Fast Diploidization in Close Mesopolyploid Relatives of *Arabidopsis* [W][OA]

Terezie Mandáková,[a] Simon Joly,[b] Martin Krzywinski,[c] Klaus Mummenhoff,[d] and Martin A. Lysak[a,1]

[a] Department of Functional Genomics and Proteomics, Institute of Experimental Biology, Masaryk University, Kamenice 5, CZ-625 00 Brno, Czech Republic

[b] Institut de Recherche en Biologie Végétale, Université de Montréal and Montreal Botanical Garden, 4101 Sherbrooke East, Montreal, Quebec, Canada H1X 2B2

[c] Canada's Michael Smith Genome Sciences Center, Vancouver, British Columbia, Canada V5Z 4S6

[d] FB Biologie/Chemie, Botanik, Universität Osnabrück, D-49069 Osnabrück, Germany

Mesopolyploid whole-genome duplication (WGD) was revealed in the ancestry of Australian Brassicaceae species with diploid-like chromosome numbers (n = 4 to 6). Multicolor comparative chromosome painting was used to reconstruct complete cytogenetic maps of the cryptic ancient polyploids. Cytogenetic analysis showed that the karyotype of the Australian Camelineae species descended from the eight ancestral chromosomes (n = 8) through allopolyploid WGD followed by the extensive reduction of chromosome number. Nuclear and maternal gene phylogenies corroborated the hybrid origin of the mesotetraploid ancestor and suggest that the hybridization event occurred ∼6 to 9 million years ago. The four, five, and six fusion chromosome pairs of the analyzed close relatives of *Arabidopsis thaliana* represent complex mosaics of duplicated ancestral genomic blocks reshuffled by numerous chromosome rearrangements. Unequal reciprocal translocations with or without preceeding pericentric inversions and purported end-to-end chromosome fusions accompanied by inactivation and/or loss of centromeres are hypothesized to be the main pathways for the observed chromosome number reduction. Our results underline the significance of multiple rounds of WGD in the angiosperm genome evolution and demonstrate that chromosome number per se is not a reliable indicator of ploidy level.

INTRODUCTION

Hybridization and polyploidization (whole-genome duplication [WGD]) are important evolutionary forces driving genetic diversification and speciation in land plants. In angiosperms, numerous cases of speciation events following this pattern were documented (Grant, 1981; Soltis et al., 2007; Hegarty and Hiscock, 2008). Comparative research using whole-genome and EST sequence data sets uncovered compelling evidence of multiple ancient WGD events in the ancestry of angiosperm lineages (De Bodt et al., 2005; Cui et al., 2006; Soltis et al., 2009; Jaillon et al., 2009). In crucifers (Brassicaceae), the analysis of the *Arabidopsis thaliana* genome sequence (Arabidopsis Genome Initiative, 2000) suggested the existence of three paleopolyploid WGDs (α, β, and γ; Bowers et al., 2003). Whereas the phylogenetic placement of the oldest event (γ) is still debated (Soltis et al., 2009), the two more recent WGDs have been shown to postdate the split between Caricaceae and Brassicaceae (Tang et al., 2008). The most recent (α) duplication apparently occurred only within Brassicaceae, and it is equivalent to the whole-genome triplication in the sister family Cleomaceae (Schranz and Mitchell-Olds, 2006; Barker et al., 2009). The exact position of the β event within the order Brassicales has to be elucidated by further research (Barker et al., 2009).

The α WGD has been dated to occur ∼23 to 43 million years ago (mya) (Barker et al., 2009; Fawcett et al., 2009), after the β duplication postdating the Brassicaceae-Caricaceae divergence 72 mya (Ming et al., 2008). The three rounds of paleopolyploid WGDs (α, β, and γ) were followed by diploidization (fractionation; Thomas et al., 2006) toward diploid-like genomes marked by genome downsizing and chromosome rearrangements. For instance, the functionally diploid and extremely compact *Arabidopsis* genome (n = 5, 1C = 157 Mb) is a result of the extensive karyotype reshuffling of already diploidized Ancestral Crucifer Karyotype (n = 8) (Lysak et al., 2006). After a WGD, some duplicated genes are lost or subjected to sub- and neofunctionalization (Sémon and Wolfe, 2007), dosage-sensitive genes such as transcription factors are preferentially retained in the duplicated genomes, as described by the gene balance theory (Birchler and Veitia, 2007; Freeling, 2008; Veitia et al., 2008; Edger and Pires, 2009).

Multiple and often lineage-specific WGD events uncovered in Asteraceae (Barker et al., 2008), Cleomaceae (Schranz and Mitchell-Olds, 2006; Barker et al., 2009), Solanaceae (Schlueter et al., 2004), and other taxa (Soltis et al., 2009) imply a key role of WGDs in the evolution of land plants. The steadily improving knowledge of crucifer genome evolution suggests that the duplication-diploidization process is ongoing and occurred

several times across the family. In addition to the three afore-metioned paleopolyploid events, several Brassicaceae groups with diploid-like genomes have experienced additional, more recent WGD events, as exemplified by a whole-genome triplication (~8 to 15 mya) most likely promoting the diversification within the tribe Brassiceae (Lysak et al., 2005, 2007; Parkin et al., 2005; Panjabi et al., 2008). Such duplications are younger than paleopolyploid events but older than neopolyploid speciation events (e.g., the origin of *Arabidopsis suecica*, *Brassica napus*, or *Cardamine schulzii*) and can be detected by comparative genetic and cytogenetic methods.

In this study, we used comparative chromosome painting to analyze endemic Australian crucifers (*Stenopetalum nutans*, *Stenopetalum lineare*, and *Ballantinia antipoda*) with low, diploid-like chromosome numbers (n = 4, 5, and 6, respectively). Comparison with the Ancestral Crucifer Karyotype (n = 8) showed that the ancestor(s) of the analyzed species experienced a mesopolyploid WGD followed by extensive karyotype reshuffling associated with genome diploidization.

RESULTS

All or Most Genomic Blocks of the Ancestral Crucifer Karyotype Are Duplicated in the Genomes of Australian Species.

We analyzed chromosome structure and genome size of the three Australian crucifer species. Except distinct domains of pericentromeric heterochromatin, heterochromatic knobs and distal segments were detected in two species (Figures 1A to 1C). In *S. nutans* (n = 4; 477 Mb; accession 86929), chromosomes SN2 and SN4 possess terminal heterochromatic knobs on upper arms, whereas chromosome SN3 has an interstitial heterochromatic knob (hkSN3) on its bottom arm. No conspicuous heterochromatic knobs were observed in the *S. lineare* karyotype (n = 5; 763 Mb). The *B. antipoda* karyotype (n = 6; 472 Mb) is characterized by six large heterochromatic segments that extend from one-third up to the entire length of a chromosome arm (Figures 1C and 2D). In all three analyzed species, the *Arabidopsis*-type telomere repeat hybridized to chromosome termini. However, additional, remarkably strong hybridization signals were observed at centromeric regions of all *B. antipoda* chromosomes (Figure 2E).

As no genetic maps for the analyzed species were available, we used the Ancestral Crucifer Karyotype (ACK; Figure 1D) as a reference genome for the construction of molecular cytogenetic maps by comparative chromosome painting (CCP). The ACK, inferred from (cyto)genetic comparisons between genomes of *Arabidopsis*, *B. napus*, and other crucifer species, comprises eight ancestral chromosomes (AK1 to AK8) and 24 conserved genomic blocks (GBs) (Lysak et al., 2006; Schranz et al., 2006). Five hundred and forty seven chromosome-specific BAC clones of *Arabidopsis* representing all genomic blocks of the ACK have been differentially labeled and hybridized to pachytene chromosomes. Conserved or rearranged ancestral GBs, their physical size (bp), and correspondence to *Arabidopsis* BAC contigs are shown in Supplemental Table 1 online for all three species.

CCP experiments showed that most GBs, present as single copies in *Arabidopsis* and the ACK, were duplicated and hybridized to two chromosomes within pachytene complements of the Australian crucifers (Figures 1A to 1C, 2A, and 3). In *S. nutans*, blocks J, M, and R hybridized to three different chromosomes, and two copies of block U hybridized to a single chromosome (SN4). Genomic blocks R and F labeled three and block J four chromosomes in *S. lineare*. In *B. antipoda*, block R hybridized to three and blocks A and B to four different chromosomes, respectively. No obvious cross-hybridization of single *Arabidopsis* BAC clones outside the homoeologous collinear regions was observed (Figure 2A). Whereas all 24 GBs were found to be duplicated in *S. nutans* (Figures 1A, 2A, and 3A), only 22 GBs were found duplicated in *S. lineare* (Figures 1B and 3B) and 20 GBs in *B. antipoda* (Figures 1C and 3C) (as single copies were identified GBs P and Q in *S. lineare*, and D, E, P, and S in *B. antipoda*). The position of each GB is compared for each pair of the Australian crucifers in Supplemental Figure 1 online. In all species, paired GB duplicates differed slightly but consistently in length and fluorescence intensity (see Figure 2A for examples). This difference was more pronounced in the case of longer GBs (the longer and brighter copy always labeled as #1). The difference in intensity is likely due to differential origins of the two genome copies (but see Discussion). In *B. antipoda*, all distal heterochromatic segments were labeled by painting probes comparably to euchromatin regions (Figure 1C).

Ancestral ACK-like associations of many duplicated GBs were retained in the extant complements (see Supplemental Table 2 online; Figures 1A to 1C). In *S. nutans*, we found seven AK-like chromosomes [AK2(#2), AK3(#1, #2), AK4(#1), AK5(#2), AK7(#2), and AK8(#1)], 11 AK-like chromosome arms [upper arm of AK1(#1), AK2 (#1), AK4(#2), AK5(#1), AK6(#2), AK7(#1), and AK8(#2); bottom arm of AK1(AK1(#1, #2), AK6(#2), and AK8(#2)], seven preserved GBs not forming any AK-like structure, and nine split GBs. In *S. lineare*, five AK-like chromosomes [AK2(#2), AK3(#2), AK5(#1, #2), and AK7(#1)], 15 AK-like chromosome arms [upper arm of AK1(#1), AK2(#1), AK4(#1, #2), AK6(#1), AK7(#2), and AK8(#1, #2); bottom arm of AK1(#1, #2), AK2(#1), AK3(#1), AK7(#2), and AK8(#1, #2)], six preserved GBs not forming any AK-like structure, and six split GBs were identified. The *B. antipoda* complement comprises six AK-like homoeologs [AK2(#2), AK3(#1, #2), AK4(#1), AK5(#2), and AK8(#1)], 12 AK-like chromosome arms [upper arm of AK4(#2), AK5(#1), AK6(#1), AK7(#2), and AK8(#2); bottom arm of AK1(#1, #2), AK4(#2), AK5(#1) AK6(#2), AK7(#1), and AK8(#2)], four preserved GBs not forming any AK-like structure, and six split GBs. Except AK1 and AK6, at least one copy of all other ancestral chromosomes is structurally preserved within the analyzed karyotypes, viz. AK2(#2), AK3(#1, #2), AK4(#1), AK5(#1, #2), AK7(#1, #2), and AK8(#1). Homoeologs AK2(#2), AK3(#2), and AK5(#2) were shared by all three species. In all three genomes, homoeologs of AK6 and AK8 are rearranged, whereby the rearranged structure of AK8(#1) is shared by all three species (see Supplemental Figure 2 and Supplemental Table 2 online). Although numerous AK chromosomes and chromosome arms were found as conserved and duplicated, they were rearranged (fused) to build extant chromosomes of the Australian species. For instance, the bottom arm of *S. nutans* chromosome SN4 consists of three complete AK
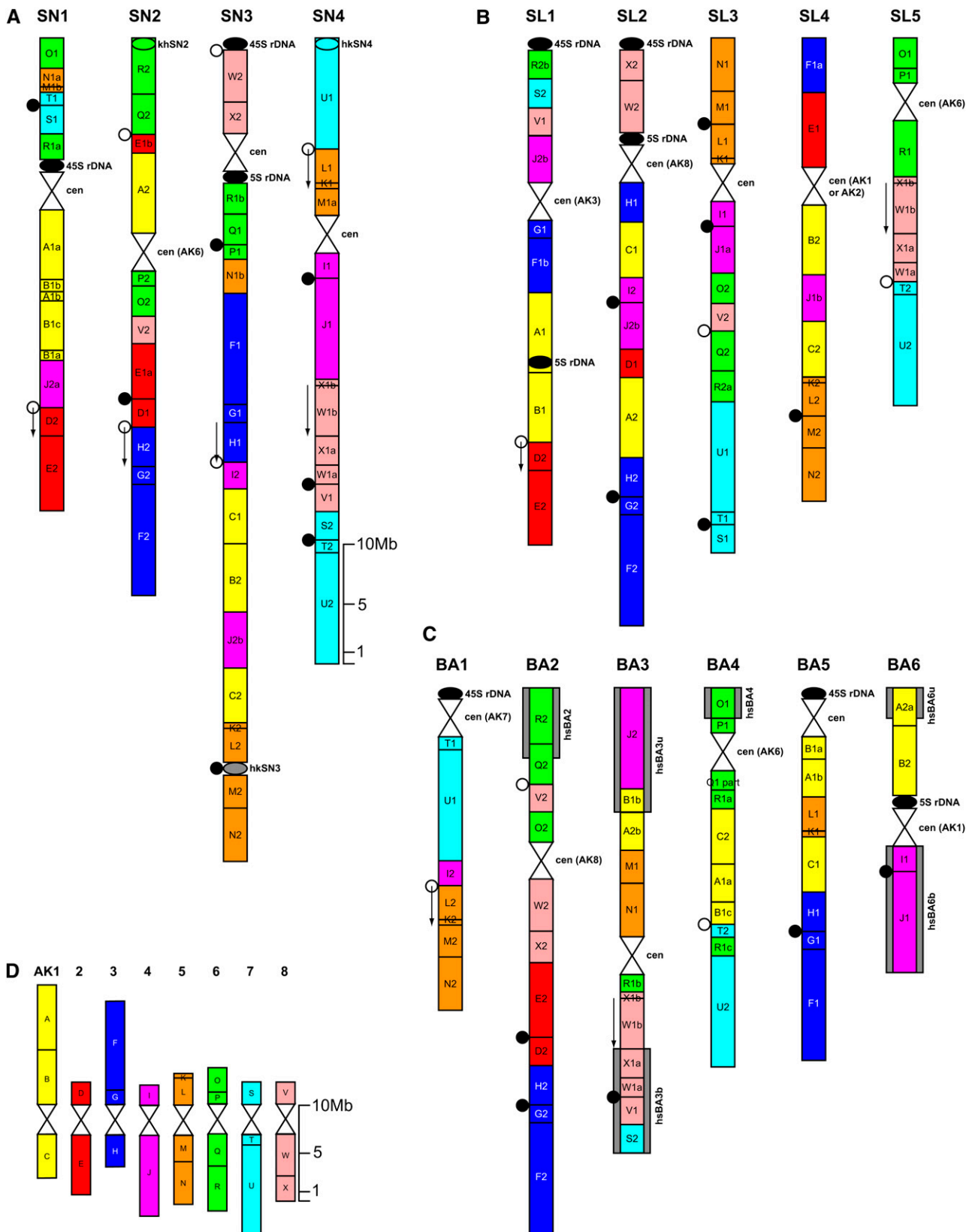
Figure 1. Cytogenetic Maps of *Stenopetalum* and *Ballantinia* Species.

homoeologs [AK4(#1), AK7(#2), and AK8(#1)], whereas the upper arm comprises GBs of two other homoeologs [block U1 of AK7(#1) and K1, L1, and M1 of AK5(#1)] (Figures 1A and 2B).

As duplicated and reshuffled karyotypes differ significantly even between congeneric taxa (*S. nutans* and *S. lineare*, Figure 1), we investigated whether the cytogenetic maps based on a single accession represent species' karyotype. Seven unique GB associations (A-B-D, C-K-L, U-K-L, V-S-T-U, R-S-O, P-O-V, and R-V-C) from all but one chromosome arms within the *S. nutans* #86929 cytogenetic map were tested in another *S. nutans* population (#76272). All GB associations were conserved between the two populations occurring >1000 km apart (see Supplemental Table 4 online). To get further insight into the origin of the revealed genome duplications, the seven unique GB associations of *S. nutans* were analyzed in two *Stenopetalum* species with chromosome numbers doubled as compared with *S. nutans* and *S. lineare*. In *Stenopetalum velutinum* (n = 8) and *Stenopetalum anfractum* (n = 10), all tested GBs were found as four genomic copies. Three and two GB associations in *S. velutinum* and *S. anfractum*, respectively, showed *S. nutans*–like patterns; remaining duplicated GBs had the deviating arrangement.

Altogether, our data indicate that *S. nutans*, *S. lineare*, and *B. antipoda* experienced a common WGD event, followed by reduction of chromosome numbers toward n = 6 to 4. We gained evidence that the common polyploid ancestor originated by a merger of two genomes resembling the ACK. The analysis of *S. velutinum* and *S. anfractum* showed that n = 8 and n = 10 chromosome complements (containing four genomic copies of GBs) are not ancestral but are derived through an additional, more recent WGD.

WGD Was Followed by Chromosome Fusions Accompanied by Genome Reshuffling and Loss or Inactivation of Centromere

Considering the low, quasidiploid chromosome numbers (n = 4 to 6) of the analyzed Australian species, it is obvious that the assumed WGD event has been followed by species-specific chromosome fusions. Chromosome number reduction was accompanied by numerous intra- and intergenomic chromosome rearrangements reshuffling ancestral genomic blocks (Figures 1 and 3; see Supplemental Table 2 online for details). In all three species, 60 to 64% of rearrangement breakpoints coincide with centromeric and terminal chromosome regions, whereas 28 to 33% of breakpoints occurred within genomic blocks splitting them into two (a and b) or three sub-blocks (a, b, and c). In *S. nutans*, out of 48 GBs, 39 (81%) blocks remained intact, whereas

nine (19%) GBs (A1, B1, E1, J2, N1, M1, R1, W1, and X1) were split. In *S. lineare*, out of 46 GBs, 40 (87%) were conserved and six (13%) blocks (F1, J1 and J2, R2, W1, and X1) were partitioned. In *B. antipoda*, out of 44 GBs, 38 (86%) blocks remained intact and six (14%) (A1 and A2, B1, R1, W1, and X1) were split into sub-blocks.

The AK8(#1) homoeolog exhibits a specific rearrangement of blocks W1 and X1 shared by all three taxa (Figures 1A to 1C; see Supplemental Figure 2 and Supplemental Table 2 online). This reshuffling of the AK8-like homoeolog has been analyzed also in two other Australian crucifers, *Arabidella eremigena* (n = 5) and *Blennodia canescens* (n = 7). Although these two species have pachytene complements not amenable to comprehensive cytogenetic analysis, we could show that they also possess the rearrangement and four other GBs duplicated. These data suggest that the Australian crucifers share a common polyploid ancestor, but with most genomic rearrangements being species specific.

Comparing the position and orientation (collinearity) of ancestral genomic blocks between the ACK and genomes of Australian species, we attempted to reconstruct the fate of centromeres in *Stenopetalum* and *Ballantinia* species (Figure 1). Of the 15 active centromeres in *S. nutans*, *S. lineare*, and *B. antipoda*, we could trace the putative origin of nine centromeres. Besides functional centromeres, we tentatively identified in total 12 positions of ancestral centromeres that were most likely lost via a series of pericentric inversions followed by unequal reciprocal translocations (white circles in Figures 1A to 1C). Furthermore, 18 ancestral GB associations containing an inactivated and/or lost ancestral centromere have been discerned (black circles in Figures 1A to 1C). With the exception of the heterochromatic knob hkSN3 on the bottom arm of SN3, whose position corresponds to the AK5 ancestral centromere (Figures 1A and 2C), GBs bordering the ancestral centromeres in the ACK were adjacent to each other, without discernible heterochromatic domains resembling centromeres (Figures 1A to 1C, 2B, and 2C).

Phylogenetic Analysis Corroborates the WGD in Australian Crucifers

To elucidate their polyploid origin and phylogenetic relationship, we sequenced *S. nutans*, *S. velutinum*, *B. antipoda*, and *A. eremigena* for three nuclear genes (chalcone synthase [*CHS*], malate synthase [*MS*], and cinnamyl alcohol dehydrogenase 5 [*CAD5*]; GenBank accession numbers in Supplemental Table 5 online) and four maternally inherited genes (chloroplast genes *rbcL*, *matK*, and *ndhF* and the mitochondrial *nad4* intron 1;

Figure 1. (continued).

S. nutans **(A)**, *S. lineare* **(B)**, *B. antipoda* **(C)**, and ACK **(D)** as a reference genome. The size of GBs A to X corresponds to the size of homoeologous blocks in *Arabidopsis* (http://www. Arabidopsis.org). Color coding of GBs A to X reflects their position on the eight ancestral chromosomes AK1 to AK8 of the ACK. Downward-pointing arrows denote the inverse orientation of GBs compared with their position in the ACK. Active centromeres are depicted as white double-triangle structures, and their presumed origins from AK chromosomes are given. White and black circles left of the chromosomes refer to ancestral centromeres presumably lost by a pericentric inversion followed by reciprocal translocation (white circles) or to centromere inactivation and/or loss (end-to-end fusion; black circles). rDNA loci shown as black ovals, heterochromatic knobs and segments shown as gray ovals and rectangles, respectively, and within the painted GBs in *S. nutans*.
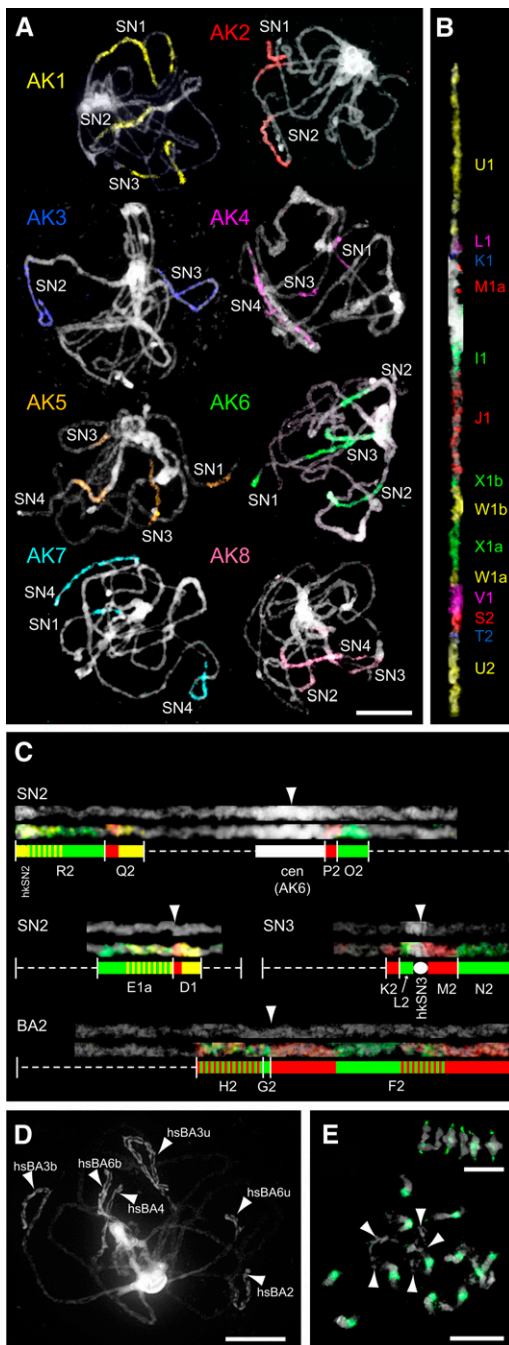
GenBank accession numbers in Supplemental Table 6 online). These genes were analyzed along with representatives from the three major crucifer lineages, with a particular focus on Lineage I as some of the analyzed species were tentatively assigned to the tribe Camelineae (Al-Shehbaz et al., 2006; see Supplemental Text 1 online).

The nuclear phylogenies (Figure 4; see Supplemental Figure 3 online) showed evidence of two different gene copies in the genome of the Australian species. The Camelineae gene copy (a1) is nested in the tribe Camelineae with affinities to the Boechereae and with close relationship with one genome copy of allopolyploid *Pachycladon* species endemic to New Zealand. The basal copy (a2) branches at the base of the Camelineae, close to Lepidieae, and with close affinities to the other genome copy of *Pachycladon* (see Supplemental Text 1 online for more details). The position of three paralogous copies in *S. velutinum* is congruent with the described pattern. Together, the presence of two or three gene copies in Australian species for most nuclear genes and their consistent phylogenetic position among trees suggest an allopolyploid origin of these species. The maternal phylogeny (see Supplemental Figure 4 online) agrees well with the nuclear gene trees and identifies the maternal parental genome as being closely related to New Zealand (*Pachycladon*) and Eurasian Camelineae taxa as well as to North American tribes Boechereae and Halimolobeae of Lineage I (the genome copy a1 in Figure 4).

Based on CCP data, we expected to find two copies of each nuclear gene in *S. nutans*, *B. antipoda*, and *A. eremigena* and four copies in the neopolyploid *S. velutinum*. Concordant with these expectations, the nuclear phylogenies (Figure 4; see Supplemental Figure 3 online) suggest that *B. antipoda* and *A. eremigena* possess two gene copies of *CHS* and *CAD5* and one copy of *MS*, whereas *S. nutans* has two copies of *CAD5* and one of *CHS* and *MS*. The presence of a single copy for some genes in the mesotetraploid species could imply a gene loss following the WGD, although it is not possible to rule out the possibility that these copies are present but that they were not sampled by the cloning procedure. In *S. velutinum*, three gene copies of *CHS* and *CAD5* and one *MS* copy were identified. The presence of three gene copies for *CHS* and *CAD5* is consistent with two WGD events. The presence of three gene copies instead of four could be explained by a loss of the fourth copy after the second WGD or

---

Figure 2. CCP in *Stenopetalum* and *Ballantinia* Species.

(A) Examples of CCP in *S. nutans*. The eight ancestral chromosomes AK1 to AK8 represented by corresponding *Arabidopsis* BAC contigs revealed homoeologous chromosome regions in *Stenopetalum* pachytene complements.

(B) Straightened pachytene *S. nutans* chromosome SN4 painted using differentially labeled *Arabidopsis* BAC contigs corresponding to 14 complete or partial genomic blocks of the ACK (Figure 1A).

(C) Ancestral centromeres in *S. nutans* and *B. antipoda* analyzed by CCP and straightened. Active and inactive ancestral centromere on chromosome SN2, inactive centromere on SN3 coinciding with the heterochro-

matic knob hkSN3, and inactive centromere on BA2 are shown. Genomic blocks of the four ancestral chromosomes are labeled, and the position of active/inactive ancestral centromere is indicated by arrowheads (see Figures 1A and 1C for further details).

(D) Six heterochromatic segments in *B. antipoda* visible as condensed 4',6-diamidino-2-phenylindole–stained structures at distal ends of pachytene chromosomes.

(E) In situ localization of the *Arabidopsis*-like telomere repeat (green) on meiotic (metaphase I, top) and mitotic chromosomes of *B. antipoda*. Some heterochromatic segments indicated by arrowheads. The telomere repeat showed preferential hybridization to centromeres; minor signals at chromosome termini are less prominent. Chromosomes counterstained by 4',6-diamidino-2-phenylindole.
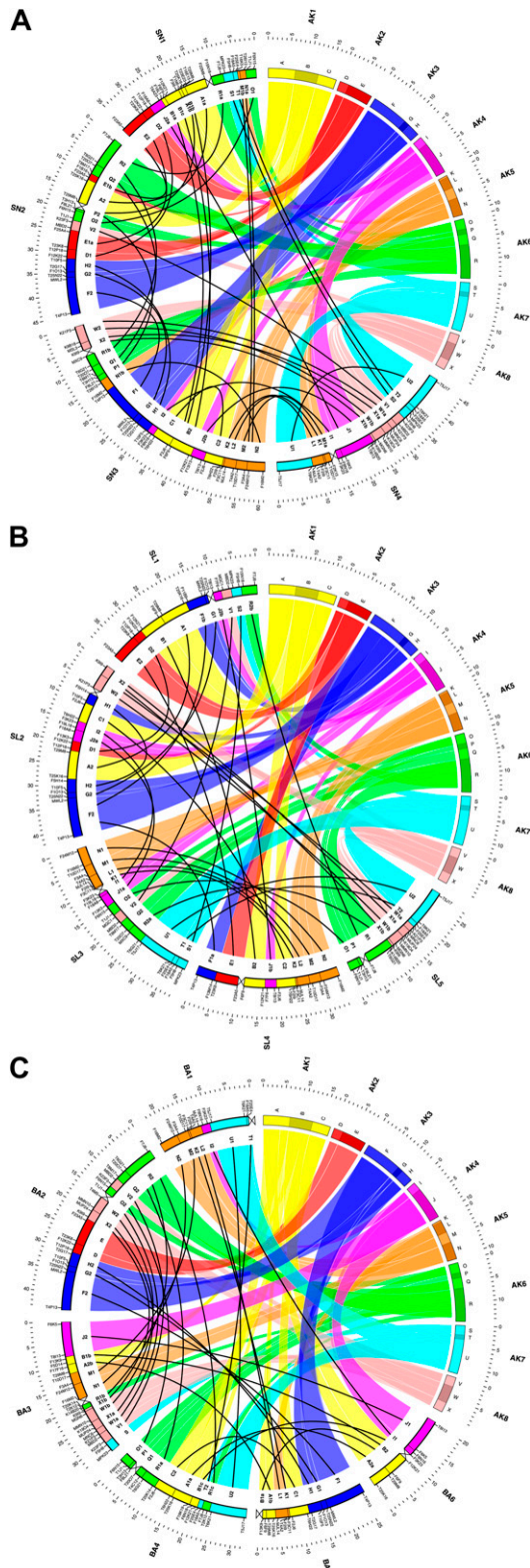
Bars = 10 μm.

Figure 3. Collinear Relationships between ACK and Modern Karyotypes of *S. nutans*, *S. lineare*, and *B. antipoda*, Based on the Cytogenetic Maps in Figure 1.

one mesotetraploid with two copies and another with one copy could be in involved in the second WGD. Also, the fourth gene copy could be present but may not have been sampled. Overall, these observations corroborate the assumed mesotetraploid (*S. nutans*, *B. antipoda*, and *A. eremigena*) and neo-octoploid (*S. velutinum*) status of the analyzed species.

Hypothesis Testing and Divergence Time Estimates

We tested different evolutionary hypotheses regarding the possible number of the allopolyploidization event(s) at the origin of the Australian species using Bayes Factors. The analyses were performed with a particular emphasis on *Pachycladon* species (Camelineae; Al-Shehbaz et al., 2006), which were also shown to have an allopolyploid origin (Joly et al., 2009).

The Bayes factor analyses could not reject the hypothesis that the Australian species and *Pachycladon* were formed from a single WDG event (see Supplemental Table 3 online). Although further data could eventually support more than one WDG event, we assumed a single origin when estimating the divergence times.

The number of synonymous mutations in the nuclear genes suggests that the WGD at the origin of the Australian polyploids occurred ~5.9 mya (median range 3.7 to 8.7; see Supplemental Figure 5 online). This estimate is similar to that obtained by a relaxed clock phylogenetic analysis of the maternal data sets (8.5 mya, 95% credible interval: 4.7 to 12.3 mya; see Supplemental Figure 4 online).

Karyotype Analysis of Purported Parental Species of Australian Mesopolyploids

Based on our phylogenetic hypotheses, comparative karyotypes of *Crucihimalaya wallichii* (*n* = 8) and *Transberingia bursifolia* (*n* = 8) were reconstructed (see Supplemental Figure 6 online). Both Camelineae genera were identified as harboring putative parental genomes, being closely associated with one paralogous gene copy of the Australian and *Pachycladon* species (Figure 4; see Supplemental Figures 3 and 4 online; Joly et al., 2009). CCP analysis showed that chromosome complements of both species are almost completely collinear with the ACK (Figure 1D). Both species differ from the ACK by a large pericentric inversion on the AK1 homoeolog and by another smaller pericentric inversion, specific for each species. Although neither *Crucihimalaya* nor *Transberingia* species can be considered as direct

---

S. nutans **(A)**, *S. lineare* **(B)**, and *B. antipoda* **(C)**. In each panel, ancestral and modern chromosomes are organized circularly as ideograms annotated with their names and megabase scales. The relative scale for ACK and modern karyotype has been adjusted so that ACK ideograms occupy one-third of the figure. Ancestral chromosomes (AK1 to AK8) are further subdivided into 24 (A to X) GBs. Syntenic relationships between regions of ancestral and modern chromosomes are identified by connecting ribbons, whose colors correspond to the ancestral chromosomes. Labels inner to the modern chromosome ideograms correspond to the GBs identified in Figure 1. Outer labels identify *Arabidopsis* BAC clones that demarcate the boundaries of the GBs (see Supplemental Table 1 online). Duplications within the modern karyotypes are shown by thin black lines connecting the duplicated regions.
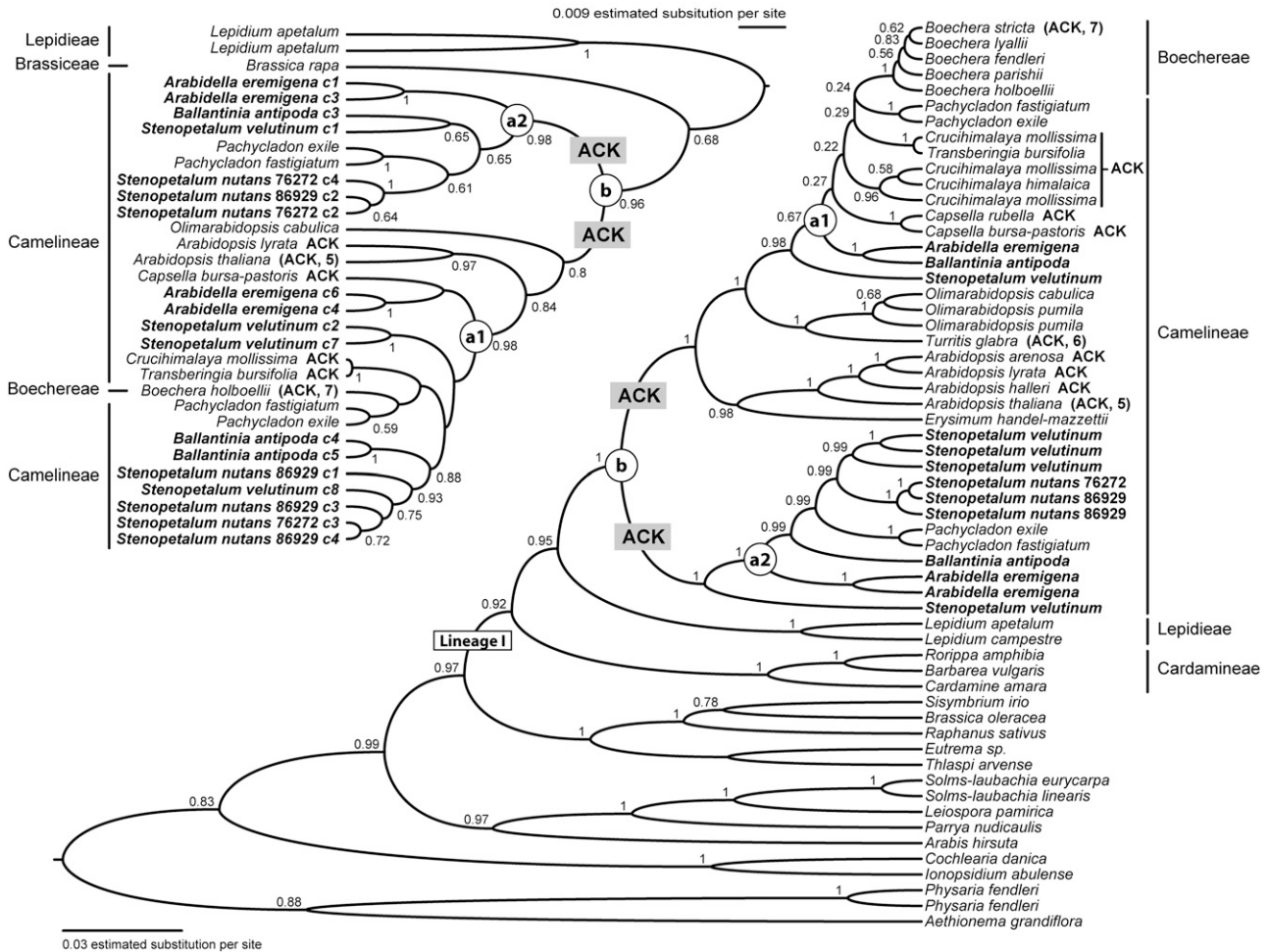
Figure 4. Phylogenies of the Nuclear Genes *CAD5* (TrN + Γ) and *CHS* (GTR + Γ + I) Showing the Position of Sequences from the Australian Species (in Bold) in the Context of Other Brassicaceae Taxa.

Clade posterior probabilities < 0.5 are not shown. "a1" and "a2" identify the two genome copies present in the genomes of Australian Brassicaceae species, whereas "b" indicates the most recent common ancestor for the two genome copies. Species within the crucifer Lineage I with karyotype resembling the ACK or karyotype derived from the ACK (in parenthesis with the extant chromosome number) are mapped onto the phylogenies (Lysak et al., 2006; Schranz et al., 2007; Roosens et al., 2008; see Supplemental Figure 6 online for karyotypes of *T. bursifolia* and of *C. wallichii* as a representative of the genus *Crucihimalaya*). Tribal classification follows Al-Shehbaz et al. (2006).

ancestors of the Australian crucifers based on the CCP data, these results further corroborate the ACK as a presumable ancestral genome of all Camelineae species.

DISCUSSION

Mesopolyploid WGD

CCP analysis with genomic blocks of the Ancestral Crucifer Karyotype (*n* = 8) as probes revealed that all (*S. nutans*) or most (*B. antipoda* and *S. lineare*) blocks exist as two paralogous copies in genomes of the three Australian crucifers. This is

compelling evidence that the Australian species, despite being genetically diploid and having low chromosome numbers (*n* = 4, 5, and 6) experienced a WGD event postdating the three (α, β, and γ) paleopolyploid WGDs in the ancestry of Brassicaceae (Bowers et al., 2003; Barker et al., 2009). We showed that multiple paleopolyploid duplications can be followed by evolutionary younger WGD event(s) masked due to massive genome repatterning and descending dysploidy. As shown for the Australian cruciferous species, even annual, self-compatible, bona fide diploid species with low chromosome number (*n* ≤ 6) might have experienced several rounds of hidden polyploidy. Hence, diploid-like chromosome number cannot be taken as a reliable indicator of true genome diploidy if, for example, the polyploid

incidence is appraised (e.g., Wood et al., 2009). Conversely, polyploid-like increases of chromosome number can be caused by serial centric fissions as reported for some orchids (Leitch et al., 2009) and cycads (Caputo et al., 1996).

Considering the age of WGDs, polyploid species are traditionally classified as paleopolyploid and neopolyploid. Recent polyploids with the increased genome size, chromosome number, gene copy number, and extant diploid ancestors are described as neopolyploids. As paleopolyploids were traditionally described, polyploid relics with high polyploid chromosome numbers and extinct diploid ancestors (Guerra, 2008). This classification has been refined by Favarger (1961) who in addition to paleopolyploids and neopolyploids recognized also mesopolyploid species of an intermediate age and with diploid ancestors in the same or closely related genus. Within the last decade, the definition of paleopolyploidy has changed and the term is currently used for genetically diploid species that experienced one or more rounds of ancient WGD and their polyploid past is becoming apparent only after a careful sequence analysis (e.g., Jaillon et al., 2009; Soltis et al., 2009). To differentiate between these differently aged WGD events, we resurrect and modify the polyploidy classification of Favarger (1961). Figure 5 displays a model of multiple waves of WGD followed by genome diploidization and their impact on the extant genome structure detected through comparative genetic mapping and cytogenetic techniques (CCP in particular). Mesopolyploid species exhibit diploid-like meiosis, disomic inheritance, and diploidized genomes up to quasidiploid complements with a very low number of chromosomes; however, the parental subgenomes are still discernible by comparative (cyto)genetic and phylogenetic methods. In paleopolyploids, the long-term amalgamation of parental genomes is hampering their identification by these methods, and ancient paleopolyploid WGD events can only be uncovered by comparsion of orthologous sequences (Figure 5).

The unexpectedly detected WGD shared by the diploid Australian crucifer species is analogous to the mesopolyploid whole-genome triplication discerned in the tribe Brassiceae by comparative genetic mapping (Lagercrantz and Lydiate, 1996; Parkin et al., 2005) and by comparative cytogenetic analysis (Lysak et al., 2005, 2007; Ziolkowski et al., 2006). Several lineage-specific WGD events thus far documented across the angiosperms (Tang et al., 2008; Jaillon et al., 2009; Soltis et al., 2009) represent only the tip of the iceberg, as potential other WGD events remain uncovered. We envisage that more mesopolyploid WGDs will be revealed in Brassicaceae and other families, such as Asteraceae, where differently aged polyploidization events (Barker et al., 2008) and descending dysploidy have a prominent role in the genome evolution.

A Common Allopolyploid Ancestor?

Associations of genomic blocks corresponding to seven, five, and six AK-like chromosomes and to 11, 15, and 72 AK-like whole-arms preserved in the karyotypes of S. nutans, S. lineare, and B. antipoda, respectively, strongly suggest the ACK ($n = 8$) as an ancestral genome of the analyzed species. Phylogenetic reconstructions showed that both ACK-like parental genomes (i.e., paralogous gene copies a1 and a2) of the Australian

mesopolyploids are nested within the paraphyletic tribe Camelineae and the closely related Boechereae. It was shown that karyotypes of all so far analyzed taxa from three tribes of Lineage I (Camelineae, Boechereae, and Descurainieae) resemble or were derived from the ACK (Figure 4; Lysak et al., 2006 and references therein; Schranz et al., 2007; Roosens et al., 2008). Here, we showed that species of two other Camelineae genera (*Crucihimalaya* and *Transberingia*), considered as potential parental genomes of the Australian species, descended from the ACK. The identification of the ACK as an ancestral karyotype of the closest relatives of the Australian Camelineae species further corroborate our conclusion that these allopolyploids originated from hybridization between taxa with ACK-derived complements. Although a recurrent origin of allopolyploid species was reported for *Arabidopsis kamchatica* (Shimizu-Inatsugi et al., 2009), *Helianthus* (Schwarzbach and Rieseberg, 2002), *Tragopogon* (Lim et al., 2008), *Persicaria* (Kim et al., 2008), and other genera, we are unable to find clear evidence for recurrent polyploid formation from our data. On the contrary, two paracentric inversions on one of the AK8-like homoeologs shared by five analyzed species (see Supplemental Figure 2 online) as well as the estimated large evolutionary distance between the parental genomes ($\sim$4 million years old at the time of the allopolyploidization; see Supplemental Figure 5 online) argue in favor of a single origin.

The fact that any two paralogous copies of GBs differed consistently in size (and fluorescence intensity) could imply that the WGD was an allopolyploidization event. Alternatively, the deviating size and fluorescence intensity of the two subgenomes could be caused by a preferential fractionation (i.e., differentiation) of paralogous sequences belonging to only one subgenome. Biased fractionation has been observed in *Arabidopsis* after the last (α) WGD event, targeting preferentially only one of the two paralogs and retaining mainly dose-sensitive regulatory (connected) genes (Thomas et al., 2006). Such patterns support the gene balance hypothesis (e.g., Birchler and Veitia, 2007; Freeling, 2008; Veitia et al., 2008). As the paralogous gene copies cluster with two distinct species groups within the nuclear gene phylogenies, we infer an allopolyploid origin of the Australian species. The fact that the other sequenced Brassicaceae species did not exhibit multiple gene copies makes the hypotheses that this pattern is caused by gene duplication or lineage sorting untenable. Analysis of maternally inherited genes showed that the Camelineae copy (a1) has been contributed by the maternal parent, whereas the basal copy (a2) is derived from the paternal ancestor.

Evolutionary Scenario of the Mesopolyploid WGD Event

As the structure of all analyzed karyotypes is complex, only a tentative scenario of this polyploid event can be proposed. The simplest model assumes a merger of two ACK-like genomes $\sim$6 to 9 mya followed by a reduction from $n = 16$ toward $n = 6$ to 4. However, the two parental genomes while having the ACK-like structure could already have deviated by chromosome number. Chromosome number reductions from $n = 8$ to $n = 7$ to 4 occurred repeatedly in Brassicaceae (Lysak et al., 2006; Schranz et al., 2007; Mandáková and Lysak, 2008; this study).
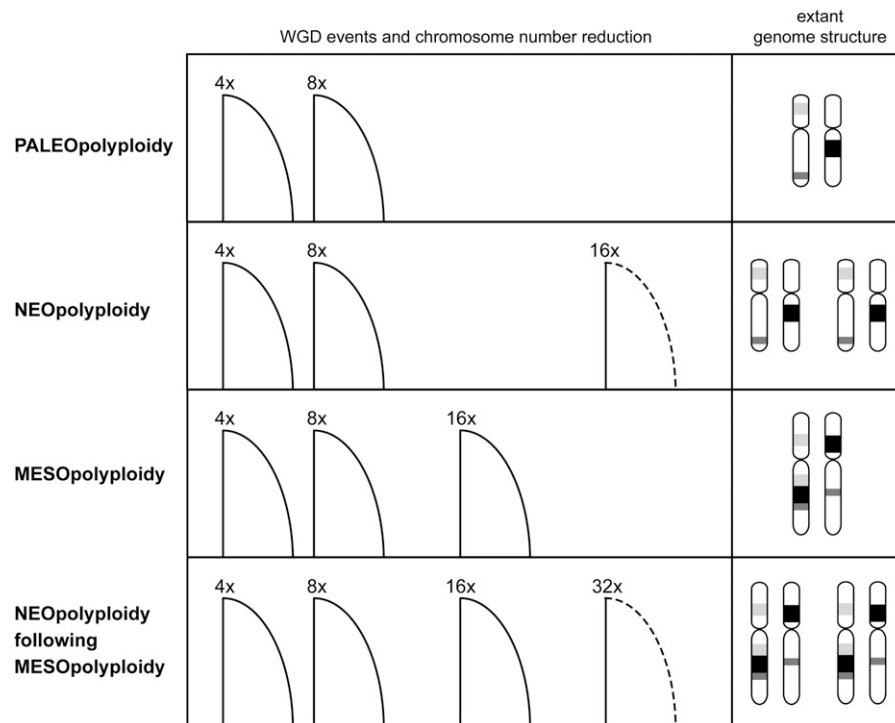
Figure 5. A Model of Genome Evolution through Multiple WGD Events Followed by Diploidization.

Subsequent WGD events of different age shown as ploidy level increases (4x to 32x) are followed by genome diploidization associated with descending dysploidy toward quasidiploid chromosome complements. Diploidization expected to occur in neopolyploids is shown by dashed lines. Chromosomes on the right represent duplicated and diploidized complements with three ancestral GBs as revealed by comparative genetic and cytogenetic techniques. From the top to bottom: Paleopolypoidy is represented by two WGD events, paralogous regions not detectable by (cyto)genetic analysis (example: *A. thaliana*, *n* = 5). Neopolyploidy: two paleopolyploid WGDs followed by a recent neopolyploid event; chromosome number unreduced and most duplicated GBs not yet reshuffled (example: *A. suecica*, *n* = 13). Mesopolyploidy: a WGD event following two paleopolyploid WGDs. Descending dysploidy results in a quasidiploid number of fusion chromosomes; duplicated GBs reshuffled by intra- and intergenomic rearrangements (example: *S. lineare*, *n* = 5). Paleo-, meso-, and neopolyploid WGD events resulting in a duplicated mesopolyploid genome (example: *S. anfractum*, *n* = 10). Possible rearrangements of GBs after the last (neopolyploid) WGD not shown.

As the WGD was detected only in the endemic Australian genera, we believe that it could have taken place in Australia after a long-distance dispersal of the parental genomes from Eurasia or North America. Alternatively, the mesopolyploid event could have occurred in Eurasia/North America and the allopolyploid alone migrated to Australia. Several cases of intercontinental long-distance dispersals were documented (Mummenhoff and Franzke, 2007), including the crucifer genera *Cardamine* (Carlsen et al., 2009) and *Lepidium* (Dierschke et al., 2009). A single or multiple polyploidization event(s) was followed by species radiation and reduction of chromosome number. More recently, reshuffled mesopolyploid genomes were involved in yet another round of auto- or allopolyploidy, as exemplified by four ACK-derived genomes identified in the neopolyploid species *S. anfractum* (*n* = 10) and *S. velutinum* (*n* = 8).

Present phylogenetic analyses uncovered the close relationship between the Australian mesopolyploids and New Zealand *Pachycladon* allopolyploids (Figure 4). Our data strongly advocate *Pachycladon* as a member of the paraphyletic tribe Camelineae as proposed by Al-Shehbaz et al. (2006) and do not support the position of the maternal parent (the *Brassica* gene copy) close to the split of Lineage I and II (Joly et al., 2009). Although it is not possible to reject the null hypothesis of a single allopolyploid origin for *Pachycladon* and the Australian endemics, less diploidized (*n* = 10) *Pachycladon* species appear to originate and radiate on the South Island of New Zealand more recently (1.6 to 0.8 mya; Joly et al., 2009). Further studies of genome evolution within the genus *Pachyladon* are needed to elucidate its origin and relationship to endemic Australian crucifer species.

Species-Specific Dysploidy Following the WGD

Multiple WGDs and subsequent chromosome number reduction resulting in a mosaic of parental genomic blocks are well documented in grasses (Salse et al., 2008; Luo et al., 2009) and in the allopolyploid *B. napus* (Lysak et al., 2005; Parkin et al., 2005; Schranz et al., 2006).

A full reconstruction of modes underlying chromosome number reduction in the three analyzed species is not yet feasible.

The high numbers of evolutionary conserved chromosome arms (Figures 7A to 7C; see Supplemental Table 2 online) suggest that rearrangement breakpoints were predominantly located in centromeric and terminal regions and that whole-arm rearrangements played a prevalent role in the reduction of chromosome number. We attempted to get a deeper insight into karyotype reshuffling by scoring ancestral (AK-like) associations of genomic blocks and their orientation compared with the ACK (ancestral versus inverted) and position of centromeres in the context of GB associations. AK-like association of GBs with one arm in inverted orientation compared with the ACK but without an active centromere was interpreted as the result of chromosome fusion through a pericentric inversion followed by reciprocal translocation involving terminal breakpoints and the loss of one of the two resulting products (the minichromosome). This mechanism has been described for several other crucifer species (Lysak et al., 2006; Schranz et al., 2006; Schubert, 2007; Mandáková and Lysak, 2008). This type of rearrangement accounts for 72 chromosome fusion and centromere loss events in all three karyotypes. We speculate that the *B. antipoda* karyotype, containing two NOR-bearing telocentric chromosomes known only in *Neslia paniculata* (*n* = 7) (Lysak et al., 2006), might represent an evolutionary transient karyotype prone to further dysploidy through translocation-mediated fusions involving the telocentrics.

Seven, six, and five ancestral centromere positions were identified in karyotypes of *S. nutans*, *S. lineare*, and *B. antipoda*, respectively, within entirely or partly preserved ancestral chromosomes (Figures 7A to 7C). It cannot be ruled out that these block associations were generated by the same type of rearrangement as described above but involving another inversion (Lysak et al., 2006; Schubert, 2007). The frequent occurrence of this relatively complicated rearrangement in all three karyotypes is possible since the clustering of breakpoints around centromeres and chromosome termini indicates a preferential involvement of repetitive sequences in erroneous repair of double strand breaks. Alternatively, the occurrence of entire ancestral chromosomes within the fusion chromosomes might result from tandem end-to-end translocation, loss of (sub)telomeric repeat tracts at the breakpoints, and centromere inactivation/loss on one of the chromosomes. A similar telomere-to-telomere translocation has been assumed for an origin of the large metacentric chromosome of the ant *Myrmecia pilosula* (*n* = 1; Imai and Taylor, 1989) and of the human chromosome 2 (Ijdo et al., 1991b). In Brassicaceae, end-to-end chromosome fusions were not described previously (Lysak et al., 2006; Schranz et al., 2006; Schubert, 2007; Mandáková and Lysak, 2008).

Tandem chromosome fusions produce dicentric chromosomes that have to be stabilized by inactivation of one centromere. The centromere inactivation will ensure regular meiotic segregation of the fusion chromosome and can result in evolutionary fixed reduction of chromosome number. Several cases of inactivated ancestral centromeres as well as the emergence of neocentromeres (centromere repositioning) were reported in mammalian species (e.g., Ferreri et al., 2005; Ventura et al., 2007), but only a few examples of centromere inactivation (and reactivation; F. Han et al., 2009) in plants, including dicentric chromosomes of Trititiceae (Sears and Camara, 1952; Luo et al., 2009), maize (*Zea mays*) B chromosomes (F. Han et al., 2006,

2009), and potentially also chromosomes of two cucurbit species (Y. Han et al., 2009). Except for the heterochromatic knob on chromosome SN3 corresponding to the AK5 centromere, no heterochromatin was observed at sites of the other 17 presumably inactivated centromeres. Illegitimate recombination between (peri)centromeric repeats is supposed to gradually remove the repeats and heterochromatin from these regions (Ventura et al., 2004). Similarly, centromere inactivation accompanied by the loss of heterochromatin has been recently envisaged for cucumber (*Cucumis sativus*) chromosome 6 (Y. Han et al., 2009). The knob hkSN3 might represent a recently inactivated ancestral centromere with discernible heterochromatin still present.

Diploidization: Retention versus Loss of Genomic Blocks

As diploidization (fractionation; Thomas et al., 2006) are described processes gradually transforming a polyploid into a quasidiploid genome. In the analyzed mesopolyploid species, diploidization could be documented at the chromosomal level and at the gene level. As all reconstructed chromosomes are complex mosaics of ancestral genomic blocks, we assume that the WGD was followed by lineage(species)-specific reductions of chromosome number accompanied by the extensive restructuring and, to a lesser extent, by the loss of duplicated GBs. Both inter- and intragenomic translocations between nonhomoeologous chromosomes prevail in the three diploidized species. Similar large-scale chromosome reshuffling was reported, for example, in natural and synthetic polyploid *B. napus* (Parkin et al., 2005; Gaeta et al., 2007), *Tragopogon* allopolyploids (Lim et al., 2008), and in fertile Festulolium hybrids (Kopecký et al., 2008).

Considering the different chromosome numbers (*n* = 4, 5, and 6), the most severe chromosome reshuffling and loss of genomic blocks are to be expected in *S. nutans* (*n* = 4), followed by *S. lineare* (*n* = 5) and *B. antipoda* (*n* = 6). However, all 24 duplicated ancestral GBs and seven conserved AK chromosomes were surprisingly found in the most reduced karyotype of *S. nutans*. In *S. lineare*, five AK chromosomes are preserved and two GBs lost (1.3 and 2.6 Mb in the *Arabidopsis* genome), and in *Ballantinia*, six AK chromosomes are conserved and four GBs were completely (1.3, 2.3, 2.4, and 6.2 Mb) or partly (0.9 Mb) lost. In all three species, the high level of retained duplicated blocks (100 to 83%) contrasts with the reduction in chromosome number. Though we do not have a plausible explanation for these findings, the different degree of repatterning argues for independent ways and different speed of descending dysploidy (diploidization), even within a well-defined genus. Future analysis of more endemic Australian species should reveal further patterns of chromosome number reduction.

Comparing the level of genome fractionation found in the three species with another group of polyploids is problematic as the age and character of genome duplication (duplication versus triplication), phylogenetic positions, and chromosome numbers have to be considered. The 6- to 9-million-year-old WGD characterizing the Australian Camelineae species can be compared with the 8- to 15-million-year-old whole-genome triplication in the ancestor of the tribe Brassiceae (Lysak et al., 2005). Although

both mesopolyploid events are presumably of a comparable age, the level of genome redundancy should be different in the duplicated Camelineae versus triplicated Brassiceae genomes. Hence, assuming similar mechanisms and rate of chromosome evolution in both groups, more extensive genome reshuffling is to be expected in Brassiceae. Triplication of 24 ancestral GBs theoretically resulted in 72 GBs comprising quasidiploid Brassiceae genomes. Out of the 24 GBs, only 13 (54%) (Parkin et al., 2005; Schranz et al., 2006) and 14 (58%) (Panjabi et al., 2008) blocks were found as three or more copies in the *rapa* (A) genome of *B. napus* or *B. juncea*. In the *nigra* (B) and *oleracea* (C) genomes, only seven (29%) (Panjabi et al., 2008) and eight (33%) (Kaczmarek et al., 2009) GBs were reported as three or more homoeologous copies in *B. juncea* and *B. oleracea*, respectively. Comparing these data with 83 to 100% of duplicated GBs retained in the Australian species, triplicated Brassiceae genomes show faster diploidization and/or originated earlier (13 to 17 mya suggested by Yang et al., 2006). A loss of paralogous gene copies was documented herein for all three single-copy nuclear genes and all analyzed species. Extensive gene loss was also found in the mesopolyploid genomes of *B. rapa* (Yang et al., 2006) and *B. oleracea* (Town et al., 2006).

Since the discovery of extensive segmental duplications in the *Arabidopsis* genome (Arabidopsis Genome Initiative, 2000), multiple paleo- and mesopolyploid WGD events have been revealed across angiosperms (e.g., Soltis et al., 2009). Our data suggest that ancient paleopolyploid genomes have often undergone more recent, mesopolyploid WGDs, which are discernible by genetic, molecular phylogenetic, and cytogenetic methods.

This CCP analysis showed that the diploid-like genomes ($n = 4$ to 6) of Australian Camelineae species experienced a mesopolyploid WGD blurred by fast, species-specific, and extensive chromosome repatterning. Chromosome numbers become reduced and the polyploid character of the corresponding genomes masked through inversions followed by reciprocal translocations with breakpoints in pericentric and terminal repeat-rich regions and by reciprocal translocations within terminal repeat arrays with a subsequent centromere inactivation/loss. Present results lay a foundation for detailed whole-genome sequence analyses of the ongoing diploidization processes in this crucifer group.

METHODS

Plant Material and Chromosome Preparation

For the origin of the analyzed species accessions, see Supplemental Table 4 online. Herbarium vouchers were deposited in the herbarium of the Masaryk University. Inflorescences of the analyzed accessions were fixed in ethanol:acetic acid (3:1) overnight and stored in 70% ethanol at −20°C. Selected inflorescences were rinsed in distilled water and in citrate buffer (10 mM sodium citrate, pH 4.8; 2 × 5 min) and incubated in an enzyme mix (0.3% cellulase, cytohelicase, and pectolyase; all Sigma-Aldrich) in citrate buffer at 37°C for 3 to 6 h. Individual flower buds were disintegrated on a microscope slide in a drop of citrate buffer and 15 to 30 μL of 60% acetic acid. The suspension was spread on a hot plate at 50°C for 0.5 to 2 min. Chromosomes were fixed by adding 100 μL of ethanol:acetic acid (3:1) fixative. The slide was dried with a hair dryer, postfixed in 4% formaldehyde dissolved in distilled water for 10 min, and air-dried.

Prior to fluorescence in situ hybridization, ready-to-use slides were treated with pepsin (0.1 mg/mL; Sigma-Aldrich) in 0.01 M HCl for 3 to 6 h, postfixed in 4% formaldehyde in 2× SSC (1× SSC: 0.15 M NaCl and 0.015 M sodium citrate) for 10 min, and dehydrated in an ethanol series (70, 80, and 96%).

Fluorescence in Situ Hybridization and CCP

For localization of 5S and 45S rDNA loci, clone pCT 4.2 corresponding to 500-bp 5S rRNA repeat (M65137) and *Arabidopsis thaliana* BAC clone T15P10 (AF167571) were used, respectively. Telomeric probe was prepared according to Ijdo et al. (1991a). A total of 547 *Arabidopsis* BAC clones were assembled to represent genomic blocks of the ACK (Schranz et al., 2006). To further characterize the orientation and internal structure of particular genomic blocks, the respective BAC contigs were arbitrarily broken into subcontigs and differentially labeled. Supplemental Table 1 online shows the structure and position of ancestral genomic blocks in the karyotypes of *Stenopetalum nutans*, *Stenopetalum lineare*, and *Ballantinia antipoda* as well as corresponding *Arabidopsis* BAC contigs. DNA probes were labeled with biotin-dUTP, Cy3-dUTP, digoxigenin-dUTP, and with diethylaminocoumarin (DEAC)-dUTP and Alexa Fluor 488-dUTP (*S. nutans* chromosome SN4; Figure 2B) by nick translation and ethanol precipitated (Mandáková and Lysak, 2008). The hybridization probes and chromosomes were denatured together on a hot plate at 80°C for 2 min and incubated in a moist chamber at 37°C overnight or for 48 h.

Posthybridization washing was performed in 50% formamide (rDNA and telomere probes) or 20% formamide (BAC contigs) in 2× SSC at 42°C. Biotin-dUTP–labeled probes were detected by avidin~Texas Red (Vector Laboratories), goat anti-avidin~biotin (Vector Laboratories), and avidin~Texas Red; digoxigenin-dUTP was detected by mouse antidigoxigenin (Jackson Immunoresearch) and goat anti-mouse~Alexa Fluor 488 (Molecular Probes). Chromosome SN4 was painted using *Arabidopsis* BAC clones labeled with Cy3-dUTP, DEAC-dUTP, and Alexa Fluor 488-dUTP, and biotin-dUTP detected by avidin~Texas Red and digoxigenin-dUTP detected by mouse antidigoxigenin and goat anti-mouse conjugated with Alexa Fluor 647. Chromosomes were counterstained with 4',6-diamidino-2-phenylindole (2 mg/mL) in Vectashield (Vector Laboratories).

Fluorescence signals were analyzed with an Olympus BX-61 epifluorescence microscope and AxioCam CCD camera (Zeiss). The monochromatic images were pseudocolored and merged using Adobe Photoshop CS2 software (Adobe Systems). Pachytene chromosomes were straightened using the "straighten-curved-objects" plugin in the Image J software (Kocsis et al., 1991). Circular visualization of ancestral genomic blocks in the ACK and genomes of the Australian crucifers (Figure 3; see Supplemental Figure 1 online) was prepared using Circos (Krzywinski et al., 2009). Comparison of the position for each genomic block and each pair of the analyzed species (see Supplemental Figure 1 online) was created by a customized version of Circos, provided by M.K.

Phylogeny Reconstruction

We sequenced *Arabidella eremigena*, *B. antipoda*, *S. nutans*, and *Stenopetalum velutinum* for three nuclear and four maternal genes, respectively. Because of the limited supply of material, we were unable to include three species (partly) analyzed by CCP (*Blennodia canescens*, *Stenopetalum anfractum*, and *S. lineare*). DNA extractions, PCR amplifications, cloning, and sequencing followed standard procedures (Franzke et al., 2009; Joly et al., 2009). Clone sequences that had >99% similarity with each others were assumed to come from a single gene copy, and only one representative clone was retained for the analyses (the sequence showing the fewer autapomorphies). By doing so, we hoped to minimize the impact of allelic variation and PCR-induced mutations on the results. Nuclear introns were removed because they were too divergent across the Brassicaceae. Data sets were aligned with

MUSCLE (Edgar, 2004), and substitution models were selected with the AIC in Modeltest 3.7 (Posada and Crandall, 1998) on a GTR+Γ+I phyml tree (see Supplemental Data Sets 1, 2, and 3 online). Phylogenies were estimated in BEAST 1.4.8 (Drummond and Rambaut, 2007) using the best fit substitution model, a lognormal relaxed molecular clock, and a birth and death model. The root of the phylogeny was estimated using a relaxed molecular clock model as part of the phylogenetic analysis. For the maternal data set, a partitioned analysis was performed where each gene had its own substitution model. Two independent MCMC chains were run for 10^8 generations (2×10^8 for *CHS*), sampling trees every 5000 generations and discarding 10% of the samples as burnin for the Bayes factor analyses and for building the maximum sum of clade credibility tree (the chains always reached their stationary phase well before this point). For the *CHS* and the maternal data set, a weight of 50 was given to the operators internalNodeHeights, subtreeSlide, and narrowExchange and 40 for wideExchange and wilsonBalding. Convergence was always reached for all parameters among independent runs and estimated sample sizes were always >>100. Evolutionary hypotheses were tested using Bayes Factors (following 8) by performing BEAST analyses with appropriate monophyly constraints. Uncertainty around the log likelihood harmonic means was assessed with 1000 bootstrap replicates.

Divergence Time Estimates

For the nuclear genes, minimal age estimates for the WGD was approximated by estimating the age of the most recent common ancestor (MRCA) of all Australian species and *Pachycladon* for each gene paralog. This was done by comparing the divergence between sequences that branched back to these MRCA. We assumed a molecular clock with a mutation rate of 1.5×10^{-8} synonymous (syn.) substitutions per syn. site (d_S) per year (Koch et al., 2000). d_S was estimated following Goldman and Yang (1994) in paml (Yang, 1997). Divergence times (T) were obtained given that syn. substitution rate = $d_S/2T$.

Because syn. substitution rates for chloroplast genes are not well documented in Brassicaceae, divergence times were estimated using soft calibration points within the BEAST phylogenetic analysis (see Supplemental Figure 4 online). Three calibration points were used: (1) a normal prior of mean 89.5 and SD of 1 was given to the crown Brassicales group to reflect a Turonian *Dressiantha* fossil ($\approx$89.5 mya; Gandolfo et al., 1998); (2) a normal prior with a mean 35 and SD 6 was enforced to the crown Brassicaceae node to reflect that the oldest reliable Brassicaceae fossil occurs in Oligocene deposits (22 to 34 mya; Cronquist, 1981); (3) a normal prior of mean 6 and SD 2 was given to the MRCA of *Rorippa* and its closest relative to reflect *Rorippa* fossils in Pliocene deposits (2 to 5 mya; Mai, 1995).

Accession Numbers

Sequence data from this article can be found in the GenBank data libraries under accession numbers GQ926501 to GQ926568 (see Supplemental Tables 5 and 6 online).

Supplemental Data

The following materials are available in the online version of this article.

Supplemental Figure 1. A Three-Way Comparison of the Relative Position of Corresponding Synteny Blocks of *Stenopetalum nutans* (SN), *S. lineare* (SL), and *Ballantinia antipoda* (BA) Relative to the Reference Ancestral Crucifer Karyotype (ACK).

Supplemental Figure 2. The Unique Rearrangement of the AK8(#1)-Like Homoeolog Shared by All Analyzed Species.

Supplemental Figure 3. Phylogeny of the Malate Synthase (*MS*) (TrN + Γ + I) Showing the Position of Sequences from the Australian Species (in Bold) in the Context of Other Brassicaceae Taxa.

Supplemental Figure 4. Maternal Phylogeny (Chronogram) of Australian Genera *Stenopetalum*, *Ballantinia*, and *Arabidella* in the Context of Other Brassicaceae Taxa Resulting from the Partitioned Analysis of the Genes *rbcL* (K81uf + Γ + I), *nad4* (TVMef + Γ + I), *matK* (TVM + Γ + I), and *ndhF* (TVM + Γ + I) in BEAST.

Supplemental Figure 5. Comparison of the Estimated Time of Original Divergence Between the Parents That Were Involved in the Allopolyploid Event (A) and the Estimated Time of the Event (B) Based on the Analysis of Three Nuclear Genes.

Supplemental Figure 6. Reconstructed Karyotypes of *Transberingia bursifolia* ($n = 8$) and *Crucihimalaya wallichii* ($n = 8$).

Supplemental Table 1. Ancestral Genomic Blocks (GB) of the Ancestral Crucifer Karyotype (ACK) Identified on Chromosomes of *Stenopetalum nutans* (chromosomes SN1-SN4), *S. lineare* (SL1-SL5), and *Ballantinia antipoda* (BA1-BA6); Details on Corresponding BAC Contigs of *Arabidopsis thaliana* Are Given.

Supplemental Table 2. Structure and Position of the Eight Ancestral Chromosomes (AK1 to AK8) and 24 Genomic Blocks (GBs) of the Ancestral Crucifer Karyotype ($n = 8$) within Duplicated Genomes of *Stenopetalum* and *Ballantinia* Species.

Supplemental Table 3. Bayes Factor (BF) Scores for Different Evolutionary Hypotheses Regarding the Number of Independent Polyploid Events at the Origin of the Australian Species Relative to the Null Hypothesis of Four Independent Origins.

Supplemental Table 4. Collection Data of Australian Crucifer Species Used in the Present Study.

Supplemental Table 5. GenBank Accession Numbers for Sequences Used in the Phylogenetic Analyses of the Nuclear Genes *CHS*, *CAD5*, and *MS*.

Supplemental Table 6. GenBank Accession Numbers Lineage and Tribal Assignments for Taxa Included in the Maternal Phylogenetic Analysis.

Supplemental Data Set 1. Text File of the *CAD5* Alignment (Figure 4).

Supplemental Data Set 2. Text File of the *CHS* Alignment (Figure 4).

Supplemental Data Set 3. Text File of the *MS* Alignment (Supplemental Figure 3).

Supplemental Data Set 4. Text File of Alignments of Chloroplast Genes *rbcL*, *matK*, and *ndhF* and the Mitochondrial *nad4* Intron 1 (Supplemental Figure 4).

Supplemental Text 1. Phylogenetic Relationships and Comparison with Previous Studies.

REFERENCES

Al-Shehbaz, I.A., Beilstein, M.A., and Kellogg, E.A. (2006). Systematics and phylogeny of the Brassicaceae (Cruciferae): An overview. Plant Syst. Evol. **259:** 89–120.

Arabidopsis Genome Initiative (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature **408:** 796–815.

Barker, M.S., Kane, N.C., Matvienko, M., Kozik, A., Michelmore, R.W., Knap, S.J., and Rieseberg, L.H. (2008). Multiple paleopolyploidizations during the evolution of the Compositae reveal parallel patterns of duplicate gene retention after millions of years. Mol. Biol. Evol. **25:** 2445–2455.

Barker, M.S., Vogel, H., and Schranz, M.E. (2009). Paleopolyploidy in the Brassicales: Analyses of the *Cleome* transcriptome elucidate the history of genome duplications in *Arabidopsis* and other Brassicales. Genome Biol. Evol. **1:** 1–9.

Birchler, J.A., and Veitia, R.A. (2007). The gene balance hypothesis: from classical genetics to modern genomics. Plant Cell **19:** 395–402.

Bowers, J.E., Chapman, B.A., Rong, J., and Paterson, A.H. (2003). Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. Nature **422:** 433–438.

Caputo, P., Cozzolino, S., Gaudio, L., Moretti, A., and Stevenson, D.W. (1996). Karyology and phylogeny of some Mesoamerican species of *Zamia* (Zamiaceae). Am. J. Bot. **83:** 1513–1520.

Carlsen, T., Bleeker, W., Hurka, H., Elven, R., and Brochmann, C. (2009). Biogeography and phylogeny of *Cardamine* (Brassicaceae). PLoS ONE **96:** 215–236.

Cronquist, A. (1981). An Integrated System of Classification of Flowering Plants. (New York: Columbia University Press).

Cui, L., et al. (2006). Widespread genome duplications throughout the history of flowering plants. Genome Res. **16:** 738–739.

De Bodt, S., Maere, S., and Van der Peer, Y. (2005). Genome duplication and the origin of angiosperms. Trends Ecol. Evol. **20:** 591–597.

Dierschke, T., Mandáková, T., Lysak, M.A., and Mummenhoff, K. (2009). A bicontinental origin of polyploid Australian/New Zealand *Lepidium* species (Brassicaceae)? Evidence from genomic *in situ* hybridization. Ann. Bot. (Lond.) **104:** 681–688.

Drummond, A.J., and Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. **7:** 214.

Edgar, R.C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. **32:** 1792–1797.

Edger, P.P., and Pires, J.C. (2009). Gene and genome duplications: The impact of dosage-sensitivity of the fate of nuclear genes. Chromosome Res. **17:** 699–717.

Favarger, C. (1961). Sur l'emploi des nombres chromo-somiques en géographie botanique historique. Ber. Geobot. Inst. Rübel **32:** 119–146.

Fawcett, J.A., Maere, S., and Van de Peer, Y. (2009). Plants with double genomes might have had a better chance to survive the Cretaceous-Tertiary extinction event. Proc. Natl. Acad. Sci. USA **106:** 5737–5742.

Ferreri, G.C., Liscinsky, D.M., Mack, J.A., Eldridge, M.D.B., and O'Neill, R.J. (2005). Retention of latent centromeres in the mammalian genome. J. Hered. **96:** 217–224.

Franzke, A., German, D., Al-Shehbaz, I.A., and Mummenhoff, K. (2009). *Arabidopsis* family ties: Molecular phylogeny and age estimates in the Brassicaceae. Taxon **58:** 425–437.

Freeling, M. (2008). The evolutionary position of subfunctionalization, downgraded. Genome Dyn. **4:** 25–40.

Gaeta, R.T., Pires, J.C., Iniguez-Luy, F., Leon, E., and Osborn, T.C. (2007). Genomic changes in resynthesized *Brassica napus* and their effect on gene expression and phenotype. Plant Cell **19:** 3403–3417.

Gandolfo, M.A., Nixon, K.C., and Crepet, W.L. (1998). A new fossil flower from the Turonian of New Jersey: *Dressiantha bicarpellata* gen. et sp. nov. (Capparales). Am. J. Bot. **85:** 964–974.

Goldman, N., and Yang, Z. (1994). A codon-based model of nucleotide substitution for protein-coding DNA sequences. Mol. Biol. Evol. **11:** 725–736.

Grant, V. (1981). Plant Speciation. (New York: Columbia University Press).

Guerra, M. (2008). Chromosome numbers in plant cytotaxonomy: Concepts and implications. Cytogenet. Genome Res. **120:** 339–350.

Han, F., Gao, Z., and Birchler, J.A. (2009). Reactivation of an inactive centromere reveals epigenetic and structural components for centromere specification in maize. Plant Cell **21:** 1929–1939.

Han, F., Lamb, J.C., and Birchler, J. A. (2006). High frequency of centromere inactivation resulting in stable dicentric chromosomes of maize. Proc. Natl. Acad. Sci. USA **103:** 3238–3243.

Han, Y., Zhang, Z., Liu, C., Liu, J., Huang, S., Jiang, J., and Lin, W. (2009). Centromere repositioning in cucurbit species: Implication of the genomic impact from centromere activation and inactivation. Proc. Natl. Acad. Sci. USA **106:** 14937–14941.

Hegarty, M.J., and Hiscock, S.J. (2008). Genomic clues to the evolutionary success of polyploid plants. Curr. Biol. **18:** 435–444.

Ijdo, J.W., Wells, R.A., Baldini, A., and Reeders, S.T. (1991a). Improved telomere detection using a telomere repeat probe (TTAGGG)n generated by PCR. Nucleic Acids Res. **19:** 4780.

Ijdo, J.W., Wells, R.A.W., Baldini, A., and Reeders, S.T. (1991b). The origin of human chromosome 2: An ancestral telomere-telomere fusion. Proc. Natl. Acad. Sci. USA **88:** 9051–9055.

Imai, H.T., and Taylor, R.W. (1989). Chromosomal polymorphisms involving telomere fusion, centromeric inactivation and centromere shift in the ant *Myrmecia* (*pilosula*) n = 1. Chromosoma **98:** 456–460.

Jaillon, O., Aury, J.M., and Wincker, P. (2009). "Changing by doubling", the impact of whole genome duplications in the evolution of eukaryotes. C. R. Biol. **332:** 241–253.

Joly, S., Heenan, P.B., and Lockhart, P.J. (2009). A Pleistocene intertribal allopolyploidization event precedes the species radiation of *Pachycladon* (Brassicaceae) in New Zealand. Mol. Phylogenet. Evol. **51:** 365–372.

Kim, S., Sultan, S.E., and Donoghue, M.J. (2008). Allopolyploid speciation in *Persicaria* (Polygonaceae): Insights from a low-copy nuclear region. Proc. Natl. Acad. Sci. USA **105:** 12370–12375.

Kaczmarek, M., Koczyk, G., Ziolkowski, P.A., Babula-Skowronska, D., and Sadowski, J. (2009). Comparative analysis of the *Brassica oleracea* genetic map and the *Arabidopsis thaliana* genome. Genome **52:** 620–633.

Koch, M.A., Haubold, B., and Mitchell-Olds, T. (2000). Comparative evolutionary analysis of chalcone synthase and alcohol dehydrogenase loci in *Arabidopsis, Arabis*, and related genera (Brassicaceae). Mol. Biol. Evol. **17:** 1483–1498.

Kocsis, E., Trus, B.L., Steer, C.J., Bisher, M.E., and Steven, A.C. (1991). Image averaging of flexible fibrous macromolecules: The clathrin triskelion has an elastic proximal segment. J. Struct. Biol. **107:** 6–14.

Kopecký, D., Lukaszewski, A.J., and Doležel, J. (2008). Cytogenetics of Festulolium (*Festuca* × *Lolium* hybrids). Cytogenet. Genome Res. **120:** 370–383.

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S.J., and Marra, M.A. (2009). Circos: An information aesthetic for comparative genomics. Genome Res. **19:** 1639–1645.

Lagercrantz, U., and Lydiate, D. (1996). Comparative genome mapping in *Brassica*. Genetics **144:** 1903–1910.

Leitch, I.J., Kahandawala, I., Suda, J., Hanson, L., Ingrouille, M.J., Chase, M.W., and Fay, M.F. (2009). Genome size diversity in orchids: Consequences and evolution. Ann. Bot. (Lond.) **104:** 469–481.

Lim, K.Y., Soltis, D.E., Soltis, P.S., Tate, J., Matyasek, R., Srubarova, H., Kovarik, A., Pires, J.C., Xiong, Z., and Leitch, A.R. (2008). Rapid chromosome evolution in recently formed polyploids in *Tragopogon* (Asteraceae). PLoS ONE **3:** e3353.

Luo, M.C., et al. (2009). Genome comparisons reveal a dominant mechanism of chromosome number reduction in grasses and accelerated genome evolution in Triticeae. Proc. Natl. Acad. Sci. USA **106:** 15780–15785.

Lysak, M.A., Berr, A., Pecinka, A., Schmidt, R., McBreen, K., and Schubert, I. (2006). Mechanisms of chromosome number reduction in *Arabidopsis thaliana* and related *Brassicaceae* species. Proc. Natl. Acad. Sci. USA **103:** 5224–5229.

Lysak, M.A., Cheung, K., Kitschke, M., and Bureš, P. (2007). Ancestral chromosomal blocks are triplicated in Brassiceae species with varying chromosome number and genome size. Plant Physiol. **145:** 402–410.

Lysak, M.A., Koch, M.A., Pecinka, A., and Schubert, I. (2005). Chromosome triplication found across the tribe Brassiceae. Genome Res. **15:** 516–525.

Mai, D.H. (1995). Tertiäre vegetationsgeschichte Europas. (Jena, Stuttgart, New York: Gustav Fisher).

Mandáková, T., and Lysak, M.A. (2008). Chromosomal phylogeny and karyotype evolution in x=7 crucifer species (Brassicaceae). Plant Cell **20:** 2559–2570.

Ming, R., et al. (2008). The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus). Nature **452:** 991–996.

Mummenhoff, K., and Franzke, A. (2007). Gone with the bird: Late Tertiary and Quaternary intercontinental long-distance dispersal and allopolyploidization in plants. Syst. Biodivers. **5:** 255–260.

Panjabi, P., Jagannath, A., Bisht, N.C., Padmaja, K.L., Sharma, S., Gupta, V., Pradhan, A.K., and Pental, D. (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 Genomics **9:** 113.

Parkin, I.A.P., Gulden, S.M., Sharpe, A.G., Lukens, L., Trick, M., Osborn, T.C., and Lydiate, D.J. (2005). Segmental structure of the *Brassica napus* genome based on comparative analysis with *Arabidopsis thaliana*. Genetics **171:** 765–781.

Posada, D., and Crandall, K.A. (1998). Modeltest: Testing the model of DNA substitution. Bioinformatics **14:** 817–818.

Roosens, N.H.C.J., Willems, G., Gode, C., Courseaux, A., and Saumitou-Laprade, P. (2008). The use of comparative genome analysis and syntenic relationships allows extrapolating the position of Zn tolerance QTL regions from *Arabidopsis halleri* into *Arabidopsis thaliana*. Plant Soil **306:** 105–116.

Salse, J., Bolot, S., Throude, M., Jouffe, V., Piegu, B., Quraishi, U.M., Calcagno, T., Cooke, R., Delseny, M., and Feuillet, C. (2008). Identification and characterization of shared duplications between rice and wheat provide new insight into grass genome evolution. Plant Cell **20:** 11–24.

Schlueter, J.A., Dixon, P., Granger, C., Grant, D., Clark, L., Doyle, J.J., and Schoemaker, R.C. (2004). Mining EST databases to resolve evolutionary events in major crop species. Genome **47:** 868–876.

Schranz, M.E., Lysak, M.A., and Mitchell-Olds, T. (2006). The ABC's of comparative genomics in the Brassicaceae: Building blocks of crucifer genomes. Trends Plant Sci. **11:** 535–542.

Schranz, M.E., and Mitchell-Olds, T. (2006). Independent ancient polyploidy events in the sister families Brassicaceae and Cleomaceae. Plant Cell **18:** 1152–1165.

Schranz, M.E., Windsor, A.J., Song, B.-H., Lawton-Rauh, A., and Mitchell-Olds, T. (2007). Comparative genetic mapping in *Boechera stricta*, a close relative of Arabidopsis. Plant Physiol. **144:** 286–298.

Schubert, I. (2007). Chromosome evolution. Curr. Opin. Plant Biol. **10:** 109–115.

Schwarzbach, A.E., and Rieseberg, L.H. (2002). Likely multiple origins of a diploid hybrid sunflower species. Mol. Ecol. **11:** 1703–1715.

Sears, E.R., and Camara, A. (1952). A transmissible dicentric chromosome. Genetics **37:** 125–135.

Sémon, M., and Wolfe, K.H. (2007). Rearrangement rate following the whole-genome duplication in teleosts. Mol. Biol. Evol. **24:** 860–867.

Shimizu-Inatsugi, R., Lihova, J., Iwanaga, H., Kudoh, H., Marhold, K., Savolainen, O., Watanabe, K., Yakubov, V.V., and Shimizu, K.K. (2009). The allopolyploid *Arabidopsis kamchatica* originated from multiple individuals of *Arabidopsis lyrata* and *Arabidopsis halleri*. Mol. Ecol. **18:** 4024–4048.

Soltis, D.E., Albert, V.A., Leebens-Mack, J., Bell, C.D., Paterson, A.H., Zheng, C., Sankoff, D., dePamphilis, C.W., Wall, P.K., and Soltis, P.S. (2009). Polyploidy and angiosperm diversification. Am. J. Bot. **96:** 336–348.

Soltis, D.E., Soltis, P.S., Schemske, D.W., Hancock, J.F., Thompson, J.N., Husband, B.C., and Judd, W.S. (2007). Autopolyploidy in angiosperms: Have we grossly underestimated the number of species? Taxon **56:** 13–30.

Tang, H., Bowers, J.E., Wang, X., Ming, X., Alam, M., and Paterson, A.H. (2008). Synteny and collinearity in plant genomes. Science **320:** 486–488.

Thomas, B.C., Pedersen, B., and Freeling, M. (2006). Following tetraploidy in an *Arabidopsis* ancestor, genes were removed preferentially from one homeolog leaving clusters enriched in dose-sensitive genes. Genome Res. **16:** 934–946.

Town, C.D., et al. (2006). Comparative genomics of *Brassica oleracea* and *Arabidopsis thaliana* reveal gene loss, fragmentation, and dispersal after polyploidy. Plant Cell **18:** 1348–1359.

Veitia, R.A., Bottani, S., and Birchler, J.A. (2008). Cellular reactions to gene dosage imbalance: Genomic, transcriptomic and proteomic effects. Trends Genet. **24:** 390–397.

Ventura, M., Antonacci, F., Cardone, M.F., Stanyon, R., D'Addabbo, P., Cellamare, A., Sprague, L.J., Eichler, E.E., Archidiacono, N., and Rocchi, M. (2007). Evolutionary formation of new centromeres in macaque. Science **316:** 243–246.

Ventura, M., et al. (2004). Recurrent sites for new centromere seeding. Genome Res. **14:** 1696–1703.

Wood, T.E., Takebayashi, N., Barker, M.S., Mayrose, I., Greenspoond, P.B., and Rieseberg, L.H. (2009). The frequency of polyploid speciation in vascular plants. Proc. Natl. Acad. Sci. USA **106:** 13875–13879.

Yang, T.J., et al. (2006). Sequence-level analysis of the diploidization process in the triplicated *FLOWERING LOCUS C* region of *Brassica rapa*. Plant Cell **18:** 1339–1347.

Yang, Z. (1997). PAML: A program package for phylogenetic analysis by maximum likelihood. Comput. Appl. Biosci. **13:** 555–556.

Ziolkowski, P.A., Kaczmarek, M., Babula, D., and Sadowski, J. (2006). Genome evolution in Arabidopsis/*Brassica*: Conservation and divergence of ancient rearranged segments and their breakpoints. Plant J. **47:** 63–74.

Fast Diploidization in Close Mesopolyploid Relatives of Arabidopsis

Terezie Mandáková, Simon Joly, Martin Krzywinski, Klaus Mummenhoff and Martin A. Lysak

This information is current as of January 6, 2011