varieties with diversified phenotypes, including leafy heads, enlarged roots, other enlarged organs of stems or inflorescences, oilseeds, sarsons, and even ornamental features. In *Brassica rapa*, heading Chinese cabbage and pak choi are consumed as leafy vegetables. Chinese cabbage is specific for its large leafy head, whereas pak choi has relatively smaller leaves and does not form heading leaves. Turnip has an enlarged root that is edible or occasionally used as fodder. Caixin and purple caitai bolt rapidly and generate long, tender stems used as food. Morphotypes of oilseed *B. rapa* produce large, full seeds for oil extraction, and sarsons grow seed pods that are eaten in India. Some morphotypes of *B. rapa* develop beautiful leaf patterns or colors, and are thus used as ornamental plants. *Brassica oleracea* also has rich morphotypes. Heading *B. oleracea* is also used as a leaf vegetable, whereas oilseed *B.*

*oleracea* produces edible oil. Cauliflower and broccoli, special morphotypes of *B. oleracea*, develop enlarged inflorescences that are eaten as vegetables. Other *Brassica* crops, such as *Brassica juncea*, have even greater numbers of rich morphotypes than *B. rapa* and *B. oleracea*. In addition to these cultivated crops, there are many wild relatives of the species in U-triangle that have greatly diversified phenotypes, further extending the diversity of *Brassica*. Additionally, many morphotypes or phenotypes shared among *Brassica* developed independently and in parallel, such as the heading leaves in *B. rapa* and *B. oleracea*, and enlarged roots in *B. rapa* and *B. juncea*.

The whole-genome triplication (WGT) that occurred in the common ancestor of *Brassica* crops played an important role in the diversification of *Brassica* plants. It was observed that new plant species always evolve after polyploidization. There are many such events in the evolutionary history of the plant kingdom, such as the diversification of early core eudicots (Jiao et al. 2102) (Lysak et al. 2005; Van de Peer et al. 2009). This process of diversification is also related to the events after whole-genome

duplication (WGD), including migrations, the fluctuations in environments, and/or human cultivation/selection (Proost et al. 2011), but polyploidization provides plants with the ability to be diversified and respond to changed habitats. Furthermore, from the aspect of genes, the multicopy genes generated from WGD could develop new functions (gene subfunctionalization or neofunctionalization); thus new traits or morphotypes of plants could evolve. *Brassica* species shared an additional common feature that they all experienced an extra WGT event, which occurred approximately 9–15 million years ago (MYA) (Beilstein et al. 2010; Wang et al. 2011b) or even approximately 28 MYA (Lukens et al. 2004; Lysak et al. 2005; Arias et al. 2014). The WGT event is important for the speciation and the morphotype expansion of genus *Brassica*. The subsequent genomic rearrangements and gene evolution initiated by WGT contributed to the booming of a variety of *Brassica* plants.

## 10.2 Chromosome Evolution after WGT and the *Brassica* Speciation

*Brassica* shared a common WGT event, which has been confirmed by different aspects, such as a comparative genomic analysis (Cheng et al. 2013). Genomic synteny between *B. rapa* and *Arabidopsis thaliana* clearly revealed the WGT event experienced by *B. rapa* (Wang et al. 2011b). Most genes inherited from their nearest common ancestor were shared by *B. rapa* and *A. thaliana* (80.2 and 73.8 % for *B. rapa* or *A. thaliana*, respectively) (Cheng et al. 2012a, b). Although big genomic fragments were fractionated during the rediploidization of *B. rapa*, the local gene order was conserved and syntenic fragments can be identified between *B. rapa* and *A. thaliana*. For each genomic fragment of *A. thaliana*, three syntenic copies were found in *B. rapa*. These three genome copies were generated from the WGT event (Wang et al. 2011b; Cheng

et al. 2012b). Furthermore, a synteny analysis between *B. oleracea* and *A. thaliana* also showed that *B. oleracea* had good genomic collinearity with *A. thaliana*, and *B. oleracea* shared the same WGT event as that of *B. rapa* (Liu et al. 2014). Meanwhile, comparative studies of *B. juncea* and other *Brassica* using information from linkage maps showed that *B. nigra* also shared the same WGT event (Panjabi et al. 2008). Furthermore, previous works, including genomic structure analyses, paleocentromere evolution, and phylogenetic studies, among multiple genomes of Brassicaceae evidenced that the diploid ancestor of *B. rapa* resembled the block arrangement of translocation Proto-Calepineae Karyotype (tPCK), which has seven chromosomes (Cheng et al. 2013). Finally, based on this information, we established that all *Brassica* crops referred to in U-triangle evolved from a common diploid tPCK genome that experienced a WGT event.

Chromosomal reduction and rearrangement, accompanied with paleocentromere descended from the hexaploid ancestor (tPCK  $\times$  3,  $n = 21$ ), were important for the speciation of *Brassica* plants. After WGT, extensive chromosome reshuffling during rediploidization lead to the origin of closely related species in *Brassica*. In polyploids, it is understandable that having more than two copies of homologous chromosomes at the synapsis stage of meiosis will result in abnormal synaptonemal complexes, thereby decreasing the fertility of gametes. Logically, natural selection drives the rediploidization process with chromosomal rearrangement that all eliminate the extra homologous chromosomes. Further rounds of genomic reshuffling in the rediploid ancestor at different evolutionary time points then created different species of *Brassica* (Fig. 10.1). In the genome of *B. rapa*, chromosomes and paleocentromeres were reduced from 21 to 10 through multichromosome translocation, fusion, and inter/intrachromosomal recombination. These genomic reshuffling events should also have occurred in the origin of other *Brassica* species.

### 10.3 Gene Evolution after WGT and the Evolution of Rich Morphotypes in *Brassica*

Biased gene retention after WGT may promote the diversification of *Brassica* plants. Phytohormones, especially auxin, play important roles in the morphogenesis of plants (Santner and Estelle 2009). The genes that are involved in phytohormone signaling are thus important for the formation of diversified morphotypes (Gazzarrini and McCourt 2003; Santner and Estelle 2009). By comparing gene categories between *B. rapa* and other genomes, such as *A. thaliana*, *Carica papaya* or *Vitis vinifera*, it was found that auxin-related genes were expanded in *B. rapa* (Wang et al. 2011b). Furthermore, analysis on gene categories that retained only one or multiple copies found that genes involved in the response to almost all kinds of phytohormone signaling were significantly over-retained in the genomes of *B. rapa* (Wang et al. 2011b), as well as in *B. oleracea* (Liu et al. 2014). These redundant phytohormone related genes should contribute to the morphotype diversification of *Brassica* plants.

Subgenome dominance was observed among the three subgenomes of *B. rapa* (Cheng et al. 2012b; Tang et al. 2012). The subgenome dominance effect drove the differentiation of paralogous genes. The following differences related to subgenome dominance were found. (1) One subgenome retained more genes than the other two. It is clearly observed by counting the number of genes within the three reconstructed tPCK subgenomes, since one subgenome has approximately 1.5-times more genes than the other two subgenomes (Wang et al. 2011b; Cheng et al. 2012b). (2) There are more genes located in the over retained subgenome expressed at higher levels than in their paralogs. Using mRNA-Seq data generated for different organs of *B. rapa*, a comparison of paralogous gene pairs showed that a greater number of genes located in the over retained subgenome are expressed at higher levels than their paralogs in the other two

subgenomes (Cheng et al. 2012b). (3) Genes in the dominant subgenome accumulated fewer functional mutations than those of the other subgenomes (Cheng et al. 2012b). The resequencing of different *B. rapa* morphotypes showed that genes located in the dominant subgenome accumulated fewer functional mutations than those located in the other two subgenomes, which was also considered as the ongoing biased fractionation in *B. rapa* (Cheng et al. 2012b).

WGT provided redundant genes as materials, or a buffer pool, for multicopy genes to evolve new functions, and the subgenome dominance effect facilitated the process by differentiating the multicopy genes that were located in different subgenomes of *B. rapa*. The newly evolved functions may promote the evolution of different morphotypes of *Brassica* species. In *A. thaliana*, many duplicated genes were subfunctionalized or neofunctionalized after several rounds of whole genome polyploidization, known as the  $\alpha$ ,  $\beta$ , and  $\gamma$  duplications. For example, genes from extra duplications have become subfunctionalized compared with those in *C. papaya*, such as the enzymes *CYP79A* and *CYP79B* who catalyze the first step of glucosinolate synthesis (Bekaert et al. 2012). Meanwhile, some genes in *A. thaliana* have become neofunctionalized to develop extra biosynthetic pathways for indole and methionine-derived aliphatic glucosinolates, which do not exist in *C. papaya*. Glucosinolate genes in *B. rapa* showed strong over retention after WGT (Wang et al. 2011a). These redundant genes could be subfunctionalization or neofunctionalization to develop new functions of glucosinolate metabolism in *B. rapa*, as in *A. thaliana*. It is expected that there are many more over-retained genes in *B. rapa*. The subgenome dominance effect may promote the evolutionary process of these over retained genes by conserving one copy of them and differentiating the other copies to develop new functions. Finally, these differentiated genes will contribute to the evolution of rich *B. rapa* varieties and similar processes can also occur in other *Brassica* species.

## 10.4 Conclusions

A WGT event promoted the diversification of *Brassica* from two levels represented by speciation and the expansion of rich morphotypes. First, WGT drove the rediploidization process to stabilize the hexaploid genome. Genomic reshuffling and chromosome reduction gave rise to the speciation of diploid *Brassica* plants as *B. rapa*, *B. nigra*, and *B. oleracea*. The genomic differentiation of the three basic genomes in the U-triangle then generated the stable allotetraploid species such as *B. carinata*, *B. napus*, and *B. juncea*. Second, subgenome differentiation, biased gene retention, and gene subfunctionalization and/or neofunctionalization, after WGT promoted the parallel evolution of many special morphotypes for each *Brassica* species. Therefore, WGT initialized the genome- and gene-level evolution that further drove the *Brassica* speciation and generated a bulk of morphotypes for *Brassica*.

In the near future, further research should be conducted to investigate the evolution of each morphotype of *Brassica*. Previous studies of whole-genome sequences determined the genome- and gene-level evolution of only one accession of *B. rapa*. After that, additional *B. rapa* accessions or other *Brassica* species should be under extensive study to address the following questions: (1) What are the origins of different morphotypes in *Brassicaceae*? (2) What are the mechanisms of parallel evolution of certain morphotypes that developed independently in different *Brassica* species? (3) Which genes were involved in the development of each morphotype and in the regulation of important agronomic traits in *Brassica* crops? Answering these questions will increase our knowledge on the diversification of *Brassica* morphotypes and transfer the benefits of genomic studies to the application of genetic improvement for *Brassica* crops.

## References

- Arias T, Beilstein MA, Tang M, McKain MR, Pires JC (2014) Diversification times among *Brassica* (*Brassicaceae*) crops suggest hybrid formation after 20 million years of divergence. *Am J Bot* 101:86–91
- Beilstein MA, Nagalingum NS, Clements MD, Manchester SR, Mathews S (2010) Dated molecular phylogenies indicate a Miocene origin for *Arabidopsis thaliana*. *Proc Natl Acad Sci USA* 107:18724–18728
- Bekaert M, Edger PP, Hudson CM, Pires JC, Conant GC (2012) Metabolic and evolutionary costs of herbivory defense: systems biology of glucosinolate synthesis. *New Phytol* 196:596–605
- Cheng F, Wu J, Fang L, Wang X (2012a) Syntenic gene analysis between *Brassica rapa* and other *Brassicaceae* species. *Front Plant Sci* 3:198
- Cheng F, Wu J, Fang L, Sun S, Liu B et al. (2012b) Biased gene fractionation and dominant gene expression among the subgenomes of *Brassica rapa*. *PLoS One* 7:e36442
- Cheng F, Mandakova T, Wu J, Xie Q, Lysak MA, Wang X (2013) Deciphering the diploid ancestral genome of the Mesohexaploid *Brassica rapa*. *Plant Cell* 25:1541–1554
- Gazzarrini S, McCourt P (2003) Cross-talk in plant hormone signalling: what *Arabidopsis* mutants are telling us. *Ann Bot* 91:605–612
- Jiao Y, Leebens-Mack J, Ayyampalayam S, Bowers JE, McKain MR et al. (2010) A genome triplication associated with early diversification of the core eudicots. *Genome Biol* 13:R3
- Liu S, Liu Y, Yang X, Tong C, Edwards D et al. (2014) The *Brassica oleracea* genome reveals the asymmetrical evolution of polyploid genomes. *Nat Commun* (in press)
- Lukens LN, Quijada PA, Udall J, Pires JC, Schranz ME et al (2004) Genome redundancy and plasticity within ancient and recent *Brassica* crop species. *Biol J Linn Soc* 82:665–674
- Lysak MA, Koch MA, Pecinka A, Schubert I (2005) Chromosome triplication found across the tribe *Brassicaceae*. *Genome Res* 15:516–525
- 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
- Panjabi P, Jagannath A, Bisht NC, Padmaja KL, Sharma S et al (2008) Comparative mapping of *Brassica juncea* and *Arabidopsis thaliana* using Intron Polymorphism (IP) markers: homoeologous relationships, diversification and evolution of the A, B and C *Brassica* genomes. *BMC Genom* 9:113

- Proost S, Pattyn P, Gerats T, Van de Peer Y (2011) Journey through the past: 150 million years of plant genome evolution. *Plant J* 66:58–65
- Santner A, Estelle M (2009) Recent advances and emerging trends in plant hormone signalling. *Nature* 459:1071–1078
- Tang H, Woodhouse MR, Cheng F, Schnable JC, Pedersen BS et al. (2012) Altered patterns of fractionation and exon deletions in *Brassica rapa* support a two-step model of paleohexaploidy. *Genetics* 190:1563–1574
- Van de Peer Y, Fawcett JA, Proost S, Sterck L, Vandepoele K (2009) The flowering world: a tale of duplications. *Trends Plant Sci* 14:680–688
- Wang H, Wu J, Sun S, Liu B, Cheng F et al. (2011a) Glucosinolate biosynthetic genes in *Brassica rapa*. *Gene* 487:135–142
- Wang X, Wang H, Wang J, Sun R, Wu J et al (2011b) The genome of the mesopolyploid crop species *Brassica rapa*. *Nat Genet* 43:1035–1039