Plant Physiol. Biochem. 39 (2001) 253–262

S0981942801012396/REV

Review

Comparative structural genomics in the Brassicaceae family

Renate Schmidt*, Adile Acarkan, Karine Boivin

Max-Delbrück-Laboratorium in der Max-Planck-Gesellschaft, Carl-von-Linné-Weg 10, 50829 Cologne, Germany

Received 24 November 2000; accepted 20 December 2000

Abstract – A particular small size has made the *Arabidopsis thaliana* genome the best-studied genome of a higher plant. Comparative genetic mapping experiments have established colinearity of genomes for species of the Brassicaceae family. More recently physical mapping and sequencing experiments established conservation of gene order. Patterns of genome colinearity that have been discovered to date will be described. One of the prime goals of comparative genome studies is the exploitation of information that has been assembled for model species with small genomes for the analysis of related plants with complex genomes. Specifically, the prospects for correlating traits in the *Brassica* crops with *Arabidopsis* candidate genes will be discussed. © 2001 Éditions scientifiques et médicales Elsevier SAS

***Arabidopsis thaliana* / *Brassica* / Brassicaceae / colinearity / genetic map / genome / physical map**

BAC, bacterial artificial chromosome / EST, expressed sequence tag / PCR, polymerase chain reaction / QTL, quantitative trait loci / RFLP, restriction fragment length polymorphism

1. INTRODUCTION

Genetic linkage maps based on molecular markers are important tools to study the organisation of plant genomes and maps have been assembled for a variety of different plant species [38]. Many maps were based on RFLP markers, but PCR-based marker systems are also widely adopted. RFLP markers often correspond to gene sequences. Coding sequences are conserved in evolution, thus RFLP markers cannot only be applied in the species they are derived from but also cross-hybridised to DNA of related species. This can be exploited to compare linkage arrangements of markers in close relatives. If the same set of RFLP markers is used for genetic mapping in related species, molecular marker maps of both species can be constructed and the marker order can be directly compared. Such experiments were conducted for species belonging to the same family and revealed conservation of marker order between chromosome segments or even entire chromosomes. Only five chromosomal inversion events had to be assumed to explain differences in marker order on the twelve tomato and potato chromosomes [50]. The Poaceae family was extensively studied in respect to genome colinearity. These experiments established a remarkable degree of genome conservation. Even if species were compared which considerably differ in genome size or diverged as long as 60 million years ago, colinearity of chromosome segments was found (reviewed in [15]).

In comparative genetic mapping studies, only a very small subset of genes from a given genome is analysed. In many cases, hundreds of genes may lie between any pair of adjacent markers chosen for a particular study. To address the question as to what extent local gene order, orientation and spacing are conserved between species, large orthologous chromosomal regions that have been characterised in detail in respect to gene content and order from different species need to be compared. This can be achieved by carrying out comparative physical mapping and sequencing experiments. The first studies analysing microcolinearity of plant genomes gave hints that microstructure is not as conserved as the gross chromosomal organisation. Many small differences in respect to gene repertoire and organisation of repetitive elements were found in orthologous regions of rice, maize and sorghum [5].

Comparative genome analysis reveals valuable data about plant genome evolution. Moreover, by describ-

**Correspondence and reprints:* fax +49 221 5062 613.
E-mail address: rschmidt@mpiz-koeln.mpg.de (R. Schmidt).

ing the relationships of small well-studied genomes of model species with the often much larger genomes of related plants, it will emerge whether and to what extent data derived from genomes of model organisms can be transferred to other species.

2. THE BRASSICACEAE FAMILY

The Brassicaceae family is widely distributed and consists of approximately 340 genera and 3 350 species [40]. Due to the well-characterised genome and amenability to genetic analysis, one of its members, *Arabidopsis thaliana*, has become a very important model organism for plant molecular biology. All aspects of plant biochemistry, physiology and development are being studied by using molecular-genetic techniques [32]. Most importantly, the *Arabidopsis* genome initiative aims to analyse the complete sequence of the genome [6], which is one of the smallest known for higher plants (130 Mbp).

Numbers of haploid chromosome sets vary in Brassicaceae species. *A. thaliana* has n = 5 chromosomes; in contrast, many of its close relatives are characterised by n = 8 chromosomes. It has been proposed that base chromosome numbers lower than n = 8 are derived in tribe Arabideae, since base chromosome number reduction from eight to between seven and five occurred several times [19]. Thus, comparative genetic mapping experiments in species with different base chromosome numbers may reveal aspects of chromosome evolution.

Agricultural importance of the Brassicaceae family is mainly due to the genus *Brassica*, which is a rich source for oil seeds as well as many different vegetable and fodder crops. Six different *Brassica* species are widely cultivated. Three of these are considered as diploids, whereas the remaining ones are amphidiploid derivatives of the diploid species. Therefore, comparative experiments carried out in *Brassica* species do allow the assessment of the influence of recent polyploidy on genome evolution.

Divergence times between the *Arabidopsis* and *Brassica* lineages were estimated at 23.1–25.9 and 14.5–20.4 million years ago, respectively [20, 56]. The separation of Brassicaceae and Asteraceae occurred approximately 112–156 million years ago and the monocot-dicot split was calculated at 170–235 million years ago [56]. Thus, most crop plants other than the *Brassica* species are only distantly related to *A. thaliana*. Comparative studies in the Brassicaceae family are therefore of particular interest. It can be tested whether and to what extent information and resources derived from the *Arabidopsis* genome project can be exploited for the analysis and improvement of traits in the most closely related crop plants, the *Brassica* species.

3. THE *ARABIDOPSIS* GENOME

A small size of 130 Mbp facilitated the assembly of comprehensive genetic (http://www.arabidopsis.org) and physical maps [30, 33] for the *A. thaliana* genome. Most importantly, the nucleotide sequence has been determined and annotated for the majority of the genome (http://www.arabidopsis.org; [29, 31]). The sequence maps are extensively cross-referenced with genetic and physical chromosome maps (http://www.arabidopsis.org).

Annotation of genes relies on predictions with appropriate computer algorithms, but comparisons with coding sequences from *Arabidopsis* and other organisms provide important additional information. EST and cDNA sequences are especially valuable for gene annotation. It is therefore of importance that currently more than 110 000 EST sequences are available [17, 34, 41].

The *Arabidopsis* genome is characterised by a high density of genes. On chromosome 2, for example, a gene occurs on average every 4.4 kbp [29]. In the heterochromatic region, the density of genes is much lower, whereas the frequency of repeated sequences is increased [29, 31]. A large proportion of genes is present in more than one copy. It has been estimated that up to 60 % of the genome is present in large segmental duplications [7]. Furthermore, clusters of related gene sequences are frequently observed [29, 31].

4. THE *BRASSICA* GENOMES

Extensive genetic and molecular analyses have been undertaken for the six cultivated *Brassica* species. Their relationship could be clarified by cytological analyses [54]. The three species, which are considered as diploids differ in chromosome number. *B. nigra*, *B. oleracea* and *B. rapa* (syn. *campestris*) have 2n = 16, 18 and 20 chromosomes, respectively. Pair-wise combinations of the diploid species *B. rapa* (AA), *B. nigra* (BB) and *B. oleracea* (CC) have resulted in the amphidiploid species *B. juncea* (AABB), *B. napus* (AACC) and *B. carinata* (BBCC).

For *B. nigra*, the smallest genome size has been reported to be 470 Mbp. Estimates for different *B. rapa* subspecies range from 470 to 520 Mbp and the ones for *B. oleracea* subspecies from 600 to 670 Mbp. Genome sizes of the amphidiploid species are expected to be larger and indeed values of 1 105–1 235 Mbp have been reported for *B. napus* and *B. juncea* [2].

Many genetic maps have been assembled for *Brassica* species (reviewed in [39]). It has been noted in the first genetic mapping experiments that a high proportion of the *B. oleracea* [47], *B. rapa* [49] and *B. nigra* [52] genomes are duplicated. The analysis of a particularly polymorphic cross in *B. nigra* revealed eight chromosomal segments, each present in three copies (*figure 1A*). An almost complete conservation of gene repertoire in *B. nigra*, *B. oleracea* and *B. rapa* allowed a detailed comparative analysis of these genomes. Colinear segments involving the whole of each of the three *B. nigra* genomes could be identified in *B. oleracea* and *B. rapa*. This led to the proposal that all three *Brassica* genomes studied have a triplicated genome, thus they could be derived from a common hexaploid ancestor [25]. Despite this extensive colinearity, the chromosomal structures of the three species are differentiated by many rearrangements [25, 28, 53].

Comparative genetic mapping experiments showed that the chromosome maps of synthetic *B. napus* are essentially unchanged when compared to the ones of its progenitors, *B. oleracea* and *B. rapa* [8, 37]. The same was observed when the genetic map of amphidiploid *B. juncea* was aligned with maps of *B. nigra* and *B. rapa*, its diploid parental species [3].

5. THE GENE REPERTOIRE IS CONSERVED BETWEEN DIFFERENT GENERA OF THE BRASSICACEAE FAMILY

Comparisons of gene sequences derived from different species of the Brassicaceae family reveal conservation of exon sequences. For thirteen different *A. thaliana* genes and their homologues in *B. napus*, an average identity of coding sequences of 87 % was established, at the nucleotide and amino acid level [11]. Very similar values were obtained by comparing protein sequences encoded by a set of another six genes [10]. *Brassica* genes coding for hypothetical proteins homologous to the disease resistance genes *RPM1* and *RPS2* also show amino acid identities of 81 and 86 %, respectively, when compared to the *Arabidopsis* orthologues [16, 55]. In contrast, the family of *CONSTANS LIKE* genes evolves more rapidly within the Brassicaceae family [24].

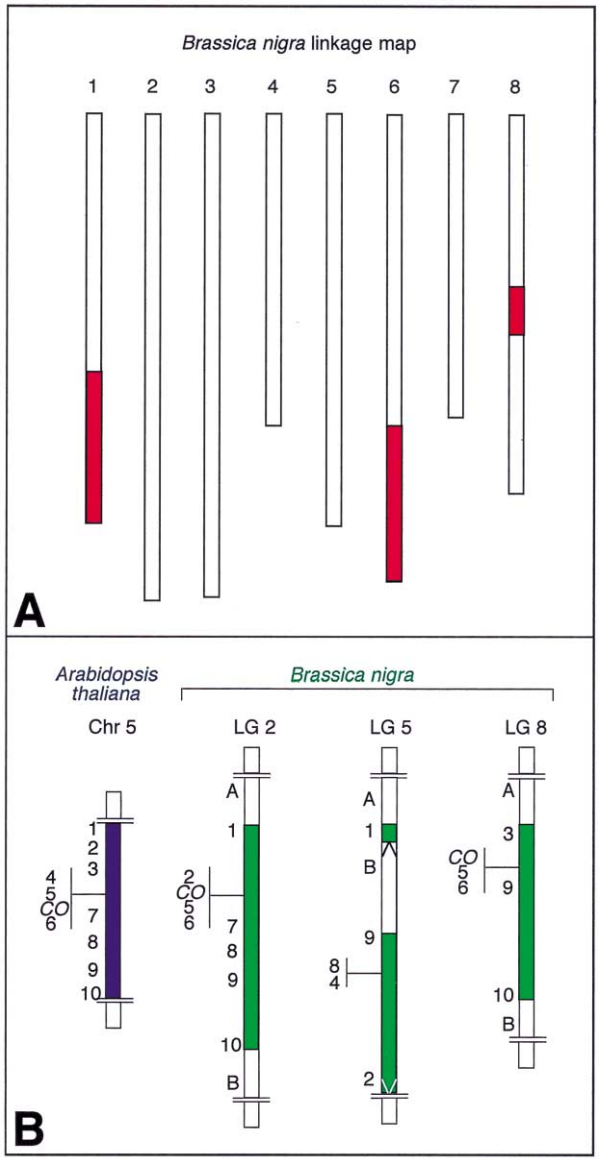

Figure 1. Evidence for triplicated segments in the *B. nigra* genome. **A**, Each bar represents a linkage group of the *B. nigra* molecular marker map. Three chromosomal segments shaded in red harbour common sets of homologous loci. These highlighted segments exemplify that the majority of the *B. nigra* genome shows a triplicated structure (adapted from [25]). **B**, A segment of *Arabidopsis* chromosome 5 harbouring the *CO* gene is compared to its counterparts in *B. nigra*. A and B represent *Brassica* markers, whereas markers 1–10 and *CO* designate *Arabidopsis* markers. Genetic mapping revealed three regions in *B. nigra* which correspond to the *CO* region. The regions show colinearity, but the segment located on linkage group 5 in *B. nigra* is characterised by an inversion event (adapted from [26]).

Sequence identity values of 92 and 93 % at the nucleotide and the amino acid level, respectively, were found by comparing coding sequences of four orthologous genes of *A. thaliana* and *C. rubella*. These results are consistent with the more recent divergence of *Arabidopsis* and *Capsella* compared to the one of *Arabidopsis* and *Brassica*. Furthermore, exon length and intron positions are conserved [1].

Due to this high sequence similarity, *A. thaliana* cDNA clones or genomic DNA fragments that correspond to exon sequences can be used as probes on Southern blots of *Capsella* genomic DNA to assess the similarity of gene repertoire in both species. Fifty (94 %) of the 53 different fragments tested clearly hybridised to *Capsella* DNA, indicating a highly similar but not identical set of genes in *Arabidopsis* and *Capsella*. In contrast, weaker or no hybridisation signals at all were found for *A. thaliana* fragments which do not harbour exon sequences ([1]; A. Acarkan and R. Schmidt, unpubl. results).

Similar results have been described for *Arabidopsis* and *Brassica*. Conner et al. [12] established that eighteen (86 %) out of twenty-one different *Arabidopsis* cDNAs tested hybridised to *B. campestris* genomic DNA. Based on the annotation of gene sequences for a region of *A. thaliana* chromosome 4, nineteen DNA fragments were selected for cross-hybridisation studies in *Brassica*. All nineteen predicted gene sequences were found to be present in *B. oleracea* [35].

6. COMPARATIVE GENETIC MAPPING EXPERIMENTS IN THE BRASSICACEAE FAMILY REVEAL COLINEAR CHROMOSOME SEGMENTS

It is expected that at least some chromosome rearrangements distinguish the genetic maps of *A. thaliana* and *Capsella* since their base chromosome numbers are five and eight, respectively. *A. thaliana* cDNA sequences and RFLP markers distributed along chromosome 4 of *A. thaliana* were used for genetic linkage analyses in *Capsella*. In *Capsella* two linkage groups could be established. A comparison of the marker order determined for the *Capsella* linkage groups and the one for *Arabidopsis* chromosome 4 revealed two large colinear chromosome segments representing the majority of this *A. thaliana* chromosome ([1]; *figure 2*). The *Arabidopsis* and *Capsella* maps are, however, not completely colinear. They are distinguished for example by an inversion. Markers mi233 and 11177 are located on *A. thaliana* chromosome 4, but corresponding sequences are absent from the *Capsella* genome. Marker ATTS3374 corresponds to a single-copy gene on *A. thaliana* chromosome 4, whereas in *Capsella* at least two unlinked loci can be detected with this marker. This demonstrates that the genome arrangement of the two species is distinguishable by, in addition to few large chromosomal rearrangements, small rearrangements, such as deletions/insertions, duplications and translocations of gene sequences (*figure 2*).

The first evidence for conserved linkage arrangements in *Arabidopsis* and *Brassica* was found by mapping a small number of genes in both *A. thaliana* and *B. rapa* [51]. More extensive comparative genetic mapping experiments between *A. thaliana* and *B. oleracea* corroborated the presence of colinear chromosome segments, which encompassed between 3.7 and 49.6 cM [21]. At least nineteen structural rearrangements differentiate the *B. oleracea* and *A. thaliana* chromosomes [28]. For a detailed comparison of the *A. thaliana* and *B. nigra* genomes, the average length of colinear segments was estimated at 8 cM. It was calculated that approximately ninety rearrangements have taken place since the divergence of these two species [23]. Thus, the colinear segments of the *Arabidopsis* and *Brassica* genomes are on average shorter than the ones which could be established for *Arabidopsis* and *Capsella*. Comparative genetic mapping experiments did not reveal recent duplications of chromosome segments in *Capsella*. In contrast, most *Arabidopsis* chromosome segments are present in three copies in *B. nigra*, providing further evidence for hexaploid ancestry of this species ([23]; *figure 1B*).

Comparative analysis of small genomic regions in *Brassica* and *Arabidopsis* has confirmed the triplicated nature of *Brassica* genomes. Colinear arrangement of loci detected by eleven *A. thaliana* fragments located on a 1.5-Mbp segment surrounding the *CO* gene on chromosome 5 could be established in three regions of *B. nigra*. In one of these *B. nigra* segments, however, colinearity is interrupted by an inversion event ([26]; *figure 1B*). The *CO* region was also found to be triplicated in *B. oleracea* [9]. A 30-cM segment of *A. thaliana* chromosome 3 containing the *FAD2*, *FAD5* and *FAD7* genes is colinear with six regions in *B. napus*, three segments correspond to the linkage groups derived from *B. oleracea* and *B. rapa* [44]. Likewise, a 30-cM segment of *A. thaliana* chromosome 4 corresponds to six regions in the oilseed rape genome, two of which display an internal inversion [11].

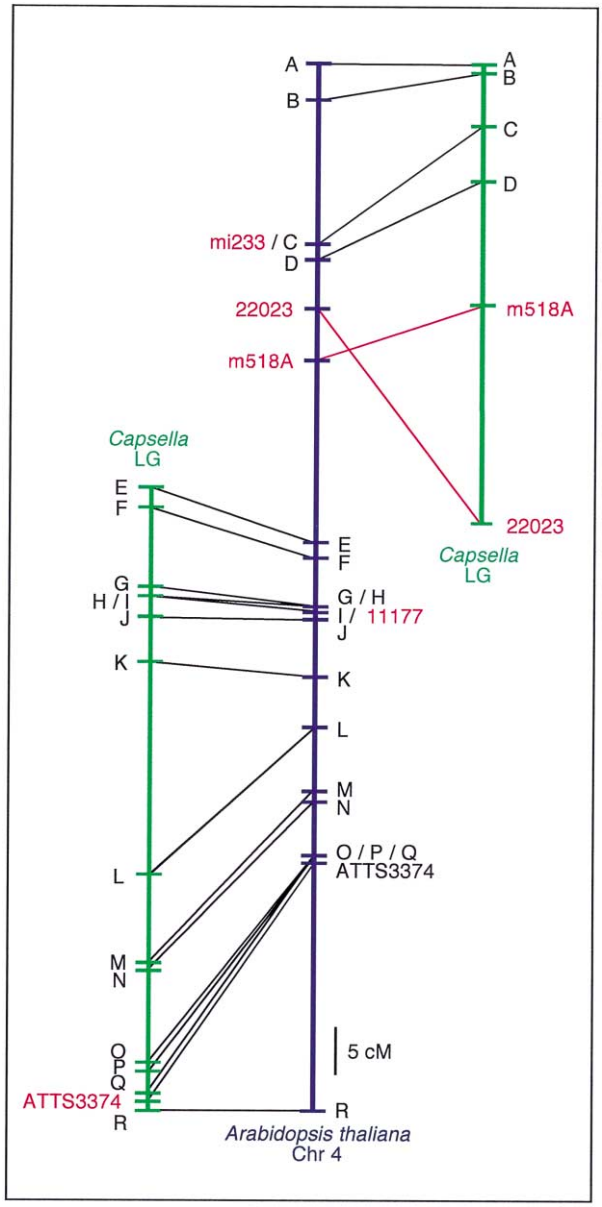

Figure 2. Comparative genetic mapping experiments reveal colinear segments in the *Arabidopsis* and *Capsella* genomes. RFLP markers and EST sequences located on *Arabidopsis* chromosome 4 (A-R, mi233, m518A, 22023, 11177 and ATTS3374) were used for genetic mapping experiments in *Capsella*. The resulting linkage groups (LG) are compared to the genetic map of *A. thaliana* chromosome 4 (Chr 4). Markers shown in red represent deviations from colinearity. Markers mi233 and 11177 are present in the *A. thaliana* genome but they do not hybridise to *Capsella* DNA. The order of markers m518A and 22023 is inverted in *Capsella* when compared to the *Arabidopsis* map. ATTS3774 represents a single-copy sequence in the *Arabidopsis* genome, whereas in *Capsella* two unlinked loci can be detected, one of which maps on the linkage group shown. Mapping data are adapted from Acarkan et al. [1].

7. MICROSYNTENY STUDIES

Microsynteny studies in the Brassicaceae family rely on comparative physical mapping studies. Homeologous regions of the *Arabidopsis* and *Brassica* genomes are characterised in respect to gene content and compared, taking advantage of libraries of cloned genomic DNA fragments, pulsed field gel electrophoresis or fluorescence in situ hybridisations.

Genes flanking the self-incompatibility genes of *B. campestris* were used as probes to identify the homeologous region in *A. thaliana*. Genes are organised in colinear fashion in both species, but evidence for deviations from microcolinearity was found. Based on the results of hybridisation experiments similar gene sequences were not detectable in *B. campestris* for three out of twenty-one genes which are located in the 275-kbp *A. thaliana* region. Sequences similar to the *B. campestris SLG* and *SRK* genes are absent from the homeologous *A. thaliana* region. In addition to these gene deletion/insertion events evidence for a small inversion and translocations was detected [12]. Most of the single-copy genes located in this region of the *A. thaliana* genome are also single-copy in *B. campestris*. These results contrast with other comparative mapping studies between *A. thaliana* and *Brassica* species. A 431-kbp contiguous segment of *A. thaliana* chromosome 2 was detected for example on several *Brassica rapa* chromosomes using fluorescence in situ hybridisations [18]. An *A. thaliana* segment carrying three single-copy genes, *GTP*, *RPM1* and *M4*, is present in six copies in the *B. napus* genome. Only two of these loci, however, contain a copy of *RPM1* (*figure 3B*). It was concluded that the absence of the *RPM1* gene in four of the loci is the result of gene deletion events [16].

Differences in gene content seem to be frequently observed in homeologous segments of the *Brassica* genomes. A 15-kbp segment of *A. thaliana* chromosome 3 encompassing five genes is conserved on a single linkage group in each of the three diploid *Brassica* species, *B. rapa*, *B. oleracea* and *B. nigra*, but evidence for partial clusters was also discovered in all three species [43]. Similar results were observed if the organisation of a 30-kbp segment of *A. thaliana* chromosome 4 carrying six genes was compared to the corresponding regions in *B. nigra*. In the *Brassica* genome, the complex was found to be multiplied and conserved in gene content at different levels with only one region carrying all six genes [42].

A detailed study corroborated these findings in comparisons between *A. thaliana* and *B. oleracea*.

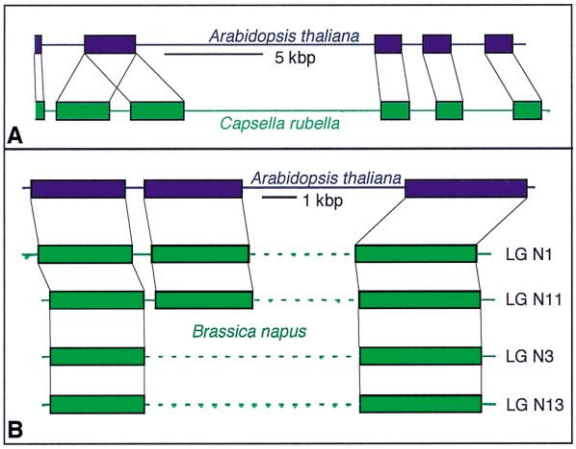

Figure 3. Exceptions from microcolinearity. **A**, A sequence comparison of orthologous regions in *A. thaliana* and *C. rubella* reveals that gene sequences, shown as boxes, are arranged in the same order in both species. The genes shown to the left are not fully represented in the drawing. Lines connect orthologous genes. One of the genes is present as a single-copy sequence in *Arabidopsis*, whereas in *Capsella* two tandemly repeated copies are found [1]. **B**, A region of *A. thaliana* chromosome 3 harbouring three genes, shown as boxes, is compared to corresponding segments in the *B. napus* genome. Lines connect homologous gene sequences. Only four of the six homeologous *B. napus* segments are shown. The *B. napus* regions located on linkage groups (LG) N1 and N11 are colinear with the *Arabidopsis* region, whereas in the segments mapping to linkage groups N3 and N13 one of the genes has been deleted. Dashed lines indicate that the exact distance between genes is not known (adapted from [16]).

Brassica BAC contigs were identified using nineteen different *Arabidopsis* gene sequences as probes. In *A. thaliana*, these genes are located on a 222-kbp segment of chromosome 4, additionally a subset is present in colinear arrangement on chromosome 5. Seven *B. oleracea* BAC contigs were identified which carried sequences homologous to the *A. thaliana* genes, three corresponded to the *Arabidopsis* chromosome 5 region whereas the remaining four were counterparts of the segment on *A. thaliana* chromosome 4. The gene content in the three different *B. oleracea* contigs corresponding to *A. thaliana* chromosome 5 region is not identical, only the gene repertoire of the three copies taken together is equivalent to all *Arabidopsis* genes analysed. With the exception of those genes missing in one or even two of the triplicated segments, all three regions are colinear with the *Arabidopsis* counterpart. Likewise, the *B. oleracea* homeologues of the *A. thaliana* chromosome 4 segment show differences in gene repertoire. Furthermore, evidence for a translocation and an inversion event was detected ([35]; *figure 4*).

Homeologous segments of *Arabidopsis* and *Brassica* are of similar size in some of the regions analysed, but in several cases an increase of size was noted for a *Brassica* region when compared to its *Arabidopsis* counterpart [12, 16, 18, 35, 42, 43, 45].

Complete colinearity was observed for a comparison of homeologous regions in *Arabidopsis* and *C. rubella*. A 60-kbp region harbouring at least ten different genes in *A. thaliana* was conserved in respect to gene repertoire in *C. rubella* [1]. For five of the genes, it could be established that their order and orientation is identical in both species; however, one of the genes is tandemly duplicated in *C. rubella* but not in *A. thaliana*. The genomic regions are very similar in size in both species (*figure 3A*).

8. PATTERNS OF GENOME EVOLUTION

Comparative genetic mapping experiments revealed colinear chromosome segments if Brassicaceae species were compared, albeit of different length. This provides evidence for the conservation of gross chromosomal structure in closely related species. If colinearity is assessed at the microlevel, deviations are found much more frequently. Gene deletion and duplication events, probably resulting from unequal crossover, are particularly common. Inversions and translocations of single genes or small groups of genes are also detected.

From all synteny studies carried out to date between species belonging to the Brassicaceae family, it appears that *Arabidopsis* and *Capsella* display more pronounced colinearity than *Arabidopsis* and *Brassica* species. This result is consistent with the more recent separation of the *Arabidopsis*/*Capsella* lineages compared to the one of *Arabidopsis*/*Brassica*. The divergence time of *Arabidopsis* and *Capsella* has been estimated at 6.2–9.8 million years ago, whereas values of 12.2–19.2 million years ago were determined for the species pair *Arabidopsis* and *Brassica* [1].

The study of triplicated segments of the *Brassica* genomes reveals striking differences with respect to gene repertoire. Only the genes of the triplicated regions taken together make up the complement in the corresponding *Arabidopsis* region; in any one of the triplicated regions, one or several homologues of the *Arabidopsis* genes may be missing [35]. These results provide evidence for the frequent occurrence of gene deletion events in multiplied regions of a genome. Thus, the proposed polyploid ancestry of the *Brassica* lineage may be an important factor contributing to the less pronounced colinearity seen for *Arabidopsis* and *Brassica* when compared to *Arabidopsis* and *Capsella*.

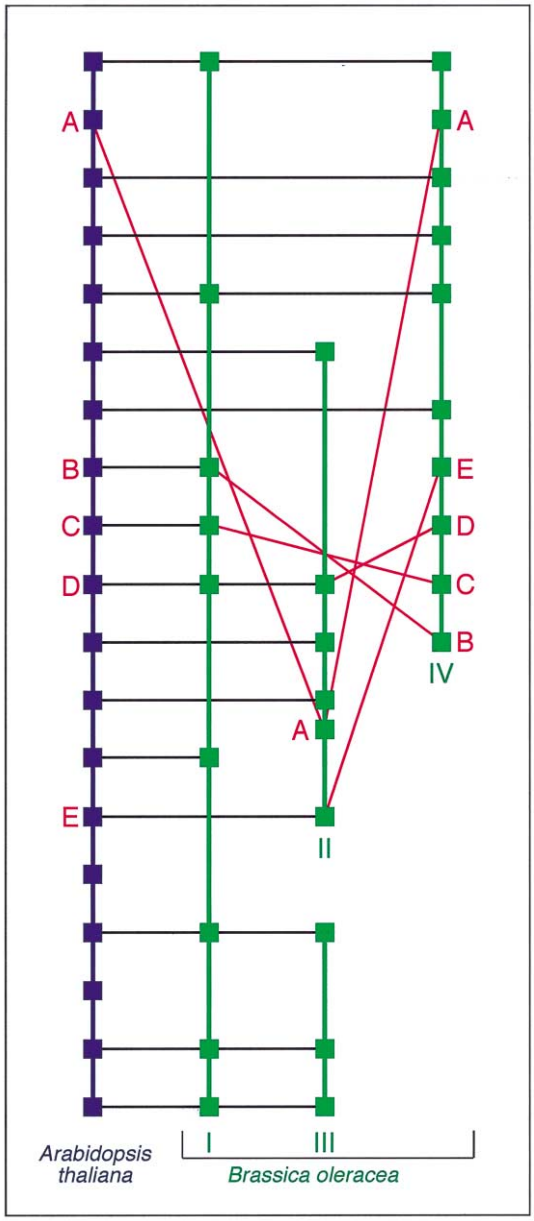

Figure 4. Homeologous *B. oleracea* regions differ in gene repertoire. A region of *A. thaliana* chromosome 4, shown on the left, is compared to corresponding segments in *B. oleracea* (adapted from [35]). *Arabidopsis* gene sequences, shown as blue squares, have been used as probes in hybridisation experiments to identify *B. oleracea* BAC clones and four BAC contigs could be assembled (I–IV). Contigs are shown as green lines. A blue square that is connected with a green square by a black or red line indicates that a particular *Arabidopsis* gene sequence hybridises to a *B. oleracea* BAC contig. Sequences homologous to gene A are found in a colinear position in contig IV when compared to the *A. thaliana* region, whereas in contig II they are residing in a position indicative of a translocation event. The order of genes B–E is inverted in contig IV in respect to the *A. thaliana* segment.

Analysis of the *Arabidopsis* genome sequence data revealed large segmental duplications [7, 29, 31, 39], the extent of which suggests a polyploid ancestry for *Arabidopsis* [4]. These segmental duplications also show differences in their gene repertoire [4, 7, 29, 31, 39].

Interestingly, the results of a comparison of a 105-kbp tomato segment and corresponding regions of the *Arabidopsis* genome are consistent with a model of frequent gene deletion events subsequent to large-scale duplications. This particular region of tomato chromosome 2 is related to four chromosomal regions in *Arabidopsis thaliana*. The observed complex relationships between segments of the *Arabidopsis* and tomato genomes are compatible with assuming two consecutive rounds of duplications in the *Arabidopsis* lineage with subsequent extensive loss of genes in the duplicated regions [22]. The earlier round of duplications has been estimated to have occurred approximately 112 million years ago [22], thus predating the divergence of *Arabidopsis* and *Brassica*. This is in accordance with colinearity studies between *B. oleracea* and *A. thaliana* which revealed that at least the segmental duplication analysed was present in the progenitor of these species [35].

In a study of synthetic polyploids of *Brassica*, evidence for extensive and rapid genome change could be observed. Different processes, such as chromosome rearrangement, point mutation, gene conversions or DNA methylation could bring about the described alterations [48]. However, polyploidisation events are not necessarily followed by large changes in chromosome structure, since the genomes of amphidiploid *B. napus* and *B. juncea* are essentially unaltered when compared to their progenitor genomes [3, 8, 37].

9. THE *ARABIDOPSIS* GENOME AS A TOOL FOR GENE ISOLATION IN OTHER SPECIES OF THE BRASSICACEAE FAMILY

The identification of genes encoding for agronomically important traits in *Brassica* is an important goal. Comparative mapping experiments may facilitate the recognition of putative orthology of *Brassica* QTL or monogenic traits and characterised genes in *Arabidopsis*. In that respect, it is important that the function of many *Arabidopsis* genes has been identified using molecular genetic techniques.

Different strategies can be followed to correlate *Brassica* QTL with *Arabidopsis* genes, all of which make use of reciprocal mapping experiments. On the one hand, putative candidate genes from *Arabidopsis*

can be used as markers on *Brassica* mapping populations and it can be evaluated whether the segregation of RFLPs detected by these genes coincide with QTL of interest. On the other hand, *Brassica* markers in the vicinity of a particular QTL can be used for genetic mapping experiments in *Arabidopsis* to identify the corresponding genomic region. Alternatively, the sequences of molecular markers flanking the locus of interest in *Brassica* can be determined and aligned with the sequence of the *A. thaliana* genome. Due to the high sequence identity between *Arabidopsis* and *Brassica* genes, there is a high likelihood that any *Brassica* marker representing exon sequences can be directly placed on the *A. thaliana* chromosome maps. If a region has been identified in the *Arabidopsis* genome which coincides with a particular QTL in the crop plant, the availability of the annotated *Arabidopsis* sequence offers unique opportunities to refine its positioning and ultimately identify the locus encoding for the trait of interest.

For example, the control of flowering time in *Brassica* is being characterised by using information on *Arabidopsis* genes that have been implicated in this mechanism. It could be shown that homologues of *CO*, one of the *Arabidopsis* genes influencing flowering time, coincide with loci regulating this process in *B. nigra* [26]. Three out of six quantitative trait loci affecting flowering time in *B. rapa* are situated in regions that show considerable homology with each other. Most importantly, these segments are colinear with those areas of the *B. nigra* genome that are associated with this trait and with a segment of *Arabidopsis* chromosome 5 harbouring a number of genes influencing this character [9]. For QTL conferring vernalisation-responsive flowering time in *B. napus* and *B. rapa* colinear regions in *Arabidopsis* could also be detected where flowering time genes are located [36]. Similarly, a number of QTL associated with early-flowering in *B. oleracea* fall in regions corresponding to segments of the *A. thaliana* genome that harbour mutations affecting this trait [27].

A *B. napus* variant with petalless flowers also exhibits a change in leaf morphology. Both characters co-segregate and are controlled by an epistatic interaction between two loci. In *A. thaliana*, the *CURLY LEAF* (*CLF*) gene has been found to influence flower as well as leaf morphologies. A co-segregation analysis could show that an oilseed rape homologue of *CLF* coincides with one of the two loci controlling the pleiotropic effects over flower and leaf morphology in *B. napus*, thus identifying *CLF* as a candidate gene for this trait [14]. A similar strategy was used to correlate homologues of the *Arabidopsis* fatty acid elongase (*FAE1*) gene with two loci controlling erucic acid content in oilseed rape [13].

In order to provide a source of candidate-resistance genes for *B. napus*, *Arabidopsis* ESTs and *Brassica* sequences with homology to cloned plant resistance genes were mapped in *B. napus*. The resulting map positions can now be integrated with the disease resistance loci that have been placed on the oilseed rape genome [46].

From the results of comparative genome analysis studies carried out to date, it emerges that the availability of the well-characterised *Arabidopsis* genome presents a unique resource for the identification of candidate genes coding for economically relevant traits in *Brassica*.

10. CONCLUSION

Comparative genetic mapping experiments have revealed extensive genome colinearity between species belonging to the Brassicaceae family. These studies have unveiled the triplicated nature of the genomes of the modern diploid *Brassica* species. Microsynteny studies have confirmed that at the level of genes, colinearity is also generally observed, but most importantly they have disclosed the rapid changes occurring in multiplied chromosome segments that complicate colinearity relationships.

The degree of genome colinearity observed at the chromosomal and at the fine structure level is high enough to allow in many cases the efficient exploitation of *A. thaliana* genome data for the analysis of traits in *Brassica* species. This should ultimately contribute to the improvement of these important crop plants.

Acknowledgments. We thank Dr Burkhard Schulz for commenting on the manuscript.

REFERENCES

[1] Acarkan A., Rossberg M., Koch M., Schmidt R., Comparative genome analysis reveals extensive conservation of genome organisation for *Arabidopsis thaliana* and *Capsella rubella*, Plant J. 23 (2000) 55–62.

[2] Arumuganathan K., Earle E.D., Nuclear DNA content of some important plant species, Plant Mol. Biol. Rep. 9 (1991) 208–218.

[3] Axelsson T., Bowman C.M., Sharpe A.G., Lydiate D.J., Lagercrantz U., Amphidiploid *Brassica*

juncea contains conserved progenitor genomes, Genome 43 (2000) 679–688.
[4] Bancroft I., Insights into the structural and functional evolution of plant genomes afforded by the nucleotide sequences of chromosomes 2 and 4 of *Arabidopsis thaliana*, Yeast 17 (2000) 1–5.
[5] Bennetzen J.L., Comparative sequence analysis of plant nuclear genomes: microcolinearity and its many exceptions, Plant Cell 12 (2000) 1021–1029.
[6] Bevan M., Ecker J., Theologis S., Federspiel N., Davis R., McCombie D., Martienssen R., Chen E., Waterston B., Wilson R., Rounsley S., Venter C., Tabata S., Salanoubat M., Quetier F., Cherry J.M., Meinke D., Objective: the complete sequence of a plant genome, Plant Cell 9 (1997) 476–478.
[7] Blanc G., Barakat A., Guyot R., Cooke R., Delseny M., Extensive duplication and reshuffling in the *Arabidopsis* genome, Plant Cell 12 (2000) 1093–1101.
[8] Bohuon E.J.R., Keith D.J., Parkin I.A.P., Sharpe A.G., Lydiate D.J., Alignment of the conserved C genomes of *Brassica oleracea* and *Brassica napus*, Theor. Appl. Genet. 93 (1996) 833–839.
[9] Bohuon E.J.R., Ramsay L.D., Craft J.A., Arthur A.E., Marshall D.F., Lydiate D.J., Kearsey M.J., The association of flowering time quantitative trait loci with duplicated regions and candidate loci in *Brassica oleracea*, Genetics 150 (1998) 393–401.
[10] Brunel D., Froger N., Pelletier G., Development of amplified consensus genetic markers (ACGM) in *Brassica napus* from *Arabidopsis thaliana* sequences of known biological function, Genome 42 (1999) 387–402.
[11] Cavell A.C., Lydiate D.J., Parkin I.A.P., Dean C., Trick M., Collinearity between a 30-centimorgan segment of *Arabidopsis thaliana* chromosome 4 and duplicated regions within the *Brassica napus* genome, Genome 41 (1998) 62–69.
[12] Conner J.A., Conner P., Nasrallah M.E., Nasrallah J.B., Comparative mapping of the *Brassica S* locus region and its homeolog in *Arabidopsis*: implications for the evolution of mating systems in the Brassicaceae, Plant Cell 10 (1998) 801–812.
[13] Fourmann M., Barret P., Renard M., Pelletier G., Delourme R., Brunel D., The two genes homologous to *Arabidopsis FAE1* co-segregate with the two loci governing erucic acid content in *Brassica napus*, Theor. Appl. Genet. 96 (1998) 852–858.
[14] Fray M.J., Puangsomlee P., Goodrich J., Coupland G., Evans E.J., Arthur A.E., Lydiate D.J., The genetics of stamenoid petal production in oilseed rape (*Brassica napus*) and equivalent variation in *Arabidopsis thaliana*, Theor. Appl. Genet. 94 (1997) 731–736.
[15] Gale M.D., Devos K.M., Comparative genetics in the grasses, Proc. Natl. Acad. Sci. USA 95 (1998) 1971–1974.
[16] Grant M.R., McDowell J.M., Sharpe A.G., de Torres Zabala M., Lydiate D.J., Dangl J.L., Independent deletions of a pathogen-resistance gene in *Brassica* and *Arabidopsis*, Proc. Natl. Acad. Sci. USA 95 (1998) 15843–15848.
[17] Höfte H., Desprez T., Amselem J., Chiapello H., Rouzé P., Caboche M., et al., An inventory of 1152 expressed sequence tags obtained by partial sequencing of cDNAs from *Arabidopsis thaliana*, Plant J. 4 (1993) 1051–1061.
[18] Jackson S.A., Cheng Z.K., Wang M.L., Goodman H.M., Jiang J.M., Comparative fluorescence *in situ* hybridization mapping of a 431-kb *Arabidopsis thaliana* bacterial artificial chromosome contig reveals the role of chromosomal duplications in the expansion of the *Brassica rapa* genome, Genetics 156 (2000) 833–838.
[19] Koch M., Bishop J., Mitchell-Olds T., Molecular systematics and evolution of *Arabidopsis* and *Arabis*, Plant Biol. 1 (1999) 529–537.
[20] Koch M.A., Haubold B., Mitchell-Olds T., Comparative evolutionary analysis of chalcone synthase and alcohol dehydrogenase loci in *Arabidopsis*, *Arabis*, and related genera (Brassicaceae), Mol. Biol. Evol. 17 (2000) 1483–1498.
[21] Kowalski S.P., Lan T.H., Feldmann K.A., Paterson A.H., Comparative mapping of *Arabidopsis thaliana* and *Brassica oleracea* chromosomes reveals islands of conserved organization, Genetics 138 (1994) 499–510.
[22] Ku H.M., Vision T., Liu J., Tanksley S.D., Comparing sequenced segments of the tomato and *Arabidopsis* genomes: large-scale duplication followed by selective gene loss creates a network of synteny, Proc. Natl. Acad. Sci. USA 97 (2000) 9121–9126.
[23] Lagercrantz U., 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 (1998) 1217–1228.
[24] Lagercrantz U., Axelsson T., Rapid evolution of the family of *CONSTANS LIKE* genes in plants, Mol. Biol. Evol. 17 (2000) 1499–1507.
[25] Lagercrantz U., Lydiate D., Comparative genome mapping in Brassica, Genetics 144 (1996) 1903–1910.
[26] Lagercrantz U., Putterill J., Coupland G., Lydiate D., Comparative mapping in *Arabidopsis* and *Brassica*, fine scale genome collinearity and congruence of genes controlling flowering time, Plant J. 9 (1996) 13–20.
[27] Lan T.H., Paterson A.H., Comparative mapping of quantitative trait loci sculpting the curd of *Brassica oleracea*, Genetics 155 (2000) 1927–1954.
[28] Lan T.H., DelMonte T.A., Reischmann K.P., Hyman J., Kowalski S.P., McFerson J., Kresovich S., Paterson A.H., An EST-enriched comparative map of *Brassica oleracea* and *Arabidopsis thaliana*, Genome Res. 10 (2000) 776–788.
[29] Lin X.Y., Kaul S.S., Rounsley S., Shea T.P., Benito M.I., Town C.D., et al., Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*, Nature 402 (1999) 761–768.
[30] Marra M., Kucaba T., Sekhon M., Hillier L., Martienssen R., Chinwalla A., Crockett J., Fedele J., Grover H., Gund C., McCombie W.R., McDonald K., McPher-

son J., Mudd N., Parnell L., Schein J., Seim R., Shelby P., Waterston R., Wilson R., A map for sequence analysis of the *Arabidopsis thaliana* genome, Nat. Genet. 22 (1999) 265–270.
[31] Mayer K., Schüller C., Wambutt R., Murphy G., Volckaert G., Pohl T., et al., Sequence and analysis of chromosome 4 of the plant *Arabidopsis thaliana*, Nature 402 (1999) 769–777.
[32] Meyerowitz E.M., Somerville C.R., *Arabidopsis*, Cold Spring Harbor Laboratory Press, New York, 1994.
[33] Mozo T., Dewar K., Dunn P., Ecker J.R., Fischer S., Kloska S., Lehrach H., Marra M., Martienssen R., Meier-Ewert S., Altmann T., A complete BAC-based physical map of the *Arabidopsis thaliana* genome, Nat. Genet. 22 (1999) 271–275.
[34] Newman T., de Bruijn F.J., Green P., Keegstra K., Kende H., McIntosh L., Ohlrogge J., Raikhel N., Somerville S., Thomashow M., Retzel E., Somerville C., Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous *Arabidopsis* cDNA clones, Plant Phys. 106 (1994) 1241–1255.
[35] O'Neill C.M., Bancroft I., 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 (2000) 233–243.
[36] Osborn T.C., Kole C., Parkin I.A.P., Sharpe A.G., Kuiper M., Lydiate D.J., Trick M., Comparison of flowering time genes in *Brassica rapa*, *B. napus* and *Arabidopsis thaliana*, Genetics 146 (1997) 1123–1129.
[37] Parkin I.A.P., Sharpe A.G., Keith D.J., Lydiate D.J., Identification of the A and C genomes of amphidiploid *Brassica napus* (oilseed rape), Genome 38 (1995) 1122–1131.
[38] Paterson A.H., Wing R.A., Genome mapping in plants, Curr. Opin. Biotechnol. 4 (1993) 142–147.
[39] Paterson A.H., Bowers J.E., Burow M.D., Draye X., Elsik C.G., Jiang C.X., Katsar C.S., Lan T.H., Lin Y.R., Ming R., Wright R.J., Comparative genomics of plant chromosomes, Plant Cell 12 (2000) 1523–1539.
[40] Price R.A., Palmer J.D., Al-Shebaz I.A., Systematic relationships of Arabidopsis: a molecular and morphological perspective, in: Meyerowitz E.M., Somerville C.R. (Eds.), Arabidopsis, Cold Spring Harbor Laboratory Press, New York, 1994, pp. 7–19.
[41] Quackenbush J., Liang F., Holt I., Pertea G., Upton J., The TIGR gene indices: reconstruction and representation of expressed gene sequences, Nucleic Acids Res. 28 (2000) 141–145.
[42] Sadowski J., Quiros C.F., Organization of an *Arabidopsis thaliana* gene cluster on chromosome 4 including the *RPS2* gene in the *Brassica nigra* genome, Theor. Appl. Genet. 96 (1998) 468–474.
[43] Sadowski J., Gaubier P., Delseny M., Quiros C.F., Genetic and physical mapping in *Brassica* diploid species of a gene cluster defined in *Arabidopsis thaliana*, Mol. Gen. Genet. 251 (1996) 298–306.
[44] Scheffler J.A., Sharpe A.G., Schmidt H., Sperling P., Parkin I.A.P., Lühs W., Lydiate D.J., Heinz E., Desaturase multigene families of *Brassica napus* arose through genome duplication, Theor. Appl. Genet. 94 (1997) 583–591.
[45] Schmidt R., Acarkan A., Koch M., Roßberg M., A strategy for comparative physical mapping in cruciferous plants, in: van Raamsdonk L.W.D., den Nijs J.C.M. (Eds.), Proceedings of the VIIth Symposium of the International Organization of Plant Biosystematists, Plant Evolution in Man-Made Habitats, Hugo de Vries Laboratory, University of Amsterdam, Amsterdam, 1999, pp. 183–196.
[46] Silito D., Parkin I.A.P., Mayerhofer R., Lydiate D.J., Good A.G., *Arabidopsis thaliana*: a source of candidate disease-resistance genes for *Brassica napus*, Genome 43 (2000) 452–460.
[47] Slocum M.K., Figdore S.S., Kennard W.C., Suzuki J.Y., Osborn T.C., Linkage arrangement of restriction fragment length polymorphism loci in *Brassica oleracea*, Theor. Appl. Genet. 80 (1990) 57–64.
[48] Song K.M., Lu P., Tang K.L., Osborn T.C., Rapid genome change in synthetic polyploids of *Brassica* and its implications for polyploid evolution, Proc. Natl. Acad. Sci. USA 92 (1995) 7719–7723.
[49] Song K.M., Suzuki J.Y., Slocum M.K., Williams P.H., Osborn T.C., A linkage map of *Brassica rapa* (syn. *campestris*) based on restriction fragment length polymorphism loci, Theor. Appl. Genet. 82 (1991) 296–304.
[50] Tanksley S.D., Ganal M.W., Prince J.P., de Vicente M.C., Bonierbale M.W., Broun P., Fulton T.M., Giovannoni J.J., Grandillo S., Martin G.B., Messeguer R., Miller J.C., Miller L., Paterson A.H., Pineda O., Röder M.S., Wing R.A., Wu W., Young N.D., High density molecular linkage maps of the tomato and potato genomes, Genetics 132 (1992) 1141–1160.
[51] Teutonico R.A., Osborn T.C., Mapping of RFLP and qualitative trait loci in *Brassica rapa* and comparison to the linkage maps of *B. napus*, *B. oleracea*, and *Arabidopsis thaliana*, Theor. Appl. Genet. 89 (1994) 885–894.
[52] Truco M.J., Quiros C.F., Structure and organization of the B genome based on a linkage map in *Brassica nigra*, Theor. Appl. Genet. 89 (1994) 590–598.
[53] Truco M.J., Hu J., Sadowski J., Quiros C., Inter- and intra-genomic homology of the Brassica genomes: implications for their origin and evolution, Theor. Appl. Genet. 93 (1996) 1225–1233.
[54] U N., Genomic analysis in Brassica with special reference to the experimental formation of *B. napus* and peculiar mode of fertilisation, Jpn. J. Bot. 7 (1935) 389–452.
[55] Wroblewski T., Coulibaly S., Sadowski J., Quiros C.F., Variation and phylogenetic utility of the *Arabidopsis thaliana Rps2* homolog in various species of the tribe Brassiceae, Mol. Phylogenet. Evol. 16 (2000) 440–448.
[56] Yang Y.W., Lai K.N., Tai P.Y., Li W.H., Rates of nucleotide substitution in angiosperm mitochondrial DNA sequences and dates of divergence between *Brassica* and other angiosperm lineages, J. Mol. Evol. 48 (1999) 597–604.