Theor Appl Genet (1990) 79:497–506

THEORETICAL AND APPLIED GENETICS

Brassica taxonomy based on nuclear restriction fragment length polymorphisms (RFLPs)

3. Genome relationships in *Brassica* and related genera and the origin of *B. oleracea* and *B. rapa* (syn. *campestris*)

K. Song [1, 2], T.C. Osborn [2] and P.H. Williams [1]

[1] Department of Plant Pathology, 1630 Linden Drive, University of Wisconsin-Madison, WI 53706, USA
[2] Department of Agronomy, 1575 Linden Drive, University of Wisconsin-Madison, WI 53706, USA

Received July 10, 1989; Accepted December 12, 1989
Communicated by J. Beckmann

Summary. RFLPs were used to study genome evolution and phylogeny in *Brassica* and related genera. Thirty-eight accessions, including 10 accessions of *B. rapa* (syn. *campestris*), 9 cultivated types of *B. oleracea*, 13 nine-chromosome wild brassicas related to *B. oleracea*, and 6 other species in *Brassica* and allied genera, were examined with more then 30 random genomic DNA probes, which identified RFLPs mapping to nine different linkage groups of the *B. rapa* genome. Based on the RFLP data, phylogenetic trees were constructed using the PAUP microcomputer program. Within *B. rapa*, accessions of pak choi, narinosa, and Chinese cabbage from East Asia constituted a group distinct from turnip and wild European populations, consistent with the hypothesis that *B. rapa* had two centers of domestication. A wild *B. rapa* accession from India was positioned in the tree between European types and East Asian types, suggesting an evolutionary pathway from Europe to India, then to South China. Cultivated *B. oleracea* morphotypes showed monophyletic origin with wild *B. oleracea* or *B. alboglabra* as possible ancestors. Various kales constitute a highly diverse group, and represent the primitive morphotypes of cultivated *B. oleracea* from which cabbage, broccoli, cauliflower, etc. probably have evolved. Cauliflower was found to be closely related to broccoli, whereas cabbage was closely related to leafy kales. A great diversity existed among the 13 collections of nine-chromosome wild brassicas related to *B. oleracea*, representing various taxonomic states from subspecies to species. Results from these studies suggested that two basic evolutionary pathways exist for the diploid species examined. One pathway gave rise to *B. fruticulosa*, *B. nigra*, and *Sinapis arvensis*, with *B. adpressa* or a close relative as the initial ancestor. Another pathway gave rise to *B. oleracea* and *B. rapa*, with *Diplotaxis erucoides* or a close relative as the initial ancestor. *Raphanus sativus* and *Eruca sativus* represented intermediate types between the two lineages, and might have been derived from introgression or hybridization between species belonging to different lineages. Molecular evidence for an ascending order of chromosome numbers in the evolution of *Brassica* and allied genera was obtained on the basis of RFLP data and phylogenetic analysis.

Key words: *Brassica* – Molecular taxonomy – Genome evolution – Phylogenetic analysis – Restriction fragment length polymorphism

Introduction

Genome relationships in *Brassica* and allied genera have attracted considerable interest over the past 60 years. Most of the information on these relationships has been based on interspecific hybridizations and cytogenetic studies (Attia and Robbelen 1986; Mizushima 1980; Morinaga 1929; Prakash and Hinata 1980; Quiros et al. 1986, 1988; U 1935; Wills 1966). From analysis of chromosome morphology, Robbelen (1960) reported that there were six basic chromosome types in *Brassica rapa* (syn. *campestris*, n = 10), *B. nigra* (n = 8), and *B. oleracea* (n = 9). Evidence for genome duplication has been obtained from isozyme studies in two n = 7 species, *Brassica adpressa* and *Diplotaxis erucoides* (Quiros et al. 1988). Recent work on construction of genetic maps of *B. oleracea* and *B. rapa* using more than 200 RFLP markers revealed that there were large duplicated chromosome fragments within both genomes (Slocum et al. 1990; K.M. Song, T.C. Osborn, J.Y. Suzuki, M.K. Slocum, P.H. Williams, unpublished data). These and other results support the hypothesis that the genomes of diploid species in *Brassica* and allied genera have evolved in

Table 1. Plant materials screened for RFLPs

Abbreviation	Species and subspecies	Accession name or collection no.	Sources[a]
B. rapa cultivars			
A1	*B. rapa* (syn. *campestris*)	Flowering pak choi	Sakata
A2		Pak choi	CrGC
A3		Narinosa	CrGC
A4		Chinese cabbage	CrGC
A5		Turnip (PTWG)	CrGC
B. rapa wild population			
A6		1742	UPM
A7		166063	NCRPIS (India)
A8		179181	NCRPIS (Turkey)
A9		4685	UPM
A10		5903	UPM
B. oleracea cultivars			
C2	*B. oleracea*	Broccoli (Nine Star Perennial)	WGB
C3		Broccoli (Packman)	Olds
C4		Cabbage (Brunswick)	Olds
C6		Thousand-head kale	WGB
C8		Portugese tree kale	CrGC
C12		Chinese kale	CrGC
C15		Kohlrabi	Olds
C19		Borecole	CrGC
C23		Cauliflower (All Year Round)	WGB
Nine chromosome wild brassicas			
Bia	*B. incana*	6560	UPM
Br	*B. rupestris*	6580	UPM
Bd	*B. drepanensis*	3821	UPM
Bis	*B. insularis*	3814	UPM
Bc1	*B. cretica* ssp. *atlantica*	1952	UPM
Bc2	*B. cretica*	6025	UPM
Bc3	*B. cretica* ssp. *laconica*	6802	UPM
Bc4	*B. cretica* ssp. *nivea*	6020	UPM
Bv	*B. villosa*	6582	UPM
Bma	*B. macrocarpa*	6585	UPM
Bmo	*B. montana*	5975	UPM
Bal	*B. alboglabra*	6820	UPM
Bol	*B. oleracea*	6824	UPM
Other species in *Brassica* and related genera			
Bf	*B. fruticulosa*	–	CrGC
De	*Diplotaxis erucoides*	1235	UPM
Es	*Eruca sativus*	–	CrGC
Rs	*Raphanus sativus*	Chinese radish	Sakata
Bt	*B. tournifortii*	–	CrGC
Bn	*B. nigra*	WPBS	CrGC

[a] WGB – Wellsbourne Gene Bank, Wellsbourne, England; CrGC – Crucifer Genetics Cooperative, University of Wisconsin, Madison/WI, USA; Olds – Olds Seed Co., Madison/WI, USA; UPM – Professor C. Gomez-Campo, Polytechnical University, Madrid, Spain; NCRPIS – North Central Regional Plant Introduction Station, Ames/I, USA; Sakata – Sakata Seed Inc., Japan. Country in parenthesis represents original collection site

ascending chromosome number originating from an n = 6 prototype (Prakash and Hinata 1980).

Based on evidence from RFLP analysis (Song et al. 1988a), we previously hypothesized that three cultivated *Brassica* diploid species evolved in two pathways, with *B. nigra* derived from one pathway and *B. rapa* and *B. oleracea* derived from another. We also proposed that *B. oleracea* and *B. rapa* have a common origin, possibly from a nine-chromosome wild species. Our RFLP data (Song et al. 1988b) provided preliminary information on the phylogeny of various cultivated morphotypes within *B. oleracea* and *B. rapa*. These results suggested that two centers of diversity exist for *B. rapa*: one center is in Europe, represented by turnip, and another is in South

China, represented by pak choi. The data also indicated that cultivated *B. oleracea* could be divided into three subgroups represented by cabbage, broccoli, and kale. However, since these studies were based on a limited number of accessions and DNA probes, some of the genome relationships of species in *Brassica* and related genera, as well as the evolutionary pathways of the diploid species, remain unclear. Also, little is known about the relationships between cultivated types and wild forms and the evolutionary pathways of various morphotypes within *B. oleracea* and *B. rapa*.

This study was conducted using more accessions and DNA probes, in order to provide new insight on the genome evolution of diploid species in *Brassica* and related genera, on the phylogenetic relationships between cultivated *B. oleracea* and wild n = 9 brassicas related to *B. oleracea*, and on the evolutionary pathways within *B. oleracea* and *B. rapa*. We also considered the effects of using data from varying numbers of probes, and one versus two restriction enzymes on the phylogenetic analysis.

Materials and methods

Thirty-eight accessions, including 10 accessions of *B. rapa* (syn. *campestris*), 9 cultivated types of *B. oleracea*, 13 nine-chromosome wild brassicas related to *B. oleracea*, and 6 other species in *Brassica* and related genera, were selected (Table 1). For each accession, a bulk-leaf sample from 24 plants was harvested, and DNA was isolated from the bulk sample.

The methods used for DNA isolation and detection of nuclear RFLPs in selected accessions were described previously (Osborn et al. 1987), except that the final wash of blots was carried out under stringency conditions of 0.25 × SSC at 65 °C. Thirty-three random nuclear DNA sequences and one seed protein gene were used as probes. The random nuclear DNA probes used in this study were selected from a PstI genomic DNA library (Song et al. 1988a). These probes represented different types of DNA sequences, which varied in copy number and degree of polymorphism detected (Table 2), and they hybridized to RFLPs that mapped to nine different linkage groups of *B. rapa* (K.M. Song, T.C. Osborn, J.Y. Suzuki, M.K. Slocum, P.H. Williams, unpublished data). For 29 out of the 34 nuclear DNA probes, plant DNA samples were digested with a single restriction enzyme, either EcoRI or HindIII, and for the remaining five probes, plant DNA samples were digested separately with EcoRI and HindIII (Table 2), in order to determine if additional information could be obtained by using two restriction enzymes. Data were obtained using a total of 39 probe-enzyme combinations.

Eleven cloned chloroplast DNA sequences from *Petunia* and one chloroplast DNA clone from Chinese cabbage were screened for detection of variation in cytoplasms among the selected accessions. With restriction enzymes EcoRI and HindIII, only the probe from Chinese cabbage, p21B5, detected polymorphism between species and, thus, it was used to help distinguish different cytoplasms, as described previously (Song et al. 1988a)

RFLP data was analyzed by both the genetic distance method (FITCH microcomputer programm, developed by J. Felsenstein, Department of Genetics, University of Washington, Seattle/WA) and the Wagner parsimony method (PAUP program, developed by D.L. Swofford, Illinois Natural History Survey, Champaign/IL). For the PAUP analysis, *Diplotaxis erucoides* was used as the outgroup, since this species has the lowest chromosome number (n = 7) among the selected taxa. Particular RFLP patterns and pairwise comparisons between taxa also were used whenever ambiguity arose.

Table 2. Probes and enzymes used for detection of RFLPs

Probe[a]	Probe source	Plant source	Copy no.[b]	Enzyme[c]
pC1	Cruciferin gene	*B. napus*	Middle	H
WR1F02	Nuclear DNA	Chinese cabbage	Single	H
WR1A12	Nuclear DNA	Chinese cabbage	Single	H
WR1EO3	Nuclear DNA	Chinese cabbage	Middle	H
WR1H02	Nuclear DNA	Chinese cabbage	Single	H, R
WR2E07	Nuclear DNA	Chinese cabbage	Low	R
WR2F03	Nuclear DNA	Chinese cabbage	Single	H
WR2A05	Nuclear DNA	Chinese cabbage	Low	H
WR2A01	Nuclear DNA	Chinese cabbage	Single	R
WR2F04	Nuclear DNA	Chinese cabbage	Middle	H
WR2A07	Nuclear DNA	Chinese cabbage	Single	H
WR1G08	Nuclear DNA	Chinese cabbage	Low	H
WR2F07	Nuclear DNA	Chinese cabbage	Single	H
WR1G09	Nuclear DNA	Chinese cabbage	Single	R
WR2E09	Nuclear DNA	Chinese cabbage	Single	H
WR1G04	Nuclear DNA	Chinese cabbage	Single	H
WR1B09	Nuclear DNA	Chinese cabbage	Single	H
WR1D12	Nuclear DNA	Chinese cabbage	Single	H, R
WR2C07	Nuclear DNA	Chinese cabbage	Low	R
WR2G09	Nuclear DNA	Chinese cabbage	Low	H
WR2C01	Nuclear DNA	Chinese cabbage	Low	H
WR1G10	Nuclear DNA	Chinese cabbage	Single	R
WR2B11	Nuclear DNA	Chinese cabbage	Low	R
WR1B03	Nuclear DNA	Chinese cabbage	Single	H
WG5A02	Nuclear DNA	Cabbage	Single	H, R
WG4E07	Nuclear DNA	Cabbage	Low	H
WG2E09	Nuclear DNA	Cabbage	Low	H
WG5B03	Nuclear DNA	Cabbage	Single	R
WG1A01	Nuclear DNA	Cabbage	Single	H, R
WG2D04	Nuclear DNA	Cabbage	Low	H
WG5A05	Nuclear DNA	Cabbage	Low	H
WG4A10	Nuclear DNA	Cabbage	Low	H
WG3F04	Nuclear DNA	Cabbage	Low	H
K4A3	Nuclear DNA	Cabbage	Middle	H, R
p21B5	Chloroplast DNA	Chinese cabbage		R

[a] Request for all of the nuclear DNA probes except K4A3 should be sent to M. K. Slocum NPI, 417 Wakara Way, Salt Lake City/UT 84108, USA

[b] The copy numbers are verified by F_2 segregation pattern and/or single-plant RFLP pattern under our stringency conditions. Single – maximum 2 bands per plant; low – max. 3–5 bands per plant; middle – max. 6–10 bands per plant

[c] H = HindIII, R = EcoRI

Results and discussion

Phylogenetic analysis

Four hundred and thirty-nine restriction fragments were scored across all accessions using 39 probe-enzyme com-

binations. Among the 439 fragments recorded, 6 fragments (1.4%) were common for all taxa, 92 (20.9%) were unique, and 341 (77.7%) were phylogenetically informative in that an individual fragment was shared by at least two accessions but not by all accessions. For the FITCH program, all of the 439 fragments were used in constructing an unrooted network and in assigning the genetic distances. For the PAUP program, only the informative fragments were used in construction of phylogenetic trees. Although both the FITCH and PAUP methods gave similar clustering patterns, the detailed topology between the unrooted network and the phylogenetic trees was different. In this report, only the results from the PAUP analysis are discussed since some problems, such as lack of protection from convergent mutations, are associated with the FITCH analysis using RFLP data (Fink 1986). The PAUP analysis can eliminate most convergent mutations in a data matrix by finding the most parsimonious tree. Also, using the PAUP analysis allowed us to directly compare the present results with results from our previous studies.

In the PAUP analysis, we simply treated each fragment as a unit character and the most parsimonious or the shortest tree was constructed using 'Hold=10, Swap=global, Mulpars, Blrange' functions (Fig. 1). Potential bias may arise by assigning a unit character to different fragments because: (1) we do not know how many mutational differences occur between RFLPs, and (2) some fragments may be phylogenetically more informative than others. However, we believe that these potential problems did not severely bias our conclusion. First, the reliability of our phylogenetic analysis is essentially dependent upon the number and genome coverage of the probes used, the polymorphism they detect, and the representativeness of overall genome changes in the taxa screened. When many probes are used, the effect of an individual probe on the pooled information will be relatively small and the effect of individual fragments will be even less. In this study, we used 34 probes that hybridize to more than 400 fragments, including RFLPs from nine linkage groups in *B. rapa*. Thus, potential bias due to uncertainty in the number of mutational differences for the fragments would be minimized in our analyses using information combined for all probes. Second, we tested the effects of weighting a few fragments detected by one probe that appears to be phylogenetically more informative in distinguishing species. This probe, K4A3, was found previously to hybridize to highly conserved RFLPs in *Brassica* species (Song et al. 1988a). We changed the weighting of these few fragments from one unit character to five unit characters to determine if this would affect the interpretation of phylogenetic relationships. The resulting phylogenetic tree (Fig. 2b) is very similar to the original tree (Fig. 2a) in overall tree topology, though there is a slight difference in cluster pattern

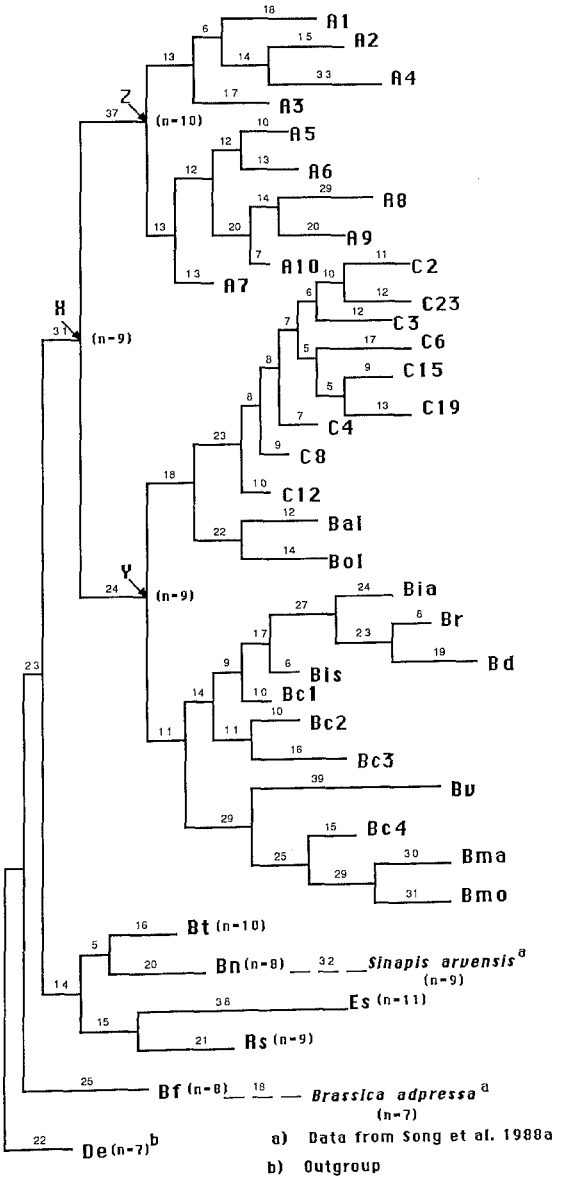
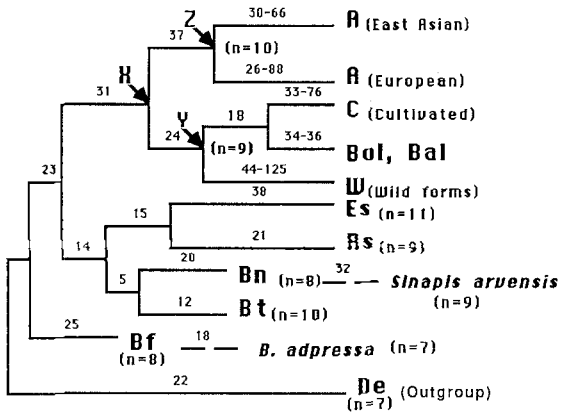
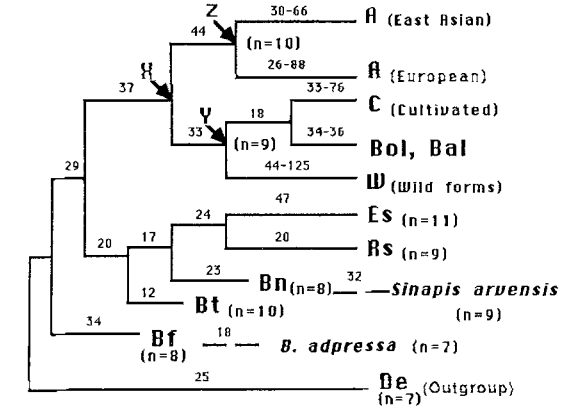


Fig. 1. The shortest phylogenetic tree using 'Hold=10, Swap=global, Mulpars and BLrange' functions of PAUP. No weighting was applied to any fragment. Letters at the end of a branch indicate the accession (Table 1), and numbers on branches indicate the unit distances (PAUP program defines these numbers as minimum number of mutation steps, but since we do not know the exact causes of a single RFLP, the term 'unit distance' was used instead of mutation steps). Point X (*arrow*) indicates the hypothetical common ancestor of *B. rapa* and *B. oleracea*, point Y indicates the hypothetical common ancestor of cultivated and wild relatives of *B. oleracea*, and point Z indicates the hypothetical common ancestor of *B. rapa* (see text)

and unit distances between species. Under certain circumstances, manipulation of weighting may make a phylogenetic tree more meaningful biologically, but it is not necessary in most cases and would not be recommended unless one has extensive information on particular probes and/or fragments.



a (unweighted)



b (weighted)

Fig. 2a and b. Phylogenetic trees showing the interrelationships between species. **a** A simplified tree from Fig. 1 in which 'A (East Asian)' represents A1 to A4 in Fig. 1, and 'A (European)' represents A5 to A10; 'C (cultivated)' represents all of the cultivated *B. oleracea* examined, and 'W (wild forms)' represents 11 wild n = 9 brassicas. **b** A simplified tree constructed by weighting a few fragments detected with probe K4A3 as five-unit characters. This tree separates the species *B. tournifortii* (Bt) and *B. nigra* (Bn) from each other with increased unit distance (see text). Only the distances between species are changed, and there are no changes within species

One important question related to our phylogenetic analysis is whether the same general conclusions would be reached if more probes were used. Although a detailed answer to this question is obviously beyond the scope of this report, we addressed this question in a preliminary fashion by comparing phylogenetic trees based on subsets of our data. Forty data sets, including subsets of 5, 10, 20, and 30 probes, were created by randomly selecting probes from the whole data set, and 40 phylogenetic trees were generated from these data sets using the PAUP program. These trees were then compared to the phylogenetic tree in Fig. 1, which was assigned as the standard tree. We divided the standard tree into groups and subgroups of accessions based on the separation of accessions by the branching patterns of the tree. Phylogenetic trees generated from data sets using subsets of probes were scored for the percentage of groups and subgroups that contained the same accessions as in the standard tree. An average similarity of 32.4% to the standard tree was observed for the 10 phylogenetic trees from data sets of 5 probes, 54.3% for the 10 trees based on 10 probes, 69.5% for the 10 trees based on 20 probes, and 83.5% for the 10 trees based on 30 probes. From analysis of variance, we found that the number of probes in a data set did have significant ($p < 0.001$) influence on the overall tree topology, and that adding more probes increased similarity to the standard tree. However, in comparing the phylogenetic trees based on 10 probes with the trees based on 5 probes, use of 5 additional probes increased the percentage of similarity to the standard tree from 32.4% to 54.3% (4.4% per probe), whereas comparing the phylogenetic trees based on 30 probes with the trees based on 20 probes, an addition of 10 probes only improved the similarity from 69.5% to 83.5% (1.4% per probe). This observation suggests that when a relatively large number of probes (more than 20) is used to construct a phylogenetic tree, each additional probe will provide less information than when a small number of probes (less than ten) is used in the original data set.

We also found that there is no significant difference among the phylogenetic trees from 10, 20, and 30 probes for separation of the large diverse groups represented by *B. rapa*, *B. oleracea* cultivars, nine-chromosome wild brassicas, and the six other species (Fig. 1). For separation of accessions within these groups, a significant difference was observed only within cultivated *B. oleracea*, which contains the least divergent accessions of all four groups based on RFLP patterns. These observations suggest that fewer probes are needed if great genetic diversity exists among the selected taxa.

Another question related to the phylogenetic analysis is whether the use of more probes with one enzyme would provide more information than the use of fewer probes with multiple enzymes. In this study, we compared the phylogenetic trees based on 5 probes with one enzyme to the tree based on the same 5 probes with two enzymes and to the trees based on 10 probes with one enzyme. The two phylogenetic trees based on 5 probes with either EcoRI or HindIII had an average similarity of 22.6% to the standard tree, whereas the tree based on information from the same 5 probes with both enzymes had a similarity of 58.1%, which was close to the average for 10 probes (54.3%). The additional information obtained was likely due to an increase in the phylogenetic value of a given probe by using an additional enzyme. The results of this test suggest that when few probes are used in an analysis, use of an additional enzyme can provide as much new phylogenetic information as use of additional probes. However, we do not know whether or not this observa-

tion will hold true for use of multiple enzymes with the same set of probes or for use of different probe sets.

In summary, there is no doubt that the phylogenetic tree in Fig. 1 would be modified or improved by using more probes and/or enzymes. However, based on the above analyses, our present data allows us to draw general conclusions with reasonable reliability about the phylogeny of the selected accessions. We expect that with addition of more probes and/or enzymes, the basic relationships among large groups and subgroups in Fig. 1 would remain almost the same, and that most of the changes would occur within the cultivated *B. oleracea* accessions.

Phylogeny of B. rapa

The RFLP data and phylogenetic tree derived from the data (Fig. 1) showed that turnip (A5) and most of the wild populations (A6, A8, A9, A10) comprised one group (European group), whereas pak choi (A1, A2), narinosa (A3), and Chinese cabbage (A4) comprised another distinct group (East Asia group). Within the East Asian group, the pak choi accession A2, which is a widely grown type with large edible leaves, was clustered in the phylogenetic tree with a landrace of Chinese cabbage (A4), suggesting that Chinese cabbage originated from a type of pak choi. However, the large unit distance of 48 between A2 and A4 indicated that Chinese cabbage has diverged considerably from pak choi. Within the European group, turnip (A5) is closely related to the wild accession A6, whereas other wild accessions (A8, A9, A10) are more diverse from each other (Fig. 1). A wild accession collected from India, A7, is close to the progenitor of the two groups and has an intermediate geographic location.

Based on our present data and previous studies, the geographic distribution and a possible evolutionary pathway of various morphotypes within *B. rapa* are proposed in Fig. 3. According to the proposal, wild *B. rapa* originated in Europe from the common ancestor of *B. rapa* and *B. oleracea*. Primitive *B. rapa* was then disseminated through the middle East to India and South China, and formed different wild populations. Wild *B. rapa* in Europe was domesticated as turnip and turnip rape. Sarson and toria types of India could have been derived either from a wild population in India or developed from turnip rape through introduction. Pak choi and narinosa probably were domesticated independently from one of the wild populations that reached South China. Chinese cabbage was derived from pak choi, and heading types were developed later by adaptation to cool temperatures through selection (Li 1980). Various Japanese vegetables, such as japonica, could have been derived originally from pak choi, but have diverged considerably through geographic isolation and intensive selection.

Phylogeny of cultivated B. oleracea

In a preliminary survey of 18 cultivated *B. oleracea* accessions, we reported that different morphotypes could be divided into three groups represented by kales, cabbage, and broccoli (Song et al. 1988b). However, since we did not include wild forms as references, some ambiguities remained regarding the origin of cultivated *B. oleracea* and the assignment of individual accessions to various groups. In our present study, 9 cultivated types and 13

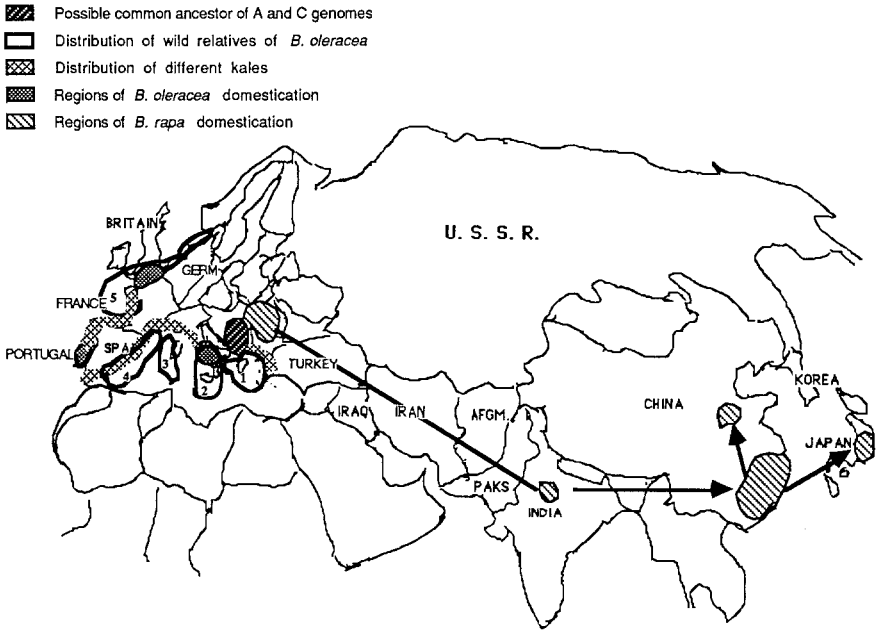


Fig. 3. Geographic distribution, and hypothetical origin and evolutionary pathways of *B. oleracea* and *B. rapa*. *B. oleracea* and *B. rapa* might have derived from a common ancestral species in Europe. *B. rapa* was then disseminated southeast and formed different centers of diversity, whereas *B. oleracea* spread out along the Mediterranean coasts to England and France. The numbers in white boxes indicate natural distribution of wild n=9 brassicas used in this study (according to Snogerup 1980): 1. *B. cretica;* 2. *B. rupestris-incana* complex; 3. *B. insularis;* 4. *B. montana;* and 5. *B. oleracea*

n=9 wild brassicas related to *B. oleracea* were analyzed for more RFLPs, to determine the origin of cultivated *B. oleracea*. All of the 9 cultivated accessions representing highly diverse morphotypes were clustered together and separated from the nine-chromosome wild brassicas (Fig. 1), suggesting a monophyletic origin of cultivated morphotypes of *B. oleracea*. If the cultivated forms of *B. oleracea* had multiple origins as proposed by Mithen et al. (1987), we would expect to find some parallel clusters of particular cultivated types and n=9 wild brassicas, but we did not observe any such clusters. Eleven of 13 wild accessions comprised a distinct group for which there was no overlap between the cultivated and wild accessions examined. The accessions of wild *B. oleracea* (Bol) and *B. alboglabra* (Bal) were positioned between cultivated types and other n=9 wild brassicas (Fig. 1), indicating that wild *B. oleracea* and *B. alboglabra* are the closest ancestors of cultivated *B. oleracea*.

Among the cultivated types, Chinese kale (C12) and Portuguese tree kale (C8) seemed to be the accessions most closely related to wild *B. oleracea* and *B. alboglabra* (Fig. 1). Cabbage (C4) was found to be closely related to Portuguese kale (C8) and Chinese kale (C12), whereas broccoli (C2, C3) and cauliflower (C23) were clustered together. These observations clarified the ambiguity in our previous study regarding the origin of cauliflower (Song et al. 1988b). Kohlrabi (C15) and borecole (C19) were clustered together but with a large unit distance of 35. Thousand-head kale (C6) seems to be a distinct morphotype.

These results, combined with our previous survey on 18 cultivated types (Song et al. 1988b) and other studies, have led us to hypothesize the evolutionary pathway in Fig. 3. We propose that primitive *B. oleracea* originated from the same ancestor as *B. rapa*, and then different wild forms developed from this primitive type. Cultivated morphotypes of *B. oleracea* have originated from a single ancient progenitor that was similar to wild *B. oleracea* and *B. alboglabra*. The earliest cultivated *B. oleracea* was likely a leafy kale from which a variety of kales became widely spread along the coasts of the Mediterranean and North Atlantic from Greece to Wales (Fig. 3). Specialized forms of *B. oleracea* have evolved in different areas through selection and adaptation to various climates. Ancient broccoli might be derived from one type of kale grown in Italy that had specialized inflorescence, and cauliflower was later developed from broccoli. Similarly, ancient cabbage probably evolved from a leafy kale, and the Portuguese cabbage may be an intermediate type from which common cabbage was developed.

Phylogeny of n=9 wild brassicas related to B. oleracea

The wild brassicas from the Mediterranean area with chromosome numbers of n=9 have been proposed as the ancestors of cultivated *B. oleracea* (Snogerup 1980). However, little is known about the genetic relationships among these wild brassicas and their relationships to cultivated *B. oleracea*. Based on morphological evidence and geographic distribution, Snogerup (1980) divided these n=9 wild brassicas into several groups represented by the *B. rupestris-B. incana* complex, *B. cretica*, *B. insularis*, *B. macrocarpa*, *B. montana*, and *B. oleracea*. In this study, 13 wild accessions representing the above groups were analyzed using RFLP data. All wild accessions examined were found to be closely related to each other and to the cultivated forms (Fig. 1). In Fig. 1, the 13 nine-chromosome wild brassicas are divided into four groups. *B. rupestris* (Br), *B. drepanensis* (Bd) and *B. incana* (Bia) constitute one group, corresponding to the *B. rupestris-incanca* complex described by Snogerup (1980). However, Br and Bd are much closer to each other than either is to Bia. Also, probe K4A3, which hybridized to conserved fragments (Song et al. 1988a), detected a different RFLP pattern in Br and Bd compared to Bia (Fig. 4). Three accessions of *B. cretica* (Bc1, Bc2 and Bc3) are closely related and form a second distinct group. *B. insularis* (Bis) seems to be related to this group but showed a different RFLP pattern with probe K4A3 (Fig. 4). *B. villosa* (Bv), *B. cretica* ssp. *nivea* (Bc4), *B. macrocarpa* (Bma), and *B. montana* (Bmo) constitute a loosely related third group, among which Bma and Bmo are most closely related to each other. The large unit distances between the four accessions and differences in particular RFLP patterns (Fig. 4) indicate that these accessions have diverged considerably. *B. oleracea* (Bol) and *B. alboglabra* (Bal) are clustered together forming a distinct group that was different from the other three groups. These two

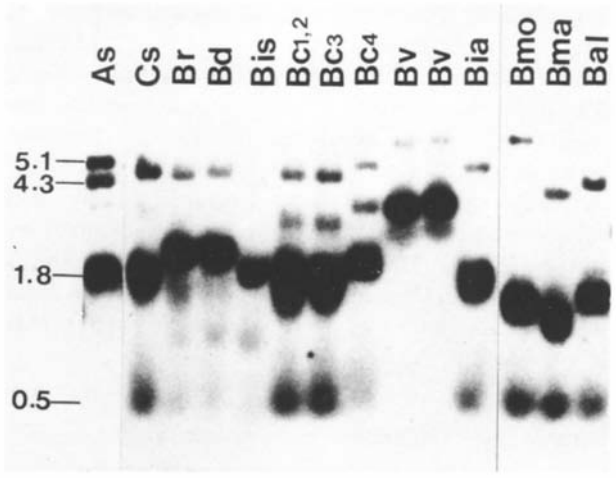


Fig. 4. Autoradiographs of *Eco*RI-digested DNAs probed with K4A3, showing the differences in RFLP patterns within cultivated types and the 13 wild n=9 brassicas. 'As' represents all of the accessions of *B. rapa* examined, and 'Cs' represents all of the cultivated types of *B. oleracea*. Other designations are the same as in Table 1 (see text for details)

accessions are located between cultivated *B. oleracea* and other wild n=9 brassicas, and have the same RFLP pattern as that of cultivated types using probe K4A3 (Fig. 4). Based on the cluster patterns, unit distances, and RFLP patterns, we suggest that all of the wild accessions examined may belong to a single species. However, tremendous differentiation has occurred among these wild brassicas, so that some of them are becoming new species. Among these wild brassicas, *B. oleracea* and *B. alboglabra* seem to be the closest to cultivated forms and, thus, they likely are most closely related to the ancestors of cultivated *B. oleracea*

Genome evolution and interrelationships in Brassica and related genera

Based on cytogenetic evidence, researchers have previously proposed that the progenitor of *Brassica* contained a chromosome number of n=6, and that species with higher chromosome numbers have evolved as aneuploids in an ascending order from the n=6 progenitor (Attia and Robbelen 1986; Prakash and Hinata 1980; Robbelen 1960). In this study, species with chromosome numbers ranging from n=7 to n=11 were examined. Based on the weighted tree (Fig. 2b), which appears to represent the relationships between some of the species examined better than the unweighted tree (Fig. 2a), *Diplotaxis erucoides* (De, n=7) has unit distances of 59 to *B. fruticulosa* (Bf, n=8), 114 to *B. nigra* (Bn, n=8), 135 to *Raphanus sativus* (Rs, n=9), 165 to *Eruca sativus* (Es, n=11), 127 to the hypothetical ancestor of *B. oleracea* (point Y in the tree, n=9), and 134 to the hypothetical ancestor of *B. rapa* (point Z in the tree, n=10). These numbers reflect a remarkable correlation between increases in chromosome number and changes in restriction fragments. Basically, the n=7 chromosome species De is more closely related to the n=8 chromosome species Bf and Bn than to the n=9 chromosome species, etc. *B. tournifortii* (Bt, n=10) presents an exception that we will discuss later. Similar results were observed in our previous study, in which *B. adpressa* (n=7) showed unit distances of 18 to *B. fruticulosa*, 47 to *B. nigra*, 74 to *B. oleracea*, and 100 to *B. rapa* (Song et al. 1988a, Fig. 2). Assuming *D. erucoides* and *B. adpressa*, with n=7, are the closest to the n=6 progenitor species, an ascending order of evolution in chromosome number seems likely for *Brassica* and related genera.

The overall phylogeny of different species in the weighted tree (Fig. 2b) revealed that *B. oleracea* and *B. rapa* are more closely related to each other with unit distance of 77 (point Y to point Z) than is either to the other six species, with unit distances ranging from 113 to 189 for *B. rapa* and from 102 to 178 for *B. oleracea*. *R. sativus* (Rs) and *E. sativus* (Es) were clustered together, though there was a large distance of 67 between the two

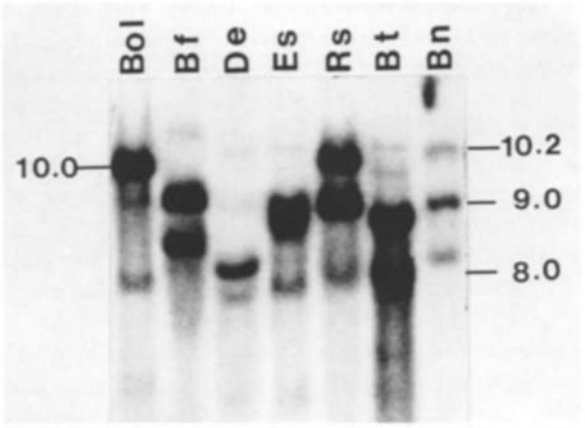


Fig. 5. Autoradiographs of *Hin*dIII-digested DNAs probed with K4A3, showing the variation among species examined. The designations of species are the same as in Table 1. *R. sativus* (Rs) shares a 10.0 kb fragment with *B. oleracea* (Bol) and a 9.0 kb fragment with *B. nigra* (Bn), suggesting the hybdrid origin of *R. sativus* (see text)

species. It is interesting to recognize that Rs and Es have a number of morphological similarities, such as flower characteristics. Rs and Es also were found to have the same RFLP pattern when probed with the chloroplast DNA probe p21B5. Thus, Rs and Es might be derived from a common ancestor. Rs also showed a close relationship to *B. nigra* (Bn), with a unit distance of 61. The conservative probe K4A3 revealed that Rs shared a 9.0-kb HindIII fragment with Bn and a 10.0-kb HindIII fragment with Bol (Fig. 5). According to a previous study of cpDNA of *Brassica* (Palmer et al. 1983), *R. sativus* had an intermediate type of cytoplasm between *B. nigra* and *B. oleracea*. The intermediate state of *R. sativus* in both cytoplasm and nuclear DNA suggests a hybrid origin for *R. sativus*

B. tournifortii (Bt) was an interesting species in our phylogenetic analysis. Although there is a high degree of similarity in nuclear DNA fragments between Bt and Bn (*B. nigra*), with a unit distance of 52 (Fig. 2b), Bt seems unlikely to have come directly from Bn. First, *B. tournifortii* has two more chromosomes than *B. nigra* and contains a different chloroplast pattern detected by the probe p21B5. Second, if Bt came from Bn, the distance between De and Bt should be larger than that between De and Bn. However, the opposite situation was observed. Thus, it is possible that *B. tournifortii* has evolved from a different lineage with recent introgression from *B. nigra*. We also noticed that in both unweighted and weighted trees (Fig. 2a and b), Bt has the shortest distance to the hypothetical common ancestor of *B. oleracea* and *B. rapa* (point X), with unit distance of 69, compared to other species having unit distances ranging from 91 to 145 (Fig. 2b). On the other hand, Bt has a shorter distance of 86 to De than either *B. rapa* (point Z) to De (unit dis-

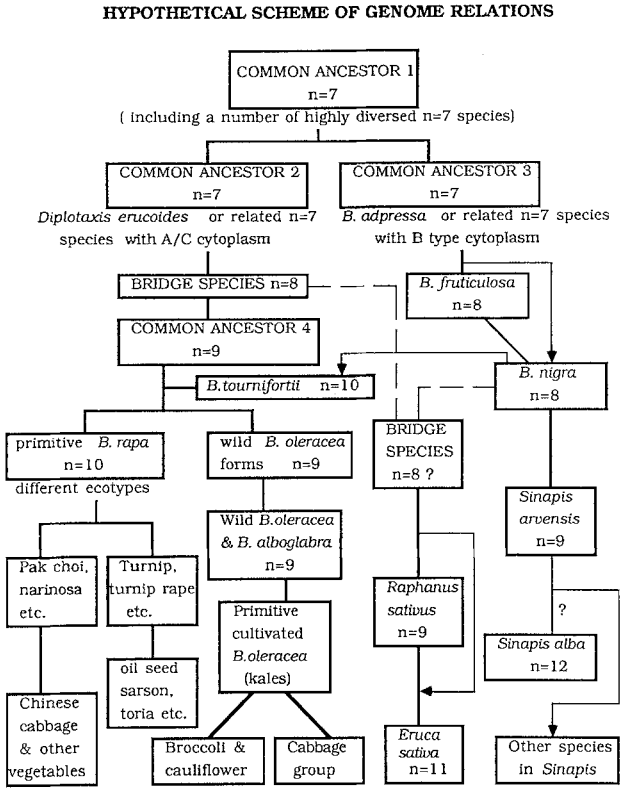


Fig. 6. A hypothetical scheme of genomic relations of *Brassica* and related genera based mainly on analysis of accessions used in this study. The *solid thick lines* indicate the main directions of genome evolution. The *solid thin lines* with *arrows* indicate the possible alternative pathways or introgression. *Dash lines* indicate possible hybridization (see text for details). A/C and B cytoplasms were determined by the same criteria as in our previous report (Song et al. 1988a)

tance = 135) or *B. oleracea* (point Y) to De (unit distance = 127). Considering the chromosome number, the cytoplasmic pattern, and the particular position in the phylogenetic tree, *B. tournifortii* seems to be closely related to the common ancestor of *B. oleracea* and *B. rapa*, which was probably a nine-chromosome species (Song et al. 1988a).

Figure 6 depicts a hypothetical scheme that summarizes the evolution of diploid species in *Brassica* and related genera based on evidence from RFLP studies and previous cytogenetic studies. Basically, the scheme presents the evolutionary pathways of the species examined in an ascending order of chromosome number. According to this hypothesis, species with chromosome numbers of n = 7 compose the most ancient groups, which probably were derived from the prototype species of n = 6. Different n = 7 species evolved into different groups of species with ascending chromosome numbers. The accessions examined seem to have been derived from two evolutionary pathways: *B. adpressa*, *B. fruticulosa*, *B. nigra*, *Sinapis arvensis*, etc. compose one group, having ascending chromosome numbers with *B. adpressa* or a close relative as the primary ancestor, whereas *D. erucoides*, *B. oleracea*, and *B. rapa* compose another ascending group with *D. erucoides* or a close relative as the primary progenitor. In our previous study (Song et al. 1988a) we arbitrarily placed *S. arvensis* before *B. nigra*. However, with new information, *S. arvensis* appears more likely to have evolved after *B. nigra*.

Some evidence from other sources support our hypothesis in Fig. 6. Based on analysis of a repeat sequence (satellite DNA), Lakshmikumaran et al. (1988) suggested that *D. erucoides* was more closely related to *B. oleracea* and *B. rapa* than was *B. nigra* to *B. oleracea* and *B. rapa*. Delourme et al. (1989) and Jahier et al. (1989) reported the results from development of *B. napus-D. erucoides* addition lines and *B. napus-nigra* addition lines, respectively. They found that the F_1 of *B. napus-D. erucoides* (ACDe) often showed five to six univalents during meiosis, whereas the F_1 of *B. napus-B. nigra* (ACB) often showed eight univalents. Also, the backcross progeny of *B. napus-D. erucoides* (AACC + De) gave a higher average number of multivalents than the backcross progeny of *B. napus-B. nigra* (AACC + B). These data indicated that the *D. erucoides* genome (De genome) has higher homology to the *B. napus* genome (A and C genomes) than the *B. nigra* genome (B genome) has to the *B. napus* genome. Furthermore, A.M. Chevre (personal communication) found that it was very difficult to make crosses between *B. napus* and *S. alba*, and the somatic hybrids between *B. napus* and *S. alba* always showed 12 univalents, which corresponds to the *S. alba* chromosome number. This result suggests that the *S. alba* genome is more distant to *B. napus* than is the *B. nigra* genome to *B. napus*. The differences in genome composition between *B. napus* and other species described above can be explained by their positions and relative distances in our hypothetical scheme. For example, *B. napus* and *S. alba* should be located at the bottoms of two lineages and, thus, have a very long genetic distance to each other (Fig. 6).

Hybridization or introgression between species belonging to different lineages may occur occasionally, giving rise to new species. For example, *R. sativus*, as mentioned before, might have derived from a bridge species that was probably a hybrid between *B. nigra* and another n = 8 species in the *B. oleracea* lineage (Fig. 6). *B. tournifortii* might come from the same lineage as that of *B. oleracea* and *B. rapa*, but the high degree of similarity to *B. nigra* suggests that introgression or hybridization might have occurred between these two species. Although the above results indicate some of the possible evolutionary lineages in *Brassica* and related genera, further analysis of both chloroplast and nuclear genomes is needed on more species representing different chromosome groups.

Acknowledgements. The authors are grateful to Prof. C. Gomez-Campo for providing most of the plant materials, to Dr. J. D. Palmer for providing *Petunia* chloroplast DNA clones, to Dr. Martha Crouch for providing the cruciferin clone, and to Dr. K. J. Sytsma and an anonymous reviewer for valuable comments on the manuscript. Funding was provided by the College of Agricultural and Life Sciences, University of Wisconsin and by the National Kraut Packers Association.

References

Attia T, Robbelen G (1986) Cytogenetic relationship within cultivated *Brassica* analyzed in amphidiploids from the three diploid ancestors. Can J Genet Cytol 28:323–329

Delourme R, Eber F, Chevre AM (1990) Intergeneric hybridization of *Diplotaxis erucoides* with *Brassica napus*. Euphytica (in press)

Fink WL (1986) Microcomputers and phylogenetic analysis. Science 234:1135–1139

Jahier J, Chevre AM, Tanguy AM, Eber F (1989) Extraction of disomic addition lines of *Brassica napus-B. nigra*. Genome 32:408–413

Lakshmikumaran M, Ranade S, Gupta V, Agnihotri A, Dhillon HS, Jagannathan V (1988) Sequence analysis and evolution of a tandemly repeated DNA of *Brassica campestris*. 2nd Int Cong Plant Mol Biol, Jerusalem, Nov. 13–18, Abstr 599

Li CW (1980) The origin, evolution, taxonomy, and hybridization of Chinese cabbage. In: Talekar NS, Griggs TD (eds) Chinese cabbage. Asian Vegetable Research Center, Taiwan, pp 1–10

Mithen RF, Lewis BG, Heaney RK, Fenwick GR (1987) Glucosinolates of wild and cultivated *Brassica* species. Phytochemistry 26:1969–1973

Mizushima U (1980) Genome analysis in *Brassica* and allied genera. In: Tsunoda S, Hinata K, Gomez-Campo C (eds) *Brassica* crops and wild allies. Japanese Scientific Society Press, Tokyo. pp 89–105

Morinaga T (1929) Interspecific hybridization in *Brassica*. 1. The cytology of F_1 hybrids of *B. napella* and various other species with ten chromosomes. Cytologia 1:16–27

Osborn TC, Alexander DC, Fobes JF (1987) Identification of restriction fragment length polymorphism linked to genes controlling soluble solids content in tomato fruit. Theor Appl Genet 73:350–356

Palmer JD, Shield CR, Cohen DB, Orton TJ (1983) Chloroplast DNA evolution and the origin of amphidiploid *Brassica* species. Theor Appl Genet 65:181–189

Prakash S, Hinata K (1980) Taxonomy, cytogenetics, and origin of crop *Brassica*, a review. Opera Bot 55:1–57

Quiros CF, Ochoa O, Kianian SF, Douches D (1986) Evolutionary trends in *Brassica*: Gathering evidence from chromosome addition lines. Cruciferae Newslett 11:22–23

Quiros CF, Ochoa O, Douches D (1988) Exploring the role of x=7 species in *Brassica* evolution: Hybridization with *B. nigra* and *B. oleracea*. J Hered 79:351–358

Robbelen G (1960) Beiträge zur Analyse des *Brassica*-Genomes. Chromosoma II:205–228

Slocum MK, Figdore SS, Kennard WC, Suzuki JY, Osborn TC (1990) The linkage arrangement of restriction fragment length polymorphism loci in *Brassica oleracea*. Theor Appl Genet (in press)

Snogerup S (1980) The wild forms of the *Brassica oleracea* group (2n=18) and the possible relations to the cultivated ones. In: Tsunoda S, Hinata K, Gomez-Campo C (eds) *Brassica* crops and wild allies. Japanese Scientific Society Press, Tokyo, pp 121–132

Song KM, Osborn TC, Williams PH (1988a) *Brassica* taxonomy based on nuclear restriction fragment length polymorphisms (RFLPs). 1. Genome evolution of diploid and amphidiploid species. Theor Appl Genet 75:784–794

Song KM, Osborn TC, Williams PH (1988b) *Brassica* taxonomy based on nuclear restriction fragment length polymorphisms (RFLPs). 2. Preliminary analysis of subspecies within *B. rapa* (syn. *campestris*) and *B. oleracea*. Theor Appl Genet 76:593–600

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

Wills AB (1966) Meiotic behavior in the Brassiceae. Caryologia 19:103–116