I. Lysak M.A., **Mandáková T.** 2013. Analysis of plant meiotic chromosomes by chromosome painting. In: Pawlowski W.P, Grelon M. (Eds.). Plant Meiosis: Methods and protocols. Humana Press. NY.

Chapter 2

Analysis of Plant Meiotic Chromosomes
by Chromosome Painting

Martin A. Lysak and Terezie Mandáková

Abstract

Chromosome painting (CP) refers to visualization of large chromosome regions, entire chromosome arms, or entire chromosomes via fluorescence in situ hybridization (FISH). For CP in plants, contigs of chromosome-specific bacterial artificial chromosomes (BAC) from the target species or from a closely related species (comparative chromosome painting, CCP) are typically applied as painting probes. Extended pachytene chromosomes provide the highest resolution of CP in plants. CP enables identification and tracing of particular chromosome regions and/or entire chromosomes throughout all meiotic stages as well as corresponding chromosome territories in premeiotic interphase nuclei. Meiotic pairing and structural chromosome rearrangements (typically inversions and translocations) can be identified by CP. Here, we describe step-by-step protocols of CP and CCP in plant species including chromosome preparation, BAC DNA labeling, and multicolor FISH (Fig. 1).

Keywords Chromosome painting, Fluorescence in situ hybridization, BAC FISH, Pachytene chromosomes, DNA labeling

1   Introduction

The chromosome painting (CP) technique in human and animal cytogenetics serves in situ identification of whole chromosomes and specific chromosome regions using flow sorted, Degenerate Oligonucleotide Primed PCR (DOP-PCR)-amplified, and fluorescently labeled chromosomes or chromosome segments. Due to abundant and diverse DNA repeats homogeneously distributed across chromosome complements (1), flow sorted or micro-dissected chromosomes are not suitable for preparation of chromosome-specific painting probes in plants. Instead, chromosome-specific high-capacity DNA vectors have become widely utilized in plant cytogenetics since the mid-1990. Specifically, individual bacterial artificial chromosome (BAC) clones and BAC contigs (continuous sets of BAC clones) are most frequently used
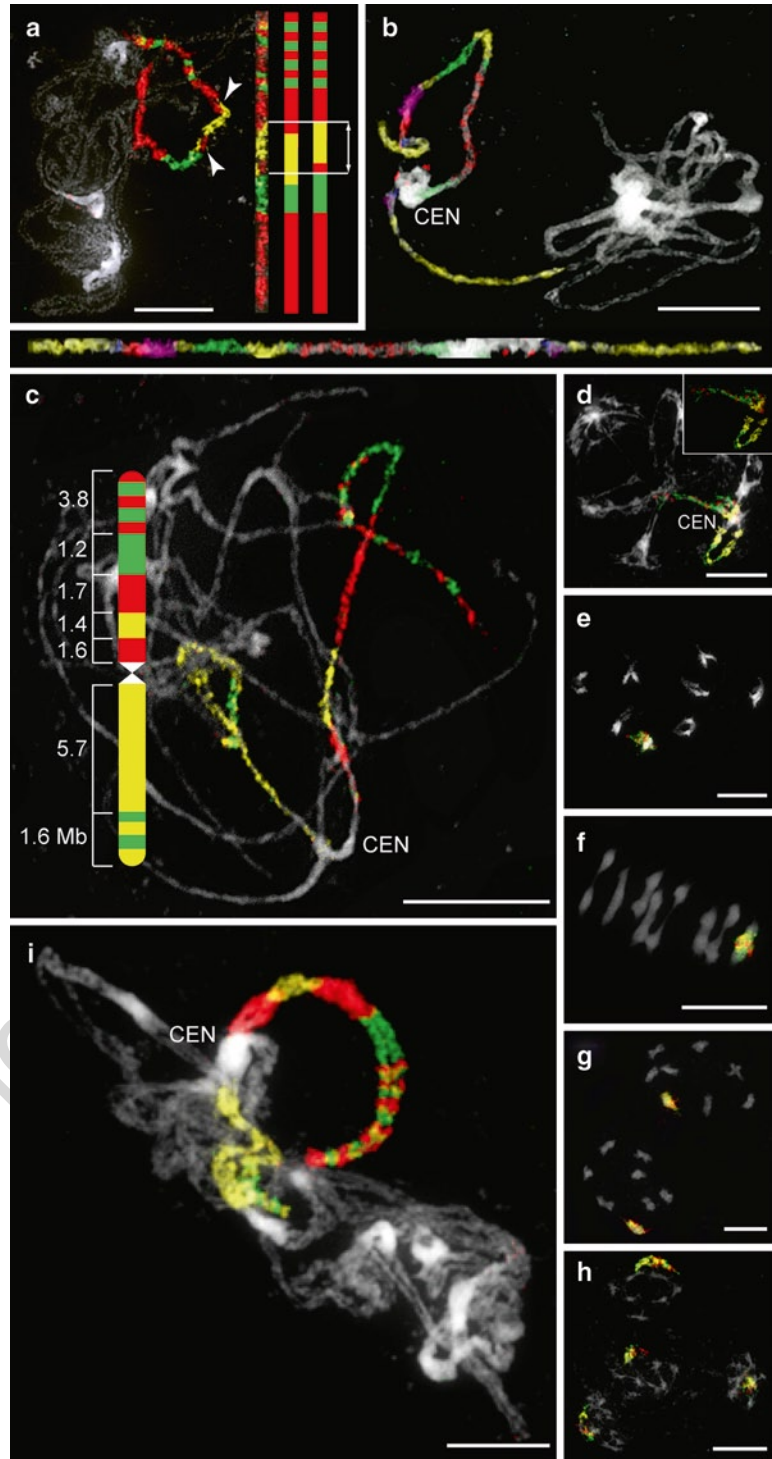
Fig. 1 Application of comparative multicolor chromosome painting for meiotic studies. (**a**) Chromosome painting in Shahdara × Columbia (Sha × Col) hybrid of *Arabidopsis thaliana* ($n = 5$, At1–At5). A 2.2 Mb paracentric inversion on the top arm of chromosome At3 is specific for Sha and absent in Col. Paired Sha and Col

as chromosome-specific probes. Fluorescence in situ hybridization 32
(FISH) of single or several BAC clones is referred to as BAC FISH, 33
whereas BAC painting or chromosome painting applies to in situ 34
hybridization of BAC contigs covering larger chromosome regions 35
(e.g., chromosome arms) or whole chromosomes (Fig. 1). 36
Chromosome-specific BAC contigs are used as painting probes 37
either in the same species (2–5) or in species with sufficient chro- 38
mosome homeology (cross-species or comparative chromosome 39
painting, CCP) (3, 6–8). Recently, reciprocal BAC painting and 40
multi-species BAC painting (using painting probes of two or more 41
species to chromosomes of another species) on pachytene chromo- 42
somes has been established (9). 43

Although CP enables identification of chromosome regions 44
and whole chromosomes during all (pre)meiotic stages (2), 45
pachytene bivalents and multivalents usually offer the highest reso- 46
lution. Here, we provide a protocol to paint meiotic and mitotic 47
chromosomes of plants using chromosome-specific BAC clones 48
and BAC contigs. These procedures are essentially based on our 49
long-term experience with CP and CCP in crucifer species 50
(*Brassicaceae*) (2, 3, 6, 8, 9). 51

[AU1]

2 Materials 52

Prepare and store all reagents at room temperature (unless indi- 53
cated otherwise). 54

2.1 Collection of Floral Material

1. Freshly prepared Carnoy's I fixative: 3 parts ethanol, 1 part 55
   glacial acetic acid; or Carnoy's II fixative: 6 parts ethanol, 3 56
   parts chloroform, 1 part glacial acetic acid (see Note 1). 57

◄

Fig. 1 (continued) homologues were identified by differentially labeled *A. thaliana* BAC contigs (inversion marked by *arrowheads*).(**b–i**) Comparative chromosome painting in *Brassicaceae* species using chromosome-specific *A. thaliana* BAC contigs. (**b**) Pachytene chromosome Sn4 of *Stenopetalum nutans* (*n* = 4, Sn1–Sn4) painted by 14 differentially labeled *A. thaliana* BAC contigs. Image of the same chromosome digitally straightened using the "straighten-curved-objects" plugin in the Image J software (12). (**c–h**) CCP of chromosome Ca1 in diploid *Cardamine amara* (*n* = 8, Ca1–Ca8) at pachytene (**c**), diplotene (**d**), diakinesis (**e**), metaphase I (**f**), telophase I (**g**), and anaphase/telophase II (**h**). (**i**) CCP of a pachytene tetravalent of Ca1 in autotetraploid *C. amara* (*n* = 16, Ca1–Ca16). Chromosome-specific *Arabidopsis* BAC contigs were labeled by biotin-dUTP (*red*) and digoxigenin-dUTP (*green*), and immuno-detected using antibodies coupled to Texas Red and Alexa Fluor 488, respectively. *Yellow* signals correspond to Cy3-dUTP-labeled contigs. In addition, BAC clones hybridized to chromosome Sn4 in (**b**) were also labeled by DEAC-dUTP (*blue*) and DNP-dUTP (Cy5, *magenta*). CENs refer to centromeres (not labeled by BAC clones). Size of BAC contigs in Mb according to http://www.arabidopsis.org. All meiotic chromosomes were isolated from anthers and counterstained with DAPI. *Bars* = 10 μm

2. 70% ethanol.

3. Forceps.

4. Glass vials or microcentrifuge tubes.

2.2 Chromosome Preparation

1. 10× citrate buffer: 40 mL of 100 mM citric acid and 60 mL of 100 mM trisodium citrate, adjust pH to 4.8; store at 4°C.

2. Pectolytic enzyme mixture in 1× citrate buffer: 0.3% pecto-lyase, 0.3% cellulase, 0.3% cytohelicase (Sigma Aldrich, St. Louis, MO, USA) prepared from frozen 1% stock solutions in 1× citrate buffer (see Note 2).

3. 60% glacial acetic acid in distilled water (see Note 3).

4. Freshly prepared ice-cold Carnoy's I fixative.

5. 4% freshly prepared formaldehyde in distilled water.

6. "Assistant" staining blocks with glass cover (Karl Hecht Assistant, Sondheim/Rhön, Germany).

7. Humid box.

8. Incubator at 37°C.

9. Stereomicroscope, microscope with phase contrast.

10. SuperFrost microscope slides (Fisher Scientific, Suwanee, GA, USA).

11. Dissection needles.

12. Fine forceps.

13. Glass Pasteur pipette.

14. Heating block.

15. Hair dryer.

2.3 Chromosome Preparation Pretreatment

1. 20× Saline Sodium Citrate (SSC): 3 M sodium chloride, 300 mM trisodium citrate, pH 7.0.

2. RNase: 100 μg/mL DNase-free ribonuclease A (AppliChem, St. Louis, MO, USA). Make stock of 1 mg/mL in distilled water. Store aliquots at –20°C.

3. Pepsin from porcine gastric mucosa (Sigma Aldrich, St. Louis, MO, USA) 0.1 mg/mL in 10 mM HCl. Prepared from 100 mg/mL stock in 10 mM HCl. Store aliquots at –20°C.

4. 4% freshly prepared formaldehyde in 2× SSC.

5. 70%, 80%, and 96% ethanol.

6. 4′, 6-diamidino-2-phenylindole (DAPI; Sigma Aldrich, St. Louis, MO, USA): 2 μg/mL in Vectashield antifade (Vector Laboratories, Burlingame, CA, USA). Store at 4°C.

7. Coverslips (24 × 24 mm and 24 × 50 mm).

8. Coplin or Hellendahl jar (see Note 4).

9. Humid box. 97

10. Incubator and water bath at 37°C. 98

11. Plastic tube rack. 99

2.4 Probe Labeling

1. 10× NT buffer: 500 mM Tris–HCl pH 7.5, 50 mM $MgCl_2$, 0.05% bovine serum albumin. 100 101

2. Nucleotide mixture: 2 mM dATP, dCTP, dGTP, and 400 mM dTTP (Roche Applied Science, Indianapolis, IN, USA). 102 103

3. 1 mM x-dUTP (x stands for biotin, digoxigenin, Cy3 or other hapten/fluorochrome; see Note 5). 104 105

4. 0.1 M β-mercaptoethanol. 106

5. DNase I (Roche Applied Science, Indianapolis, IN, USA). Use a 1:250 dilution of a 1 mg/mL DNase stock in 0.15 M NaCl in 50% glycerol. 107 108 109

6. DNA polymerase I (10 U/μL; Fermentas, Glen Burnie, MA, USA). 110 111

7. 0.5 M EDTA, pH 8.0. 112

8. 100 bp DNA ladder. 113

9. 0.5 mL microcentrifuge tubes. 114

10. Thermocycler for 15°C and 60°C. 115

11. Electrophoresis system. 116

12. 1% agarose gel. 117

2.5 Probe Preparation and In Situ Hybridization

1. 3 M sodium acetate, pH 5.2. 118

2. 70% ethanol, 96% ice-cold ethanol. 119

3. Hybridization buffer: 50% formamide, 10% dextran sulfate in 2× SSC. 120 121

4. 2 mL microcentrifuge tubes. 122

5. Refrigerated centrifuge. 123

6. Vacuum desiccator or SpeedVac. 124

7. Thermomixer (a heating block with exact temperature control) at 80°C. 125 126

8. 24 × 24 mm, 22 × 22 mm, or 24 × 32 mm coverslips. 127

9. Rubber cement. 128

10. Humid box. 129

11. Incubator at 37°C. 130

2.6 Fluorescence Detection of Hybridized Probes

1. 2× SSC. 131

2. 50% or 20% deionized formamide in 2× SSC, pH 7.0 (see Note 6). 132

3. 4T buffer: 4× SSC pH 7.0, 0.05% Tween-20. 133

4. Blocking solution: 5% bovine serum albumin, 0.2% Tween-20 in 4× SSC.

5. TNT buffer: 100 mM Tris–HCl pH 7.5, 150 mM NaCl, 0.05% Tween-20.

6. TNB buffer: 100 mM Tris–HCl pH 7.5, 150 mM NaCl, 0.5% blocking reagent (Roche Applied Science, Indianapolis, IN, USA).

7. Antibodies: avidin–Texas Red (Vector Laboratories, Burlingame, CA, USA), goat anti-avidin–biotin (Vector Laboratories, Burlingame, CA, USA), mouse anti-digoxigenin (Jackson ImmunoResearch Laboratories, West Grove, PA, USA), goat anti-mouse–Alexa Fluor 488 (Invitrogen, Carlsbad, CA, USA).

8. DAPI (2 µg/mL) in Vectashield antifade.

9. 70%, 80% and 96% ethanol.

10. Coplin or Hellendahl jars (see Note 2).

11. Water bath shaker at 42°C.

12. Coverslips (24×32 mm and 24×50 mm).

13. Humid box.

14. Incubator at 37°C.

15. Epifluorescence microscope equipped with optical filters for DAPI and other fluorochromes, a digital charge-coupled device (CCD) camera, and image acquisition software.

3 Methods

Carry out all procedures at room temperature unless otherwise specified. To minimize fluorochrome bleaching, avoid overexposure to light during procedures 3.4–3.6.

3.1 Collection of Floral Material

1. Fix entire inflorescences, individual flower buds, or anthers in Carnoy's fixative at room temperature or at 4°C overnight; change the fixative several times.

2. Exchange the fixative for 70% ethanol and store the fixed material at –20°C (see Note 7).

3.2 Chromosome Preparation

1. Rinse floral material with distilled water in a staining block or small petri dish for 10 min. Under a stereomicroscope, remove and discard unwanted parts (e.g., yellow anthers containing pollen).

2. Replace water with 1× citrate buffer and wash two times, 5 min each wash.

3. Incubate the material in ~1 mL of pectolytic enzyme mixture in a humid box at 37°C for 3 h (see Note 8). 172 173

4. Replace the enzymes with 1× citrate buffer and keep the material on ice or at 4°C until use (see Note 9). 174 175

5. Put a single flower bud/anther on a microscope slide using a Pasteur pipette, remove the excess fluid and add ~20 µL of 60% acetic acid. Disintegrate the bud by dissection needles until a fine suspension is formed. 176 177 178 179

6. Place the slide on a heating block (50°C) and spread the suspension by careful circular stirring with a needle for ~30 s (see Note 10). 180 181 182

7. Fix the chromosomes by pipetting 100 µL of Carnoy's I fixative around the suspension drop. Discard the fluid by tilting the slide and dry using a hair dryer. 183 184 185

8. Using a phase-contrast light microscope, examine the preparation for specific meiotic stages and the amount of cytoplasm. 186 187

9. Post-fix the slides in a Coplin or Hellendahl jar with 4% formaldehyde in distilled water for 10 min and leave to air-dry. Carry out this step in a fume hood. 188 189 190

10. Store dried slides in a dust-free box at 4°C (see Note 11). 191

3.3 Chromosome Preparation Pretreatment

1. Wash slides two times in 2× SSC in a Coplin jar, 5 min each wash. 192 193

2. Add 100 µL of RNase solution, cover with 24×50 mm coverslip and incubate the slides in a humid box at 37°C for 1 h (see Note 12). 194 195 196

3. Tilt slides to let the coverslip fall off and wash slides as in step 1. 197

4. To remove cytoplasm, treat slides with pepsin at 37°C for 5 min in a Coplin jar placed in a water bath. 198 199

5. Wash slides as in step 1. 200

6. Dehydrate slides in an ethanol series (70%, 80%, and 96% ethanol, 3 min each) and leave them to air-dry. 201 202

7. Apply 15 µL of Vectashield with DAPI to a slide and cover it with a 24×24 mm coverslip. Check the slide with a fluorescence microscope. Chromosomes should be undamaged and free of cytoplasm. When cytoplasm is persistent, remove the coverslip using running water and repeat steps 4–7. 203 204 205 206 207

8. Remove the coverslip using running water flow and wash slides as in step 1. 208 209

9. Post-fix slides in a Coplin or Hellendahl jar with 4% formaldehyde in 2× SSC for 10 min. Carry out this and the following step in a fume hood. 210 211 212

10. Wash slides as in step 1. 213

214
215

11. Dehydrate slides in an ethanol series (70%, 80%, and 96% ethanol, 3 min each) and leave them to air-dry.

216 **3.4 Probe Labeling**
217 **by Nick Translation**
218 **(See Note 13)**

1. Combine in an 0.5 mL microcentrifuge tube: 1 µg of BAC DNA in 32 or 29 µL distilled water, 5 µL of 10× NT buffer, 5 µL of nucleotide mixture, 1 µL of 1 mM comercial x-dUTP or 4 µL of 1 mM custom-made x-dUTP, 5 µL of 0.1 M β-mercaptoethanol, 1 µL of DNase I, and 1 µL of DNA polymerase I. Vortex gently and spin briefly (see Note 14).

222

2. Incubate at 15°C for 90 min.

223
224
225

3. Transfer the tube on ice and load 5 µL of the reaction volume on a 1% agarose gel along with a 100-bp DNA ladder and run the electrophoresis.

226
227
228
229
230

4. When the smear of labeled fragments is ~200–500 bp in size, stop the nick translation by adding 1 µL of 0.5 M EDTA and heating at 60°C for 10 min. When fragments are longer than 500 bp, extend the incubation at 15°C for further 30 min and repeat steps 3 and 4.

231

5. Store the probe at –20°C until use.

232 **3.5 Probe**
233 **Preparation and In**
234 **Situ Hybridization**

1. Pool individually labeled BAC clones by pipetting 5 µL (~100 ng of DNA) of each BAC into a 2 mL microcentrifuge tube.

235
236
237
238

2. To reduce the probe volume and remove unincorporated nucleotides, precipitate the DNA by adding 0.1 volume of 3 M sodium acetate and 2.5 volume of ice-cold 96% ethanol. Vortex and keep on ice or at –20°C for at least 30 min.

239

3. Centrifuge at $13,000 \times g$ at 4°C for 30 min.

240

4. Carefully discard the supernatant.

241

5. Add 500 µL of 70% ethanol and centrifuge again for 5 min.

242

6. Repeat step 4.

243

7. Dry the pellet using a desiccator or SpeedVac.

244
245

8. Resuspend the pellet in 20 µL of hybridization buffer and incubate at 37°C in a thermomixer (see Note 15).

246
247
248

9. Add 20 µL of probe to the chromosome preparation (see Subheading 3.3), cover with a cover slip and seal with rubber cement around the edges.

249
250

10. Denature the probe and the chromosomal DNA by placing the slide on a heating block at 80°C for 2 min.

251
252

11. Incubate the slide in a humid box at 37°C overnight (see Note 16).

3.6 Fluorescence Detection of Hybridized Probes

Signal detection and amplification is described here for hapten-labeled probes (biotin- and digoxigenin-dUTP) visualized by indirect immunofluorescence via Texas Red and Alexa Fluor 488, respectively (dinitrophenyl (DNP)-dUTP can also be used as a third hapten). This protocol can be also used with fluorochrome-labeled probes (e.g., Cy3- or DEAC-dUTP) that require only post-hybridization washing prior to microscopic evaluation (steps 1–4, and 14). All washing steps are carried out at 42°C in Coplin or Hellendahl jars placed in a water bath shaker. All incubation steps are carried out at 37°C on microscope slides covered with 24×50 mm coverslips and placed in a humid box. To minimize unspecific background, do not let the slides dry during the entire procedure.

1. Remove the rubber cement frame with forceps and let the coverslip fall off.

2. Wash slides in 2× SSC for 2 min.

3. Wash slides three times in 50% or 20% formamide, 5 min each wash (see Note 6).

4. Wash slides in 2× SSC for 2 min. If fluorochrome-labeled probes are used, proceed with step 14.

5. Wash slides in 4T buffer for 5 min.

6. Incubate slides in 100 µL of blocking solution for 30 min.

7. Wash slides in 4T buffer 2×5 min.

8. Incubate slides in 100 µL of avidin–Texas Red in TNB buffer (1:1,000) for 30 min.

9. Wash slides in TNT buffer 2×5 min.

10. Incubate slides in 100 µL of mixed goat anti-avidin–biotin antibody (1:200) and mouse anti-digoxigenin antibody (1:250) in TNB buffer for 30 min (see Note 17).

11. Repeat step 9.

12. Incubate slides in 100 µL of mixed avidin–Texas Red (1:1,000) and goat anti-mouse Alexa Fluor 488-coupled antibody (1:200) in TNB buffer for 30 min.

13. Repeat step 9.

14. Dehydrate the slides in an ethanol series (70%, 80%, and 96% ethanol, 3 min each) and leave them to air-dry.

15. Apply 15 µL of Vectashield with DAPI to each slide and cover it with a 24×32 mm coverslip.

16. Observe and photograph slides under a fluorescence microscope equipped with appropriate optical filters, CCD camera, and image acquisition software.

294 ## 4 Notes

295 1. Carnoy's II is believed to be more suitable than the Carnoy's
296   I fixative for fixation of floral material. However, we did not
297   observe a significant difference between the two fixatives and
298   use the 3:1 fixative routinely.

299 2. The enzyme mixture can be reused several times (store at –20°C).
300   Digestion time might need to be increased after each use.

301 3. Glacial acetic acid can be used at concentrations from 45%
302   to 60%.

303 4. A Coplin jar can hold up to nine slides, Hellendahl jar up to 15
304   slides.

305 5. Labeled nucleotides are available commercially or can be syn-
306   thesized by coupling allylamine-dUTP to succinimidyl-ester
307   derivatives of haptenes or fluorochromes (10).

308 6. 50% formamide should be used for stringent post-hybridization
309   washing after same-species (homologous) in situ hybridization;
310   20% formamide should be used in case of cross-species (home-
311   ologous) in situ hybridization.

312 7. Fixed material can be used after several months or years of
313   storage. However, best quality preparations are usually obtained
314   from freshly fixed material.

315 8. The amount of enzyme mixture should be proportional to the
316   amount of digested material (tissue should be submerged).
317   The incubation time of 3 h may not be appropriate to all spe-
318   cies/types of floral material. After the incubation, we make a
319   test preparation following protocol (3.2, steps 5–8). If the tissue
320   is hard to disintegrate and there is a large quantity of cytoplasm
321   and tissue fragments, we extend enzyme incubation for another
322   30 min or longer.

323 9. Digested floral tissue can be stored in 1× citrate buffer at 4°C
324   overnight or longer. Overnight storage further softens the
325   digested material.

326 10. The amount of acetic acid can be increased to 40–60 μL and
327   the duration of suspension spreading can be prolonged. The
328   duration of spreading can be shortened in case of very small
329   flower buds or anthers. The needle should not touch the slide
330   surface.

331 11. Post-fixed slides can be stored at 4°C for several weeks or
332   months. However, freshly prepared or few-days old slides
333   guarantee better results.

334 12. Omitting the RNase treatment may result in a longer pepsin
335   treatment used in Subheading 3.4, step 4 and/or increased
336   background.

13. As *Brassicaceae* possess very small genomes and low amount of repeats localized in pericentromeric heterochromatin, chromosome-specific BAC clones from euchromatic regions can be used without the need to block repetitive sequences from cross-hybridization. For chromosome painting in other plant taxa, BAC clones have to be screened for the presence of dispersed repeats or genomic/Cot DNA needs to be applied for blocking.

14. DNA of large-insert clones (usually BAC clones) can be isolated by use of a standard alkaline lysis protocol (11) or using a DNA isolation kit (e.g., the Qiagen Plasmid Midi Kit; Qiagen, Hilden, Germany). A larger amount (4 μL) of custom-made x-dUTPs must be used as custom dUTPs label DNA less efficiently than commercial nucleotides (10). DNA of several probes/BAC clones can be pulled and labeled together in a single nick translation reaction. However, in our laboratory we label each probe individually.

15. Time needed to resuspend the probe is dependent on the initial amount of precipitated DNA. DNA of several BAC clones is well dissolved after several minutes or hours. Complex probes comprising several dozen or hundred clones should be incubated overnight.

16. Overnight incubation is usually sufficient for in situ hybridization of homologous (same species) sequences. Longer hybridization times (48–72 h) may be required to ensure hybridization of homeologous (cross-species) sequences.

17. If the first detection step for hapten-labeled probes (see Subheading 3.6, step 8) yields a sufficiently strong signal, further signal amplification (e.g., by goat anti-avidin–biotin antibody in this protocol) can be omitted.

Acknowledgments

References

1. Schubert I, Fransz PF, Fuchs J, de Jong JH (2001) Chromosome painting in plants. Methods Cell Sci 23:57–69

2. Lysak MA, Fransz PF, Ali HBM, Schubert I (2001) Chromosome painting in *Arabidopsis thaliana*. Plant J 28:689–697

3. Lysak M, Fransz P, Schubert I (2006) Cytogenetic analyses of Arabidopsis. In: Salinas J, Sanchez-Serrano JJ (eds) Methods in molecular biology, vol 323, Arabidopsis protocols. Humana, Totowa, NJ, pp 173–186

4. Szinay D, Chang S-B, Khrustaleva L, Peters S, Schijlen E, Bai Y et al (2008) High-resolution chromosome mapping of BACs using multi-colour FISH and pooled-BAC FISH as a backbone for sequencing tomato chromosome 6. Plant J 56:627–637

5. Febrer M, Goicoechea JL, Wright J, McKenzie N, Song X, Lin J et al (2010) An integrated physical, genetic and cytogenetic map of *Brachypodium distachyon*, a model system for grass research. PLoS One 5:e13461

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

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

8. Mandáková T, Joly S, Krzywinski M, Mummenhoff K, Lysak MA (2010) Fast diploidization in close mesopolyploid relatives of *Arabidopsis*. Plant Cell 22:2277–2290

9. Lysak MA, Mandáková T, Lacombe E (2010) Reciprocal and multi-species chromosome BAC painting in crucifers (*Brassicaceae*). Cytogenet Genome Res 129:184–189

10. Henegariu O, Bray-Ward P, Ward DC (2000) Custom fluorescent nucleotide synthesis as an alternative method for nucleic acid labeling. Nat Biotechnol 18:345–348

11. Sambrook J, Russell DW (2001) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

12. Kocsis E, Trus BL, Steer CJ, Bisher ME, Steven AC (1991) Image averaging of flexible fibrous macromolecules: the clathrin triskelion has an elastic proximal segment. J Struct Biol 107:6–14

II. Lysak M.A., **Mandáková T.**, Lacombe E. 2010. Reciprocal and multi-species chromosome BAC painting in crucifers (Brassicaceae). Cytogenetic and Genome Research 129: 184-189.

Reciprocal and Multi-Species Chromosome BAC Painting in Crucifers (Brassicaceae)

M.A. Lysak[a]   T. Mandáková[a]   E. Lacombe[b]

[a]Department of Functional Genomics and Proteomics, Institute of Experimental Biology, Faculty of Science, Masaryk University, Brno, Czech Republic; [b]Laboratoire de Biochimie et Physiologie Moleculaire des Plantes, UMR Université Montpellier 2-CNRS-INRA-Montpellier SupAgro, Montpellier, France

Abstract

In crucifer cytogenomics, BAC contigs of *Arabidopsis thaliana* have been used as probes for comparative chromosome painting among species. Here we successfully tested chromosome-specific BAC contigs of *A. thaliana* (n = 5) and *A. halleri* (n = 8) as probes for reciprocal BAC painting. Furthermore, BAC contigs of both *Arabidopsis* species were applied as multi-species painting probes to a third crucifer species, *Noccaea caerulescens* (n = 7), revealing their shared chromosome homeology. Specifically, we found homeology across portions of chromosomes At2, Ah4, and Nc4, which reflects their shared common origin with chromosome AK4 of the Ancestral Crucifer Karyotype (n = 8). We argue that multi-species and multi-directional painting will significantly expedite comparative cytogenomics in Brassicaceae and other plant families.

Copyright © 2010 S. Karger AG, Basel

Identification of plant chromosomes by chromosome-specific painting probes is hampered by a wide spectrum of dispersed repetitive elements equally distributed over chromosomes of a given complement. Hence, all flow-sorted or microdissected, and DOP-PCR amplified, DNA probes yielded unspecific cross-hybridization signals [reviewed by Schubert et al., 2001]. An alternative strategy to paint chromosomes of plants has been found in the application of contigs and supercontigs of chromosome-specific BAC clones. This approach relies on the availability of chromosome-specific BAC libraries and a low amount of repetitive elements in the donor and target genomes.

Arabidopsis thaliana (n = 5) has become the first plant species with all chromosomes painted using chromosome-specific BAC supercontigs [Lysak et al., 2001; Pecinka et al., 2004]. Later the very same BAC contigs were applied as painting probes for comparative chromosome painting (CCP) across the mustard family (Brassicaceae) [e.g. Lysak et al., 2003, 2006; Mandáková and Lysak, 2008]. Recently, cross-species BAC FISH was successfully developed for sorghum and maize (Poaceae) [Amarillo and Bass, 2007], tomato and potato (Solanaceae) [Iovene et al., 2008; Tang et al., 2008] as well as for cucurbit species (Cucurbitaceae) [Han et al., 2009].

Martin A. Lysak
Department of Functional Genomics and Proteomics
Institute of Experimental Biology, Masaryk University
Kamenice 5, Building A2, CZ–625 00 Brno (Czech Republic)
Tel. +420 549 494 154, Fax +420 549 492 654, E-Mail lysak@sci.muni.cz

In Brassicaceae, BAC clones of *A. thaliana* were exclusively used as painting probes for cross-species CCP as chromosome-specific BAC libraries for other crucifer species were not developed. This is changing as genomes of an increasing number of cruciferous species are being sequenced and genomic resources become available. Chromosome-specific BAC contigs were assembled for *Brassica rapa* [Mun et al., 2008] or *A. halleri*, a model species of heavy metal accumulation [Lacombe et al., 2008]. In upcoming years we will witness a steadily increasing number of whole-genome sequencing projects including important crop and model crucifer species such as *A. lyrata*, *Boechera* spp., *Capsella rubella*, *Eutrema* (*Thellungiella*) *halophila* (DOE Joint Genome Institute; http://www.jgi.doe.gov), and *B. rapa* (http://www.brassica.info). These sequencing efforts will generate new cytogenomic resources and significantly facilitate family-wide comparative research.

In anticipation of newly available genomic resources for crucifer cytogenomics, we have tested chromosome-specific BAC contigs of *A. halleri* (n = 8) and *A. thaliana* as probes for reciprocal CCP. Furthermore, the BAC contigs of both *Arabidopsis* species were applied simultaneously as multi-species painting probes to reveal shared chromosome homeology in a third species.

Material and Methods

Plant Material
Inflorescences of *A. halleri* subsp. *tatrica* (Pawł.) Kolník (2n = 16) were collected from a single wild population (Slovakia, Belianske Tatry Mts., Tatranská Javorina, the Zadné Meďodoly valley). *A. thaliana* accession C24 (2n = 10) was used in the present study. Flower buds of *Noccaea caerulescens* (2n = 14) were collected from a population in the village of Kořenec, Czech Republic [Mandáková and Lysak, 2008].

Preparation of Pachytene Chromosomes
Entire inflorescences were fixed in freshly prepared ethanol:acetic acid (3:1) fixative overnight and stored in 70% ethanol at –20°C until use. Selected inflorescences were rinsed in distilled water and citrate buffer (10 mM sodium citrate, pH 4.8) and subsequently incubated in the 0.3% enzyme mix (cellulase, cytohelicase and pectolyase; all Sigma-Aldrich) in citrate buffer at 37°C for 3–4 h. After digestion, individual flower buds were disintegrated by a needle in a small drop of citrate buffer and the material spread in 20 μl of 60% acetic acid on a hot plate (50°C). The chromosomes were fixed using 100 μl of ice-cold ethanol:acetic acid fixative, then the slides were dried using a hair dryer. The preparations were staged using a phase contrast microscope and suitable slides were post-fixed in 4% formaldehyde in distilled water for 10 min and air-dried.

Painting Probes
A. thaliana BAC clones were obtained from the *Arabidopsis* Biological Resource Center (ABRC, Columbus, OH). The 3 BAC contigs used [F3K23 (AC006841) – F8N16 (AC005727), F16P2 (AC004561) – F7F1 (AC004669), and F13P17 (AC004481) – T8I13 (AC002337)] represent an almost complete BAC tiling path of *A. thaliana* chromosome At2, corresponding to the collinear region on the ancestral chromosome AK4 of the Ancestral Crucifer Karyotype [Schranz et al., 2006]. From this tiling path each third BAC was isolated and labeled.

A. halleri BAC clones were selected from the genomic BAC library [Lacombe et al., 2008] and 2 BAC contigs were built based on the chromosome collinearity existing between *A. thaliana* and *A. lyrata* [Kuittinen et al., 2004]. The A contig (from AT2G31130 to AT2G32520) comprises BAC clones XIII-P-6, II-F-23, XIV-I-23, II-K-14, V-A-2, I-E-5, I-K-8, and I-I-19. The B contig (from AT2G32840 to AT2G33510) is composed of BAC clones IV-K-16, III-M-8, and II-F-13. Furthermore, 2 repeat-rich BAC clones IV-E-9 and IX-B-21 [Lacombe et al., 2008] were tested (fig. 1A, D).

DNA of individual BAC clones was isolated using a standard alkaline extraction omitting the phenol:chloroform purification step. BAC DNA was labeled by biotin-, digoxigenin-, and Cy3-dUTP via nick translation as follows: 1 μg of BAC DNA diluted in distilled water to 29 μl, 5 μl of nucleotide mix (2 mM dATP, dCTP, and dGTP, 400 μM dTTP; all Roche), 5 μl of 10× NT buffer (0.5 M Tris-HCl, pH 7.5; 50 mM MgCl$_2$, 0.05% BSA), 4 μl of 1 mM x-dUTP (in which x was biotin, digoxigenin or Cy3), 5 μl of 0.1 M b-mercaptoethanol, 1 μl of DNase I (Roche), and 1 μl of DNA polymerase I (Fermentas). The nick translation mixture was incubated at 15°C for 90 min (or longer) to obtain fragments of ~200 to 500 bp. The reaction was stopped by adding 1 μl of 0.5 M EDTA, pH 8.0, followed by incubation at 65°C for 10 min. Labeled DNA of individual BAC clones was stored at –20°C until use.

Chromosome Painting: BAC FISH
The selected slides were treated by RNase (AppliChem; 100 μg/ml in water) for 1 h at 37°C, and washed in 2× SSC for 2–5 min. To remove cytoplasm, the slides were treated with pepsin (Sigma-Aldrich; 0.1 mg/ml) in 0.01 M HCl at 38°C for 10 min, followed by a wash in 2× SSC for 2–5 min. Subsequently, the slides were post-fixed in 4% formaldehyde in 2× SSC for 10 min, washed in 2× SSC (2–5 min), and dehydrated in an ethanol series (70, 80, and 96%).

Labeled BAC DNAs were pooled and precipitated to reduce the probe volume. For a single slide, the probe was dissolved in 20 μl of hybridization mix (50% formamide, 10% dextran sulfate in 2× SSC) at 37°C overnight. Probe and chromosomes were denatured together on a hot plate at 80°C for 2 min and incubated in a moist chamber at 37°C for 48 h. Post-hybridization washing was performed in 50% or 20% (CCP in *N. caerulescens*) formamide in 2× SSC for 3–5 min at 42°C. DNA labeled by biotin-dUTP was detected using avidin~Texas Red (Vector Laboratories) and amplified by goat anti-avidin~biotin (Vector Laboratories) and avidin~Texas Red. Probes labeled by digoxigenin-dUTP were visualized by mouse anti-digoxigenin (Jackson ImmunoResearch) and goat anti-mouse~Alexa Fluor 488 (Molecular Probes). Chromosomes were counterstained with DAPI (2 μg/ml) in Vectashield (Vector Laboratories).
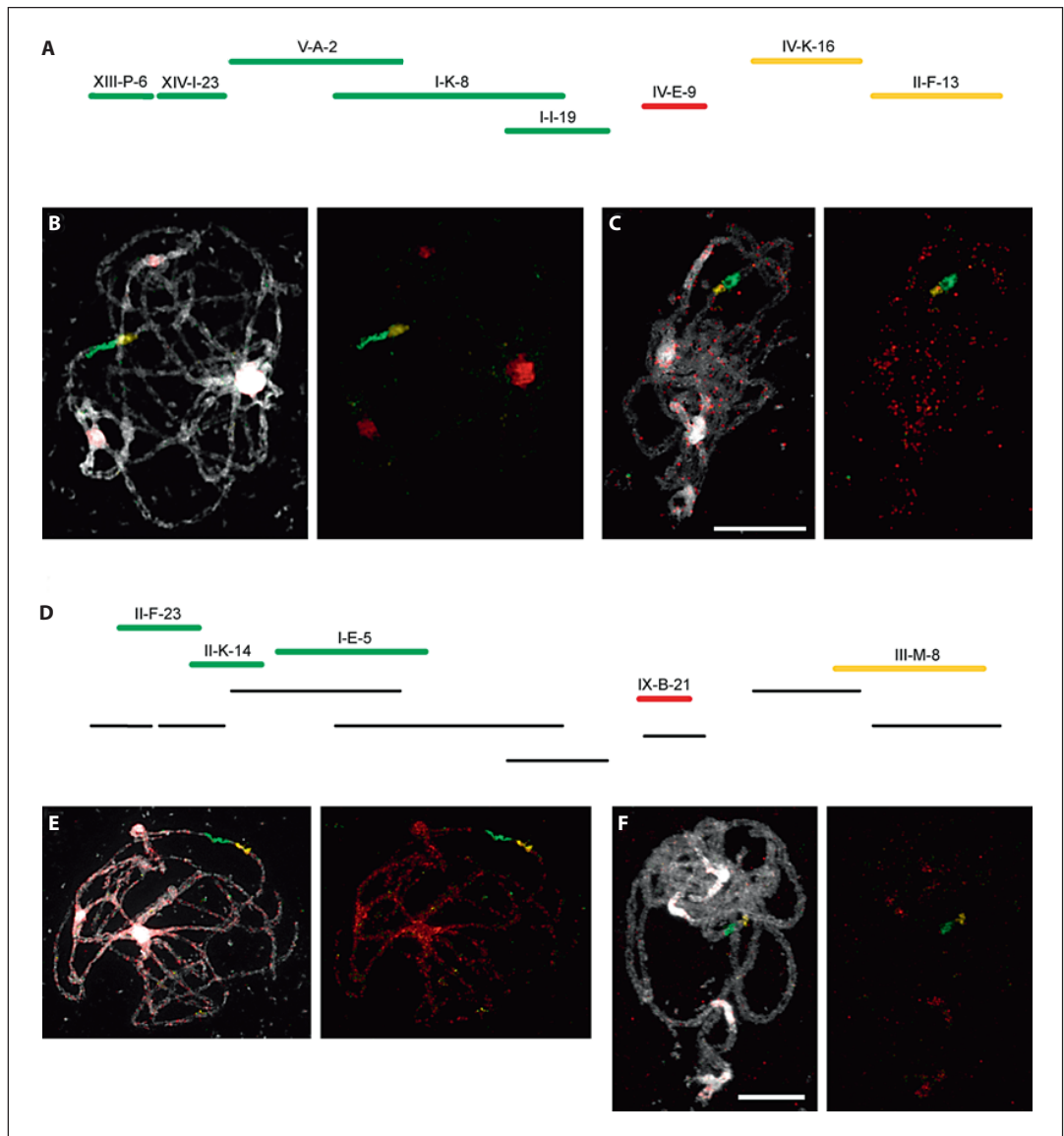
Fig. 1. BAC painting in *Arabidopsis halleri* and *A. thaliana*. **A** Labelling scheme of *A. halleri* BAC contigs A and B, and BAC clone IV-E-9. **B**, **C** In situ hybridization of the labeled probes (**A**) to pachytene chromosomes of *A. halleri* (**B**) and *A. thaliana* (**C**), respectively. **D** Labelling scheme of *A. halleri* BAC contigs A and B, and BAC clone IX-B-21. Black lines correspond to BAC clones labeled in (**A**). **E**, **F** In situ hybridization of the labeled probes (**D**) to pachytene chromosomes of *A. halleri* (**E**) and *A. thaliana* (**F**), respectively. Bar (10 μm) has the same magnification in (**B**) and (**C**), and in (**E**) and (**F**), respectively.

Preparations were observed using an Olympus BX-61 epifluorescence microscope, and images were acquired separately for all fluorochromes using appropriate excitation and emission filters (AHF Analysentechnik) using an AxioCam CCD camera (Zeiss). The 4 monochromatic images were pseudocolored and merged using the Adobe Photoshop CS2 software (Adobe Systems).

Results and Discussion

We have hybridized 2 *A. halleri* BAC contigs to pachytene chromosomes of *A. halleri* and *A. thaliana*. The 8 and 3 BAC clones of the A (c. 0.55 Mb in *A. thaliana*) and
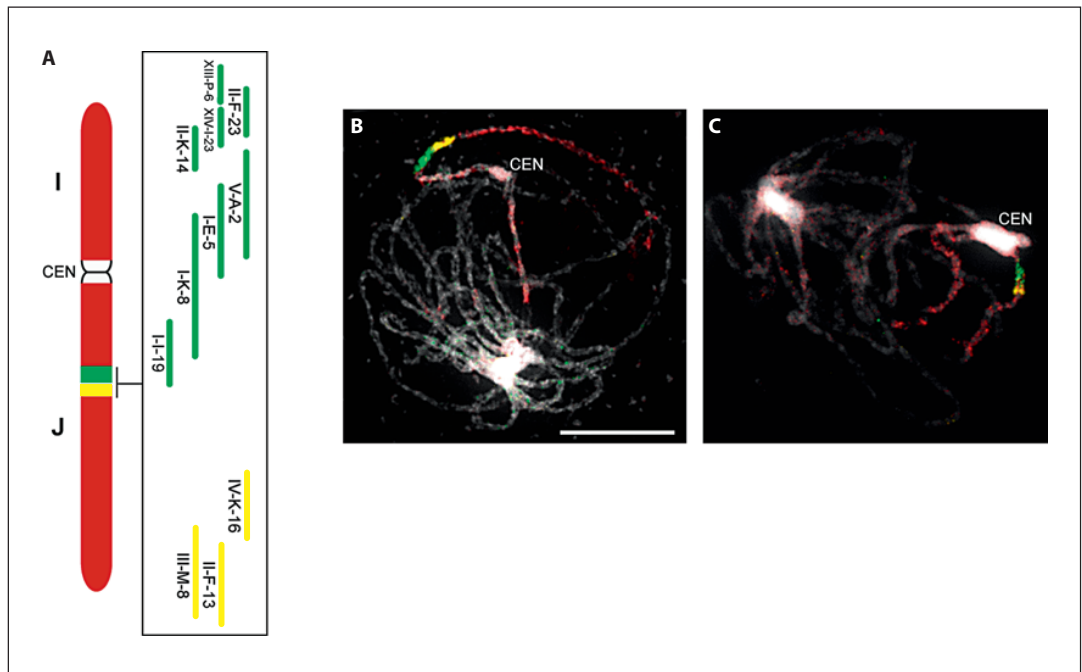
Fig. 2. BAC chromosome painting in *Arabidopsis halleri* and multi-species chromosome painting in *Noccaea caerulescens*. **A** Labeling scheme of *A. halleri* BAC contigs A and B (green, yellow), and *A. thaliana* supercontigs corresponding to genomic blocks I and J (red) of the Ancestral Crucifer Karyotype [Schranz et al., 2006]. Simultaneous in situ hybridization of *A. halleri* and *A. thaliana* BAC contigs to pachytene chromosome of *A. halleri* (**B**) and *N. caerulescens* (**C**). Bar (10 μm) has the same magnification in (**B**) and (**C**).

B (c. 0.27 Mb) contigs, respectively, have been differentially labeled (fig. 1A, D), and tested in both *Arabidopsis* species in the absence of blocking DNA. In *A. halleri*, the digoxigenin- (green) and Cy3-dUTP (yellow) labeled contigs A and B identified the expected chromosome region on the bottom arm of chromosome Ah4 (fig. 1B, E). The same contigs cross-hybridized to *A. thaliana* pachytene chromosomes provided a comparably strong signal on the bottom arm of chromosome At2 (fig. 1C, F), corresponding to the collinear region of Ah4 [Kuittinen et al., 2004; Schranz et al., 2006].

We have also tested 2 BAC clones positioned between the contigs A and B and presumably containing repetitive elements [Lacombe et al., 2008]. In *A. halleri*, BAC clones IV-E-9 and IX-B-21 hybridized to all centromeres and pericentromeres, respectively (fig. 1B, E). In *A. thaliana*, red signals of the IV-E-9 were irregularly scattered along all chromosomes with some preferential hybridization at pericentromeric regions (fig. 1C), whereas BAC IX-B-21 (red) hybridized to pericentromeric regions, though the

hybridization signal intensity was weaker than in *A. halleri* (fig. 1F). The hybridization of BAC IV-E-9 to centromeric heterochromatin is due to *gypsy*- and CACTA-like mobile elements, whereas the clone IX-B-21 contains presumably a *copia*-like retrotransposon(s) [Lacombe et al., 2008].

Furthermore, we demonstrate how BAC contigs of both *Arabidopsis* species can be combined and simultaneously used to paint specific chromosome regions in a third species (fig. 2A–C). The 2 *A. halleri* contigs along with 3 *A. thaliana* BAC supercontigs homeologous to the genomic blocks I and J on ancestral chromosome AK4 [Schranz et al., 2006], except the part covered by *A. halleri* BACs, have been hybridized to pachytene chromosomes of *A. halleri* and *Noccaea (Thlaspi) caerulescens*. The latter species is phylogenetically closely associated with the crucifer Lineage II [Mandáková and Lysak, 2008], which is relatively distantly related to the genus *Arabidopsis* of the Lineage I. We assumed that the *A. halleri* karyotype resembles that of *A. lyrata* (n = 8) [Kuittinen et al., 2004],

and hence the Ah4 chromosome will have a structure similar to Al4 of *A. lyrata*. The overall conserved collinearity between 2 other *A. halleri* chromosomes (Ah6 and Ah7) and chromosomes Al6 and Al7 of *A. lyrata* has been confirmed previously [Lysak et al., 2003]. In *N. caerulescens* (n = 7), chromosome Nc4 also resembles the ancestral structure of AK4 [Mandáková and Lysak, 2008]. In both analyzed species, a strong green/yellow signal of the *A. halleri* contigs was observed. Somewhat weaker red fluorescence of the *A. thaliana* supercontigs labeled the remaining parts of chromosome Ah4 and Nc4, respectively (fig. 2B, C). The weaker hybridization signal intensity of *A. thaliana* contigs reflected the selective use of approximately each third clone of the BAC tiling path. Multi-species painting of chromosomes Ah4 and Nc4 along the entire length confirmed their conserved AK4-like structure.

We have shown that BAC clones of *A. halleri* can be successfully used as reliable chromosome painting probes in this species, congeneric *A. thaliana* as well as in the more distantly related species *N. caerulescens*. Comparative cytogenetics in the Brassicaceae has been until now limited to the use of chromosome-specific BAC supercontigs of *A. thaliana*. As the genome of *A. halleri* (0.24 pg C^{-1}) [Lysak et al., 2009] is 1.5-fold larger than the *A. thaliana* genome (0.16 pg C^{-1}) [Bennett et al., 2003], it can be expected to find a higher percentage of repetitive elements in the *A. halleri* genome. Indeed, the sequence data analysis of Lacombe et al. [2008] corroborated in the present study identified 2 BAC clones containing transposon and retrotransposon elements. These data suggest that the selection of BAC clones suitable for chromosome painting in *A. halleri* and other crucifer species with a similar or larger genome must be carried out with even more caution than in *A. thaliana* [see Lysak et al., 2003]. In species with larger genomes, undesirable cross-hybridization of repetitive sequences could be reduced through the application of a highly repetitive Cot-100 genomic fraction or sheared genomic DNA as shown for Solanaceae species [Iovene et al., 2008; Tang et al., 2008]. Repetitive elements present in 2 *A. halleri* BAC clones (IV-E-9 and IX-B-21) are shared by the 2 *Arabidopsis* species, though the weaker and more scattered hybridization signals in *A. thaliana* suggest that the repeats are less abundant and/or more divergent in *A. thaliana* as compared to *A. halleri*.

As the karyotype structure of *A. thaliana* is relatively complex due to chromosome rearrangements accompanying the chromosome number reduction (from n = 8 to n = 5) in this species [Lysak et al., 2006; Schranz et al., 2006], chromosome-specific BAC (super)contigs of closely related n = 8 species would be more suitable as CCP probes. The present data suggest that *A. halleri* BAC contigs can be successfully employed as CCP probes, and even more extensive resources should become available for *A. lyrata* and *C. rubella* (both n = 8) within the DOE genome sequencing program (http://www.jgi.doe.gov). Moreover, by hybridizing *A. halleri* and *A. thaliana* BAC contigs to chromosomes of *N. caerulescens*, we showed for the first time that multi-species comparative chromosome painting in a plant species is feasible. Current experiments prove that (i) painting BAC probes derived from a pair of species can be used reciprocally in the respective species (*A. halleri* ↔ *A. thaliana*), and (ii) BAC contigs of one species can be substituted or supplemented by painting probes of another closely related taxon (*A. halleri* + *A. thaliana* → *N. caerulescens*). We anticipate that the latter approach, namely multi-species CCP, will significantly facilitate and expedite comparative reconstructions of karyotype evolution in Brassicaceae and other plant families.

As previously reported [Lacombe et al., 2008], physical mapping based on the conserved synteny among *Arabidopsis* species can be used to assemble BAC contigs. However, the presence of repetitive elements such as transposons or retrotransposons hampers the use of this strategy for building supercontigs and constructing complete physical maps. In the present study we showed that cytogenetic mapping can anchor individual BAC clones and contigs on chromosomes, build supercontigs and identify gaps between them. Therefore, BAC painting is an efficient alternative to chromosome walking in constructing genome-wide physical maps.

References

Amarillo FIE, Bass HW: A transgenomic cytogenetic sorghum *(Sorghum propinquum)* bacterial artificial chromosome fluorescence in situ hybridization map of maize *(Zea mays* L.) pachytene chromosome 9, evidence for regions of genome hyperexpansion. Genetics 177:1509–1526 (2007).

Bennett MD, Leitch IJ, Price HJ, Johnston JS: Comparisons with *Caenorhabditis* (~100 Mb) and *Drosophila* (~175 Mb) using flow cytometry show genome size in *Arabidopsis* to be ~157 Mb and thus 25% larger than the *Arabidopsis* genome initiative estimate of ~125 Mb. Ann Bot 91:547–557 (2003).

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

Iovene M, Wielgus SM, Simon PW, Buell CR, Jiang J: Chromatin structure and physical mapping of chromosome 6 of potato and comparative analyses with tomato. Genetics 180:1307–1317 (2008).

Kuittinen H, de Haan AA, Vogl C, Oikarinen S, Leppala J, et al: Comparing the linkage maps of the close relatives *Arabidopsis lyrata* and *A. thaliana*. Genetics 168:1575–1584 (2004).

Lacombe E, Cossegal M, Mirouze M, Adam T, Varoquaux F, et al: Construction and characterisation of a BAC library from *Arabidopsis halleri*: evaluation of physical mapping based on conserved synteny with *Arabidopsis thaliana*. Plant Sci 174:634–640 (2008).

Lysak MA, Fransz PF, Ali HBM, Schubert I: Chromosome painting in *Arabidopsis thaliana*. Plant J 28:689–697 (2001).

Lysak MA, Pecinka A, Schubert I: Recent progress in chromosome painting of *Arabidopsis* and related species. Chromosome Res 11:195–204 (2003).

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

Lysak MA, Koch MA, Beaulieu JM, Meister A, Leitch IJ: The dynamic ups and downs of genome size evolution in Brassicaceae. Mol Biol Evol 26:85–98 (2009).

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

Mun JH, Kwon SJ, Yang TJ, Kim HS, Choi BS, et al: The first generation of a BAC-based physical map of *Brassica rapa*. BMC Genomics 9:280 (2008).

Pecinka A, Schubert V, Meister A, Kreth G, Klatte M, et al: Chromosome territory arrangement and homologous pairing in nuclei of *Arabidopsis thaliana* are predominantly random except for NOR-bearing chromosomes. Chromosoma 113:258–269 (2004).

Schranz ME, Lysak MA, Mitchell-Olds T: The ABC's of comparative genomics in the Brassicaceae: building blocks of crucifer genomes. Trends Plant Sci 11:535–542 (2006).

Schubert I, Fransz PF, Fuchs J, de Jong JH: Chromosome painting in plants. Meth Cell Sci 23:57–69 (2001).

Tang X, Szinay D, Lang C, Ramanna MS, van der Vossen EAG, et al: Cross-species bacterial artificial chromosome-fluorescence in situ hybridization painting of the tomato and potato chromosome 6 reveals undescribed chromosomal rearrangements. Genetics 180:1319–1328 (2008).

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

RESEARCH ARTICLES

Chromosomal Phylogeny and Karyotype Evolution in x=7 Crucifer Species (Brassicaceae) [W]

Terezie Mandáková and Martin A. Lysak[1]

Department of Functional Genomics and Proteomics, Institute of Experimental Biology, Masaryk University, Brno CZ-625 00, Czech Republic

Karyotype evolution in species with identical chromosome number but belonging to distinct phylogenetic clades is a long-standing question of plant biology, intractable by conventional cytogenetic techniques. Here, we apply comparative chromosome painting (CCP) to reconstruct karyotype evolution in eight species with x=7 (2n=14, 28) chromosomes from six Brassicaceae tribes. CCP data allowed us to reconstruct an ancestral Proto-Calepineae Karyotype (PCK; n=7) shared by all x=7 species analyzed. The PCK has been preserved in the tribes Calepineae, Conringieae, and Noccaeeae, whereas karyotypes of Eutremeae, Isatideae, and Sisymbrieae are characterized by an additional translocation. The inferred chromosomal phylogeny provided compelling evidence for a monophyletic origin of the x=7 tribes. Moreover, chromosomal data along with previously published gene phylogenies strongly suggest the PCK to represent an ancestral karyotype of the tribe Brassiceae prior to its tribe-specific whole-genome triplication. As the PCK shares five chromosomes and conserved associations of genomic blocks with the putative Ancestral Crucifer Karyotype (n=8) of crucifer Lineage I, we propose that both karyotypes descended from a common ancestor. A tentative origin of the PCK via chromosome number reduction from n=8 to n=7 is outlined. Comparative chromosome maps of two important model species, *Noccaea caerulescens* and *Thellungiella halophila*, and complete karyotypes of two purported autotetraploid Calepineae species (2n=4x=28) were reconstructed by CCP.

INTRODUCTION

Comparative linkage mapping of shared genetic markers reveals the extent of interspecies chromosome collinearity, as shown for grasses (Devos, 2005; Wei et al., 2007) and crucifers (reviewed by Koch and Kiefer, 2005). Cytogenetic evidence for collinear chromosome segments among closely related plants is difficult to obtain because ubiquitous and abundant dispersed repeats usually prevent sufficient chromosome specificity of DNA probes (Schubert et al., 2001). Crucifers (Brassicaceae) represent the only plant family to date in which homeologous chromosome regions/chromosomes have successfully been analyzed with chromosome-specific painting probes. Comparative chromosome painting (CCP) in Brassicaceae is feasible due to the availability of virtually repeat-free *Arabidopsis thaliana* BAC clones and the preferential location of highly repetitive DNA sequences at pericentromeric heterochromatin in most crucifer species. Therefore, CCP with selected repeat-free *Arabidopsis* BAC contigs can reveal collinear chromosome regions in other crucifer species (Jiang and Gill, 2006; Lysak and Lexer, 2006).

Although CCP reveals the extent of chromosome collinearity between *A. thaliana* and other species, it provides only limited

information on the direction of karyotype evolution. To ascertain the evolutionary progression of karyotypic alterations, it is necessary to distinguish ancestral from derived karyotypes. Comparative genetic and cytogenetic analyses showed that the compact *A. thaliana* genome is characterized by a highly reshuffled and derived karyotype (Boivin et al., 2004; Kuittinen et al., 2004; Koch and Kiefer, 2005; Lysak et al., 2006), making this species inappropriate as a reference point for comparative studies across Brassicaceae (Schranz et al., 2006). Therefore, an Ancestral Crucifer Karyotype (ACK) with eight chromosomes (AK1 to AK8) and some 24 conserved genomic blocks (GBs; A to X) has been proposed (Figure 1A) (Lysak et al., 2006; Schranz et al., 2006). The ACK is based on the fact that x=8 is the most common base number in the family (Warwick and Al-Shehbaz, 2006) and that the eight chromosomes of *Arabidopsis lyrata* and *Capsella rubella* represent nearly identical linkage groups (Boivin et al., 2004; Kuittinen et al., 2004; Koch and Kiefer, 2005). Also n=6 and n=7 karyotypes share ancestral chromosomes with the ACK (Lysak et al., 2006). Moreover, the five studied extant karyotypes (n=5 to n=8) (Lysak et al., 2006) as well as the allopolyploid genome of *Brassica napus* (Parkin et al., 2005) and the ACK are composed of 24 evolutionarily conserved GBs (Schranz et al., 2006).
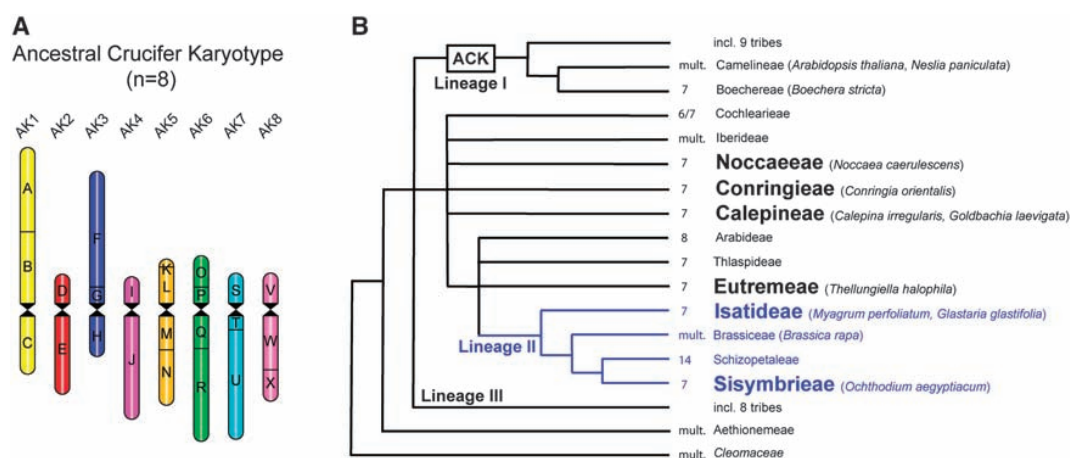
The current intrafamiliar classification of the family Brassicaceae (Al-Shehbaz et al., 2006), based on the *ndhF* phylogeny (Beilstein et al., 2006), includes at least 25 tribes clustered into three major phylogenetic lineages (Lineages I to III, with *A. thaliana* in Lineage I and *Brassica* in Lineage II) or having an unresolved position (Figure 1B). Genetic and cytogenetic

Figure 1. ACK (n=8) and Phylogenetic Relationship of Selected Tribes and Species within the Family Brassicaceae.

(A) Scheme of the ACK of crucifer Lineage I comprising eight chromosomes (AK1 to AK8) and 24 GBs (A to X). Modified from Schranz et al. (2006).
(B) Three major phylogenetic lineages (Lineages I to III) were recognized within Brassicaceae (Beilstein et al., 2006). The six analyzed x=7 tribes (in boldface) are embedded within an unresolved assemblage of tribes including Lineage II (in blue). A tentative phylogenetic position of the ACK, and species analyzed and/or discussed in this study, are given. Base chromosome number (x) is indicated for each tribe (mult., multiple base numbers). The tree is modified from Koch and Al-Shehbaz (2008).

analyses of cross-species chromosome homeology have mostly focused on close relatives of *A. thaliana* within Lineage I (Lysak et al., 2003, 2006; Koch and Kiefer, 2005; Schranz et al., 2007) and to a lesser extent on species of the tribe Brassiceae, belonging to Lineage II (Lysak et al., 2005; Parkin et al., 2005). Here we investigate (1) to what extent tribes of Lineage II and potentially affiliated tribes (Figure 1B), mostly possessing x=7 karyotypes, are congruous with the tentative ACK (n=8) of Lineage I; (2) whether the x=7 karyotypes descended from a common n=7 ancestral karyotype or originated repeatedly within this group of tribes (involving different rearrangements each time); and (3) to what degree collinear GBs are conserved among the chromosomes of x=7 species. Furthermore, we searched for clade- and tribe-specific chromosome rearrangements (cytogenetic signatures) that could help resolve the ambiguous tribal polytomy associated with Lineage II and affiliated x=7 tribes (Figure 1B) (Bailey et al., 2006; Beilstein et al., 2006; Koch et al., 2007; Koch and Al-Shehbaz, 2008).

To address the above questions, the pattern of karyotype evolution has been reconstructed for eight x=7 (2n=14, 28) species from six tribes by CCP with *Arabidopsis* BAC contigs arranged according to the eight ancestral chromosomes (AK1 to AK8) and 24 ancestral GBs (A to X) of the ACK. We analyzed three species of Lineage II (*Glastaria glastifolia*, *Myagrum perfoliatum* [both of tribe Isatideae], and *Ochthodium aegyptiacum* [assigned to Sisymbrieae; K. Mummenhoff, personal communication]) and five species of associated tribes (*Calepina irregularis*, *Goldbachia laevigata* [Calepineae], *Conringia orientalis* [Conringieae], *Noccaea* [*Thlaspi*] *caerulescens* [Noccaeeae], and *Thellungiella halophila* [=*Eutrema salsugineum*; Eutremeae]). Among these species, *N. caerulescens* is an important model of heavy metal accumulation (Peer et al., 2006) and *T. halophila* is an important model of salt and cold tolerance (Gong et al., 2005).

RESULTS

CCP Analysis Suggests a Common Ancestry of the x=7 Species and the ACK

The revealed CCP patterns are here described in relation to the ACK (n=8; Figure 1A) as a reference genome. All 24 (A to X) conserved GBs of the ACK were unambiguously identified in chromosome complements of the analyzed species by CCP with *A. thaliana* BAC contigs as probes (see Methods and Supplemental Table 1 online). In *C. orientalis* (2n=2x=14; Conringieae), painting probes for ancestral chromosomes AK1 to AK4 and AK7 labeled a single pachytene bivalent or mitotic chromosome pair each (Figure 2A). *C. irregularis* and *G. laevigata* (2n=4x=28; Calepineae) showed the karyotype structure similar to that described for *Conringia*, but all five AK-like chromosomes were present in two copies due to a putative genome duplication (Figures 2B and 2C; see Supplemental Figure 1 online). Figure 2D shows the reconstructed karyotype of *N. caerulescens* (2n=2x=14; Noccaeeae). Similar to the pattern observed in Calepineae and Conringieae, five *Noccaea* homeologues resemble AK1 to AK4 and AK7 chromosomes. However, all but one AK-like chromosomes exhibit specific inversions (Figure 2D; see Supplemental Figure 2 online). The repatterning of AK1- and AK4-like homeologues was probably caused by two subsequent pericentric inversions, and the structure of the AK3-like chromosome can be explained by a paracentric and subsequent pericentric inversion (see Supplemental Figure 2 online). The AK7-like homeologue probably underwent a single pericentric inversion. The analyzed species from the three remaining tribes (all 2n=2x=14), *M. perfoliatum* and *G. glastifolia* (both Isatideae), *O. aegyptiacum* (Sisymbrieae), and *T. halophila* (Eutremeae),
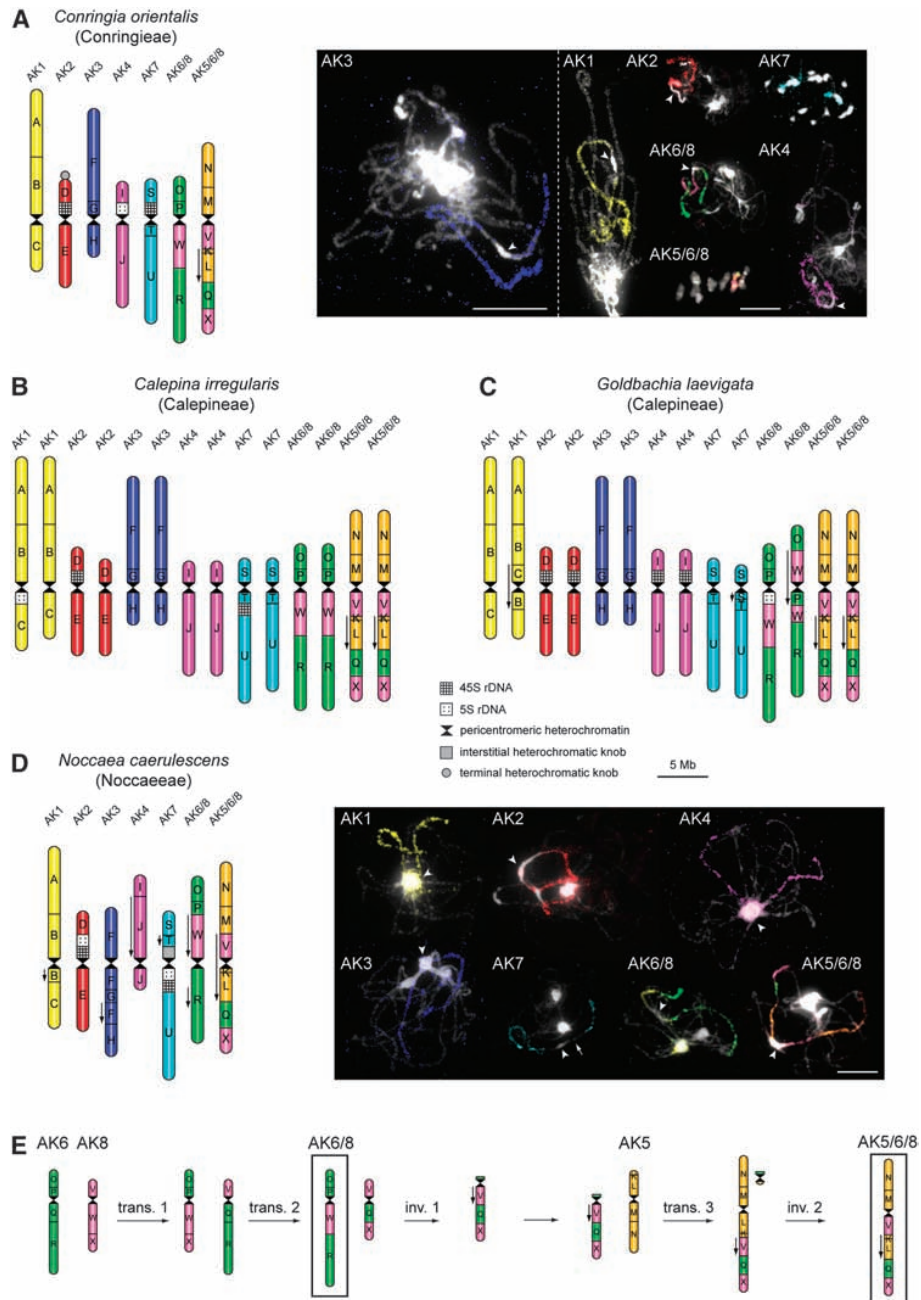
Figure 2. Reconstructed Karyotypes of Conringieae, Calepineae, and Noccaeeae Species.

(A) Karyotype of *C. orientalis* (2n=14; Conringieae) and chromosomes of this species revealed by CCP.
(B) Karyotype of *C. irregularis* (2n=28; Calepineae).
(C) Karyotype of *G. laevigata* (2n=28; Calepineae).
(D) Karyotype of *N. caerulescens* (2n=14; Noccaeeae) and *Noccaea* chromosomes after CCP.
(E) Tentative scenario of the origin of translocation chromosomes AK6/8 and AK5/6/8 from ancestral chromosomes AK5, AK6, and AK8.
The 24 GBs are indicated by uppercase letters (A to X) and colored according to their positions on chromosomes AK1 to AK8 of the ACK (Figure 1A). Downward-pointing arrows indicate the opposite orientation of GBs compared with the position in the ACK. Karyotypes in **(A)** to **(D)** are drawn to scale (bar = 5 Mb). In CCP images, arrowheads point to pericentromeric heterochromatin and the arrow indicates the heterochromatic knob. Bars = 10 μm.

share four homeologues resembling the ancestral chromosomes AK1, AK3, AK4, and AK7 (Figures 3A and 3C to 3E).

The identification of at least four AK-like homeologues and GB associations corresponding to those of the ACK across all analyzed tribes strongly suggests that the x=7 taxa and the ACK descended from a common ancestor. This was further corroborated by comparing the relative lengths of the 24 GBs among all eight species (see Supplemental Table 2 online) and with megabase size estimations for these segments in *A. thaliana* (see Supplemental Table 1 online). There was no significant difference between relative lengths of GBs among the analyzed taxa (analysis of variance; P = 0.008).
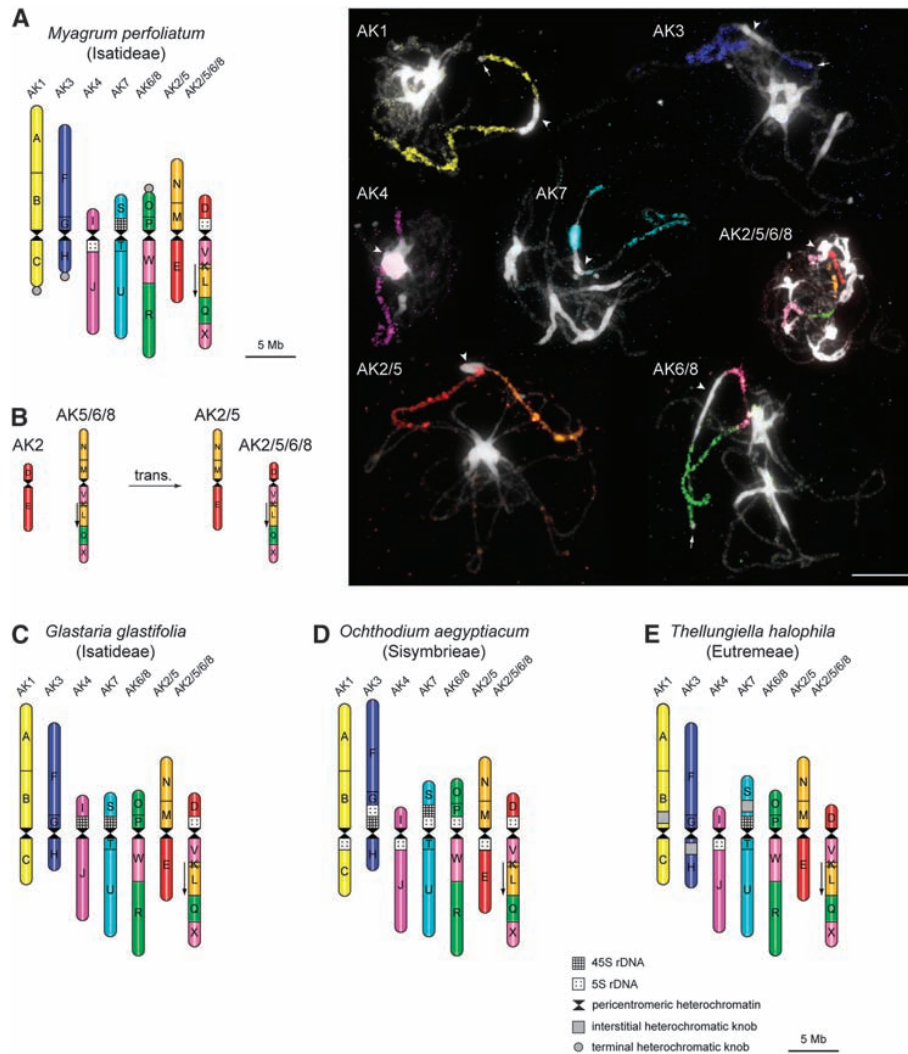
Figure 3. Reconstructed Karyotypes of Eutremeae, Isatideae, and Sisymbrieae Species.

(A) Karyotype of *M. perfoliatum* (2n=14; Isatideae) and *Myagrum* chromosomes after CCP.
(B) A tentative scenario of the origin of translocation chromosomes AK2/5 and AK2/5/6/8 from chromosomes AK2 and AK5/6/8.
(C) Karyotype of *G. glastifolia* (2n=14; Isatideae).
(D) Karyotype of *O. aegyptiacum* (2n=14; Sisymbrieae).
(E) Karyotype of *T. halophila* (2n=14; Eutremeae).
The 24 GBs are indicated by uppercase letters (A to X) and colored according to their positions on chromosomes AK1 to AK8 of the ACK (Figure 1A). Downward-pointing arrows indicate the opposite orientation of GBs compared with the position in the ACK. In *O. aegyptiacum*, it was not possible to distinguish whether the 5S rDNA loci are located within a pericentromeric region of short or long arms of the analyzed mitotic chromosomes; thus, the 5S rDNA positions are only approximate. Karyotypes in **(A)** and **(C)** to **(E)** are drawn to scale (bar = 5 Mb). In CCP images, arrowheads refer to pericentromeric heterochromatin and arrows indicate heterochromatic knobs. Bar = 10 μm.

Structure and Origin of the Rearranged Chromosomes: the Proto-Calepineae Karyotype

Besides AK-like chromosomes, two chromosomes comprising novel associations of GBs were identified in the analyzed x=7 species. The most parsimonious origin of the two chromosomes (AK6/8 and AK5/6/8) has been reconstructed assuming that the two rearranged chromosomes arose as a result of chromosome number reduction from n=8 to n=7 and that the ancestral eight chromosomes resembled AK chromosomes of the ACK.

In Conringieae (*C. orientalis*) and Calepineae (*C. irregularis* and *G. laevigata*), chromosome AK6/8 consists of AK6-derived blocks O, P, and R, with the block W assigned to AK8. The AK6/8 chromosome probably originated from two subsequent reciprocal translocations between AK6 and AK8; the AK6/8 centromere was derived from AK6 or AK8 (Figure 2E). The initial translocation event led to a whole-arm exchange followed by a reciprocal translocation involving blocks R and X. The translocation chromosome V/Q/X became acrocentric via a pericentric inversion involving its short arm (block V). Then a reciprocal translocation transferred the long arm to the short arm end of AK5. The small translocation product (a minichromosome bearing the V/Q/X centromere) was apparently meiotically unstable and thus became eliminated. The AK5/6/8 was further modified by a paracentric inversion (Figure 2E). AK5/6/8 comprises the short arm and the centromere from AK5 (blocks M/N), whereas its long arm is composed of GBs from AK5 (K/L), AK6 (Q), and AK8 (V/X) (Figure 2). In Noccaeeae (*N. caerulescens*), AK6/8 and AK5/6/8 chromosomes were altered by secondary intrachromosomal rearrangements (Figure 2D). The structure of AK6/8 can be explained by a pericentric and a subsequent paracentric inversion, whereas the AK5/6/8 chromosome probably experienced a single pericentric inversion (see Supplemental Figure 2 online).

In Eutremeae, Isatideae, and Sisymbrieae (Figure 3), the AK6/8 chromosome is identical to that described in Conringieae, Calepineae, and Noccaeeae. The translocation chromosome AK2/5/6/8 bears blocks V/K/L/Q/X on its long arm, as in all other taxa; however, the short arm comprises block D of AK2. Blocks M/N (from AK5) and E (from AK2) constitute the complementary product (AK2/5) of a reciprocal translocation between AK5/6/8 and the ancestral chromosome AK2. The centromere identity in the two translocation chromosomes cannot be inferred from the present data (Figure 3B). In *G. glastifolia*, the AK2/5/6/8 chromosome is modified by a paracentric inversion on its long arm (Figure 3C).

Based on the above data, a tentative Proto-Calepineae Karyotype (PCK; n=7) comprising five ancestral chromosomes (AK1 to AK4 and AK7) and two translocation chromosomes (AK6/8 and AK5/6/8) has been deduced. Following the most parsimonious interpretation, the PCK is likely the ancestor of all x=7 karyotypes analyzed (Figure 4).

Comparative Analysis of Heterochromatic Chromosomal Landmarks

Chromosome homeology and chromosome rearrangements revealed by CCP have been further compared with prominent heterochromatin and repeat block landmarks of the analyzed species. Heterochromatic arrays of pachytene bivalents were identified by 4′,6-diamidino-2-phenylindole staining and by fluorescence in situ hybridization localization of 5S and 45S rDNA. Species-specific heterochromatin profiles were compared with estimated genome size values.

The number and position of rDNA loci found in the x=7 species are given in Table 1 and in Figures 2A to 2D, 3A, and 3C to 3E. In all species, 45S rDNA loci (NORs) have been detected adjacent to pericentromeric heterochromatic regions. In *C. orientalis*, *C. irregularis*, *G. laevigata*, and *N. caerulescens*, a 45S rDNA locus was consistently found adjacent to block D on the AK2-like chromosome and frequently on the AK7-like chromosome. In *G. glastifolia*, *M. perfoliatum*, *O. aegyptiacum*, and *T. halophila*, AK2 arms are separated by translocation and the 45S rDNA locus was no longer found to be associated with block D on chromosome AK2/5/6/8. Another locus is always associated with block S on the AK7-like homeologue. Hence, the location of NORs on chromosomes AK2 and AK7 may represent an ancestral condition of the PCK (Figure 4).

The 5S rDNA loci are localized within pericentromeric heterochromatic regions of different homeologues in species of Conringieae, Calepineae, and Noccaeeae (Figures 2A to 2D). In tribes Eutremeae, Isatideae, and Sisymbrieae, the 5S rDNA is most frequently found on the short arm of chromosome AK2/5/6/8 (adjacent to block D) and/or on the long arm of AK4-like homeologue (Figures 3A and 3C to 3E). Remarkably, 5S rDNA loci are present within pericentromeric regions of all seven chromosome pairs in *O. aegyptiacum* (Figure 3D; see Supplemental Figure 3A online).

In tetraploid Calepineae species, rDNA loci were often located on only one of the two homeologues. In *C. irregularis*, both 5S and 45S rDNA loci were located on only one of the two homeologues (Figure 2B; see Supplemental Figure 1 online), whereas in *G. laevigata*, a 5S rDNA locus was situated within one of the two homeologues, while 45S rDNA loci occupied pericentromeric positions on both homeologues (Figure 2C; see Supplemental Figure 1 online).

Except for rDNA arrays, terminal and interstitial heterochromatic knobs were discerned in three species. In *C. orientalis*, a terminal knob is located on the short arm of AK2 homeologue (Figure 2A). Three terminal knobs (on AK1-like, AK3-like, and AK6/8 chromosomes) were revealed in *M. perfoliatum* (Figure 3A). In *T. halophila*, three large interstitial knobs are located close to centromeres of AK1-, AK3-, and AK7-like chromosomes (Figure 3E).

The monoploid genome size (C_x) varied from 0.20 pg (223 Mb) in *C. orientalis* to 0.66 pg (648 Mb) in *T. halophila* (Table 1). No apparent relationship was found between genome size variation and heterochromatin amount, chromosome homeology pattern, or the phylogenetic position of the x=7 species. Although the highest C value in *T. halophila* can be tentatively linked to the presence of three interstitial knobs and the large arrays of pericentromeric heterochromatin revealed in this species (see Supplemental Figure 3B online), a comparable genome size (0.61 pg/600 Mb) was estimated in *N. caerulescens* possessing only one knob and less extensive pericentromeric heterochromatic regions (Figure 2D).
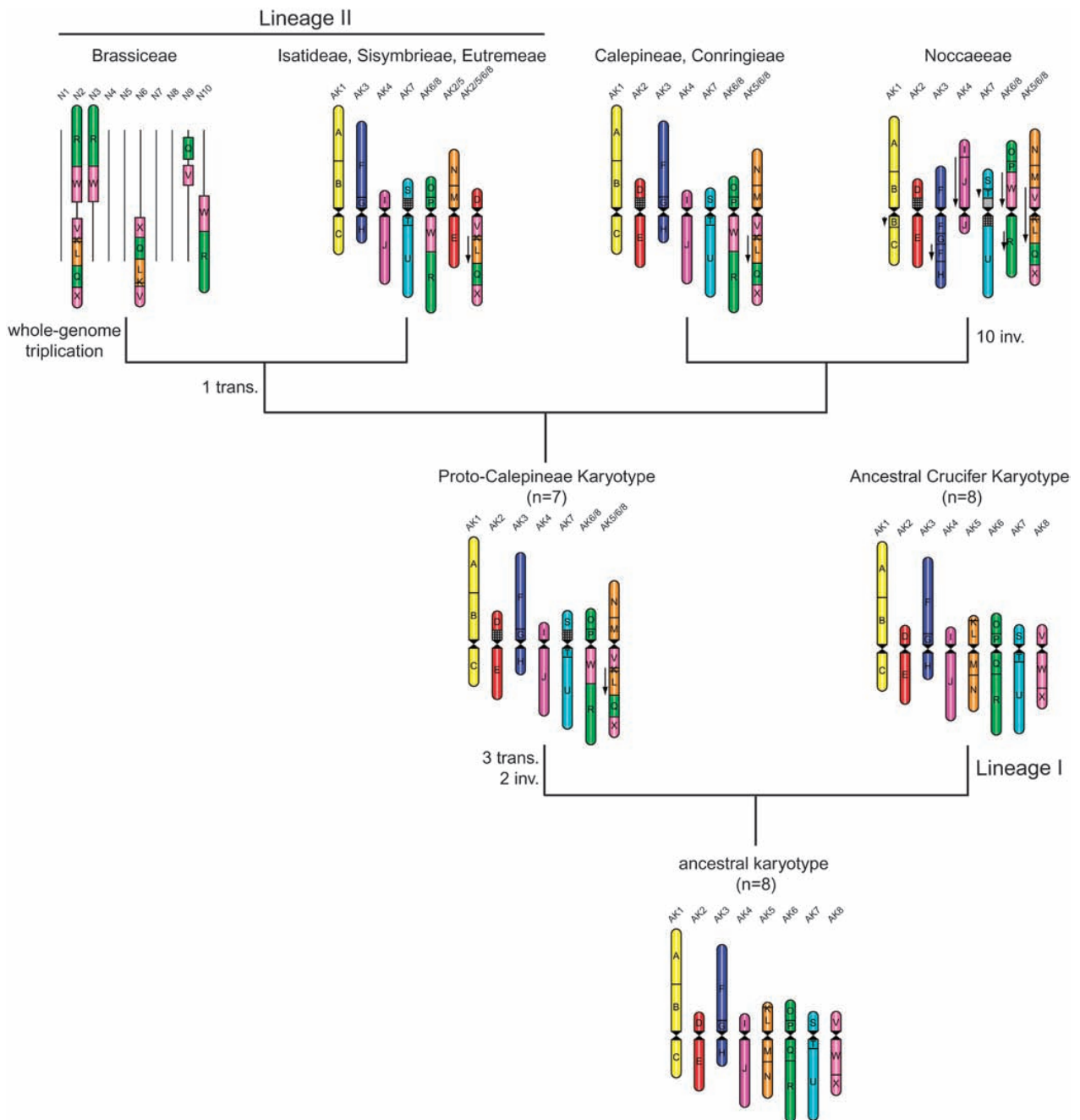
Figure 4. Reconstruction of Karyotype Evolution in Six x=7 Tribes and Brassiceae from the PCK (n=7) and an Ancestral Karyotype (n=8) Based on CCP Data.

The 24 GBs are indicated by uppercase letters (A to X) and colored according to their position on chromosomes AK1 to AK8 of the ACK (Figure 1A). Downward-pointing arrows indicate the opposite orientation of GBs compared with the position in the ACK. A tentative number of translocations (transl.) and inversions (inv.) is given at the nodes of the chromosomal phylogeny. 45S rDNA loci are shown as cross-hatched boxes. For the tribe Brassiceae (exemplified by the *B. rapa* karyotype, n=10; Parkin et al., 2005), only associations of GBs shared with the PCK are displayed; other chromosomes/chromosome regions are shown as black bars.

Table 1. Chromosome Number, Monoploid Genome Size (C_x), and Number of rDNA Loci of the x=7 Species Analyzed

Species	2n	C_x (pg/Mb)[a]	No. of 45S rDNA Loci	No. of 5S rDNA Loci
Species with two translocation chromosomes				
Conringia orientalis	14	0.228/223	4	2
Calepina irregularis	28	0.202/396	4	2
Goldbachia laevigata	28	0.218/427	8	2
Noccaea caerulescens (=*Thlaspi caerulescens*)	14	0.612/600	4	4
Species with three translocation chromosomes				
Myagrum perfoliatum	14	0.312/305	2	4
Glastaria glastifolia	14	0.425/416	4	2
Ochthodium aegyptiacum	14	0.336/329	4	14
Thellungiella halophila (=*Eutrema salsugineum*)	14	0.661/648	2	2

[a] C_x values based on the C_x value of *A. thaliana* (0.16 pg/157 Mb).

DISCUSSION

We have reconstructed the karyotype evolution of eight x=7 species from six Brassicaceae tribes, thereby reconstructing by chromosome analyses the extent of whole-genome collinearity among a group of plant species for which no comparative cytogenetic and genetic maps were previously available. CCP using *Arabidopsis* BAC contigs arranged according to their position within the putative ACK of Lineage I (Lysak et al., 2006; Schranz et al., 2006) was applied to uncover homeologous chromosome regions of the x=7 species.

The ACK and the PCK

Five chromosomes shared between the ACK (n=8) and the PCK (n=7) unequivocally argue for a common origin of the two karyotypes (Figure 4). As the phylogenetic relationship between Lineage I and tribes affiliated with Lineage II remains largely unresolved (Koch and Al-Shehbaz, 2008) (Figure 1B), two alternative scenarios reconstructing the origin of the translocation chromosomes as well as the entire PCK complement have to be considered. The first assumes that a common ancestral karyotype for both clades (i.e., Lineages I and II) possessed eight chromosome pairs, and the second scenario implies a karyotype of seven chromosome pairs (n=7). Based on our current understanding of genome evolution in Brassicaceae, the former theoretical model advocating an ancestral n=8 karyotype for both lineages is preferred. The most parsimonious scenario of chromosome number reduction from n=8 to n=7 requires only three translocations and two inversions (Figure 2E), compared with a higher number of less parsimonious steps necessary to reconstruct the ACK from the PCK. Moreover, the latter model (i.e., n=7 → n=8) would require an origin of new centromere by an unknown mechanism. Although the emergence of de novo centromeres (Nasuda et al., 2005) and centric fissions (Jones, 1998) is documented in plants, these have not been observed in crucifers as yet.

As ancestral chromosomes per se do not provide insight into the evolutionary relationship among the analyzed species, intertribal relationships have been elucidated through evolutionarily novel associations of crucifer GBs. The chromosome AK6/8 (association of O/P/W/R blocks) and the collinearity of V/K/L/Q/X blocks found in all x=7 species serve as unique cytogenetic signatures unambiguously underlying the monophyletic origin of all six tribes (Figure 4). Chromosome AK6/8 and the association of V/K/L/Q/X blocks must have been present already in the PCK. Eutremeae, Isatideae, and Sisymbrieae share a younger reciprocal whole-arm translocation between chromosomes AK5/8/6 and AK2 (Figure 4).

Cytogenetic Signatures: Phylogenetic Implications

Cytogenetic signatures (e.g., novel associations of GBs) can be used to delimit taxa, as chromosome rearrangements generally exhibit only a low level of homoplasy and thus have the power to disentangle unresolved or conflicting phylogenetic relationships. The tribe Eutremeae (~25 species) had an unresolved position within the group of x=7 tribes (Figure 1B) (Beilstein et al., 2006; Koch et al., 2007; Koch and Al-Shehbaz, 2008). We have shown that the karyotype of Eutremeae (*T. halophila*) is identical to that of Isatideae and Sisymbrieae but different from that of Calepineae, Conringieae, and Noccaeeae. Hence, this supports the inclusion of Eutremeae together with Brassiceae, Isatideae, Schizopetaleae, and Sisymbrieae in Lineage II, forming a monophyletic group (Figure 4). The data here corroborate the earlier exclusion of Calepineae and Conringieae from the Brassiceae (Lysak et al., 2005). Although Calepineae and Conringieae possess similar karyotypes, differences in morphological characters substantiate their recent circumscription as two closely related tribes (German and Al-Shehbaz, 2008). CCP data are also congruent with the close relationship between Conringieae and the tribe Noccaeeae revealed by nuclear and chloroplast gene phylogenies (Bailey et al., 2006; Beilstein et al., 2006; German and Al-Shehbaz, 2008). The reconstructed chromosomal phylogeny elucidated relationships among the analyzed x=7 tribes and provided compelling evidence that tribes of Lineage II (including Eutremeae) and Calepineae, Conringieae, and Noccaeeae have a monophyletic origin (Figure 4). Further CCP analysis of other tribes currently placed into the phylogenetic proximity of Lineage II (Figure 1B) is needed to further clarify phylogenetic relationships and genome evolution within this crucifer group.

By analyzing species with an identical chromosome number (2n=14 or 28) from six closely related but distinct tribes, we also tested for the functional role of gross karyotypic alterations during speciation. Our data indicate that even morphologically distinct groups such as Calepineae/Conringieae/Noccaeeae, when compared with tribes of Lineage II, display almost complete genome collinearity and very similar karyotypes. It remains elusive to what extent the Lineage II–specific translocation resulting in chromosomes AK2/5 and AK2/5/6/8 contributed to the divergence of Lineage II tribes. However, the resolution of CCP analysis is constrained by the size of BAC painting probes, reducing the ability to detect minor rearrangements that might be of evolutionary significance. The above findings indicate that the speciation resulting in morphologically distinct but closely related taxa (Al-Shehbaz et al., 2006; German and Al-Shehbaz, 2008) is not necessarily accompanied or determined by major karyotypic changes.

Karyotype Evolution Preceding the Whole-Genome Triplication in Brassiceae

As Isatideae, Sisymbrieae, and Eutremeae along with tribes Brassiceae and Schizopetaleae form a monophyletic group called Lineage II (Beilstein et al., 2006; Koch et al., 2007; this study), some conclusions on the early evolution of the whole lineage can be drawn from the data presented here. The chromosome homeology patterns of Isatideae, Sisymbrieae, and Eutremeae can be compared with the structure of *Brassica* genomes (tribe Brassiceae) revealed by comparative genetic analysis (Parkin et al., 2005). This comparison showed that the R/W block association (=long arm of AK6/8) as well as the V/K/L/Q/X association are present also in the *Brassica rapa* component (n=10) of the *B. napus* genome (n=19) (Figure 4) (Parkin et al., 2005; Schranz et al., 2006). Three copies of the R/W block association are located on *B. napus* chromosomes N2, N3, and N10, whereas the V/K/L/Q/X association has been found in two copies on chromosomes N2 and N6 (only blocks Q and V of the third genomic copy were detected on chromosome N9 [Figure 4]). Due to the whole-genome triplication in the tribe Brassiceae (Lysak et al., 2005; Parkin et al., 2005), three instead of two copies of the V/K/L/Q/X association are to be expected. Nevertheless, this pattern corresponds well with the fragmentary character of *Brassica* genomes, as the genomic redundancy after the whole-genome triplication was followed by the diploidization eroding the original triplicated pattern. Consequently, some GBs were retained in three copies, other GBs as well as smaller segments were preserved as one or two copies, and some were secondarily multiplied, probably through unequal homeologous recombination (Lukens et al., 2004; Parkin et al., 2005). The present structure of the *B. napus* genome does not allow us to arrive at a clear-cut conclusion whether the AK5/6/8–AK2 translocation is shared by all tribes of Lineage II or whether it is specific only for Isatideae, Sisymbrieae, and Eutremeae. Regardless of this limitation, the aforementioned cytogenetic signatures provide convincing evidence that Brassiceae and the whole Lineage II descended from the PCK (n=7). In the clade leading to Brassiceae, the PCK karyotype has undergone the whole-genome triplication. This parsimonious scenario does not

rule out the possibility that the tribe-specific triplication was preceded by another round of chromosome number reduction (n=7 → n=6), as x=6 was suggested as an ancestral base chromosome number of *Brassica* species by several authors (Catcheside, 1937; Röbbelen, 1960). However, the present data do not warrant any conclusion regarding the chromosome number of the primary hexaploid (Lysak et al., 2007).

Mechanism of Chromosome Number Reduction from n=8 to n=7

Chromosome rearrangements revealed in the x=7 taxa can be reconstructed by considering only inversions and reciprocal translocations. Chromosome number reduction from n=8 to n=7 implies the loss or inactivation of one centromere. Our data indicate that the centromere loss occurred as a result of a reciprocal translocation between an acrocentric chromosome (formed via a preceding pericentric inversion) and a (sub)metacentric chromosome. This scenario yields a translocation chromosome (AK5/6/8) and a minichromosome comprising mainly the centromere of the acrocentric chromosome (Figure 2E). The meiotically unstable minichromosome is eliminated. The same mechanism has been argued as responsible for chromosome number reduction in tribes of Lineage I (Lysak et al., 2006) and is assumed to be a prominent mechanism reducing chromosome numbers in Brassicaceae (Schranz et al., 2006; Schubert, 2007). Other chromosome rearrangements identified in the x=7 species included pericentric and paracentric inversions and whole-arm reciprocal translocations. No evidence for duplications or deletions of GBs has yet been detected.

We have scrutinized the number of rearrangement events and corresponding break points necessary to explain the structure of the extant x=7 karyotypes. The origin of the PCK (i.e., translocation chromosomes AK6/8 and AK5/6/8) from an ancestral n=8 karyotype requires five events and nine break points (five in centromeric, two in terminal, and two in interstitial regions at W/X and Q/R boundaries). The reciprocal translocation between chromosomes AK2 and AK5/6/8 yielding the Lineage II–specific chromosomes AK2/5 and AK2/5/6/8 involves another two pericentromeric break points. In *G. laevigata*, six species-specific break points have been identified (one pericentromeric and five interstitial, including one coinciding with the O/P boundary). In *Glastaria*, 11 plus 2 species-specific break points at K/L and Q/X boundaries were identified. The highest number of break points is in *N. caerulescens*, with 9 break points shared with the PCK and 20 additional break points necessary to account for species-specific pericentric and paracentric inversions (Figure 2D; see Supplemental Figure 2 online). Of the 20 break points, 10 map to pericentromeric regions, whereas 10 are at interstitial positions (two at T/U and W/R boundaries). In total, of the 11 basic break points necessary for the origin of the three translocation chromosomes, 9 map to pericentromeric or telomeric regions and only 2 are interstitial. From a total of 28 secondary species-specific break points, 11 localize to centromeres. Thus, from all 39 break points, half (20 break points) involve telomeric/pericentromeric regions and the other half (19 break points) occur at interstitial positions. Although pericentromeric and telomeric regions characterized by clusters of tandem repeats are

considered as chromosome regions prone to chromosome rearrangements, most likely by etopic (nonallelic) homologous recombination (Gaut et al., 2007; Schubert, 2007), the observed equal frequency of interstitial break points suggests that large clusters of repetitive sequence elements are not the exclusive sites of chromosomal breakage. Segmental duplications were proposed as regions facilitating chromosome rearrangements, such as inversions in primate genomes (Kehrer-Sawatzki and Cooper, 2008). In *Drosophila*, duplications of nonrepetitive sequences at the break point regions of ∼60% of the analyzed inversions were presumably caused by staggered breaks rather than by ectopic recombination between repetitive elements (Ranz et al., 2007). Future genome sequencing of the x=7 species (e.g., *N. caerulescens* and *T. halophila*) will unveil the detailed sequence structure of the break points.

Diverse Routes of Chromosome Number Reduction from n=8 to n=7

So far, data on karyotype evolution in Brassicaceae suggest that at least two major phylogenetic branches (Lineage I and II) descended from a common ancestral karyotype with eight chromosome pairs. We assume that in some crucifer taxa, the number of ancestral linkage groups was reduced toward karyotypes with n=5 to n=7. One crucial question is whether chromosome number reduction within diverse phylogenetic lineages is based on the same or on different events. Among crucifer x=7 taxa, the genome structure of only two species from Lineage I, *Neslia paniculata* and *Boechera stricta*, was analyzed with a sufficient precision (Lysak et al., 2006; Schranz et al., 2007). In *N. paniculata* (Camelineae), the reduction combined the ancestral chromosomes AK4 and AK5. Both chromosomes became acrocentric via pericentric and paracentric inversions and fused by a translocation with break points close to the centromeres. The chromosome AK4/5 comprises blocks I/J (from AK4) on the short arm and the K/L/M/N block association (from AK5) on its long arm (Lysak et al., 2006). In *B. stricta* (Boechereae), chromosome number reduction followed a more complex scenario than in *Neslia* (Schranz et al., 2007). The reduction involved translocations between five AK chromosomes and resulted in translocation chromosomes AK1/2 (blocks A/B/C/D), AK3/8 (blocks F/G/W/X), and AK3/5/8 (blocks K/L/M/N/V/H). In *Neslia* and *Boechera* (Lineage I) as well as in the x=7 tribes affiliated with Lineage II, the ancestral chromosome AK5 is always involved in chromosome number reduction; however, the changes leading to the reduction differ. In *Neslia* and *Boechera*, all AK5-derived GBs (K/L/M/N) remain associated within new chromosomes, whereas in the x=7 species analyzed here, the GBs are split up within AK5/6/8 or between chromosomes AK2/5 and AK2/5/6/8 (Figure 4). The comparison here clearly shows that chromosome number reduction in Lineage I (Camelineae and Boechereae) is different from that of the x=7 tribes associated with Lineage II. The repeated participation of certain ancestral chromosomes in diverse rearrangements was also observed in other groups such as grasses (Srinivasachary et al., 2007).

Heterochromatic Landmarks in the x=7 Species

The analyzed x=7 species show a rather uniform pattern of rDNA loci and heterochromatic knobs. In our analyzed x=7 species, all NORs were located interstitially, adjacent to pericentromeric heterochromatic regions. This is in contrast with the prevalence of terminal NORs found elsewhere in the family (Ali et al., 2005), including Camelineae and Descurainieae species analyzed by CCP (Berr et al., 2006; Lysak et al., 2006).

The number of 5S rDNA loci was two or four, except in *O. aegyptiacum*, with 14 loci (Table 1; see Supplemental Figure 3 online). The remarkable localization of 5S rDNA on all chromosomes of a complement is not yet known in higher plants (Cloix et al., 2003). This high number of heterochromatic 5S rDNA sites is most likely responsible for the severe stickiness of prophase I bivalents at pericentromeric heterochromatic regions in this species. Similarly, better spread pachytene bivalents were obtained in *A. thaliana* accessions with only two major 5S rDNA loci compared with the accessions having 5S rDNA on three chromosomes (Fransz et al., 1998). It remains unclear whether the chromosomal ubiquity of 5S rDNA in *O. aegyptiacum* was caused by the mobility of transposable elements, similar to the *En/Spm* transposon–mediated transfer of 5S rDNA observed in *Aegilops* (Raskina et al., 2004, 2008).

The frequent mobility of 45S and 5S rDNA sites (Pontes et al., 2004; Schubert, 2007; Raskina et al., 2008) precludes any plausible assumption on the ancestral distribution of rDNA sites. While the loss of some NORs can occur during chromosome number reduction events (Lysak et al., 2006; Schranz et al., 2006; Schubert, 2007) and other rDNA clusters can be amplified or transposed via ectopic recombination or due to the activity of transposable elements (Pontes et al., 2004; Raskina et al., 2004, 2008), direct evidence for the involvement of rDNA repeats in crucifer chromosome rearrangements is lacking.

Three x=7 species possess heterochromatic knobs positioned terminally or interstitially. In plants, knobs can originate through inversions, including pericentromeric heterochromatin (e.g., the hk4S knob in *A. thaliana* [Fransz et al., 2000]), the insertion and subsequent silencing of repetitive elements such as tandem repetitive transgenes (Probst et al., 2003), or can be copied from other chromosomal regions due to the physical proximity within interphase nuclei (Zhong et al., 1998). Apart from the hk4S knob in *A. thaliana* (Fransz et al., 2000) and small interstitial as well as distal knobs in *Brassica* species (Röbbelen, 1960; Lim et al., 2005; Lysak et al., 2007), large terminal heterochromatic knobs are not yet reported in Brassicaceae species. These terminal knobs resemble the knobs on chromosomes of maize (*Zea mays*; Lamb et al., 2007), rice (*Oryza sativa*; Ohmido et al., 2001), or tomato (*Solanum lycopersicum*; Zhong et al., 1998). As in these species, it is likely that knobs in crucifer species will primarily comprise tandem repeats. Satellite repeats composing distal knobs in the above species were classified as subtelomeric with a potential functional role (Lamb et al., 2007). Three large interstitial knobs within the *T. halophila* karyotype (Figure 3E) are not linked to a higher frequency of species-specific chromosome rearrangements, as seen in *N. caerulescens*, a species with only a single knob (Figure 2D). Our data suggest that interstitial as well as terminal knobs, revealed only in *C. orientalis* (Figure 2A) and *M. perfoliatum* (Figure 3A), are apparently of more recent origin and did not mediate the described chromosome repatterning.

Cytogenetic Evidence of Genome Diploidization in Autotetraploid Calepineae Species

In this study, the complete karyotype structure of two tetraploid Calepineae species was analyzed. In *C. irregularis*, diploid (2n=2x=14) and tetraploid (2n=4x=28) cytotypes are known, whereas only tetraploids are recorded for *G. laevigata* (Warwick and Al-Shehbaz, 2006). As *Calepina* is a monospecific genus (German and Al-Shehbaz, 2008), we assume that the 4x genome was derived from the 2x cytotype via autopolyploidy relatively recently. Indeed, our CCP data corroborated the autopolyploid origin of the 4x cytotype, as all seven chromosomes showed the same structure in both genomes (Figure 2B). In *G. laevigata*, one pair each of AK1, AK7, and AK6/8 differs from the other pair by pericentric inversions (Figure 2C). Although an allopolyploid origin of the tetraploid *G. laevigata* cannot be ruled out within the genus comprising five other species (German and Al-Shehbaz, 2008), autopolyploidization followed by three species-specific inversions is a more likely interpretation. No records of a diploid cytotype may indicate an earlier origin of the tetraploid *G. laevigata* compared with the 4x cytotype of *C. irregularis*. This is underlined by the pericentric inversions that differentiate three of seven homologous pairs of *G. laevigata*. Although chromosome rearrangements may reduce the risk of meiotic multivalent formation and thus contribute to the genome diploidization of autopolyploids (Ma and Gustafson, 2005), regular meiotic pairing and 14 bivalents have been observed in both species (see Supplemental Figure 1 online). This implies that the diploid-like meiosis in the two autopolyploids is determined genetically rather than by gross intrachromosomal rearrangements. A partial diploidization of meiosis was also observed in established autotetraploid lines of *A. thaliana* (Santos et al., 2003).

METHODS

Plant Material

All but one species were grown from seeds obtained from the Millennium Seed Bank, Royal Botanic Gardens, Kew (*Glastaria glastifolia*, *Ochthodium aegyptiacum*), the Botanical Garden of Bordeaux (*Calepina irregularis*, *Conringia orientalis*), the Hortus Botanicus Hauniensis, Copenhagen (*Goldbachia laevigata*, *Myagrum perfoliatum*), and from D. Baum, University of Wisconsin, Madison (*Thellungiella halophila*). Inflorescences of *Noccaea caerulescens* were collected from a single wild population (Kořenec, Czech Republic). Plants grown from seeds were cultivated in a greenhouse or growth chamber under long-day light conditions. Herbarium vouchers of all analyzed species are deposited in the herbarium of Masaryk University.

Preparation of Pachytene Chromosomes

Entire inflorescences were fixed in ethanol:chloroform:acetic acid (6:3:1) overnight and stored in 70% ethanol at $-20°C$ until use. Selected flower buds were rinsed in distilled water (2 × 5 min) 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, transferred into citrate buffer, and kept at 4°C until use (overnight to 2 d). Individual anthers were put on a microscope slide of a stereomicroscope and disintegrated by a needle in a drop of citrate buffer. Then the suspension was softened by adding 15 to 30 μL of 60%

acetic acid and spread by stirring with a needle 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 tilted to remove the fixative and dried with a hair dryer. The dried preparation was staged with a phase contrast microscope. Suitable slides were postfixed in 4% formaldehyde dissolved in distilled water for 10 min and air-dried.

Painting Probes

Arabidopsis thaliana BAC clones were obtained from the ABRC. The selection of clones suitable for chromosome painting and DNA isolation from individual BAC clones was performed as described by Lysak et al. (2003). For CCP, each third BAC was used and contigs were arranged and differentially labeled according to the GBs of the ACK (Schranz et al., 2006). For BAC contigs used for CCP, see Supplemental Table 1 online. To characterize species-specific inversions in *N. caerulescens* and *G. laevigata*, the respective BAC contigs were split into smaller subcontigs and used as CCP probes. The BAC clone T15P10 (AF167571) bearing 45S rRNA genes was used for in situ localization of NORs, and clone pCT 4.2, corresponding to a 500-bp 5S rRNA repeat (M65137), was used for localization of 5S rDNA loci.

All DNA probes were labeled by nick translation with biotin-dUTP, digoxigenin-dUTP, or Cy3-dUTP as follows: 1 μg of DNA diluted in distilled water to 29 μL, 5 μL of nucleotide mix (2 mM dATP, dCTP, and dGTP, 400 μM dTTP; all Roche), 5 μL of 10× NT buffer (0.5 M Tris-HCl, pH 7.5, 50 mM MgCl$_2$, and 0.05% BSA), 4 μL of 1 mM X-dUTP (in which X was biotin, digoxigenin, or Cy3), 5 μL of 0.1 M β-mercaptoethanol, 1 μL of DNase I (Roche), and 1 μL of DNA polymerase I (Fermentas). The nick translation mixture was incubated at 15°C for 90 min (or longer) to obtain a fragment length of ∼200 to 500 bp. The nick translation reaction was stopped by adding 1 μL of 0.5 M EDTA, pH 8.0, and incubation at 65°C for 10 min. Individual labeled probes were stored at $-20°C$ until use.

CCP

To remove cytoplasm, the slides were treated with pepsin (0.1 μg/μL; Sigma-Aldrich) in 0.01 M HCl for 3 to 6 h, postfixed in 4% formaldehyde dissolved in 2× SSC (1× SSC is 0.15 M NaCl and 0.015 M sodium citrate) for 10 min, dehydrated in an ethanol series (70, 80, and 96%), and air-dried.

Selected BAC clones were pooled and precipitated by adding one-tenth volume of 3 M sodium acetate, pH 5.2, and 2.5 volumes of ice-cold 96% ethanol, kept at $-20°C$ for 30 min, and centrifuged at 13,000*g* at 4°C for 30 min. The pellet was resuspended in 20 μL of hybridization mix (50% formamide and 10% dextran sulfate in 2× SSC) per slide. Cover slips were framed by rubber cement. The probe and chromosomes were denatured together on a hot plate at 80°C for 2 min and incubated in a moist chamber at 37°C for 39 to 63 h.

Posthybridization washing was performed in 50% formamide (*A. thaliana*, which was used as a probe quality control) or 20% formamide (all other species) in 2× SSC at 42°C. Fluorescent detection was as follows: biotin-dUTP was detected by avidin–Texas Red (Vector Laboratories) and amplified by goat anti-avidin–biotin (Vector Laboratories) and avidin–Texas Red; digoxigenin-dUTP was detected by mouse anti-digoxigenin (Jackson ImmunoResearch) and goat anti-mouse Alexa Fluor 488 (Molecular Probes), and Cy3-dUTP labeled probes were observed directly. Chromosomes were counterstained with 4′,6-diamidino-2-phenylindole (2 μg/mL) in Vectashield (Vector Laboratories).

Fluorescence signals were analyzed with an Olympus BX-61 epifluorescence microscope and AxioCam CCD camera (Zeiss). Images were acquired separately for all four fluorochromes using appropriate excitation and emission filters (AHF Analysentechnik). The four monochromatic images were pseudocolored and merged using Adobe Photoshop CS2 software (Adobe Systems).

Chromosome Measurements and Genome Size Estimation

The length of painted GBs was measured using the ImageJ program (National Institutes of Health). The nuclear DNA content has been estimated by flow cytometry at the Institute of Botany and Zoology, Masaryk University. For details, see Lysak et al. (2007).

Supplemental Data

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

Supplemental Figure 1. Examples of CCP in Tetraploid Calepineae Species (2n=4x=28).

Supplemental Figure 2. Species-Specific Chromosome Rearrangements (Inversion Events) in *N. caerulescens*.

Supplemental Figure 3. Heterochromatic Landmarks in *O. aegyptiacum* and *T. halophila*.

Supplemental Table 1. GBs of the ACK and Corresponding *A. thaliana* BAC Contigs Used as Painting Probes in This Study.

REFERENCES

Ali, H.B.M., Lysak, M.A., and Schubert, I. (2005). Chromosomal localization of rDNA in the Brassicaceae. Genome **48:** 341–346.

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.

Bailey, C.D., Koch, M.A., Mayer, M., Mummenhoff, K., O'Kane, S.L., Jr., Warwick, S.I., Windham, M.D., and Al-Shehbaz, I.A. (2006). Toward a global phylogeny of the Brassicaceae. Mol. Biol. Evol. **23:** 2142–2160.

Beilstein, M.A., Al-Shehbaz, I.A., and Kellogg, E.A. (2006). Brassicaceae phylogeny and trichome evolution. Am. J. Bot. **93:** 607–619.

Berr, A., Pecinka, A., Meister, A., Kreth, G., Fuchs, J., Blattner, F.R., Lysak, M.A., and Schubert, I. (2006). Chromosome arrangement and nuclear architecture but not centromeric sequences are conserved between *Arabidopsis thaliana* and *Arabidopsis lyrata*. Plant J. **48:** 771–783.

Boivin, K., Acarkan, A., Mbulu, R.S., Clarenz, O., and Schmidt, R. (2004). The Arabidopsis genome sequence as a tool for genome analysis in Brassicaceae: A comparison of the Arabidopsis and *Capsella rubella* genomes. Plant Physiol. **135:** 735–744.

Catcheside, D.G. (1937). Secondary pairing in *Brassica oleracea*. Cytologia Fujii Jub. 366–378.

Cloix, C., Tutois, S., and Tourmente, S. (2003). 5S rDNA and 5S RNA in higher plants. Rec. Res. Dev. Plant Mol. Biol. **1:** 207–221.

Devos, K.M. (2005). Updating the 'crop circle.' Curr. Opin. Plant Biol. **8:** 155–162.

Fransz, P., Armstrong, S., Alonso-Blanco, C., Fischer, T.C., Torres-Ruiz, R.A., and Jones, G.H. (1998). Cytogenetics for the model system *Arabidopsis thaliana*. Plant J. **13:** 867–876.

Fransz, P.F., Armstrong, S., de Jong, J.H., Parnell, L.D., van Drunen, G., Dean, C., Zabel, P., Bisseling, T., and Jones, G.H. (2000). Integrated cytogenetic map of chromosome arm 4S of *A. thaliana*: Structural organization of heterochromatic knob and centromere region. Cell **100:** 367–376.

Gaut, B.S., Wright, S.I., Rizzon, C., Dvorak, J., and Anderson, L.K. (2007). Recombination: An underappreciated factor in the evolution of plant genomes. Nat. Rev. Genet. **8:** 77–84.

German, D.A., and Al-Shehbaz, I.A. (2008). Five additional tribes (Aphragmeae, Biscutelleae, Calepineae, Conringieae, and Erysimeae) in the Brassicaceae (Cruciferae). Harv. Pap. Bot. **13:** 165–170.

Gong, Q., Li, P.H., Ma, S.S., Rupassara, S.I., and Bohnert, H.J. (2005). Salinity stress adaptation competence in the extremophile *Thellungiella halophila* in comparison with its relative *Arabidopsis thaliana*. Plant J. **44:** 826–839.

Jiang, J., and Gill, B.S. (2006). Current status and the future of fluorescence in situ hybridization (FISH) in plant genome research. Genome **49:** 1057–1068.

Jones, K. (1998). Robertsonian fusion and centric fission in karyotype evolution of higher plants. Bot. Rev. **64:** 273–289.

Kehrer-Sawatzki, H., and Cooper, D.N. (2008). Molecular mechanisms of chromosomal rearrangement during primate evolution. Chromosome Res. **16:** 41–56.

Koch, M., and Al-Shehbaz, I.A. (2008). Molecular systematics and evolution of "wild" crucifers (Brassicaceae or Cruciferae). In Biology and Breeding of Crucifers, P.K. Gupta, ed (London: Taylor and Francis Group).

Koch, M.A., Dobeš, C., Kiefer, C., Schmickl, R., Klimes, L., and Lysak, M.A. (2007). Supernetwork identifies multiple events of plastid trnF$_{(GAA)}$ pseudogene evolution in the Brassicaceae. Mol. Biol. Evol. **24:** 63–73.

Koch, M.A., and Kiefer, M. (2005). Genome evolution among cruciferous plants: A lecture from the comparison of the genetic maps of three diploid species—*Capsella rubella*, *Arabidopsis lyrata* subsp. *petraea*, and *A. thaliana*. Am. J. Bot. **92:** 761–767.

Kuittinen, H., de Haan, A.A., Vogl, C., Oikarinen, S., Leppala, J., Koch, M., Mitchell-Olds, T., Langley, C.H., and Savolainen, O. (2004). Comparing the linkage maps of the close relatives *Arabidopsis lyrata* and *A. thaliana*. Genetics **168:** 1575–1584.

Lamb, C.J., Meyer, M.J., Corcoran, B., Kato, A., Han, F., and Birchler, A.J. (2007). Distinct chromosomal distributions of highly repetitive sequences in maize. Chromosome Res. **15:** 33–49.

Lim, K.B., et al. (2005). Characterization of rDNAs and tandem repeats in the heterochromatin of *Brassica rapa*. Mol. Cells **19:** 436–444.

Lukens, L.N., Quijada, P.A., Udall, J., Pires, J.C., Schranz, M.E., and Osborn, T.C. (2004). Genome redundancy and plasticity within ancient and recent *Brassica* crop species. Biol. J. Linn. Soc. **82:** 665–674.

Lysak, M., 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.

Lysak, M.A., and Lexer, C. (2006). Towards the era of comparative evolutionary genomics in Brassicaceae. Plant Syst. Evol. **259:** 175–198.

Lysak, M.A., Pecinka, A., and Schubert, I. (2003). Recent progress in chromosome painting of *Arabidopsis* and related species. Chromosome Res. **11:** 195–204.

Ma, X.F., and Gustafson, J.P. (2005). Genome evolution of allopolyploids: A process of cytological and genetic diploidization. Cytogenet. Genome Res. **109:** 236–249.

Nasuda, S., Hudakova, S., Schubert, I., Houben, A., and Endo, T.R. (2005). Stable barley chromosomes without centromeric repeats. Proc. Natl. Acad. Sci. USA **102:** 9842–9847.

Ohmido, N., Kijima, K., Ashikawa, I., de Jong, J.H., and Fukui, K. (2001). Visualization of the terminal structure of rice chromosomes 6 and 12 with multicolor FISH to chromosomes and extended DNA fibers. Plant Mol. Biol. **47:** 413–421.

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.

Peer, W.A., Mahmoudian, M., Freeman, J.L., Lahner, B., Richards, E.L., Reeves, R.D., Murphy, A.S., and Salt, D.E. (2006). Assessment of plants from the Brassicaceae family as genetic models for the study of nickel and zinc hyperaccumulation. New Phytol **172:** 248–260.

Pontes, O., Neves, N., Silva, M., Lewis, M.S., Madlung, A., Comai, L., Viegas, W., and Pikaard, C.S. (2004). Chromosomal locus rearrangements are a rapid response to formation of the allotetraploid *Arabidopsis suecica* genome. Proc. Natl. Acad. Sci. USA **101:** 18240–18245.

Probst, A.V., Fransz, P.F., Paszkowski, J., and Mittelsten Scheid, O. (2003). Two means of transcriptional reactivation within heterochromatin. Plant J. **33:** 743–749.

Ranz, J.M., Maurin, D., Chan, Y.S., von Grotthuss, M., Hillier, L.W., Root, J., Ashburner, M., and Bergman, C.M. (2007). Principles of genome evolution in the *Drosophila melanogaster* species group. PLoS Biol. **5:** 1366–1381.

Raskina, O., Barber, J.C., Nevo, E., and Belyayev, A. (2008). Repetitive DNA and chromosomal rearrangements: Speciation-related events in plant genomes. Cytogenet. Genome Res. **120:** 351–357.

Raskina, O., Belyayev, A., and Nevo, E. (2004). Quantum speciation in *Aegilops*: Molecular cytogenetic evidence from rDNA cluster variability in natural populations. Proc. Natl. Acad. Sci. USA **101:** 14818–14823.

Röbbelen, G. (1960). Beiträge zur Analyse des *Brassica*-Genoms. Chromosoma **11:** 205–228.

Santos, J.L., Alfaro, D., Sanchez-Moran, E., Armstrong, S.J., Franklin, F.C.H., and Jones, G.H. (2003). Partial diploidization of meiosis in autotetraploid *Arabidopsis thaliana*. Genetics **165:** 1533–1540.

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., Windsor, A.J., Song, B., 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.

Schubert, I., Fransz, P.F., Fuchs, J., and de Jong, J.H. (2001). Chromosome painting in plants. Methods Cell Sci. **23:** 57–69.

Srinivasachary, Dida, M.M., Gale, M.D., and Devos, K.M. (2007). Comparative analyses reveal high levels of conserved colinearity between the finger millet and rice genomes. Theor. Appl. Genet. **115:** 489–499.

Warwick, S.I., and Al-Shehbaz, I.A. (2006). Brassicaceae: Chromosome number index and database on CD-ROM. Plant Syst. Evol. **259:** 237–248.

Wei, F., et al. (2007). Physical and genetic structure of the maize genome reflects its complex evolutionary history. PLoS Genet. **3:** 1254–1263.

Zhong, X.B., Fransz, P.F., Wennekes-van Eden, J., Ramanna, M.S., van Kammen, A., Zabel, P., and de Jong, J.H. (1998). FISH studies reveal the molecular and chromosomal organization of individual telomere domains in tomato. Plant J. **13:** 507–517.

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

ANNALS OF
BOTANY
Founded 1887

A bicontinental origin of polyploid Australian/New Zealand *Lepidium* species (Brassicaceae)? Evidence from genomic *in situ* hybridization

Tom Dierschke[1],*, Terezie Mandáková[2], Martin A. Lysak[2] and Klaus Mummenhoff[1]

[1]*Universität Osnabrück, Abteilung Botanik, Barbarastrasse 11, 49076 Osnabrück, Germany and* [2]*Masaryk University, Department of Functional Genomics & Proteomics, Institute of Experimental Biology, Kamenice 5, Building A2, CZ-62500 Brno, Czech Republic*

• *Background and Aims* Incongruence between chloroplast and nuclear DNA phylogenies, and single additive nucleotide positions in internal transcribed spacer (ITS) sequences of polyploid Australian/New Zealand (NZ) *Lepidium* species have been used to suggest a bicontinental hybrid origin. This pattern was explained by two trans-oceanic dispersals of *Lepidium* species from California and Africa and subsequent hybridization followed by homogenization of the ribosomal DNA sequence either to the Californian (C-clade) or to the African ITS-type (A-clade) in two different ITS-lineages of Australian/NZ *Lepidium* polyploids.
• *Methods* Genomic *in situ* hybridization (GISH) was used to unravel the genomic origin of polyploid Australian/NZ *Lepidium* species. Fluorescence *in situ* hybridization (FISH) with ribosomal DNA (rDNA) probes was applied to test the purported ITS evolution, and to facilitate chromosome counting in high-numbered polyploids.
• *Key Results* In Australian/NZ A-clade *Lepidium* polyploids, GISH identified African and Australian/NZ C-clade species as putative ancestral genomes. Neither the African nor the Californian genome were detected in Australian/NZ C-clade species and the Californian genome was not detected in Australian/NZ A-clade species. Five of the eight polyploid species (from 7x to 11x) displayed a diploid-like set of rDNA loci. Even the undecaploid species *Lepidium muelleriferdinandi* ($2n = 11x = 88$) showed only one pair of each rDNA repeat. In A-clade allopolyploids, *in situ* rDNA localization combined with GISH corroborated the presence of the African ITS-type.
• *Conclusions* The nuclear genomes of African and Australian/NZ C-clade species were detected by GISH in allopolyploid Australian/NZ *Lepidium* species of the A-clade, supporting their hybrid origin. The presumed hybrid origin of Australian/NZ C-clade taxa could not be confirmed. Hence, it is assumed that Californian ancestral taxa experienced rapid radiation in Australia/NZ into extant C-clade polyploid taxa followed by hybridization with African species. As a result, A-clade allopolyploid *Lepidium* species share the Californian chloroplast type and the African ITS-type with the C-clade Australian/NZ polyploid and African diploid species, respectively.

Key words: *Lepidium*, Brassicaceae, FISH, GISH, hybridization, polyploidy, long-distance dispersal, ITS, rDNA, Australia, New Zealand.

INTRODUCTION

Lepidium is one of the largest genera in the Brassicaceae, consisting of approximately 230 species distributed worldwide (Al-Shehbaz *et al.*, 2006). This genus is characterized by reduced floral organs, an autogamous mating system and mucilaginous seeds (Al-Shehbaz, 1986). Recent molecular phylogenetic studies using the nuclear ribosomal DNA (rDNA) internal transcribed spacer (ITS), non-coding chloroplast DNA (cpDNA) and single-copy nuclear DNA sequences (an intron of the PISTILLATA gene, PI), respectively, clarified only some relationships within the genus (Bowman *et al.*, 1999; Mummenhoff *et al.*, 2001; Lee *et al.*, 2002). All molecular phylogenies support three main infrageneric lineages, corresponding to (1) section *Monoploca sensu stricto* (*s.s.*) (Australia), (2) section *Lepia* with *Cardaria* included (Eurasia) and (3) *Lepidium s.s.* (referred to hereafter as *Lepidium*) representing by far the bulk of *Lepidium* species on every continent. The fossil data, easily dispersible mucilaginous seeds, common autogamous breeding systems and low

levels of sequence divergence between *Lepidium* species from different continents or islands suggest a rapid radiation of *Lepidium* by long-distance dispersal in the Pliocene/Pleistocene (Mummenhoff *et al.*, 2001). The large number of polyploid species (Warwick and Al-Shehbaz, 2006) suggests reticulate evolution in *Lepidium* and this was indicated also in the analysis of PISTILLATA intron sequences, in which many polyploid taxa harbour two or more phylogenetically distinct sequences (Lee *et al.*, 2002). Chromosome numbers have been reported for about 70 *Lepidium* species and subspecies (Warwick and Al-Shehbaz, 2006), of which all except five have $x = 8$ chromosomes (Al-Shehbaz 1986). Vaarama (1951) speculated that species with $2n = 24$ (e.g. *L. sativum*) are hexaploids based on $x = 4$, but this interpretation was not supported by any further research. More than half of *Lepidium* species studied are polyploids, and chromosome numbers vary widely ($2n = 16$, 24, 28, 32, 40, 48, 64 and 80; Warwick and Al-Shehbaz, 2006).

In Australia and New Zealand (NZ), *Lepidium* is represented by 19 and seven native species, respectively. Incongruence between the chloroplast and nuclear DNA phylogenetic

* For correspondence. E-mail Tomiworld@gmx.de

TABLE 1. *Composition of the ITS sequences of Californian, Australian/New Zealand and African* Lepidium *species*

	ITS1	ITS2
	1111111112233344 6788001344677593 6933 8347896568024179 5402	
Californian C-clade species		
L. nitidum, L. oxycarpum, L. dictyotum	CATTAGCCCGTCATCGG–TA	
Australian/NZ C-clade polyploids		
L. aschersonii	CATTATCCCGTCATCGG–TW	
L. muelleriferdinandi	CATTRCCCCGTCATCGG–TA	
Australian/NZ A-clade polyploids		
L. hyssopifolium (KM 1670)	YWYKRSSYSSWACYTA–TCT	
L. hyssopifolium (KM 1669)	TTCGGCGTGCAACCTA–TCT	
L. pseudohyssopifolium	TTCGGCGTGCAACCYA–TCT	
African A-clade species		
L. africanum, L. transvaalense	TTCGGCGTGCAACCTA–TCT	

This data matrix contains only those positions of a complete alignment (not shown) that distinguish Californian from African *Lepidium* taxa. Californian and African species are extant members of two different lineages suggested to have been involved in the hybridogenous origin of Australian/NZ *Lepidium* species. R = A and G; W = A and T; Y = C and T; K = G and T; S = G and C. Site numbers are those of the complete alignment. For details see Mummenhoff *et al.* (2004).

analyses indicated a hybridogenous genomic constitution of Australian/NZ *Lepidium* species: all 18 studied species shared a Californian cpDNA-type, while 11 Australian/NZ species appeared to harbour a Californian ITS-type (C-clade) and a group of the remaining seven species shared an African ITS-type (A-clade) (Mummenhoff *et al.*, 2004; Table 1, Fig. 1). This pattern was explained by two transoceanic dispersals of ancestral taxa from California and Africa to Australia/NZ and subsequent hybridization followed by sequence homogenization of the rDNA either to the Californian or to the African ITS-type in the two different lineages, with single additive nucleotide positions remaining, representing the ancestral parental input (Table 1, Fig. 1). However, Mummenhoff *et al.* (2004) did not assess the ploidy level of Australian/NZ species and the conclusion of a bicontinental hybridogenous genomic constitution was based only on two markers representing nuclear (ITS) and chloroplast genomes (three non-coding cpDNA regions).

The present study aimed to (1) assess ploidy level and chromosome number variation among Australian/NZ *Lepidium* species, (2) test the purported allopolyploid bicontinental origin of Australian/NZ species of the A- and C-clades using genomic *in situ* hybridization (GISH) and (3) gain insight into the physical organization of rDNA loci in the presumed parental species and their Australian/NZ hybrid derivatives.

MATERIALS AND METHODS

Plant material

Collection data of selected species of *Lepidium* L. used in the current study are given in Table 2. The presumed parental and hybrid species were used exemplarily, based on results of previous studies (Mummenhoff *et al.*, 2001, 2004, 2009).

Chromosome preparations

Plants were grown from seeds in plastic Petri dishes on sieved potting soil in a phytotron with long day illumination (16 h light at 20 °C, 8 h dark at 15 °C). Young inflorescences were fixed in ethanol/acetic acid (3 : 1, v/v) fixative for 24 h at 4 °C. Fixative was replaced by 70 % ethanol and the material stored at −20 °C until further use. Chromosome spreads were prepared as described by Lysak *et al.* (2006). Slides were examined under phase contrast for the presence of suitable mitotic metaphase spreads. Selected slides were post-fixed in 4 % formaldehyde in 2× sodium saline citrate (SSC, 5 min), washed in 2× SSC (twice for 5 min), and dehydrated in an ethanol series (70, 90 and 100 %, 2 min each). The preparations were stained with 2 μg mL^{-1} 4,6-diamino-2-phenylindole (DAPI) in Vectashield antifade (Vector Laboratories, Burlington, Ontario, Canada) and screened for the quality of metaphase and the presence of cytoplasm under an epifluorescence microscope (BX-61, Olympus, Tokyo, Japan). If appropriate, the slides were treated with pepsin (0·1 g L^{-1}) in 0·01 M HCl at 39 °C for 15–45 min, and post-fixed and dehydrated as described above.

DNA probes

The *Arabidopsis thaliana* BAC clone T15P10 (AF167571) bearing 45S rRNA gene repeats was used for *in situ* localization of 45S rDNA, and *A. thaliana* clone pCT 4·2 (M65137), corresponding to a 500-bp 5S rRNA repeat, was used for localization of 5S rDNA loci. For GISH, total genomic DNA (gDNA) was extracted from healthy young leaves according to Dellaporta *et al.* (1983) followed by RNase treatment (50 μg mL^{-1}). Extracted gDNA was checked for protein, starch or RNA contamination using a Beckmann photospectrometer and run on a 1 % (w/v) agarose gel in 1× Tris-acetate-EDTA (TAE) buffer. DNA probes were labelled either by biotin- or by digoxigenin-dUTP using the Nick Translation Mix (Roche, Mannheim, Germany) according to the manufacturer's instructions.

In situ *hybridization*

Labelled 5S and 45S rDNA probes were precipitated and the pellet was resuspended in 20 μL hybridization mixture containing 50 % formamide and 10 % dextrane sulfate in 2× SSC. The probe was applied to the slide, covered with a cover slip and denatured on a hot plate at 80 °C for 2 min. Slides were hybridized at 37 °C for approx. 12 h. To prevent drying, cover slips were framed by rubber cement. Labelled gDNA probes were prepared and denatured as described above. Hybridization time for GISH was between 12 and 48 h.

Following hybridization, stringent washing was carried out in 50 % formamide in 2× SSC (v/v) at 42 °C three times for 5 min. Detection of hybridization signals was performed according to Lysak *et al.* (2006). The biotin-labelled probes were detected by avidin–Texas Red (Vector Laboratories), and signals amplified by biotinylated goat anti-avidin (Vector Laboratories) and avidin–Texas Red. Digoxigenin-labelled probes were detected by mouse anti-digoxigenin (Roche) and goat anti-mouse–Alexa Fluor 488 antibodies
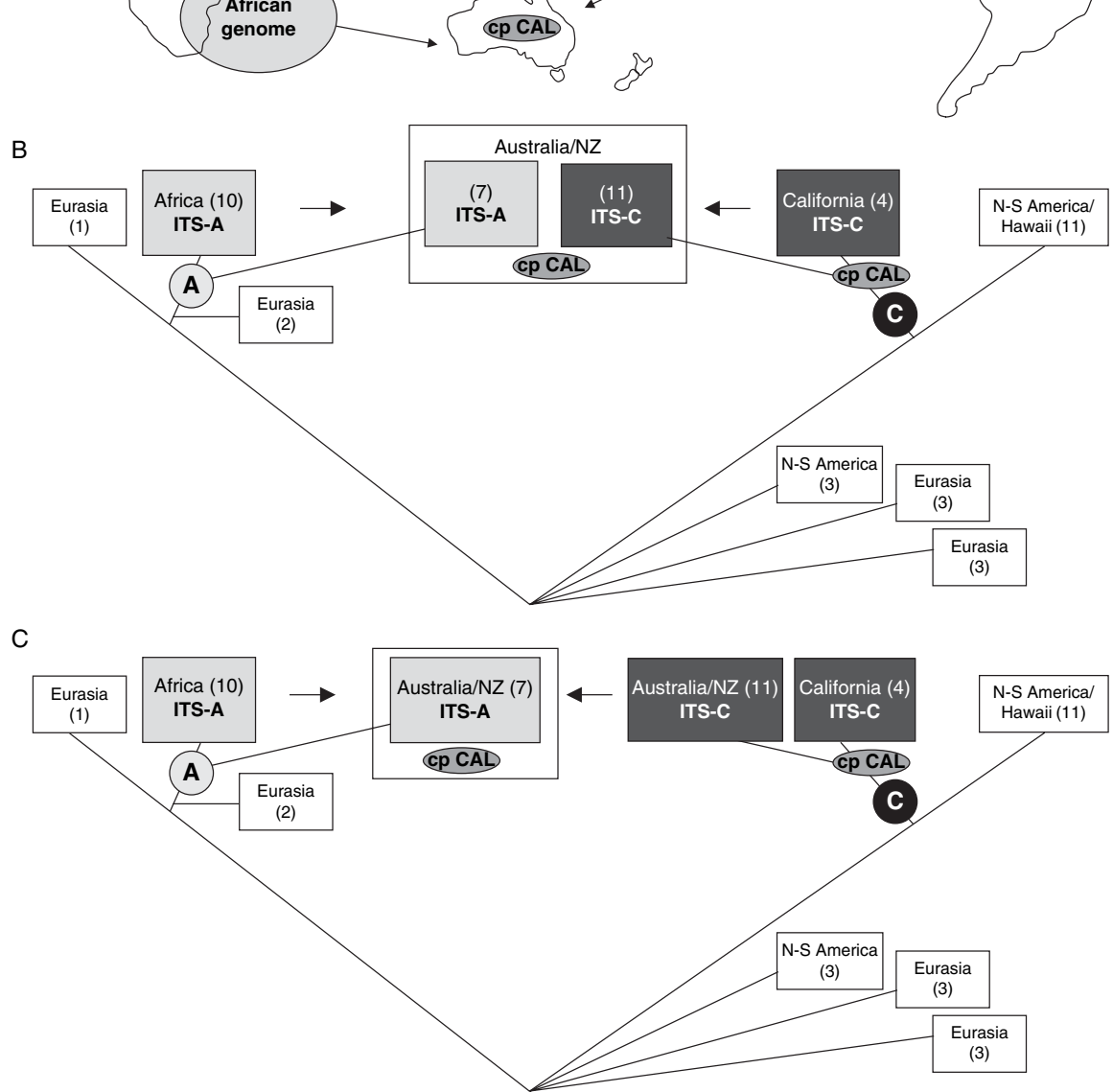
FIG. 1. Organismal and geographical context of the genome evolution of polyploid Australian/NZ A- and C-clade *Lepidium* species. (A) Geographical context of the evolution of polyploid Australian/NZ *Lepidium* species further outlined in (B) and (C). (B) Evolution of polyploid Australian/NZ A- and C-clade *Lepidium* species as inferred from ITS- and cpDNA sequence data (Mummenhoff *et al.*, 2004). All Australian/NZ species harbour a Californian chloroplast type (cp CAL), while bidirectional concerted evolution of ribosomal DNA (ITS) subsequent to hybridization between African and Californian parental species resulted in the presence of either of the parental ITS-types, i.e. African (ITS-A) and Californian (ITS-C), in the two lineages (Australian/NZ A- and C-clade species, respectively). Arrows indicate hybridization. Numbers in boxes represent number of analysed species of given geographical origin. Encircled letters A and C refer to A- and C-clade species, respectively. See text for more details. (C) Evolution of polyploid Australian/NZ A- and C-clade *Lepidium* species as inferred from ITS and cpDNA sequence data (Mummenhoff *et al.*, 2004) and current GISH experiments. In this alternative scenario polyploid Australian/NZ C-clade species are the outcome of rapid radiation of an ancient Californian ancestor in Australia/New Zealand. Australian/NZ A-clade species originated from hybridization between African and Australian/NZ C-clade species, still harbouring the ancestral Californian chloroplast type. In this scenario concerted evolution operated unidirectionally towards the African ITS-type in Australian/NZ A-clade hybrids. Arrows indicate hybridization. Numbers in boxes represent number of analysed species of given geographical origin. Encircled letters A and C refer to A- and C-clade species, respectively. See text for more details.

TABLE 2. *Origin, chromosome number, ploidy level and number of rDNA loci for* Lepidium *species studied*

Species	Accession no.	Origin/Collector	Chromosome no.	No. of 45S rDNA loci	No. of 5S rDNA loci
L.africanum (Burm.f.) DC.*	KM 1702	Australia, Melbourne, Royal Park, South of Park Street/Scarlett, N.	$2n = 2x = 16$	2	2
	KM 1793	South Africa, Eastern Cape/Clark, V. R.	$2n = 2x = 16$	2	2
	KM 1794	South Africa, Eastern Cape/Clark, V. R.	$2n = 2x = 16$	2	2
L. transvaalense Marais	KM 1792	South Africa, Eastern Cape/Clark, V. R.	$2n = 2x = 16$	2	2
L. nitidum Nutt.	KM 1711	USA, California, Tucker Herbarium at UC Davis/Bowman, J.	$2n = 7x = \text{approx. } 56$	2	2
L. oxycarpum Torr. & Gray	KM 1713	USA, California, Central Valley, Los Banos/Bowman, J.	$2n = 7x = \text{approx. } 56$	2	2
L. dictyotum Gray	KM 1714	USA, California, Carrizo plain/Bowman, J.	$2n = 7x = \text{approx. } 56$	2	4
L. hyssopifolium Desv.	KM 1669	Australia, New South Wales, Bathurst/Scarlett, N.	$2n = 9x = 72$	2	4
	KM 1670	Australia, Victoria, Belfast Lough, Port Fairy end of the Lough/Scarlett, N.	$2n = 9x = 72$	2	4
L. pseudohyssopifolium Hewson	KM 1666	Australia, Victoria, Melbourne, Fairfield, Banks of the Yarra River/Parsons, R. F.	$2n = 9x = 72$	2	2
L. ginninderrense Scarlett	KM 1673	Australia, Australian Capital Territory, Belconnen Naval Station, floodplain of Ginninderra Creek/Scarlett, N.	$2n = 14x = 112$	4	2
L. muelleriferdinandi Thell.	KM 1710	Australia, Northern Territory, Alice Springs Desert Park, outside herbarium/Albrecht, D. E.	$2n = 11x = 88$	2	2
L. aschersonii Thell.	KM 1668	Australia, Victoria, Lake Beeac/Scarlett, N.	$2n = 7x = 56$	2	2

* L. africanum is native to South Africa, and an introduced weed in Australia.

(Molecular Probes, New Haven, CT, USA). Chromosomes were counterstained with DAPI ($2\ \mu g\ mL^{-1}$) in Vectashield. Fluorescence signals were analysed with an Olympus BX-61 epifluorescence microscope and AxioCam CCD camera (Carl Zeiss, Jena, Germany). Individual images were merged and processed by using Photoshop CS software (Adobe Systems).

RESULTS

Chromosome counts

In the analysed *Lepidium* species, chromosome numbers varied from $2n = 16$ to $2n = 112$ (see Table 2 for all chromosome counts). The two African species (*L. africanum*, *L. transvaalense*) have a diploid chromosome number ($2n = 2x = 16$), whereas all Australian/NZ and Californian species are polyploid (Fig. 2). Four different ploidy levels ($7x$, $9x$, $11x$ and $14x$) were discerned in Australian/NZ species. In the three Californian species, we could not obtain suitable chromosome spreads to determine exact chromosome numbers. Observed chromosome numbers were tentatively interpreted as heptapolyploid ($2n = 7x = 56$).

Mitotic chromosomes of *Lepidium* species are small (approx. 2–$5\ \mu m$) with the bulk of heterochromatin concentrated within pericentromeres, and often fuzzy euchromatic chromosome ends (Figs 2 and 3). In the two African species, chromosomes are larger and possess a less distinct eu-/heterochromatin profile than chromosomes in the other two geographical groups. The presumably different chromosome structure of African species is particularly apparent upon GISH in the A-clade allopolyploids, whereby chromosomes of African origin show a more even labelling pattern and appear to be larger than the remaining chromosomes (Fig. 3).

Localization of rDNA loci

Most *Lepidium* species exhibit one pair of 5S and 45S rDNA loci (Table 2 and Fig. 2). This pattern was also observed in high-numbered polyploid species such as *L. nitidum*, *L. oxycarpum*, *L. pseudohyssopifolium*, *L. muelleriferdinandi* and *L. aschersonii*. Two species (*L. dictyotum* and *L. hyssopifolium*) possess two pairs of 5S rDNA, and one species (*L. ginninderrense*) possesses two pairs of 45S rDNA loci (Table 2 and Fig. 2). In *L. oxycarpum*, 5S and 45S loci were found to be partly co-localized (Fig. 2C). 5S rDNA loci are positioned on chromosomes interstitially, whereas 45S rDNA loci showed terminal localization (Fig. 2).

GISH in polyploid Australian/NZ Lepidium *species*

With the aim to discern the origin of Australian/NZ *Lepidium* polyploids, labelled gDNA of putative African and Californian parental species was hybridized to mitotic chromosomes of Australian/NZ polyploid A- and C-clade taxa. In two A-clade species (*L. hyssopifolium*, *L. pseudohyssopifolium*), gDNA of presumed African parents (*L. africanum*, *L. transvaalense*) labelled 16 of the 72 chromosomes, whereas no chromosomes were identified using gDNA of presumed Californian parental species (*L. dictyotum*, *L. oxycarpum*) (data not shown). No chromosome-specific hybridization signals were observed in
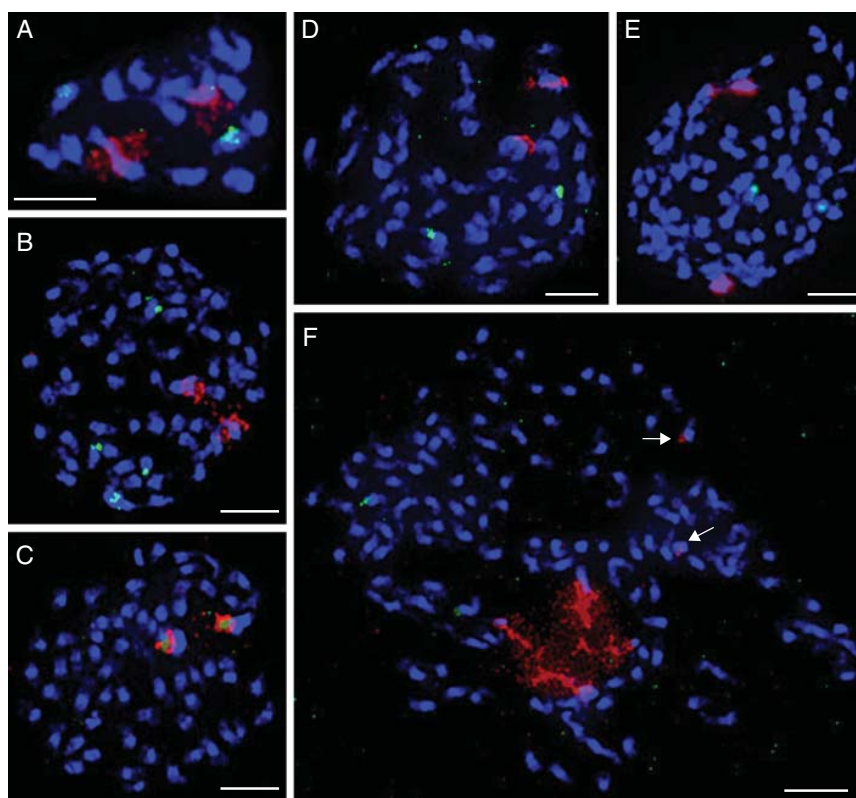
FIG. 2. DAPI-stained mitotic chromosome spreads from flower bud tissue of *Lepidium* species. Chromosomes were hybridized with 5S (green) and 45S (red) rDNA probes and counterstained with DAPI. (A) *L. africanum* ($2n = 2x = 16$), (B) *L. dictyotum* ($2n = 7x = 56$), (C) *L. oxycarpum* ($2n = 7x = 56$), (D) *L. aschersonii* ($2n = 7x = 56$), (E) *L. muelleriferdinandi* ($2n = 11x = 88$), (F) *L. ginninderrense* ($2n = 14x = 112$). In (F) one pair of major 45S rDNA loci is decondensed around the nucleolus; two minor 45S rDNA loci are indicated by arrows. Scale bars = 5 μm.
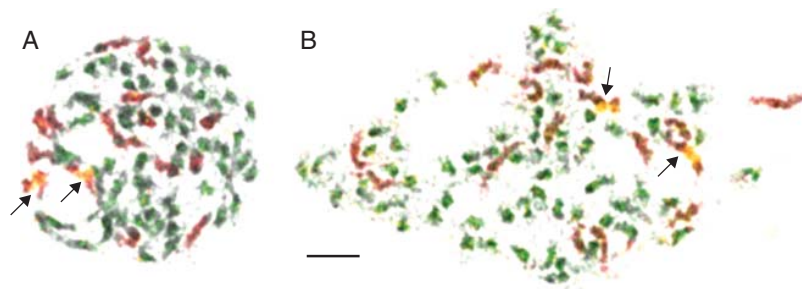
FIG. 3. GISH in two nonaploid ($2n = 9x = 72$) Australian/NZ *Lepidium* A-clade species. Chromosomes of *L. hyssopifolium* (A) and *L. pseudohyssopifolium* (B) hybridized with gDNA of *L. africanum* (16 red chromosomes) and gDNA of an Australian/NZ C-clade species, i.e. *L. muelleriferdinandi* (56 green chromosomes). Arrows mark 45S rDNA-bearing chromosomes (yellow/red). Scale bar = 5 μm.

C-clade species (*L. aschersonii*, *L. muelleriferdinandi*) using the same probes as used in the A-clade species (data not shown).

In follow-up GISH experiments, 56 of 72 chromosomes in the Australian/NZ A-clade species *L. hyssopifolium* and *L. pseudohyssopifolium* were detected using gDNA of two Australian/NZ C-clade species, i.e. *L. muelleriferdinandi* (Fig. 3) and *L. aschersonii* (data not shown). In the two A-clade allopolyploids ($2n = 72$), 16 chromosomes correspond to *L. africanum* and 56 chromosomes were labelled by gDNA of *L. aschersonii*/*L. muelleriferdinandi* (Fig. 3). This experiment also showed that 45S rDNA in *L. hyssopifolium*

and *L. pseudohyssopifolium* has probably been contributed by the African parental genome(s).

DISCUSSION

rDNA patterns in polyploid Lepidium species

The present study followed three principal goals as far as *in situ* localization of rDNA repeats is concerned. First, we were interested to see if the three geographical species groups (Africa, Australia/NZ and California) are characterized

by specific rDNA patterns. Secondly, the number of rDNA loci in Australian/NZ polyploids was compared with the number of the same loci in the presumed progenitor species, and finally, the origin of 45S rDNA loci in Australian/NZ polyploid species was inferred from GISH analysis and compared with conclusions based on ITS sequence data (Mummenhoff *et al.*, 2004).

From the limited data set (only ten of 230 *Lepidium* species were analysed) no apparent geographical pattern of rDNA variation emerges. In all but one polyploid species analysed, only a diploid-like number of 45S rDNA loci were found, with two pairs of 45S identified in the tetradecaploid *L. ginninderrense* ($2n = 112$). This finding is somewhat surprising given the high ploidy levels (from $7x$ to $14x$) of the analysed taxa. Regardless of prevailing modes of polyploid evolution (auto- vs. allopolyploidy) in the Californian and Australian/NZ groups, respectively, the present data indicate that there is a strong tendency toward diploidization of rDNA loci in both geographical species groups. Although less likely, preferential loss of rDNA-bearing chromosomes in odd-numbered (7, 9 and $11x$) *Lepidium* polyploids cannot be ruled out.

Ali *et al.* (2005) analysed three different crucifer species with the same polyploid chromosome number ($2n = 48$). In two species, the number of 45S was high (approx. 14 in *Camelina microcarpa*, and $10 + 6$ minor loci in *Olimarabidopsis cabulica*), whereas only one pair has been recorded in *Aethionema schistosum* (Ali *et al.*, 2005). It could be speculated that this limited comparison corroborates the generally accepted correlation between the number of 45S rDNA loci and the age of polyploid species as the genus *Aethionema* is sister to and older than the rest of Brassicaceae (Al-Shehbaz *et al.*, 2006). However, the *Aethionema* polyploid species may be of a similar age to two Camelineae species analysed. Hence, it appears to be difficult to establish a direct link between the age of a phylogenetic split and the diploidization rate in crucifer taxa.

With each polyploidization event the number of rDNA loci increases and some rDNA sequences become redundant. These ITS sequences can be changed by bidirectional interlocus concerted evolution, resulting in complete homogenization to either of the parental types or to mosaic-like sequences with additive nucleotides from both parents (Wendel, 2000). In the long term, genome diploidization events as discovered for ancient allopolyploid *Nicotiana* species of the section *Repandae* (4·5 Myr old) can occur, resulting in the loss of redundant rDNA loci and a diploid-like number of rDNA loci, which can be even lower than the number of 45S loci found in the progenitor species (Clarkson *et al.*, 2005). However, more recently formed *Nicotiana* allopolyploids (e.g. *N. tabacum*, 0·2 Myr old) display the sum of the rDNA loci of their progenitors (Clarkson *et al.*, 2005). As a maximum estimated age for the origin of polyploid Australian/NZ *Lepidium* species is approx. 1·3 Ma (Mummenhoff *et al.*, 2004), this or a shorter period of time has been sufficient for (nearly) complete diploidization of rDNA loci (Table 2).

Within the set of the 72 chromosomes of Australian/NZ A-clade species (*L. hyssopifolium*, *L. pseudohyssopifolium*) combined GISH and rDNA fluorescence *in situ* hybridization (FISH) localization detected two 45S rDNA loci on two

chromosomes belonging to the 16 chromosomes of African origin. The 56 remaining chromosomes displayed no 45S rDNA loci (Fig. 3). This is in agreement with the ITS phylogenetic data showing that all Australian/NZ A-clade species harbour the African ITS-type (Mummenhoff *et al.*, 2004; Fig. 1C, Table 1). For these species, physical loss of C-clade rDNA loci must be assumed subsequent to hybridization. Furthermore, the African ITS-type in polyploid Australian/NZ A-clade species can be explained by unidirectional concerted evolution to the African ITS-type with single additive nucleotides remaining (Table 1, *L. hyssopifolium* KM 1670).

Previous hypothesis on the origin of polyploid Australian/NZ Lepidium *species*

The rDNA ITS regions are the most widely used nuclear-encoded phylogenetic markers (Álvarez and Wendel, 2003). Their high evolutionary rate permits discrimination of closely related putative parental species and the identification of additive patterns in hypothesized allopolyploids (Marhold and Lihová, 2006). Although concerted evolution can erase nucleotide additivity, bidirectional, nearly complete homogenization or intergenomic recombination may result in a mosaic-like structure of ITS sequences, which has been described in the literature (Marhold and Lihová, 2006). In the Brassicaceae, such analyses unravelled hybridization events in *Thlaspi* (Mummenhoff *et al.*, 1997), *Cardamine* (Franzke and Mummenhoff, 1999; Lihová *et al.*, 2006) and *Boechera* (Koch *et al.*, 2003). Support for the bicontinental hybrid origin of polyploid Australian/NZ *Lepidium* species was provided by bidirectional, nearly complete concerted evolution of the ITS to either of the two presumed parental ITS-types (African and Californian) with single additive (parental) nucleotides remaining (Mummenhoff *et al.*, 2004; Table 1). It appears that all Australian/NZ polyploids share a Californian cpDNA-type (Mummenhoff *et al.*, 2004), but this cpDNA evidence was detected only in 60 % of the maximally parsimonious trees (Mummenhoff *et al.*, 2004). Furthermore, this scenario of a bicontinental hybridogenous genomic constitution was based on only two markers representing the nuclear (ITS) and chloroplast genome (Mummenhoff *et al.*, 2004). Thus, GISH experiments were performed to gain deeper insight into the origin and evolution of polyploid Australian/NZ *Lepidium* species.

Origin of Australian/NZ C-clade species: neither African nor Californian progenitors detected by GISH

In previous studies based on sequence analysis of nuclear ITS and non-coding cpDNA (Mummenhoff *et al.*, 2004), a bicontinental hybrid origin of polyploid Australian/NZ *Lepidium* species has been suggested. In GISH experiments, gDNA probes of presumed African (*L. africanum*, *L. transvaalense*) and Californian parental species (*L. dictyotum*, *L. oxycarpum*) did not label specific chromosomes in the Australian C-clade species. This result is in contrast to the ITS and cpDNA analysis (Mummenhoff *et al.*, 2004). The evidence of a bicontinental hybridogenous genomic constitution of C-clade species is based on weak cpDNA evidence (see above), ITS sequence homology of

Australian/NZ species with Californian taxa, and single additive nucleotide positions representing both parental ITS-types. If this additivity in single ITS nucleotide positions is not random and indicates an African input, present-day African species used for GISH experiments do not represent the original African progenitors. This means that the genome of the original African species contributing to the ITS additivity in the polyploid Australian/NZ C-clade species is too distinct from the genome of extant African taxa used in the present study (*L. africanum*, *L. transvaalense*), and thus gDNA probes of *L. africanum/L. transvaalense* could not be used to detect the original African parental genome in C-clade species.

Ancestral Californian species (migrating to Australia) may have transmitted the Californian ITS-type to the allopolyploid Australian/NZ C-clade species via hybridization with an unknown taxon, potentially of African origin. At the same time the ancestral Californian taxa might have evolved into the present-day polyploid Californian species, probably by several rounds of hybridization and polyploidization events, accompanied by the dynamic evolution of genome-specific dispersed repeats. As a consequence, the extant Californian species differ significantly in genome constitution from their presumed polyploid Australian/NZ descendants, and thus their gDNAs do not reveal the ancestral Californian genome in the polyploid Australian/NZ C-clade species. This can be analogous to the decreased efficiency of GISH and its failure in approx. 1- and 5-Myr-old *Nicotiana* allopolyploids, respectively (Lim *et al.*, 2007).

Australian/NZ C-clade species: simply descendants of Californian ancestors?

Negative results using gDNA of Californian and African species in C-clade Australian/NZ polyploids suggest long-distance dispersal of ancestral Californian species to Australia followed by radiation into extant C-clade polyploid species without hybridization with African species being involved. All Australian/NZ C-clade species appear to harbour a Californian ITS- and cpDNA-type (Mummenhoff *et al.*, 2004) and only one additive nucleotide position each (representing African and Californian ancestral nucleotides, respectively) in the ITS sequences of two C-clade species studied (*L. aschersonii*, *L. muelleriferdinandi*). This single African nucleotide in these two C-clade species could also be explained by random mutations or could represent an input by hybridization with an unknown species. The remaining 32 chromosomes of the genome of present-day *L. muelleriferdinandi* ($2n = 88$) not detected in Australian A-clade polyploids ($2n = 72$) could indeed indicate some sort of hybridization/polyploidization with unknown taxa in the evolution of *L. muelleriferdinandi* subsequent to the hybridization of an ancestor of *L. muelleriferdinandi* (presumably with $2n = 56$) with African species ($2n = 16$). GISH signals of an African genome input could not be detected, suggesting that Australian/NZ C-clade species are not the outcome of a bicontinental allopolyploidization scenario as originally described (Mummenhoff *et al.*, 2004). Calibration of molecular trees yielded ages of approx. $0 \cdot 7 - 1 \cdot 3$ and $0 \cdot 3 - 0 \cdot 55$ Ma for the Australian/NZ species of the C- and

A-clades, respectively (Mummenhoff *et al.*, 2004). These age differences might indicate that the radiation of Californian ancestors into Australian/NZ C-clade species predates the later arrival of African ancestors and the origin of A-clade allopolyploids.

Origin of Australian/NZ A-clade species: allopolyploid hybrids between African and Australian/NZ C-clade species

GISH experiments clearly detected 16 chromosomes of African origin (*L. africanum*, *L. transvaalense*) in the two polyploid Australian/NZ A-clade species analysed (*L. hyssopifolium*, *L. pseudohyssopifolium*). The remaining 56 chromosomes in the A-clade allopolyploid genomes were not labelled by any of the Californian presumed parental gDNAs used. Thus, the presumed direct Californian genomic input into the Australian/NZ A-clade *Lepidium* species could not be confirmed by GISH.

One might thus speculate that Australian/NZ C-clade taxa have contributed the Californian cpDNA-type to the Australian/NZ A-clade taxa by hybridization with African species in Australia. To test this scenario, GISH experiments were conducted using gDNAs of Australian/NZ C-clade species (*L. aschersonii* $2n = 56$, *L. muelleriferdinandi* $2n = 88$) as probes onto the Australian/NZ A-clade polyploid species *L. hyssopifolium* and *L. pseudohyssopifolium* ($2n = 72$). Of the 72 chromosomes, 56 were labelled by gDNA of Australian/NZ C-clade species and, as outlined above, the remaining 16 chromosomes were labelled by gDNA of *L. africanum/L. transvaalense*. The remaining 32 chromosomes of the genome of present-day *L. muelleriferdinandi* ($2n = 88$) not detected in Australian A-clade polyploids ($2n = 72$) could indicate some sort of hybridization/polyploidization with unknown taxa in the evolution of *L. muelleriferdinandi* subsequent to the hybridization of an ancestor of *L. muelleriferdinandi* (presumably with $2n = 56$) with African species ($2n = 16$).

This general scenario is in agreement with the cpDNA phylogeny in which the Australian/NZ A-clade hybrids harbour a Californian chloroplast type (Mummenhoff *et al.*, 2004). There is some additional evidence for this alternative scenario: one accession of an Australian/NZ A-clade polyploid (*L. hyssopifolium*, KM 1670) shows additivity in almost every diagnostic nucleotide position that differs between the C-clade and A-clade species group (Table 1). Regardless, all GISH and rDNA FISH experiments using two different accessions of *L. hyssopifolium* (KM 1670, KM 1669) show the same results. As ITS homogenization can occur very rapidly, as also observed in experiments with recent and synthetic allopolyploids where 45S rDNA homogenization was apparent within a few generations (Franzke and Mummenhoff, 1999; Skalická *et al.*, 2003; Kovarik *et al.*, 2005; Shcherban *et al.*, 2008), it is possible that additivity in almost all diagnostic nucleotide position in *L. hyssopifolium* (KM 1670) indicates recent and/or recurrent hybridization between African and Australian/NZ C-clade species in Australia.

Thus, the main conclusions from GISH analysis are as follows. The nuclear genomes of African and Australian/NZ C-clade species were detected in allopolyploid Australian/NZ *Lepidium* A-clade species. The presumed hybrid origin of

Australian/NZ C-clade taxa with African and Californian parents involved (Mummenhoff *et al.*, 2004) could not be confirmed. This does not mean that hybridization processes did not play a role in the evolution of polyploid Australian C-clade species, for example *L. muelleriferdinandi* with $2n = 88$ chromosomes. Hence, it is assumed that ancestral Californian taxa subsequent to their dispersal to Australia experienced a rapid radiation in Australia and New Zealand into extant C-clade taxa that hybridized with African species. As a result, A-clade allopolyploid *Lepidium* species in Australia/NZ share the Californian chloroplast type and the African ITS-type with polyploid Australian/NZ C-clade and diploid African species, respectively.

LITERATURE CITED

Al-Shehbaz IA. 1986. The genera of *Lepidieae* (Cruciferae; Brassicaceae) in the southeastern United States. *Journal of the Arnold Arboretum* **67**: 265–311.

Al-Shehbaz IA, Beilstein MA, Kellogg EA. 2006. Systematics and phylogeny of the Brassicaceae (Cruciferae): an overview. *Plant Systematics and Evolution* **259**: 89–120.

Ali HBM, Lysak MA, Schubert I. 2005. Chromosomal localization of rDNA in the Brassicaceae. *Genome* **48**: 341–346.

Álvarez I, Wendel JF. 2003. Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution* **29**: 417–434.

Bowman JL, Brüggemann H, Lee J-Y, Mummenhoff K. 1999. Evolutionary changes in floral structure within *Lepidium* L. (Brassicaceae). *International Journal of Plant Sciences* **160**: 917–929.

Clarkson JJ, Lim KY, Kovarik A, Chase MW, Knapp S, Leitch AR. 2005. Long-term genome diploidization in allopolyploid *Nicotiana* section *Repandae* (Solanaceae). *New Phytologist* **168**: 241–252.

Dellaporta SL, Wood J, Hicks JB. 1983. A plant DNA minipreparation: version II. *Plant Molecular Biology Reporter* **1**: 19–21.

Franzke A, Mummenhoff K. 1999. Recent hybrid speciation in *Cardamine* (Brassicaceae) conversion of nuclear ribosomal ITS sequences in *statu nascendi*. *Theoretical and Applied Genetics* **98**: 831–834.

Kovarik A, Pires JC, Leitch AR, *et al*. 2005. Rapid concerted evolution of nuclear ribosomal DNA in two Tragopogon allopolyploids of recent and recurrent origin. *Genetics* **169**: 931–944.

Koch M, Al-Shehbaz IA, Mummenhoff K. 2003. Molecular systematics, evolution, and population biology in the mustard family (Brassicaceae). *Annals of the Missouri Botanical Garden* **90**: 151–171.

Lee J-Y, Mummenhoff K, Bowman JL. 2002. Alloploidization and evolution of species with reduced floral structures in *Lepidium* L. (Brassicaceae). *Proceedings of the National Academy of Sciences* **16**: 835–840.

Lihová J, Shimizu KK, Marhold K. 2006. Allopolyploid origin of *Cardamine asarifolia* (Brassicaceae): incongruence between plastid and nuclear ribosomal DNA sequences solved by a single-copy nuclear gene. *Molecular Phylogenetics and Evolution* **39**: 759–786.

Lim KY, Kovarik A, Matyasek R, *et al*. 2007. Sequence of events leading to near-complete genome turnover in allopolyploid *Nicotiana* within five million years. *New Phytologist* **175**: 756–763.

Lysak MA, Berr A, Pecinka A, Schmidt R, McBreen K, Schubert I. 2006. Mechanisms of chromosome number reduction in *Arabidopsis thaliana* and related Brassicaceae species. *Proceedings of the National Academy of Sciences of the USA* **103**: 5224–5229.

Marhold K, Lihová J. 2006. Polyploidy, hybridization and reticulate evolution: lessons from the Brassicaceae. *Plant Systematics and Evolution* **259**: 143–174.

Mummenhoff K, Franzke A, Koch M. 1997. Molecular phylogenetics of *Thlaspi* s.l. (Brassicaceae) based on chloroplast DNA restriction site variation and sequences of the internal transcribed spacers of nuclear ribosomal DNA. *Canadian Journal of Botany* **75**: 469–482.

Mummenhoff K, Brüggemann H, Bowman JL. 2001. Chloroplast DNA phylogeny and biogeography of *Lepidium* (Brassicaceae). *American Journal of Botany* **88**: 2051–2063.

Mummenhoff K, Linder P, Friesen N, Bowman JL, Lee J-Y, Franzke A. 2004. Molecular evidence for bicontinental hybridogenous genomic constitution in *Lepidium* sensu stricto (Brassicaceae) species from Australia and New Zealand. *American Journal of Botany* **91**: 254–261.

Mummenhoff K, Polster A, Mühlhausen A, Theißen G. 2009. *Lepidium* as a model system for studying the evolution of fruit development in Brassicaceae. *Journal of Experimental Botany* **60**: 1503–1513.

Shcherban AB, Sergeeva EM, Badaeva ED, Salina EA. 2008. Analysis of 5S rDNA changes in synthetic allopolyploids *Triticum × Aegilops*. *Molecular Biology* **42**: 536–542.

Skalická K, Lim KY, Matyasek R, Koukalová B, Leitch A, Kovarik A. 2003. Rapid evolution of parental rDNA in a synthetic tobacco allotetraploid line. *American Journal of Botany* **90**: 988–996.

Vaarama A. 1952. Chromosome number and cryptic polyploidy in *Lepidium sativum*. *Hereditas* **37**: 290–292.

Warwick SI, Al-Shehbaz IA. 2006. Brassicaceae: chromosome number index and database on CD-Rom. *Plant Systematics and Evolution* **259**: 237–248.

Wendel JF. 2000. Genome evolution in polyploids. *Plant Molecular Biology* **42**: 225–249.

V.   **Mandáková T.**, Joly S., Krzywinski M., Mummenhoff K., Lysak M.A. 2010. Fast diploidization in close mesopolyploid relatives of *Arabidopsis*. Plant Cell 22: 2277-2290.

Fast Diploidization in Close Mesopolyploid Relatives of *Arabidopsis* ᵂᴼᴬ

Terezie Mandáková,ᵃ Simon Joly,ᵇ Martin Krzywinski,ᶜ Klaus Mummenhoff,ᵈ and Martin A. Lysakᵃ,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 preceding 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
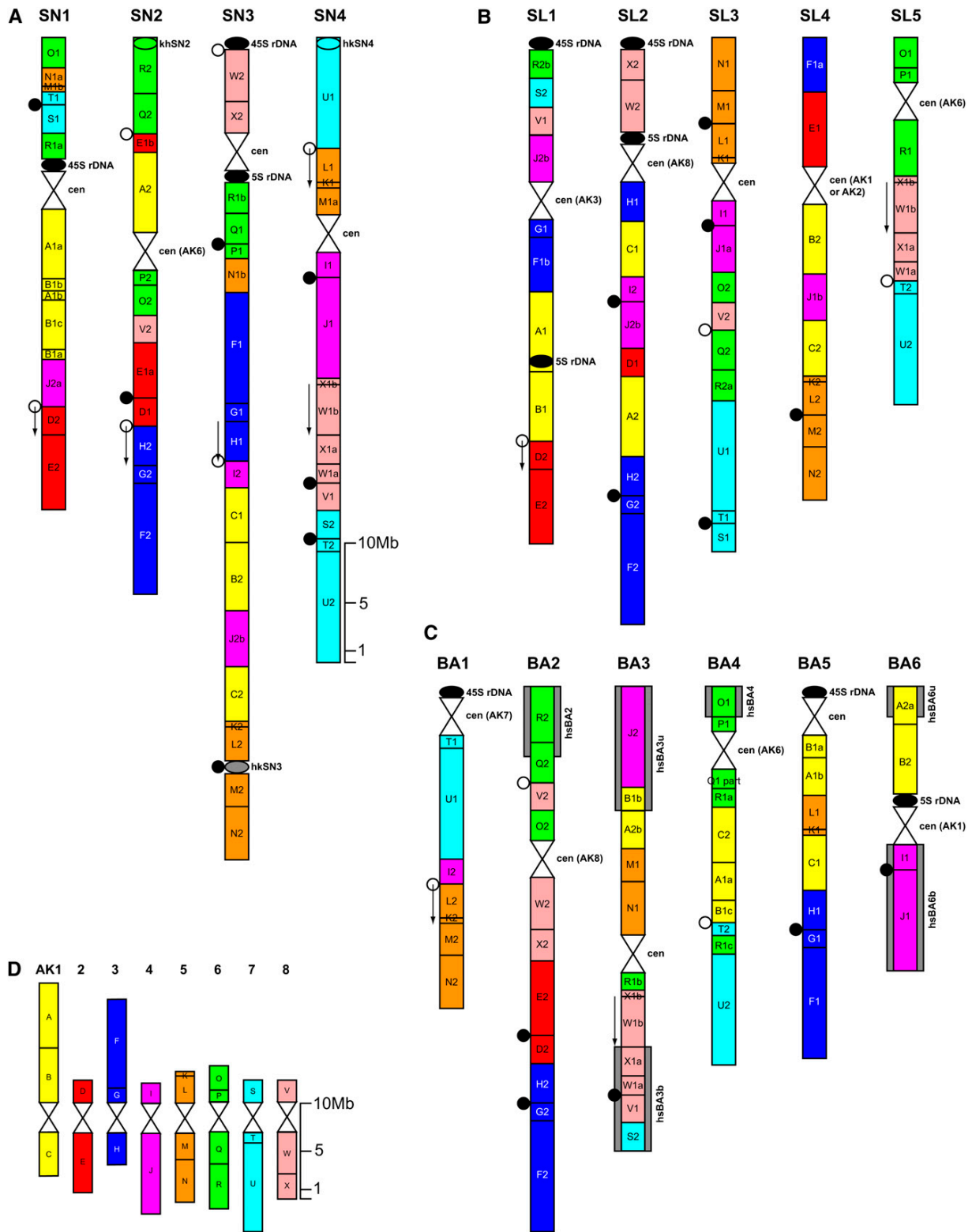
Figure 1. Cytogenetic Maps of *Stenopetalum* and *Ballantinia* Species.

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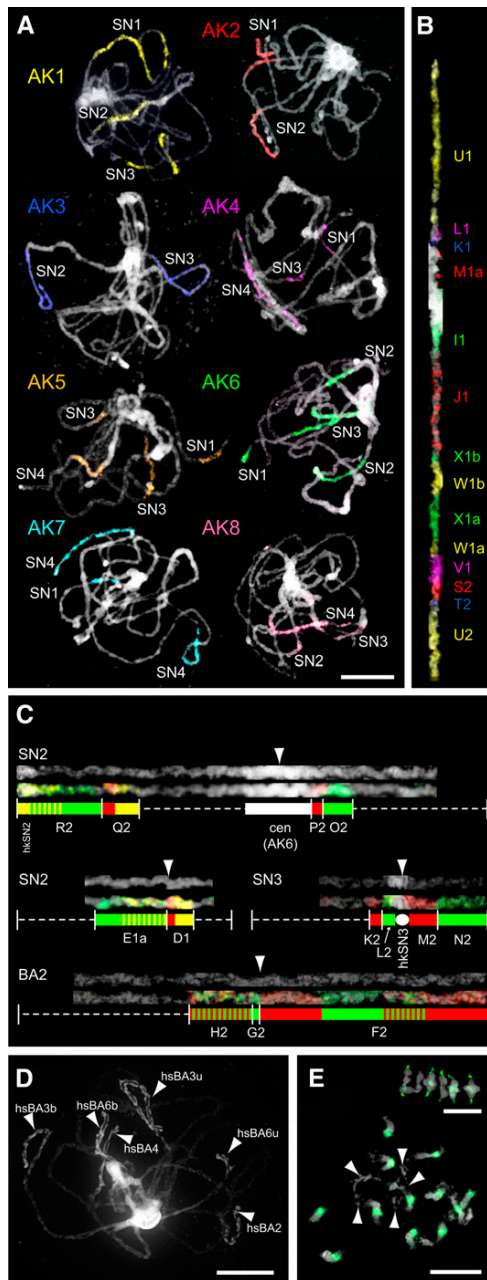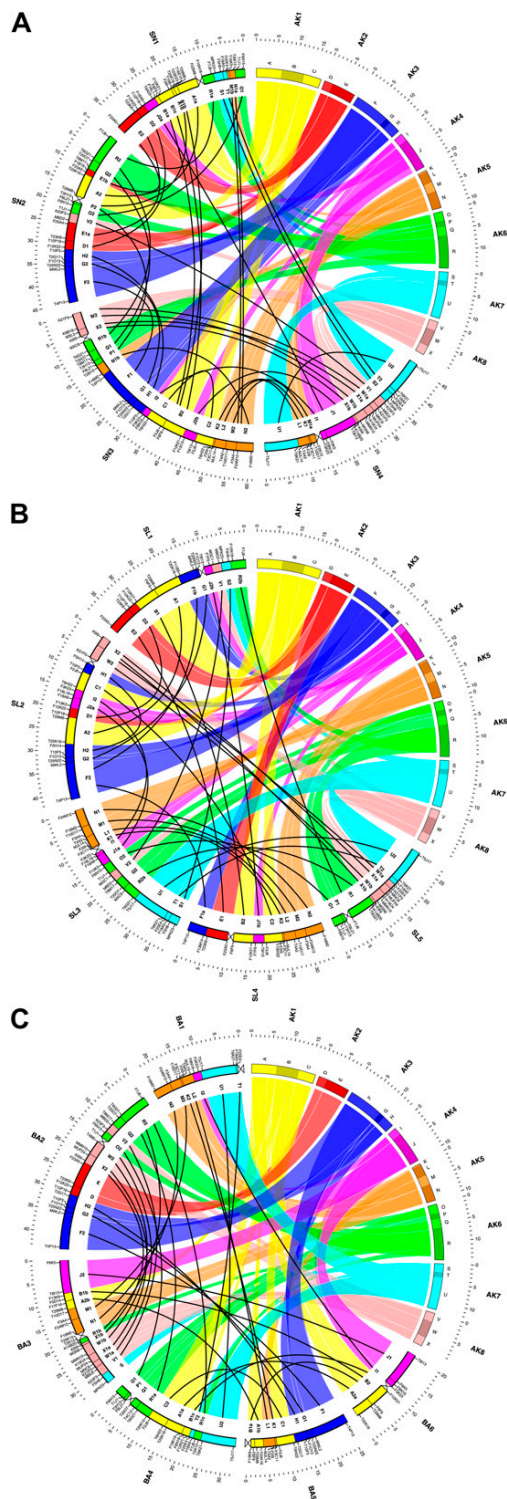
Figure 2. CCP in *Stenopetalum* and *Ballantinia* Species.

(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-

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

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.

cinnamyl alcohol dehydrogenase 5 (*CAD5*)   chalcone synthase (*CHS*)
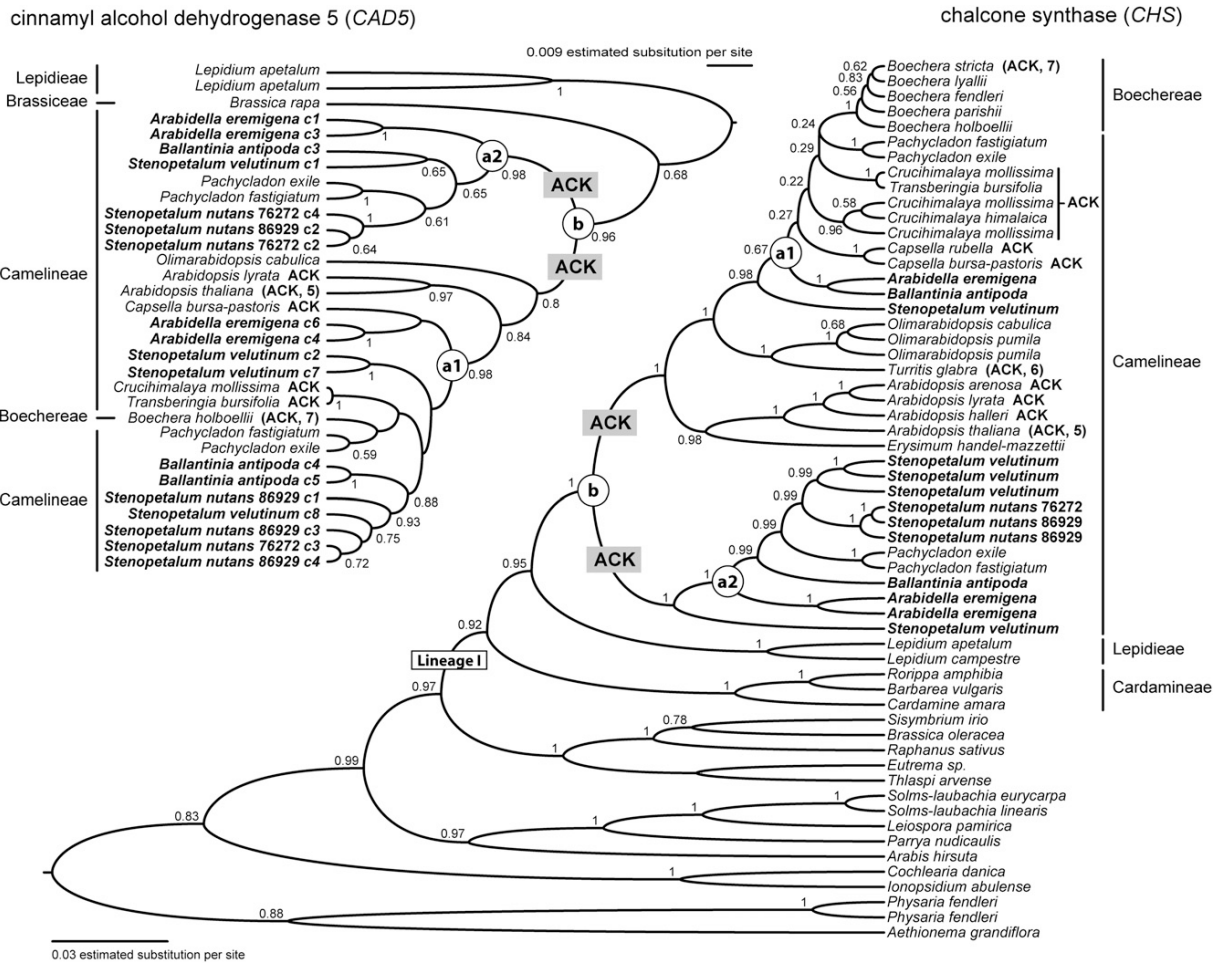
Figure 4. Phylogenies of the Nuclear Genes *CAD5* (TrN + Γ) and *CHS* (GTR + Γ + I) Showing the Position of Sequences from the Australian Species (in Bold) in the Context of Other Brassicaceae Taxa.

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).
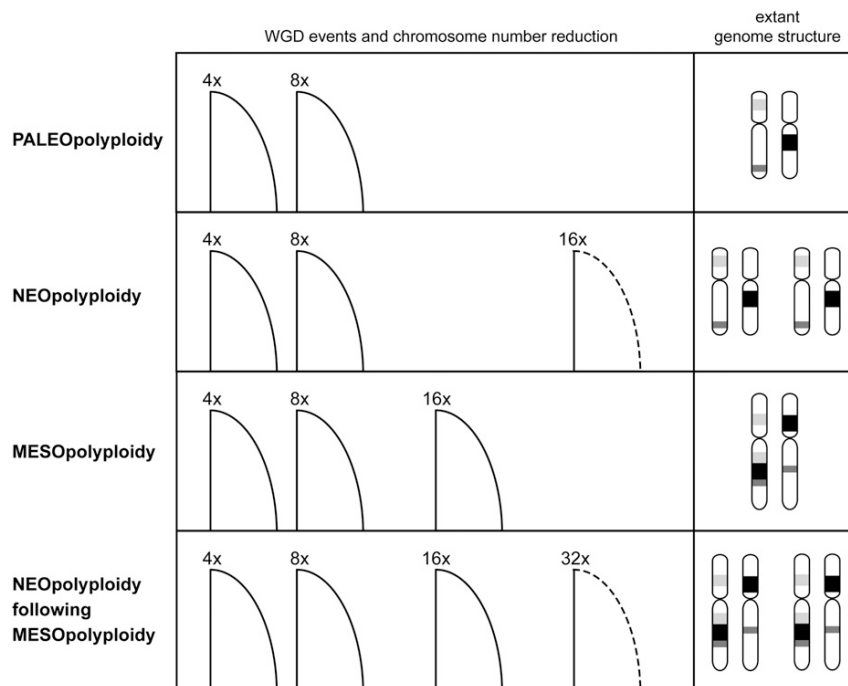
Figure 5. A Model of Genome Evolution through Multiple WGD Events Followed by Diploidization.

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 ivolving 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.

VI. **Mandáková T.**, Mummenhoff K., Al-Shehbaz I.A., Mucina L., Mühlhausen A., Lysak M.A. 2012. Whole-genome triplication and species radiation in the southern African tribe Heliophileae (Brassicaceae). Taxon 64: 989-1000.

Whole-genome triplication and species radiation in the southern African tribe Heliophileae (Brassicaceae)

Terezie Mandáková,¹ Klaus Mummenhoff,²,⁵ Ihsan A. Al-Shehbaz,³ Ladislav Mucina,⁴ Andreas Mühlhausen² & Martin A. Lysak¹,⁵

1　*RG Plant Cytogenomics, CEITEC - Central European Institute of Technology, Masaryk University, Kamenice 5, 625 00 Brno, Czech Republic*
2　*Osnabrück University, Biology Department, Botany, Barbarastrasse 11, 49076 Osnabrück, Germany*
3　*Missouri Botanical Garden, St. Louis, Missouri 63166-0299, U.S.A.*
4　*Curtin Institute for Biodiversity & Climate, Department of Environment & Agriculture, Curtin University, GPO Box U1987, Perth, 6845, Australia*
5　contributed equally to this work
Authors for correspondence:　*Martin A. Lysak, martin.lysak@ceitec.muni.cz*
　　　　　　　　　　　　　　Klaus Mummenhoff, Mummenhoff@Biologie.Uni-Osnabrueck.de

Abstract　The unigeneric tribe Heliophileae includes ca. 90 *Heliophila* species, all endemic to southern Africa. The tribe is morphologically the most diverse Brassicaceae lineage in every aspect of habit, foliage, flower and fruit morphology. Despite this diversity, virtually nothing is known about its origin and genome evolution. Here we present the first in-depth information on chromosome numbers, rDNA in situ localization, genome structure, and phylogenetic relationship within Heliophileae. Chromosome numbers determined in 27 *Heliophila* species range from $2n = 16$ to $2n = $ ca. 88, but $2n = 20$ and 22 prevail in 77% of the examined species. Chromosome-number variation largely follows three major lineages (A, B, and C) resolved in the ITS phylogeny. Clade A species mostly have a chromosome number of $2n = 20$, whereas $2n = 22$ is the dominant number in clade C ($2n = 16$ and 22 were counted in two diploid species of clade B). The number and position of 5S and 45S rDNA loci vary between species and cannot be employed as phylogenetically informative characters. Seven species with different chromosome number and from the three ITS clades were analyzed by comparative chromosome painting. In all species analyzed, 90% of painting probes unveiled three homeologous chromosome regions in *Heliophila* haploid chromosome complements. These results suggest that all *Heliophila* species, and probably the entire tribe Heliophileae, experienced a whole-genome triplication (WGT) event. We hypothesize that the mesohexaploid ancestor arose through hybridization between genomes resembling the Ancestral Crucifer Karyotype with $n = 8$. The WGT has been followed by species-specific chromosome rearrangements (diploidization) resulting in descending dysploidy towards extant quasi-diploid genomes. More recent neopolyploidization events are reflected by higher chromosome numbers ($2n = 32$–88). The WGT might have contributed to diversification and species radiation in Heliophileae. To our knowledge, this is the first study to document polyploidy as a potential major mechanism for the radiation of a Cape plant lineage.

Keywords　Cape flora; chromosome painting; comparative phylogenomics; *Heliophila*; ITS; karyotype evolution; phylogenetics; polyploidy; rDNA; whole-genome duplication

Supplementary Material　Figure S1 and Table S1 (both in the Electronic Supplement) and the alignment are available in the Supplementary Data section of the online version of this article (http://www.ingentaconnect.com/content/iapt/tax).

■ INTRODUCTION

There are multiple reasons why the unigeneric tribe Heliophileae DC. (hereafter alternatively referred to under its generic type *Heliophila* L.) is a truly remarkable group in the context of the entire Brassicaceae (Cruciferae). *Heliophila* (>90 species), the 9th-largest genus in Brassicaceae, is also one of the most complex genera from a taxonomic point of view, as suggested by more than 400 scientific names listed in the International Plant List Index (http://ipni.org), as well as 15 generic names currently placed in its synonymy (Al-Shehbaz, 2012). *Heliophila* is morphologically by far the most diverse genus in the family, especially in habit (tiny herbs to shrubs or lianas),

foliage (entire to variously dissected), floral size (1.2–25 mm) and color, petal and stamen appendages, inflorescence position (terminal vs. intercalary; Marais 1970), flower number (few to numerous), ovule number (2–80), fruit size (2–120 mm long), shape and type of fruit (silique, silicle, samara, schizocarp), fruit flattening (terete, latiseptate, angustiseptate), and dehiscence (Mummenhoff & al., 2005).

Heliophila is mainly native to South Africa; the distribution ranges of three species extend into Lesotho, seven into Namibia, and one into Swaziland. The center of the greatest diversity is the Cape Floristic Region (CFR) which has 87 species (77 endemic), and only three species (*H. formosa* Hilliard & Burtt, *H. obibensis* Marais, *H. scandens* Harv.) are endemic

to KwaZulu-Natal Province (compiled from Marais, 1970). One species (*H. pusilla* L. f.) is naturalized in Australia.

Despite two molecular phylogenetic studies (Mummen-hoff & al., 2005; Verboom & al., 2009) and recent taxonomic accounts (Marais, 1970; Al-Shehbaz & Mummenhoff, 2005), *Heliophila* remains under-explored in many respects, including genome evolution, pollination and reproductive biology, dispersal ecology, character development and evolution, rapid radiation and speciation, phylogeography, etc. Chromosome counts have been reported for only four unvouched of the 90 species (*Heliophila africana* (L.) Marais, *H. amplexicaulis* L. f., *H. crithmifolia* Willd., and *H. linearis* DC.; Jaretzky, 1932; Manton, 1932). Assessing chromosome and genome evolution in the tribe thus remained unfeasible for the last eighty years. The closest relatives of Heliophileae remain unknown, as the tribe groups with many others in a basal hard polytomy (expanded lineage II) in a multi-gene Brassicaceae phylogeny (Couvreur & al., 2010; Franzke & al., 2011). These unexplored features, substantial taxonomic complexity and the very high endemism (97%) of the genus in a relatively small CFR (only ~90,000 km²; Goldblatt & Manning, 2000), all make *Heliophila* a challenging study system.

Here we present the first in-depth information on chromosome number variation, localization of ribosomal RNA genes (rDNA), and genome structure in the light of an updated taxonomy and a comprehensive phylogeny. We characterized the extent of chromosome number variation across the tribe by analyzing approximately one third of its species. Unexpectedly, a comparative chromosome painting (CCP) analysis revealed that diploid-like genomes of *Heliophila* species have undergone a whole-genome triplication (WGT) event, which along with other factors might have promoted the unique radiation of this Brassicaceae lineage.

■ **MATERIALS AND METHODS**

Plant material. — Plant material for the phylogenetic and chromosomal analyses (including CCP) was obtained from herbarium specimens and from two field expeditions in 2008 and 2009, respectively. Ninety accessions, representing 57 *Heliophila* species sensu Marais (1970) and Al-Shehbaz & Mummenhoff (2005), were analysed. *Chamira circaeoides* (L. f.) A. Zahlbr. the only species of this genus without tribal assignment (Al-Shehbaz, 2012), was chosen as outgroup for the Bayesian analysis as done in previous molecular phylogenetic studies (Mummenhoff & al., 2005). Origin and collection data of plant material are given in Table S1 (Electronic Supplement).

Chromosome counts. — Mitotic chromosomes were counted from enzymatically digested young flower buds collected in the field. They were fixed in 3 : 1 (ethanol : acetic acid) fixative and stored in 70% ethanol at −20°C until use. For details of the protocol, see Mandáková & al. (2010a). DAPI-stained (4′,6-diamidino-2-phenylindole in Vectashield, 2 µg/ml) chromosome figures were photographed with an Olympus BX-61 epifluorescence microscope and CoolCube CCD camera (Metasystem), processed in Adobe Photoshop CS2 (Adobe

Systems), and counted. Multiple chromosome figures were analyzed per species. These chromosome spreads were probed with 5S and 45S rDNA probes to identify rRNA gene loci.

FISH: rDNA localization and comparative chromosome painting (CCP). — The *Arabidopsis thaliana* (L.) Heynh. BAC clone T15P10 (AF167571) containing 45S rRNA genes was used for fluorescence in situ localization of nucleolar organizing regions (45S rDNA). *Arabidopsis thaliana* clone pCT4.2 (M65137), corresponding to a 500-bp 5S rRNA repeat, was used to identify 5S rDNA loci. For CCP, we used *A. thaliana* chromosome-specific BAC contigs corresponding to ancestral genomic blocks on chromosomes AK1 and AK8 of the Ancestral Crucifer Karyotype (*n* = 8, Schranz & al., 2006). In the *A. thaliana* genome (http://www.arabidopsis.org/), AK1 corresponds to 17.1 Mb (block A: 6.7 Mb, B: 5.7 Mb, C: 4.7 Mb) and AK8 to 9.2 Mb (block V: 2.4 Mb, W: 4.3 Mb, X: 2.5 Mb). BAC contigs were differentially labeled with biotin-dUTP, digoxigenin-dUTP, and Cy3-dUTP by nick translation, ethanol precipitated, and hybridized to pachytene chromosomes of *Heliophila* species (for details see Mandáková & al., 2010a). Pachytene chromosome spreads were prepared as described by Mandáková & al. (2010a). DAPI-stained and painted chromosome complements were photographed, pseudo-colored and analyzed in Adobe Photoshop. Some chromosome images were straightened (Fig. 2) using the plugin 'Straighten Curved Objects' (Kocsis & al., 1991) in ImageJ v.1.38 program (Abramoff & al. 2004).

Phylogenetic analyses. — The ITS analysis presented here is based on the taxon sampling of Mummenhoff & al. (2005; 47 accessions) and 43 newly acquired accessions. Thus the current phylogeny comprises 90 accessions, representing 57 species. This represents 63% of the 90 known species.

Methods for DNA extraction, PCR, and ITS sequencing follow Mummenhoff & al. (2001, 2004). The DNA sequences were aligned automatically using the MAFFT v.6.0 multiple alignment programme (http://mafft.cbrc.jp/alignment/server/) and edited by hand. The model of nucleotide substitution for the Bayesian analysis was determined using MrModeltest v.2.2 software (Nylander, 2004; Posada, 2008) and the Akaike information criterion (AIC). Bayesian inference of phylogeny was performed using MrBayes v.3.1 (Ronquist & Huelsenbeck, 2003). Following MrModeltest, the General Time Reversible model with a proportion of invariable sites and a gamma-shaped distribution of rates across sites (GTR+R+I) was used in MrBayes. Two separate analyses were conducted, each starting from a different randomly chosen tree, on four parallel chains with a temperature for the heated chain set to 0.2. Each chain was run for 2,500,000 update cycles and every 100th cycle sampled. The initial 6250 samples (25%) were discarded as burn-in. Bayesian search results were summarized by a 85% majority-rule consensus tree. Tracer v.1.5 (Rambaut & Drummond, 2009) was used to check for convergence of the model likelihood and parameters after each run until reaching stationary status.

Molecular dating. — Node ages were estimated as follows: We incorporated the Cleomaceae/Brassicaceae split as a secondary calibration point using values of 19 Ma (Franzke

Version of Record (identical to print version).

& al., 2009) and 32 Ma (Couvreur & al., 2010) for this split. In addition, we simultaneously applied a minimum age constraint of 2.58 and 5.33 Ma for the clade containing *Rorippa* Scop. and *Cardamine* L., as this corresponds to the lower and upper age estimate of Pliocene dated *Rorippa* fruit fossils (Mai, 1995).

A relaxed-clock approach was chosen to infer the ages of the Heliophileae radiation. *Tarenaya spinosa* (Jacq.) Raf. (= *Cleome spinosa* Jacq.), *Aethionema grandiflorum* Boiss. & Hohen., *Alliaria petiolata* (M. Bieb.) Cavara & Grande, *Rorippa amphibia* (L.) Besser and *Cardamine matthioli* Moretti were used as outgroup taxa. The analyses were performed in BEAST v.1.7.1 (Drummond & Rambaut, 2012) following an approach by Drummond & al. (2006) under a relaxed-clock model. The BEAST input file was created using the BEAST user interface BEAUti v.1.7.1. In this approach, a relaxed-clock model is incorporated into a Bayesian phylogenetic inference. The analyses were conducted with a relaxed molecular clock model drawing uncorrelated rates from a log-normal distribution. We created a normal prior with an age of 32/19 Ma for the Cleomaceae/Brassicaceae split and a normal prior with an age of 2.58/5.33 Ma for the *Rorippa*/*Cardamine* split, respectively. Two runs were calibrated with two calibration points (32 Ma and 5.33 Ma; 19 Ma and 2.58 Ma) in the same analysis. All analyses were started with a random tree. We used the General Time Reversible model of nucleotide substitution with a gamma (G) distribution with four gamma categories of rates and a proportion of invariant sites (I). After tuning the operators using

the auto-optimization option, the final analysis was done with the MCMC chain-length set to thirty million. These settings resulted in ESS (effective sampling size) values, determined in Tracer, for all estimated parameters and node ages above 100, indicating a sufficient posterior distribution quality.

■ RESULTS

ITS phylogeny and dating. — Bayesian analysis of ITS data resulted in a robust phylogenetic tree of 57 *Heliophila* species, with a basal trichotomy comprising sublineages A (19 spp.), B (17 spp.) and C (19 spp.) (Fig. 1). Using different approaches to estimate and to calibrate the radiation of *Heliophila* in southern Africa, multiple previous analyses arrived at age estimations of 1.0–4.4 Ma for the crown group of Heliophileae (Table 1) (Mummenhoff & al., 2005; Verboom & al., 2009). The analysis of our new and expanded dataset using different calibration settings and BEAST confirms previous estimates, all of which date the crown group age of *Heliophila* in the Pliocene and Pleistocene (1.22–5.64 Ma).

Chromosome-number variation (Figs. 1 & 2; Table 2). — Chromosome numbers were determined in 72 populations representing 27 *Heliophila* species, representing all three infratribal clades. Chromosome numbers range from $2n = 16$ in *H. juncea* (P.J. Bergius) Druce to $2n =$ ca. 88 in *H. pubescens* Burch. (Fig. 2C, D). Twenty-one species (78%) have $2n = 20$

Table 1. Age estimates for the Heliophileae (*Heliophila*) clade and its three main lineages A, B, and C using different methods and calibration options. Numbers indicate million years ago (Ma).

		Node		
Calibration	*Heliophila*	A	B	C
Mummenhoff & al. (2005)				
Wikström & al.[a]				
NPRS[b]	4.2/4.6	3.4/3.7	2.9/3.3	3.9/4.3
NPRS jackknife[c]	3.7–5.4	2.8–4.5	2.2–3.9	n.a.
Forced clock[d]	1.9	1.3	1.0	1.7
rDNA ITS rates[e]				
Non-ultrametric distances	2.7–5.8	1.1–2.3	1.5–3.2	2.2–3.8
Forced clock[d]	2.3–4.9	1.6–3.3	1.3–2.7	2.1–4.4
Current Study (expanded dataset)				
Franzke & al. (2009)[f]	3.77	1.70	1.22	2.03
Couvreur & al. (2010)[g]	7.84	5.60	3.99	5.64

[a] The *Brassica*/*Cleome* clade was dated to 22 Ma (from Wikström & al., 2001).
[b] Tree made ultrametric using NPRS (Non-Parametric Rate Smoothing); additional calibration by constraining the *Rorippa*/*Cardamine* clade to be minimally 2.5/5.0 Ma.
[c] Estimated on jackknife re-sampled branch lengths under NPRS.
[d] Tree made ultrametric assuming a global clock.
[e] (3.9–8.3) 3×10^{-9} substitutions/site/year (from Sang & al., 1994; Zhang & al., 2001).
[f] The Brassicaceae/Cleomaceae split was dated to 19 Ma; additional calibration by constraining the *Rorippa*/*Cardamine* clade to be minimally 2.58 Ma. This age of the *Rorippa*/*Cardamine* split corresponds to the lower limit of the minimal age estimate of a *Rorippa* macrofossil (Mai, 1995).
[g] The Brassicaceae/Cleomaceae split was dated to 32 Ma; additional calibration by constraining the *Rorippa*/*Cardamine* clade to be minimally 5.33 Ma. This age of the *Rorippa*/*Cardamine* split corresponds to the upper limit of the minimal age estimate of a *Rorippa* macrofossil (Mai, 1995)

Fig. 1. Phylogeny of Heliophileae (*Helio-phila*). 85% majority-rule consensus tree from a Bayesian analysis of ITS sequence data. This is a conservative estimate of the phylogeny, where nodes with posterior probability (PP) values <0.85 were collapsed. Nodes with PP 0.95–1.0 are shown as thick lines, and normal lines represent PP 0.94–0.85. Capital letters A, B, and C refer to the main phylogenetic lineages. *Chamira circaeoides* was used as outgroup. Chromosome numbers (*2n*) and images of flowers are given for selected species.
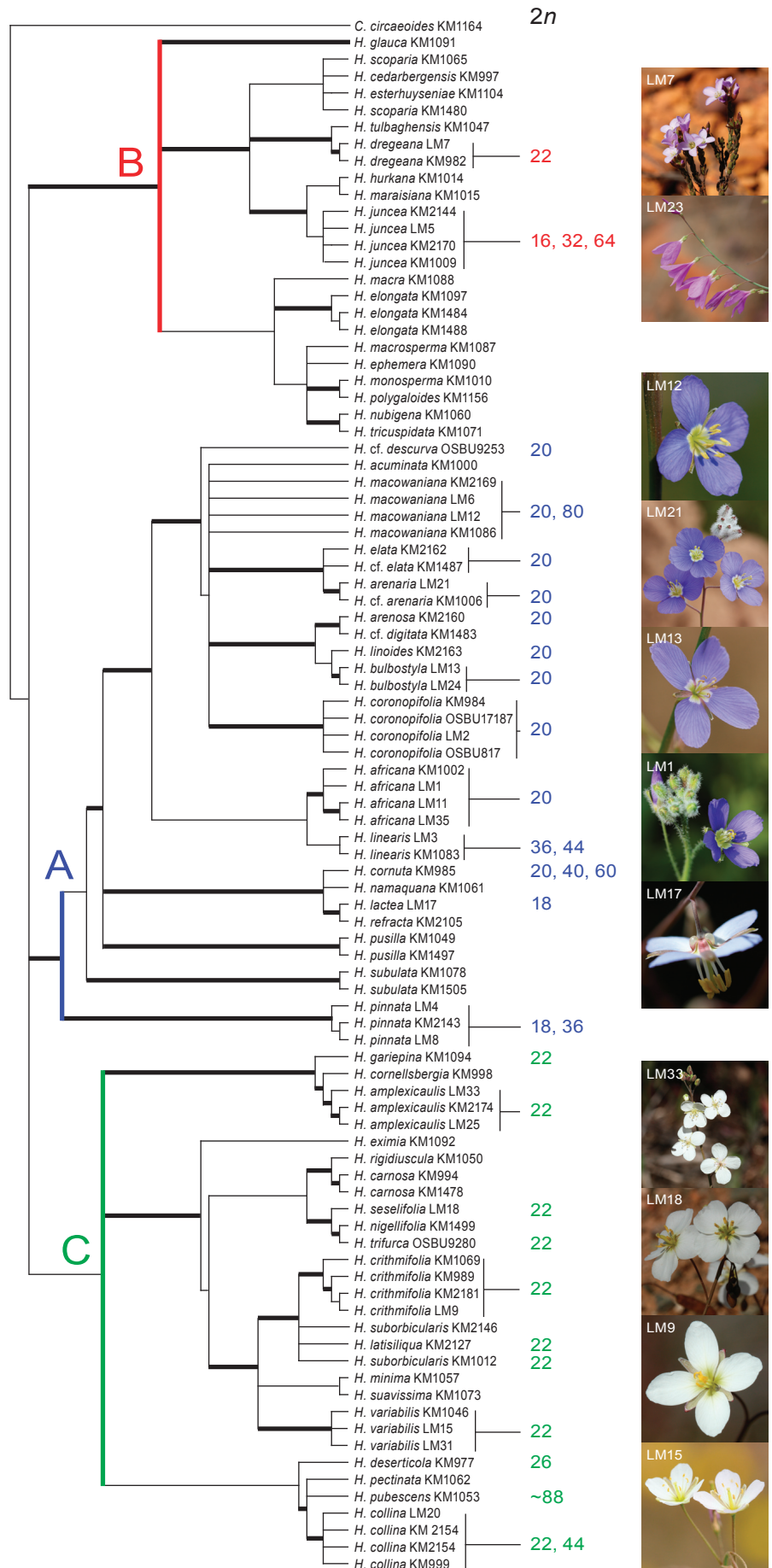
(11 spp.) or 2*n* = 22 (10 spp.) chromosomes; two species have 2*n* = 18 and one has 2*n* = 26. Five species were found to have two or three intraspecific cytotypes. For example, the morphologically well-defined *H. juncea* has diploid (2*n* = 16), tetraploid (2*n* = 32) and octoploid (2*n* = 64) cytotypes, and in *H. cornuta* Sond. diploid (2*n* = 20), tetraploid (2*n* = 40) and hexaploid (2*n* = 60) populations were found. In *H. linearis*, two chromosome counts (2*n* = 36 and 44) apparently reflect the unsettled taxonomy of this species. The size of mitotic chromosomes (ca. 2–13 μm in *H. latisiliqua* E. Mey. ex Sond.) is largely dependent on the level of chromatin compaction, whereby less condensed prophase and pre-metaphase chromosomes are longer than metaphase chromosomes (Fig. 2). In most (if not all) species, mitotic chromosome complements comprise large and small chromosomes (e.g., Fig. 2A). Heterochromatin is localized in (peri)centromeric regions and, in some species, in distinct interstitial or terminal heterochromatic knobs (Figs. 2B & 3; Fig. S1 in the Electronic Supplement). Some trends in chromosome-number variation in *Heliophila* are detected when superimposing chromosome numbers on the ITS tree

(Fig. 1). The lowest chromosome count (2*n* = 16) was found only in *H. juncea* (clade B). In clade A, 2*n* = 20 is the dominating number (78% of species analyzed), and 2*n* = 18 and 20 were found exclusively in this clade. Species of clade C mostly have 2*n* = 22 (91% of species analyzed).

rDNA (Fig. 2; Table 2). — Within diploid chromosome complements, the number of 45S rDNA loci varies from one to four pairs. Six loci were observed in tetraploid (*H. collina* O.E. Schulz, 2*n* = 44) and hexaploid (*H. cornuta*, 2*n* = 60) cytotypes, and eigth loci in tetraploid (*H. pinnata* L. f., 2*n* = 36) and octoploid (*H. macowaniana* Schltr., 2*n* = 80) cytotypes. Several clade A species have two 45S rDNA loci, whereas four loci were detected in most clade C species. However, this pattern is inconsistent between the two clades. The same is true for the number of 5S rDNA loci, which varies from 1 to 14 pairs (in *H. pinnata*, 2*n* = 36), although two pairs are found most frequently. The tetraploid cytotype of *H. pinnata* stands out from all other species as it has the highest number of both 5S and 45S rDNA loci. Interestingly, all 45S and 5S rDNA loci are localized interstitially on chromosomes in all *Heliophila* species analyzed (Fig. 2C–H). The size
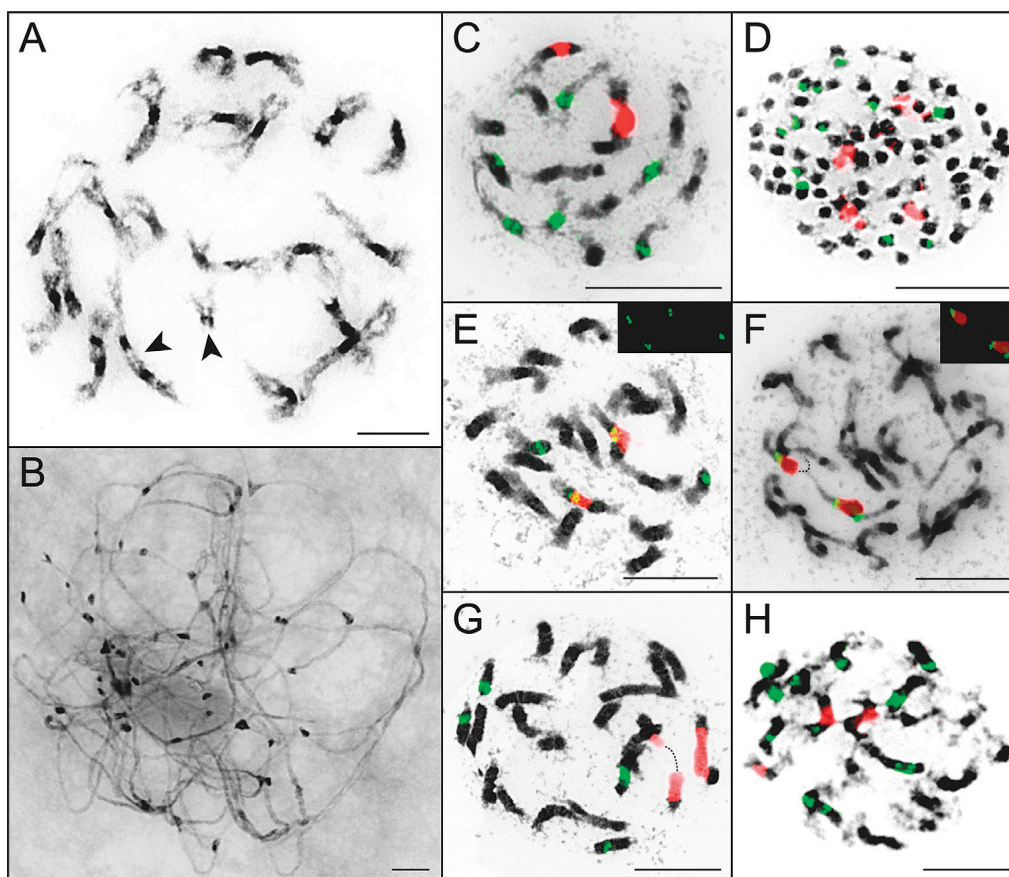
Fig. 2. Chromosome and rDNA variation in *Heliophila*. **A,** Differences in chromosome size in *H. latisiliqua* LM36 (2*n* = 22); arrowheads point to the smallest (4 μm) and largest (13 μm) chromosomes; **B,** numerous interstitial and terminal heterochromatic knobs on pachytene chromosomes of *H. linearis* LM3B (2*n* = 44); **C–H,** chromosomal localization of 5S rDNA (green) and 45S rDNA (red): **C,** one chromosome pair with interstitial 45S rDNA loci and six chromosomes with 5S rDNA in *H. juncea* KM2144 (2*n* = 16); **D,** twelve 5S rDNA loci in *H. pubescens* KM2137 (2*n* = ~88), the species with the highest chromosome number (the number of 45S rDNA loci has not been determined); **E,** *H. africana* LM11 (2*n* = 20); 5S and 45S rDNA loci are partly colocalized; **F,** *H. bulbostyla* LM24 (2*n* = 20); one chromosome pair with interstitial loci of 5S and 45S rDNA; one chromosome with a 45S locus flanked by 5S rDNA loci on both sides; **G,** unusually large 45S rDNA loci in *H. amplexicaulis* LM25 (2*n* = 22); **H,** eleven loci of 5S and three interstitial loci of 45S rDNA in *H. crithmifolia* KM2136 (2*n* = 22). — Chromosomes were counterstained with DAPI and B/W images inverted in Adobe Photoshop. Scale bars = 10 μm.

of the 45S rDNA locus can differ significantly between species (compare *H. amplexicaulis* and *H. crithmifolia*, Fig. 2G & H), but rarely differs within a chromosome complement (*H. juncea*, Fig. 2C). In two species, 5S and 45S rDNA loci are co-localized (*H. africana*, Fig. 2E) or reside in close proximity (*H. bulbostyla* P.E. Barnes, Fig. 2F). In *H. bulbostyla*, one 45S locus is flanked on both sides by 5S rDNA, probably due to an inversion within co-localized 5S and 45S rDNA loci.

Chromosome painting in diploids (Fig. 3; Fig. S1). — Two chromosomes of the Ancestral Crucifer Karyotype (ACK), AK1 and AK8, were chosen to get a glimpse of genome structure and karyotype evolution in seven *Heliophila* species. These species were chosen such that they would represent the three ITS clades and have both different ploidy levels and chromosome numbers (Fig. 1). Chromosome-specific *A. thaliana* BAC contigs, representing the two AK chromosomes and six

Table 2. Chromosome numbers and numbers of rDNA loci (vouchers at MO and OSBU).

			rDNA loci		Chromo-somes with 45S and 5S				rDNA loci		Chromo-somes with 45S and 5S
Accession	Species	2*n*	45S	5S		Accession	Species	2*n*	45S	5S	
KM2166	*H. africana*	20	2	4	2	KM2120	*H. lactea*	18	2	4	2
LM1	*H. africana*	20	2	4	2	LM17	*H. lactea*	18	2	2	2
LM11	*H. africana*	20	2	4	2	KM2127	*H. latisiliqua*	22	4	4	2
LM35	*H. africana*	20	2	4	2	LM34	*H. latisiliqua*	22	4	2	0
KM2161	*H. africana*	20	2	4	2	LM3B	*H. linearis* agg.	44	6	4	–
LM25	*H. amplexicaulis*	22	2	4	0	LM3A	*H. linearis* agg.	>30	4	10	4
LM33	*H. amplexicaulis*	22	2	4	0	KM2163	*H. linoides*	20	2	2	2
LM16	*H. arenaria*	20	2	4	2	KM2140	*H. macowaniana*	20	2	4	2
LM21	*H. arenaria*	20	2	4	2	KM2169	*H. macowaniana*	80	8	>8	0
KM2156	*H. arenaria* var. *acocksii*	20	2	4	2	LM12	*H. macowaniana*	20	2	4	2
KM2160	*H. arenosa*	20	2	2	0	LM6	*H. macowaniana*	20	2	4	2
LM13	*H. bulbostyla*	20	2	4	2	KM2143	*H. pinnata*	18	2	8	2
LM24	*H. bulbostyla*	20	2	3	2	KM2150	*H. pinnata*	18	4	>6	4
LM20	*H. collina*	22	2	4	0	LM10	*H. pinnata*	36	–	–	–
KM2101	*H. collina*	22	4	–	–	LM4	*H. pinnata*	36	8	14	8
KM2154	*H. collina*	44	6	4	0	LM8	*H. pinnata*	18	4	4	2
KM2094	*H. cornuta*	20	2	–	–	KM2137	*H. pubescens*	~88	>22	18	–
KM2117	*H. cornuta*	60	6	–	–	KM2130	*H. seselifolia*	22	4	–	–
KM2134	*H. cornuta*	40	4	–	–	KM2131	*H. seselifolia*	22	–	–	–
LM2	*H. coronopifolia*	20	2	4	2	KM2133	*H. seselifolia*	22	4	4	2
KM2116	*H. crithmifolia*	22	4	2	0	KM2135	*H. seselifolia*	22	4	–	–
KM2119	*H. crithmifolia*	22	4	4	2	KM2148	*H. seselifolia*	22	4	4	2
KM2136	*H. crithmifolia*	22	3	8	0	LM18	*H. seselifolia*	22	4	4	2
KM2147	*H. crithmifolia*	22	4	8	1	KM2138	*H. schulzii*	20	2	–	–
KM2247	*H. crithmifolia*	22	4	4	0	KM2146	*H. suborbicularis*	22	4	4	2
LM9	*H. crithmifolia*	22	4	10	0	KM2107	*H. trifurca*	22	2	4	2
LM14	*H. descurva*	20	2	4	2	KM2100	*H. variabilis*	22	4	4	–
KM2106	*H. deserticola*	26	4	4	2	KM2112	*H. variabilis*	22	4	–	–
KM2108	*H. deserticola*	26	4	–	–	KM2113	*H. variabilis*	22	4	–	–
LM7	*H. dregeana*	22	–	–	–	KM2139	*H. variabilis*	22	4	4 or 6	0
KM2162	*H. elata*	20	2	4	2	LM15	*H. variabilis*	22	4	4	0
KM2115	*H. gariepina*	22	2	–	–	LM22	*H. variabilis*	22	–	–	–
KM2121	*H. gariepina*	22	2	4	2	LM30	*H. variabilis*	22	4	6	0
KM2129	*H. juncea*	32	4	4	0	LM31	*H. variabilis*	22	4	6	0
KM2144	*H. juncea*	16	2	6	0	LM32	*H. variabilis*	22	4	4	0
LM23	*H. juncea*	32	4	9	0						
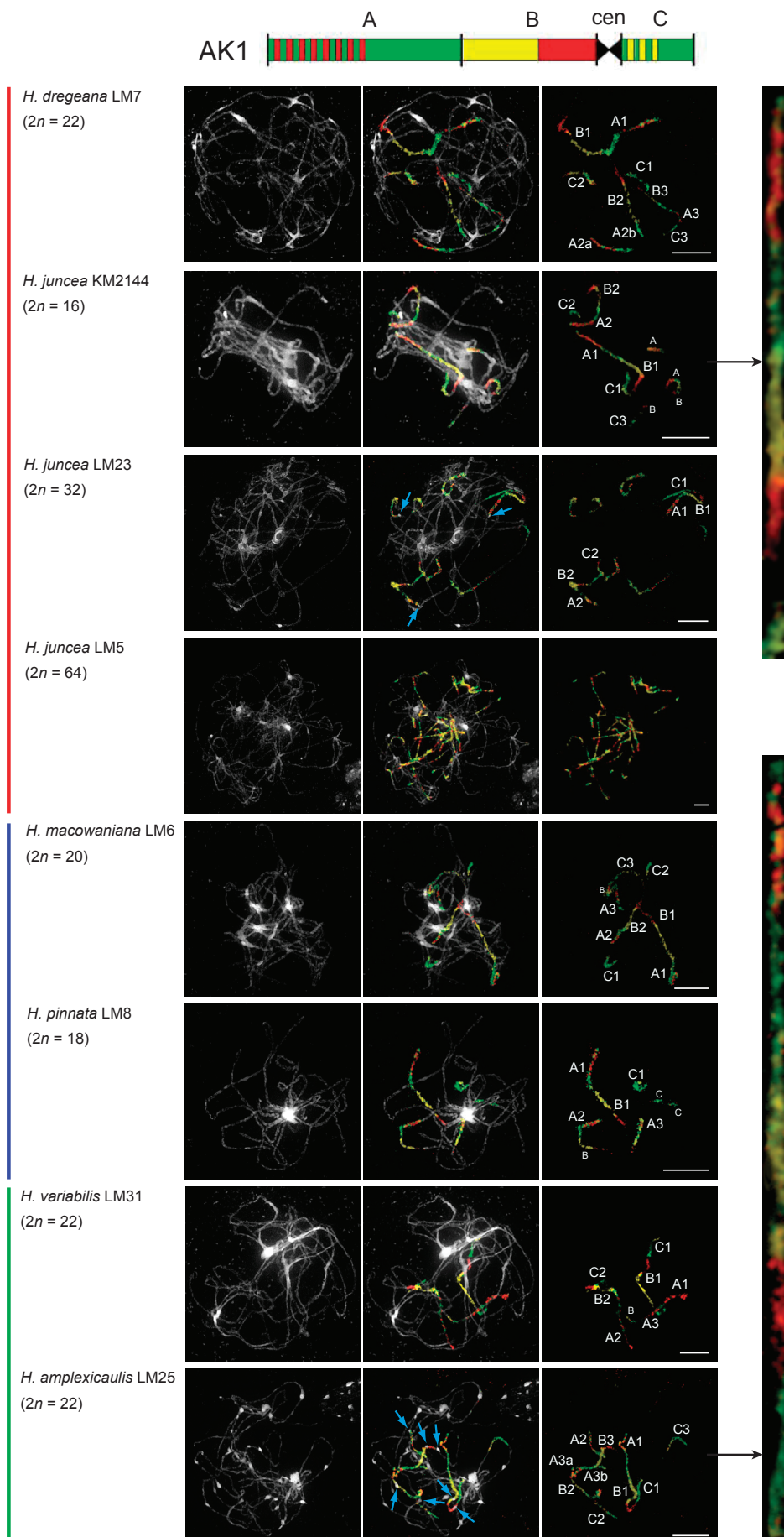
Fig. 3. Examples of comparative chromosome painting in *Heliophila* species/cytotypes with different chromosome numbers (2*n* = 16–64) and from the three infratribal clades (A–C). *Arabidopsis thaliana* BAC contigs corresponding to AK1-specific genomic blocks A, B, and C were differentially labeled and polarized, and used as painting probes on pachytene bivalents. Triplicated GBs were tentatively assigned, for example, as A1, A2, and A3. In *H. juncea* with 2*n* = 64, the complex hybridization pattern has not been analyzed. Small capitals refer to rearranged/split blocks. Blue arrows mark heterochromatic knobs. The AK1 homeologue in the diploid *H. juncea* and in *H. amplexicaulis* was straightened and shown as enlarged images on the right. Chromosomes were counterstained with DAPI. Scale bars = 10 μm.

ancestral genomic blocks (GB), were used as painting probes on *Heliophila* pachytene chromosomes. Painting probes were polarized and differentially labeled to facilitate identification of GBs building up the ancestral chromosomes (AK1: blocks A, B, and C, Fig. 3; AK8: V, W, and X, Fig. S1). The fluorescent painting signals were strong and unambiguously identified respective homeologous chromosome regions. Most GBs were found triplicated in the analysed *Heliophila* accessions with $2n = 16$, 18, 20, and 22 (100% of A and W, 97% of C, V and X, and 93% of B were triplicated). In most species, one homeologous copy was on average longer than the other two genomic copies (see the diploid *H. juncea* cytotype in Fig. 3). In 37% of AK1 copies, A and B form the ancestral association corresponding to the upper arm of AK1. Less frequently (13% of AK1 copies) the whole AK1-like chromosome structure remained preserved (see the straightened chromosomes of *H. juncea* and *H. amplexicaulis* in Fig. 3). AK8 painting probe revealed one conserved VW and one rearranged WX association, respectively, as well as partly rearranged AK8-like chromosome in *H. macowaniana*. In both the AK1- and AK8-like chromosomes, the position of the functional centromere does not match the position of the ancestral centromere in AK1 and AK8. The ancestral structure of AK1 and AK8 in *Heliophila* was frequently altered by intrachromosomal rearrangements and translocations with other (unpainted) chromosomes, whereby AK8 homeologous copies are significantly more often rearranged than segments homeologous to AK1. Interestingly, in *H. dregeana* Sond., *H. pinnata*, *H. variabilis* Burch. ex DC. and *H. amplexicaulis*, in addition to three homeologous copies of blocks X and W, one or two very short extra copies of these blocks were also found (Fig. S1).

Chromosome painting in polyploids (Fig. 3). — Patterns of inter-species chromosome homeology to AK1 were analyzed in two polyploid accessions of *H. juncea* ($2n = 4x$, $8x = 32$, 64) and in the tetraploid accession of *H. pinnata* ($2n = 4x = 36$). In the tetraploid *H. juncea* genome, one AK1-like knob-bearing chromosome and one AB association observed in the diploid cytotype were found as two identical genomic copies. In the octoploid cytotype of *H. juncea*, at least twelve homeologous copies of blocks A, B and C were found. At least six copies of these blocks were observed in the tetraploid accession of *H. pinnata*. However, the complexity of painting results prevented us from clear conclusions on the exact number and structure of homeologous segments in these polyploid accessions.

■ DISCUSSION

Infratribal phylogeny. — The present ITS study corroborates the results published by Mummenhoff & al. (2005) and demonstrates the monophyly of the unigeneric tribe Heliophileae, with *Chamira circaeiodea* as sister species. Monophyly of Heliophileae was recently also confirmed by Couvreur & al. (2010) based on a multi-gene dataset representing the nuclear, chloroplast and mitochondrial genomes. The phylogeny presented here recognizes three main lineages (A, B and C), which correspond to the main groups detected by Mummenhoff & al. (2005) based on a smaller taxon sampling. Aspects of

infrageneric classification, morphological character evolution, and eco-geographical evolution of *Heliophila* have already been discussed in detail by Mummenhoff & al. (2005).

Chromosome-number variation. — While chromosome numbers of only four *Heliophila* species have been published in the last eighty years (Jaretzky, 1932; Manton, 1932), the present study reports chromosome counts for 27 species. This still only represents less than one-third of all known Heliophileae species. Similar to many other Brassicaceae groups with small chromosomes, heterochromatic knobs and nucleolar organizing regions (often detached from chromosomes) are often erroneously counted as chromosomes or interphase chromocenters. Jaretzky (1932) already reported on difficulties with chromosome counting "in cells filled with an achromatic mass" (= chromocenters). Although chromosome numbers are known for only one-third of Heliophileae species, counts of $2n = 20$ and 22 dominate (77% of species). These chromosome numbers, though not exceptional in Brassicaceae, are restricted only to a few tribes and genera such as Anastaticeae, Brassiceae, *Iberis* L., *Leavenworthia* Torr., *Menonvillea* DC. and *Pachycladon* Hook. f. (Warwick & Al-Shehbaz, 2006). Interestingly, as in Heliophileae, $2n = 20$ and 22 are found in taxa of mesopolyploid origin, i.e., in Brassiceae (Lysak & al. 2005), *Pachycladon* (Mandáková & al. 2010a) and *Leavenworthia* (T. Mandáková and M.A. Lysak, unpub. data). Therefore, other species and genera with $n = 11$ might also have experienced mesopolyploid evolution.

Higher chromosome counts ($2n = 32$, 36, 40, 44, 60, 64, 80 and 88) represent tetra-, hexa- and octopolyploids derived from mesopolyploid diploidized cytotypes, as was shown for *H. juncea* and *H. pinnata* by our CCP results. We consider these accessions as very recent neopolyploids, representing intraspecific karyological variation. The geographic distribution of the neopolyploid cytotypes remains to be studied, and it needs to be established whether there is gene flow between different ploidy levels.

rDNA. — We analyzed the number and position of rDNA loci in 72 accessions of 27 *Heliophila* species with the aim of obtaining phylogenetically informative markers. Although some differences in the number of 45S rDNA loci between clades and individual species were observed, often identical numbers of rDNA loci preclude these markers as species-specific characters. The interstitial position of 45S rDNA, otherwise rare in Brassicaceae (Ali & al., 2005), apparently is common in Heliophileae.

Whole-genome triplication. — Triplicated homeologous chromosome regions corresponding to ancestral chromosomes AK1 and AK8 found in *Heliophila* quasi-diploid species ($2n = 16$–22) argue for a whole-genome triplication (WGT) shared by all three clades. This was an unexpected finding, as all known chromosome numbers ($2n = 20$ and 22) as well as $2n = 16$ and 18 were implicitly considered diploid. Also, the genome size (ca. 400 Mb) of two *Heliophila* species with $2n = 20$ and 22 is comparable to many truly diploid crucifer taxa (Lysak & al., 2009). As the WGT was detected by CCP in seven *Heliophila* species from all three clades, it is apparently shared by the whole tribe and predates the diversification of Heliophileae (Fig. 4). As the homeologous chromosome regions identified correspond to

Version of Record (identical to print version).

ancestral GBs and GB associations (i.e., AB of AK1 and VW of AK8) in several species, we assume that the hexaploid ancestor of Heliophileae was derived from hybridization events between two or three species with genomes resembling the Ancestral Crucifer Karyotype ($n = 8$; Schranz & al., 2006). The fact that the two homeologous copies are on average shorter than the third one, and that the chromosomes are of different size in the triplicated genomes (Jaretzky, 1932; this study), argues for an allopolyploid origin of the mesohexaploid ancestor. Similarly, the different size of homeologous chromosome segments was also reported in other mesopolyploid taxa, such as Brassiceae (Lysak & al., 2005) and Australian crucifers (Mandáková & al., 2010b). Although biased (sub)genome fractionation that usually follows ancient polyploidization events (e.g., Schnable & al., 2011) may explain the subgenome difference in *Heliophila*, we argue that a more likely scenario is that proposed for the three subgenomes in *Brassica rapa* L. (Tang & al., 2012). Analogous to *B. rapa*, the two more fractionated ("shorter") genomic copies might represent a more diploidized tetraploid ancestral genome prior the WGT, whereas the least fractionated ("longer") subgenome was brought into a hexaploid *Heliophila* by subsequent hybridization (Fig. 4). One or two very short extra copies of blocks X and W found in four diploid *Heliophila* species are more intriguing. We hypothesize that these segments may be relics of a whole-genome duplication (WGD) event predating the triplication or large segmental duplications. Clearly, more detailed analyses are needed to disentangle the genome structure of ancient diploids and the hexaploid ancestor.
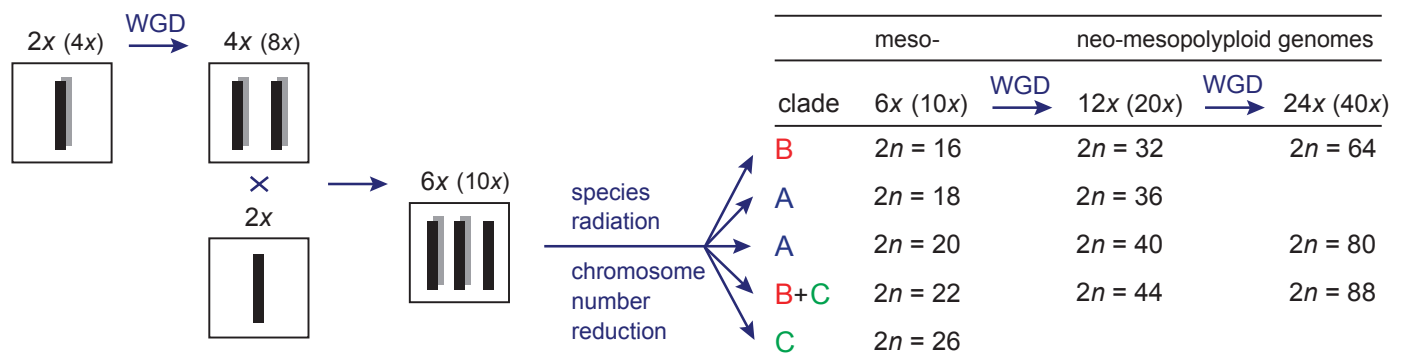
The Heliophileae-specific WGT is yet another example of a lineage-specific mesopolyploid event recently identified in Brassicaceae. Mesopolyploid triplication has been described in the ancestry of Brassiceae (Lysak & al., 2005, 2007; Tang & al., 2012), whereas mesotetraploid events mark the evolution of *Orychophragmus* Bunge (Lysak & al., 2007), Australian endemic crucifer genera (Mandáková & al., 2010b) and *Pachycladon* Hook. f. (Mandáková & al., 2010a).

WGT and karyotypic diploidization. — The WGT has been followed by descending dysploidy (chromosome-number reduction) towards the diploid-like chromosome complements found in extant *Heliophila* species. The diploidization process during which the presumably more than 40 chromosomes of the mesohexaploid ancestor were reduced to only 16 to 22 must

have been accompanied by extensive chromosome rearrangements, centromere loss, and by changes at the epigenetic and sequence levels. Similar to the Australian and New Zealand crucifers (Mandáková & al., 2010a, b), the overall collinearity of AK1 and AK8 in *Heliophila* species was retained, whereas the ancestral centromeres were deleted or lost their function. In some species analyzed, the positions of inactive paleocentromeres coincide with interstitial heterochromatic knobs (Fig. 3; Fig. S1), and knobs on several other chromosomes may also be remnants of inactive centromeres (e.g., Fig. 2B). The exact mechanism of centromere elimination in Brassicaceae taxa remains to be studied.

Has the whole-genome triplication been a driver of the species radiation in Heliophileae? — Factors that might promote speciation and/or determine speciation rates of Cape clades are intrinsic and extrinsic (Barraclough, 2006). These factors include climatic changes supposed to generate an array of responses including phenology shifts and migration facilitating both fragmentation as well as re-union of populations, potential geographical (topographic in the first place) isolating barriers (e.g., mountain ranges, deep river valleys), steepness of ecological and altitudinal gradients, regional edaphic (e.g., soil types, geology) and disturbance heterogeneity (specially fire regimes differing in spatial distribution, frequency and severity) and last but not least a plethora of biotic interactions including mycorrhiza and related phenomena and animal–plant interactions especially pollinators, pollination and dispersal (Linder, 2005; Barraclough, 2006; Verboom & al., 2009; Warren & al., 2011; Schnitzler & al., 2011; Mucina & Majer, 2012), but generally also intrinsic factors such as WGD events (Franzke & al., 2011). Consequently, the speciation rate of a particular lineage will depend on the interaction between a wide range of intrinsic and extrinsic factors.

Heliophila, as many other lineages in the CFR, has probably undergone an adaptive radiation and has diversified at the highest rates calculated for a dated Cape radiation (Warren & Hawkins, 2006). Using different methodologies and calibration concepts, the radiation was dated to 1.0–5.64 Ma (Mummenhoff & al., 2005; Verboom & al., 2009; this study, see Table 1), and all datings indicate a recent diversification against a background of increased aridity in the Late Pliocene and of alternation of cold and warm (= dry and wet, respectively)

clade	meso- 6x (10x)	WGD → 12x (20x)	WGD → 24x (40x)
B	$2n = 16$	$2n = 32$	$2n = 64$
A	$2n = 18$	$2n = 36$	
A	$2n = 20$	$2n = 40$	$2n = 80$
B+C	$2n = 22$	$2n = 44$	$2n = 88$
C	$2n = 26$		

Fig. 4. Model of genome evolution in Heliophileae (see Discussion for details). A haploid genome is shown for each putative ancestral species. Additional genome copies detected in some diploid species by CCP (Fig. S1) are depicted as grey bars and reflected by ploidy levels in parentheses.

climatic periods of the Pleistocene. As suggested by Linder (2003), the radiation of many lineages in the Cape region resembles evolutionary processes of island-species radiation. The Cape seed plant endemism of almost 70% is comparable to that found on islands. Furthermore, a high contribution by a very small number of clades to overall species richness is also typical of island floras (Linder, 2003). This indicates a great degree of isolation of the CFR. While islands are isolated by water, the CFR differs in climate, topography and soil types from the rest of southern Africa and is surrounded by oceans on three sides (Marloth, 1929; Linder, 2003). To our knowledge, no other example in Brassicaceae exhibits such a species-rich radiation confined to a relatively small region as does tribe Heliophileae.

Similar to the WGT event specific for tribe Brassiceae and its apparent impact on the infratribal diversification (238 spp. in 47 genera), we link cladogenesis, species radiation and the WGT event in Heliophileae. Ancient polyploidization and subsequent diploidization have been repeatedly shown as driving forces of genetic diversification and radiation in plant lineages (e.g., Barker & al., 2008; Soltis & al., 2009; Jiao & al., 2011). For example, gene duplication, neofunctionalization or loss may facilitate reproductive isolation and speciation. The paleopolyploid event (α) that occurred very early in the history of Brassicaceae provided the genetic raw material for radiation, and diversification rates that are among the highest reported for flowering plants (Franzke & al., 2011). We suggest that the Miocene onset of global aridity (Zachos & al., 2001; Guo & al., 2002; Liu & al., 2009) heralded the advent of a prolonged climatically unstable period punctuated by periods of enhanced aridity and later (towards Pliocene) associated with climatic oscillations. These dynamics, driven by changes in periodicity of Milankovich oscillations (Raymo & Huybers, 2008), and boasts of increased aridity during the Pleistocene (deMenocal, 2004), set the scene for the remarkable radiation of *Heliophila*. The progressing aridity presumably resulted in opening of the formerly forested landscapes (Axelrod & Raven, 1978) and the creation of more sun-exposed ('heliophilous') habitats, which underwent further changes (including alterations such as expansion, contraction, obliteration, formation) in reaction to the dry-wet cycles of the Pleistocene. Repeated isolation of populations (fragmentation) and putative reunion of migrating populations could then have acted as an 'evolutionary pump' (see also Plana, 2004 and Compton, 2011 for applications of this concept). Environmental heterogeneity, including geological and topographical complexity of the region, steep ecological gradients (see Linder, 2003; Linder & al., 2010), various large-scale disturbance factors (Rebelo & al., 2006) and small-scale animal–plant interactions (Van der Niet & al., 2006) are presumed to have been of vital importance in propelling this evolutionary pump. Phenological responses, i.e., shifts in mean flowering time (from summer towards spring) and flowering time duration, consistent with wet/dry climatic cycles, may also have contributed to the species diversity in *Heliophila* and other Cape lineages (Warren & al., 2011). We argue that the purported whole-genome triplication occurred prior to the origin of the three *Heliophila* lineages and facilitated radiation in changing environment.

Dating the radiation of Heliophileae. — Molecular dating within Brassicaceae has been limited and controversial due to a lack of fossils and calibration options. Different age estimates for the origin of Brassicaceae and its major lineages are given in Beilstein & al. (2010) and Franzke & al. (2011). Using different approaches to calibrate phylogenetic trees, or different algorithms, Mummenhoff & al. (2005) arrived at age estimates of 1.0–4.4 Ma for the crown group age of Heliophileae (Table 1). Our new and expanded ITS dataset was dated with an uncorrelated relaxed molecular clock approach (BEAST) using different calibration settings, e.g., direct or primary fossil calibration (*Rorippa*/*Cardamine* split) and secondary calibration (age of the Cleomaceae/Brassicaceae split). The age of the radiation of *Heliophila* was inferred to be ca. 3.4 (1.22–5.64) Ma, which agrees with previous studies (Table 1), all dating the diversification of *Heliophila* to the Pliocene and Pleistocene. This corresponds with findings by Warren & Hawkins (2006) who suggested that *Heliophila* diversified at the highest rates calculated for a dated Cape radiation.

Phylogenetic relationships of Heliophileae. — Nothing is known about the closest relatives of Heliophileae. However, based on the average age of the origin of *Heliophila* (1.9–7.8 Ma; Table 1), one might speculate that ancestors of *Heliophila*—like those of South African endemic *Lepidium* species—entered the African continent from southwest Asia, the presumed cradle of Brassicaceae (Franzke & al., 2011). From here they may have reached southern Africa by several putative routes, including the East African route (Hedge, 1976) and the (more probable) Arid Corridor (Verdcourt, 1969). The latter may have functioned as an intermittent link between the semi-deserts and deserts of North Africa and Arabia and the emerging arid regions of southern Africa since the Late Miocene (about 7 Ma) when the Sahara Desert started to emerge as a major arid biome (Schuster & al., 2006; Mummenhoff & al., 2001). Clearly, the closest relatives of Heliophileae are in the expanded lineage II sensu Franzke & al. (2011), and of the other 24 tribes included in that lineage, Brassiceae and Sisymbrieae, which also show some diversification in South Africa, are more likely candidates than other tribes.

Polyploidy in the evolution of Cape flora. — Based on available surveys of chromosome numbers and ploidy levels, polyploidy was considered to have only minor importance for radiations of genera centered in the CFR (Goldblatt & al., 1993; Goldblatt & Takei, 1997; Goldblatt & Manning, 2011; Prebble & al., 2011; P. Goldblatt, pers. comm.). The Heliophileae is the first group of the Cape flora where the prominent role of ancient and more recent polyploidization events has been demonstrated. Multiple base numbers in Bruniaceae (Goldblatt 1981) suggest that ancient WGD events played an important role in the evolution of this family, an endemic to the CFR. In *Oxalis* L., remarkable intraspecific ploidy variation was found in nearly 30% of Cape species investigated (Suda & al., 2012). However, it is premature to draw major conclusions on the role of WGD events in the evolution of Cape genera as karyological data are limited and no in-depth genomic studies are available. Prior to this study only four chromosome numbers were known for *Heliophila* (>90 spp.), nine species have been

counted in *Erica* (ca. 680 Cape spp., Pirie & al., 2011), and only three chromosome counts are available for the 350 species of Restionaceae (Linder, 2003 and references therein).

■ LITERATURE CITED

Abramoff, M.D., Magalhaes, P.J. & Ram, S.J. 2004. Image processing with ImageJ. *Biophotonics International* 11: 36–42.

Ali, H.B.M., Lysak, M.A. & Schubert, I. 2005. Chromosomal localization of rDNA in the Brassicaceae. *Genome* 48: 341–346.

Al-Shehbaz, I.A. 2012. A generic and tribal synopsis of the Brassicaceae (Cruciferae). *Taxon* 61: 931–954. [This issue.]

Al-Shehbaz, I.A. & Mummenhoff, K. 2005. Transfer of the South African genera *Brachycarpaea, Cycloptychis, Schlechteria, Silicularia,* and *Thlaspeocarpa* to *Heliophila* (Brassicaceae). *Novon* 15: 385–389.

Axelrod, D.I. & Raven, P.H. 1978. Late Cretaceous and Tertiary vegetation history of Africa. Pp. 77–130 in: Werger, M.J.A. (ed.), *Biogeography and ecology of southern Africa.* The Hague: Dr W. Junk Publishers.

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

Barraclough, T.G. 2006. What can phylogenetics tell us about speciation in the Cape flora? *Diversity & Distrib.* 12: 21–26.

Beilstein, M., Nagalingum, N.S., Clements, M.D., Manchester, S.R. & Mathews, S. 2010. Dated molecular phylogenies indicate a Miocene origin for *Arabidopsis thaliana. Proc. Natl. Acad. Sci. U.S.A.* 107: 18724–18728.

Compton, J.S. 2011. Pleistocene sea-level fluctuations and human evolution on the southern coastal plain of South Africa. *Quatern. Sci. Rev.* 30: 506–527.

Couvreur, T.L.P., Franzke, A., Bakker, F., Koch, M. & Mummenhoff, K. 2010. Molecular phylogenetics, temporal diversification, and principles of evolution in the mustard family (Brassicaceae). *Molec. Biol. Evol.* 27: 55–71.

deMenocal, P.B. 2004. African climate change and faunal evolution during the Pliocene-Pleistocene. *Earth Planet. Sci. Lett.* 220: 3–24.

Drummond, A.J., Ho, S.Y.W., Phillips, M.J. & Rambaut, A. 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biol.* 4(5): e88, doi:10.1371/journal.pbio.0040088.

Drummond, A.J., Suchard, M.A., Xie, D. & Rambaut, A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molec. Biol. Evol.* 29: 1969–1973.

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

Franzke, A., Lysak, M.A., Al-Shehbaz, I.A, Koch, M.A. & Mummenhoff, K. 2011. Cabbage family affairs: The evolutionary history of Brassicaceae. *Trends Pl. Sci.* 16: 108–116.

Goldblatt, P. 1981. Chromosome cytology of Bruniaceae. *Ann. Missouri Bot. Gard.* 68: 546–550.

Goldblatt, P. & Manning, J.C. 2000. *Cape plants: A conspectus of the Cape Flora of South Africa.* Pretoria: National Botanical Institute & St Louis: Missouri Botanical Garden.

Goldblatt, P. & Manning, J.C. 2011 A review of chromosome cytology in Hyacinthaceae subfamily Ornithogaloideae (*Albuca, Dipcadi, Ornithogalum* and *Pseudogaltonia*) in sub-Saharan Africa. *S. African J. Bot.* 77: 581–591.

Goldblatt, P. & Takei, M. 1997. Chromosome cytology of Iridaceae, base numbers, patterns of variation and modes of karyotype change. *Ann. Missouri Bot. Gard.* 84: 285–304.

Goldblatt, P., Takei, M. & Razzaq, Z.A. 1993. Chromosome cytology in tropical African *Gladiolus* (Iridaceae). *Ann. Missouri Bot. Gard.* 80: 461–470.

Guo, Z.T., Ruddiman, W.F., Hao, Q.Z., Wu, H.B., Qiao, Y.S., Zhu, R.X., Peng, S.Z., Wel, J.J., Yuan, B.Y. & Liu, T.S. 2002. Onset of Asian desertification by 22 Myr ago inferred from loess deposits in China. *Nature* 416: 159–163.

Hedge, I.C. 1976. A systematic and geographical survey of the Old World Cruciferae. Pp. 1–45 in: Vaughan, J.G., Macleod, A.J. & Jones, B.M.G. (eds.), *The biology and chemistry of the Cruciferae.* London: Academic Press.

Jaretzki, R. 1932. Beziehungen zwischen Chromosomenzahl und Systematik bei den Cruciferen. *Jahrb. Wiss. Bot.* 76: 485–527.

Jiao, Y., Wickett, N.J., Ayyampalayam, S., Chanderbali, A.S., Landherr, L., Ralph, P.E., Tomsho, L.P., Hu, Y., Liang, H., Soltis, P.S., Soltis, D.E., Clifton, S.W., Schlarbaum, S.E., Schuster, S.C., Ma, H., Leebens-Mack, J. & dePamphilis, C.W. 2011. Ancestral polyploidy in seed plants and angiosperms. *Nature* 473: 97–102.

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

Linder H.P. 2003. The radiation of the Cape flora, southern Africa. *Biol. Rev.* 78: 597–638.

Linder, H.P. 2005. Evolution of diversity: The Cape flora. *Trends Pl. Sci.* 10: 536–541.

Linder, H.P., Johnson, S.D., Kuhlmann, M., Matthee, C.A., Nyffeler, R. & Swartz, E.R. 2010. Biotic diversity in the Southern African winter-rainfall region. *Curr. Opin. Environm. Sustain.* 2: 109–116.

Liu, L., Eronen, J.T. & Fortelius, M. 2009. Significant mid-latitude aridity in the middle Miocene of East Asia. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 279: 201–206.

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

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

Lysak, M.A., Koch, M.A., Beaulieu, J.M., Meister, A. & Leitch, I.J. 2009. The dynamic ups and downs of genome size evolution in Brassicaceae. *Molec. Biol. Evol.* 26: 85–98.

Mai, D.H. 1995. *Tertiäre Vegetationsgeschichte Europas.* Jena: Fischer.

Mandáková, T., Heenan, P.B. & Lysak, M.A. 2010a. Island species radiation and karyotypic stasis in *Pachycladon* allopolyploids. *B. M. C. Evol. Biol.* 10: 367, doi: 10.1186/1471-2148-10-367.

Mandáková, T., Joly, S., Krzywinski, M., Mummenhoff, K. & Lysak, M.A. 2010b. Fast diploidization in close mesopolyploid relatives of Arabidopsis. *Pl. Cell* 22: 2277–2290.

Manton, I. 1932. Introduction to the general cytology of the Cruciferae. *Ann. Bot. (London)* 46: 509–556.

Marais, W. 1970. Cruciferae. Pp. 1–118 in: Codd, L.E., De Winter, B., Killick, D.J. & Rycroft, H.B. (eds.), *Flora of Southern Africa*, vol. 13. Pretoria: Government Printer.

Marloth, R. 1929. Remarks on the realm of the Cape flora. *S. African J. Sci.* 26: 154–159.

Mucina, L. & Majer, J. 2012. Ants and the origins of plant diversity in old, climatically stable landscapes: A great role for tiny players. *S. African J. Bot.* 83: 44–46.

Mummenhoff, K., Al-Shehbaz, I., Linder, P., Bakker, F. & Mühlhausen, A. 2005. Phylogeny, morphological evolution, and speciation of endemic Brassicaceae genera in the Cape flora, southern Africa. *Ann. Missouri Bot. Gard.* 92: 400–424.

Mummenhoff, K., Bowman, J.L., Linder, P., Friesen, N. & Franzke, A. 2004. Molecular evidence for bicontinental hybridogenous genomic constitution in *Lepidium* sensu stricto (Brassicaceae) species from Australia and New Zealand. *Amer. J. Bot.* 91: 252–259.

Mummenhoff, K., Brüggemann, H. & Bowman, J.L. 2001. Chloroplast DNA phylogeny and biogeography of *Lepidium* (Brassicaceae). *Amer. J. Bot.* 88: 2051–2063.

Nylander, J.A.A. 2004. MrModeltest, version 2. Program distributed by the author. Uppsala: Evolutionary Biology Centre, Uppsala University. http://www.abc.se/~nylander/mrmodeltest2/mrmodeltest2.html.

Pirie, M.D., Oliver, E.G. & Bellstedt, D.U. 2011. A densely sampled ITS phylogeny of the Cape flagship genus *Erica* L. suggests numerous shifts in floral macro-morphology. *Molec. Phylogen. Evol.* 61: 593–601.

Plana, V. 2004. Mechanism and tempo of evolution in the African Guineo-Congolian rainforest. *Philos. Trans., Ser. B* 359: 1585–1594.

Posada, D. 2008. jModelTest: Phylogenetic model averaging. *Molec. Biol. Evol.* 25: 1253–1256.

Prebble, J.M., Cupido C.N., Meudt H.M. & Garnock-Jones, P.J. 2011. First phylogenetic and biogeographical study of the southern bluebells (*Wahlenbergia*, Campanulaceae). *Molec. Phylogen. Evol.* 59: 636–648.

Rambaut, A. & Drummond, A.J. 2009. Tracer, version 1.5. Available from: http://beast.bio.ed.ac.uk/.

Raymo, M.R. & Huybers, P. 2008. Unlocking the mysteries of the ice ages. *Nature* 451: 284–285.

Rebelo, A.G., Boucher, C., Helme, N., Mucina, L., Rutherford, M.C., Smit, W.J., Powrie, L.W., Ellis, F., Lambrechts, J.J., Scott, L., Radloff, F.G.T., Johnson, S.D., Richardson, D.M., Ward, R.A., Proches, S.M., Oliver, E.G.H., Manning, J.C., Jürgens, N., McDonald, D.J., Janssen, J.A.M., Walton, B.A., Le Roux, A., Skowno, A.L., Todd, S.W. & Hoare, D.B. 2006. Fynbos Biome. Pp. 52–219 in: Mucina, L. & Rutherford, M.C. (eds.), *The vegetation of South Africa, Lesotho and Swaziland*. Pretoria: SANBI.

Ronquist, F. & Huelsenbeck, J.P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572–1574.

Sang, T., Crawford, D.J., Kim, S.-C. & Stuessy, T.F. 1994. Radiation of the endemic genus *Dendroseris* (Asteraceae) on the Juan Fernandez islands: Evidence from sequences of the ITS regions of nuclear ribosomal DNA. *Amer. J. Bot.* 81: 1494–1501.

Schnable, J.C., Springer, N.M. & Freeling, M. 2011. Differentiation of the maize subgenomes by genome dominance and both ancient and ongoing gene loss. *Proc. Natl. Acad. Sci. U.S.A.* 108: 4069–4074.

Schnitzler, J., Barraclough, T.G., Boatwright, J.S., Goldblatt, P., Manning, J.C., Powell, M.P., Rebelo, T. & Savolainen, V. 2011. Causes of plant diversification in the Cape biodiversity hotspot of South Africa. *Syst. Biol.* 60: 343–357.

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

Schuster, M., Duringer, M.P., Ghienne, J.-F., Vignaud, P., Mackaye, H.T., Likius, A. & Brunet, M. 2006. The age of the Sahara Desert. *Science* 311: 821.

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. & Soltis, P.S. 2009. Polyploidy and angiosperm diversification. *Amer. J. Bot.* 96: 336–348.

Suda, J., Krejčíková, J., Lučanová, M., Sudová, R., Trávníček, P., Urfus, T., Vít, P., Dreyer, L., Oberlander, K. & Weiss-Schneeweiss, H. 2012. Understanding whole genome processes in a world biodiversity hotspot: Flow cytometric investigations of the Cape flora. Abstract S1–4 in: Fehrer, J. & Kovařík, A. (eds.), *International Conference on Polyploidy, Hybridization and Biodiversity.* 7–10 May 2012, Průhonice.

Tang, H., Woodhouse, M.R., Cheng, F., Schnable, J.C., Pedersen, B.S., Conant, G., Wang, X., Freeling, M. & Pires, J.C. 2012. Altered patterns of fractionation and exon deletions in *Brassica rapa* support a Two-Step model of paleohexaploidy. *Genetics* 190: 1563–1574.

Van der Niet, T., Johnson, S.D. & Linder, H.P. 2006. Macroevolutionary data suggest a role for reinforcement in pollination system shifts. *Evolution* 60: 1596–1601.

Verboom, G.A., Archibald, J.K., Bakker, F.T., Bellstedt, D.U., Conrad, F., Dreyer, L.L., Forest, F., Galley, C., Goldblatt, P., Henning, J.F., Mummenhoff, K., Linder, H.P., Muasya, A.M., Oberlander, K.C., Savolainen, V., Snijman, D.A., Van der Niet, T.V. & Nowell, T.L. 2009. Origin and diversification of the Greater Cape flora: Ancient species repository, hot-bed of recent radiation, or both? *Molec. Phylogen. Evol.* 51: 44–53.

Verdcourt, B. 1969. The arid corridor between the north-east and south-west areas of Africa. *Palaeoecol. Africa* 4: 140–144.

Warren, B.H. & Hawkins, J.A. 2006. The distribution of species diversity across a flora's component lineages: Dating the Cape's 'relics'. *Proc. Roy. Soc. London, Ser. B, Biol. Sci.* 273: 2149–2158.

Warren, B.H., Bakker, F.T., Bellstedt, D.U., Bytebier, B.R., Claßen-Bockhoff, R., Dreyer, L.L., Edwards, D., Forest, F., Galley, C., Hardy, C.R., Linder, H.P., Muasya, A.M., Mummenhoff, K., Oberlander, K.C, Quint, M., Richardson, J.E., Savolainen, V., Van der Niet, T., Verboom, G.A., Yesson, C. & Hawkins, J.A. 2011. Consistent phenological shifts in the making of a biodiversity hotspot: The Cape flora. *B. M. C. Evol. Biol.* 11: 39, doi: 10.1186/1471-2148-11-39.

Warwick, S.I. & Al-Shehbaz, I.A. 2006. Brassicaceae: Chromosome number index and database on CD-Rom. *Pl. Syst. Evol.* 259: 237–248.

Wikström, N., Savolainen, V. & Chase, M.W. 2001. Evolution of the angiosperms: Calibrating the family tree. *Proc. Roy. Soc. London, Ser. B, Biol. Sci.* 268: 2211–2220.

Zachos, J., Pagani, M., Sloan, L., Thomas, E. & Billups, K. 2001. Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science* 292:686–693.

Zhang, L.-B., Comes, H.P. & Kadereit, J.W. 2001. Phylogeny and Quaternary history of the European montane/alpine endemic *Soldanella* (Primulaceae) based on ITS and AFLP variation. *Amer. J. Bot.* 88: 2331–2345.

VII. Cheng F., **Mandáková T.**, Wu J., Xie Q., Lysak M.A., Wang X. Deciphering the diploid ancestral genome of the mesohexaploid *Brassica rapa*. Submitted.

Deciphering the diploid ancestral genome of the mesohexaploid

Brassica rapa

Feng Cheng,[a] Terezie Mandáková,[b] Jian Wu,[a] Qi Xie,[c] Martin A. Lysak,[b,1] and Xiaowu Wang[a,1]

[a] Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences, Beijing 100081, China; [b] RG Plant Cytogenomics, CEITEC and Faculty of Science, Masaryk University, Kamenice 5, CZ-625 00 Brno, Czech Republic; [c] State Key Laboratory of Plant Genomics, National Center for Plant Gene Research (Beijing), Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, China

Synopsis

The ancestral genome of *Brassica* oil and vegetable crops has undergone an ancient whole-genome triplication. This work reports on the reconstruction of the three diploid *Brassica* genomes with seven chromosomes involved in the origin of the hexaploid ancestor with 42 chromosomes. The hexaploid genome evolved through extensive genome diploidization toward extant diploid-like *Brassica* genomes.

Abstract

The genus *Brassica* includes several important agricultural and horticultural crops. Their current genome structures were shaped by whole-genome triplication followed by extensive diploidization. The availability of several crucifer genome sequences, especially that of Chinese cabbage (*Brassica rapa*), enabled novel insights regarding the evolution of the mesohexaploid *Brassica* genomes from their "diploid" progenitors. We reconstructed three ancestral subgenomes of *B. rapa* ($n = 10$) by comparing its whole-genome sequence to ancestral and extant Brassicaceae genomes. All three *B. rapa* paleogenomes apparently consisted of seven chromosomes, similar to the ancestral translocation Proto-Calepineae Karyotype (tPCK, $n = 7$), which is the evolutionarily younger variant of the Proto-Calepineae Karyotype ($n = 7$). Based on comparative analysis of genome sequences or linkage maps of *B. oleracea*, *B. nigra*, *Raphanus sativus*, and other closely related species, we propose a two-step merger of three tPCK-like genomes to form the hexaploid ancestor of the tribe Brassiceae with 42 chromosomes. Subsequent diversification of the Brassiceae was marked by extensive genome reshuffling and chromosome number reduction mediated by translocation events and followed by loss and/or inactivation of centromeres. Furthermore, via interspecies genome comparison, we refined intervals for seven of the genomic blocks of the Ancestral Crucifer Karyotype ($n = 8$), thus revising the key reference genome for evolutionary genomics of crucifers.

Introduction

The genus *Brassica* currently comprises 38 species and numerous varieties, many of which are important crops or weeds. The six species that make up the so-called "Triangle of U" are of particular economic importance because of their cultivation as vegetables, condiments, and sources of oilseed (Nagaharu, 1935). These species are the "diploids" *Brassica rapa* (AA genome), *B. nigra* (BB), and *B. oleracea* (CC), and the allotetraploid species *B. juncea* (AABB), *B. napus* (AACC), and *B. carinata* (BBCC). *Brassica* is one of 47 genera in the monophyletic tribe Brassiceae. This tribe, which contains about 238 species primarily distributed in the Mediterranean region, southwestern Asia, and northern Africa (Prakash and Hinata, 1980; Al-Shehbaz, 2012), belongs to lineage II, one of three well-supported lineages recognized within the Brassicaceae family (**Figure 1A**). Within lineage II, the Brassiceae have been grouped together with tribes Sisymbrieae, Isatideae, and Thelypodieae (Schizopetaleae) (Al-Shehbaz et al., 2006; Beilstein et al., 2006; Beilstein et al., 2008; Franzke et al., 2011). However, the phylogenetic relationships within lineage II, in particular the position of the closely related tribes Eutremeae and Thlaspideae, remain unclear, and more than 20 loosely interrelated tribes have been described within an "expanded" lineage II (**Figure 1B**) (Franzke et al., 2011).
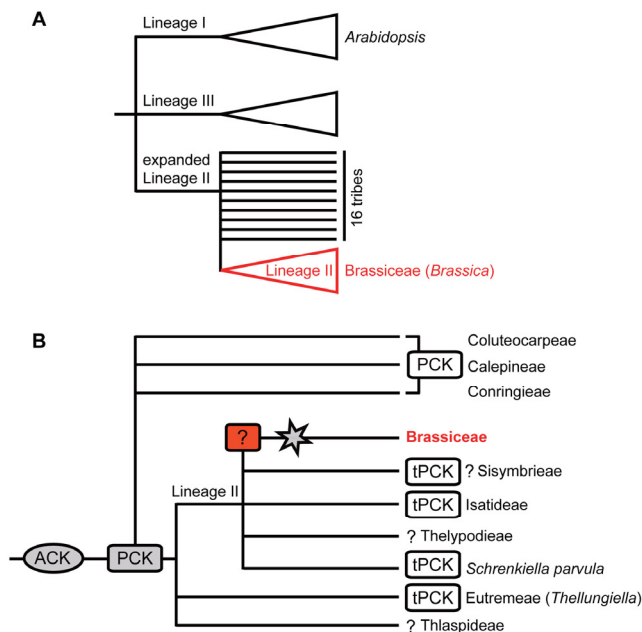
Figure 1. Currently hypothesized phylogenetic relationships among the Brassicaceae taxa discussed in this study and purported evolutionary relationships among extant and ancestral crucifer genomes (adopted and modified from Franzke et al. [2011] and Mandakova and Lysak [2008]). **(A)** Schematic phylogenetic relationships among the three major Brassicaceae lineages (I, II, and III). **(B)** Tentative genome evolution in Lineage II and five closely related tribes. The star refers to the Brassiceae-specific whole-genome triplication. Question marks indicate unknown or assumed ancestral genomes. ACK: Ancestral Crucifer Karyotype ($n = 8$); tPCK/PCK: (translocation) Proto-Calepineae Karyotype ($n = 7$).

Early chromosomal studies indicated that the diploid *Brassica* species might actually represent "balanced secondary polyploids" (e.g., Catcheside, 1934; Röbbelen, 1960). Later comparative RFLP mapping among these species and between *Brassica* species and *Arabidopsis thaliana* (Lagercrantz and Lydiate, 1996; Lagercrantz, 1998) suggested that "diploid" *Brassica* genomes descended from a hexaploid ancestor. All Brassiceae species analyzed to date contain either three or six copies of orthologous genomic regions of *A. thaliana* (Lysak et al., 2005; Parkin et al., 2005; Lysak et al., 2007). These findings provide compelling evidence for an ancestral whole-genome triplication in the ancestry of the monophyletic tribe Brassiceae. More recently, analysis of the draft genome sequence of *B. rapa* ($n = 10$) revealed the three subgenomes of the hexaploid ancestor (Wang et al., 2011). A comparison of collinearity between the 10 *B. rapa* chromosomes and the 5 *A. thaliana* chromosomes using 24 ancestral genomic blocks (GBs) of the Ancestral Crucifer Karyotype (ACK, $n = 8$; Schranz et al., 2006) revealed three syntenic copies of each *A. thaliana* GB, with only one exception, in the *B. rapa* genome (Wang et al., 2011). Because the three GB copies differ regarding their rate of gene loss (fractionation), blocks were classified according to their gene density as belonging to the least fractionated (LF), the medium fractionated (MF1), and the most fractionated (MF2) subgenome. Wang et al. (2011) suggested a two-step origin of *B. rapa* involving a tetraploidization followed by substantial genome fractionation (subgenomes MF1 and MF2), and subsequent hybridization with a third, less-fractionated genome. This two-step model gained support by comparing the frequency of deletions in exon sequences between the three *B. rapa* subgenomes (Tang et al., 2012).

Karyotype and genome evolution in the expanded lineage II were analyzed by Mandakova and Lysak (2008). Comparative chromosome painting using *A. thaliana* BAC contigs demonstrated that the six analyzed tribes (i.e., Calepineae, Coluteocarpeae, Conringieae, Eutremeae, Isatideae, and Sisymbrieae) descended from a common ancestor with the Proto-Calepineae Karyotype (PCK) ($n = 7$). While Calepineae, Coluteocarpeae (syn. Noccaeeae), and Conringieae taxa retain the PCK genome structure, the tribes Eutremeae, Isatideae, and Sisymbrieae display an additional whole-arm translocation (**Figures 1B and 2A**). This younger ancestral genome is henceforth referred to as the translocation PCK (tPCK). The tPCK genome structure was recently confirmed by whole-genome sequencing of *Schrenkiella parvula* (syn. *Thellungiella parvula*, formerly placed in the Eutremeae) (Dassanayake et al., 2011) and *Thellungiella salsuginea* (Eutremeae) (Wu et al., 2012). The PCK shares five chromosomes with the ACK and structurally closely resembles the extant karyotypes of *Arabidopsis lyrata* and *Capsella rubella* (Lysak et al., 2006; Schranz et al., 2006; Hu et al., 2011). It has therefore been suggested that either the PCK and ACK are descended from a common ancestor or, more likely, that the seven PCK chromosomes are derived from the eight chromosomes of the ACK (Mandakova and Lysak, 2008) (**Figure 1B**).

For more than half a century, the chromosome number and genomic structure of the hexaploid *Brassica* ancestor and its three parental subgenomes have been a matter of debate. Several chromosome numbers, ranging from $x = 3$ to 7, have been proposed for ancestral *Brassica* subgenomes (e.g., Röbbelen, 1960; Truco et al., 1996; reviewed by Prakash and Hinata, 1980). Based on a comparison of *B. rapa*, *B. oleracea,* and *B. nigra*

linkage maps, Truco et al. (1996) suggested that all three subgenomes were derived from an ancestor with five, six, or seven chromosomes. Although those authors favored six ancestral chromosomes (W1–W6), a seventh chromosome (W7) could not be ruled out. Comparison of the PCK with the structure of the *B. rapa* subgenome in *B. napus* (Parkin et al., 2005) revealed two shared rearrangements, thus suggesting that two or all three subgenomes of the hexaploid ancestor of the Brassiceae corresponded to either PCK or tPCK (Mandakova and Lysak, 2008).

The aim of this study was to elucidate the structure of the three ancestral subgenomes of the hexaploid ancestor of the Brassiceae. Toward that aim, structural genome analyses of *B. rapa* in comparison with ancestral and extant crucifer genomes were performed. Specifically, we tested whether the PCK and/or tPCK represent the karyotype of the diploid ancestors of the mesohexaploid *B. rapa*, the genus *Brassica*, and the tribe Brassiceae.

Results

Deciphering the Ancestral Diploid Karyotypes of *B. rapa*
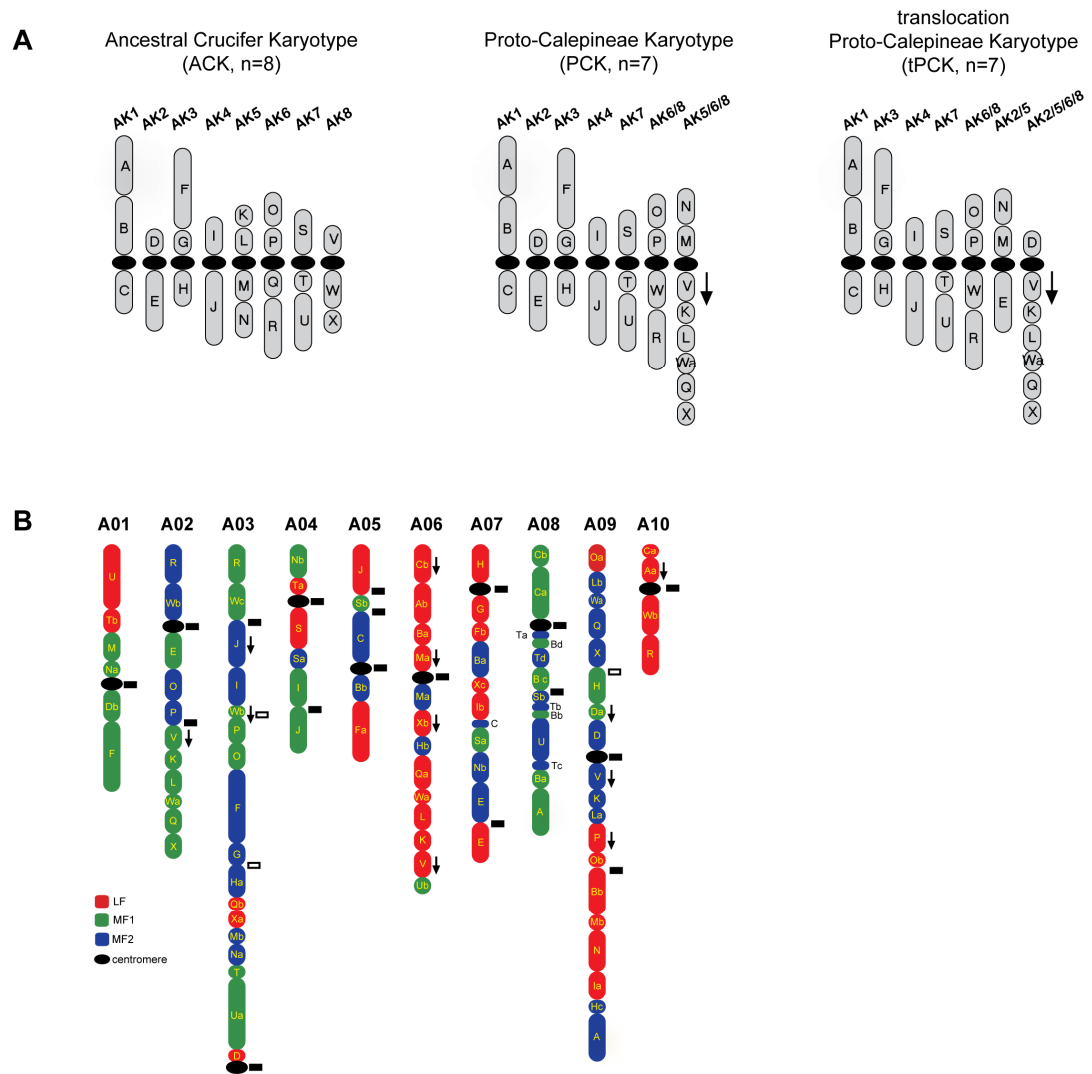
Figure 2. Ancestral Brassicaceae genomes and the distribution of ancestral genomic blocks (GBs) along 10 chromosomes of *B. rapa*. **(A)** The ancestral genomes ACK, PCK, and tPCK, each comprising 24 ancestral GBs. In the PCK and tPCK, chromosome labels reflect the presumed origin of the two karyotypes from the more ancestral ACK genome. **(B)** Ten *B. rapa* chromosomes (A01–A10) consisting of 24 triplicated GBs (71 blocks in total, as one copy of block G was lost) assigned to subgenomes least fractionated (LF, red), medium fractionated (MF1, green), and most fractionated (MF2, blue). Two or more segments of a single block are labeled by lowercase letters (a, b, etc.). Downward pointing arrows are adjacent to GBs that were inverted relative to other block(s) which originated from a single ACK chromosome. Detected and undetected ancestral centromeres in the *B. rapa* genome are shown as black and white small rectangles, respectively.

Based on the previously determined syntenic relationship of the three *B. rapa* subgenomes (LF, MF1, and MF2) to the *A. thaliana* genome (Wang et al., 2011; Cheng

et al., 2012b), 71 of 72 expected GBs (3 × 24) were detected in the *B. rapa* genome (**Figure 2B**, **Suppl. Table S1**). Of the 24 blocks, 23 were present in three copies and one (block G) in only two copies. Among the 71 GBs, 48 (67.6%) exhibited continuous sequence collinearity between *B. rapa* and *A. thaliana*, 21 (29.6%) were disrupted into at least two segments, and two (2.8%; one copy each of blocks E and T) contained sequence deletions (**Suppl. Tables S1** and **S2**). Analysis of GB integrity did not reveal significant differences in the degree of rearrangements among the three *B. rapa* subgenomes (**Suppl. Tables S3** and **S4**).

We compared the structure of the ACK, PCK, and tPCK (**Figure 2A**) with that of the three *B. rapa* subgenomes (**Figure 2B**). Using gene densities characteristic for each of the three *B. rapa* subgenomes and methods previously described (Wang et al., 2011; Cheng et al., 2012b), we then reconstructed three tentative *B. rapa* ancestral subgenomes by projecting the 71 GBs onto the ACK genome and the seven chromosomes of the PCK and tPCK. When a modern genome is compared with ancestral ones, the more recent the ancestor, the greater the expected degree of genomic continuity with the present-day genome. To examine genome continuity, we defined two adjacent genomic blocks as a GB association and compared all GB associations between *B. rapa* and the ancestral genomes ACK, PCK, and tPCK. There are 16 GB associations in the ACK genome and 17 associations in the PCK and tPCK genomes (**Table 1**, **Figure 2A**). In *B. rapa*, 10, 15, and 16 GB associations were identified between ACK, PCK, and tPCK, respectively. Of the 10 ACK-specific block associations, 4 were present in three copies in *B. rapa*, while 6 were present in two copies. Of the 15 PCK-specific GB associations, 8 were present in three copies, 6 in two, and 1 as a single copy. Of the 16 tPCK-specific GB associations, 8 were present in three copies, 6 in two copies, and 2 as single copies (**Table 1**, **Suppl. Table S5**).

Block associations V/K/L/Q/X and O/P/W/R, specific for both PCK and tPCK, were identified in all three *B. rapa* subgenomes (**Figure 2B, Table 1**). Three genomic copies of the V/K/L/Q/X association were found on chromosomes A02 (MF1), A06 (LF), and A09 (MF2). Two copies of the O/P/W/R association were rearranged on chromosomes A02 (MF2) and A03 (MF1). The third O/P/W/R copy was split between chromosomes A09 (blocks O/P) and A10 (W/R).

tPCK contained two specific GB associations—D/V and M/E—resulting from a whole-arm reciprocal translocation between chromosomes AK2 and AK5/6/8 within the PCK genome (**Figure 2A**; Mandakova and Lysak, 2008). In the MF2 subgenome, the D/V association was identified on chromosome A09. In the LF subgenome, blocks D and V were split between chromosomes A03 (MF1) and A06 (LF) due to an intersubgenomic translocation (**Suppl. Figure S1A**). The M/E association is not present in *B. rapa*. We propose, however, that the present-day block associations $M_{(LF)}/M_{(MF2)}$ (A06) and $E_{(LF)}/E_{(MF2)}$ (A07) are the products of intersubgenomic translocation between two ancestral M/E associations (**Suppl. Figure S1B**). Furthermore, evidence that the MF1 subgenome also originated from the tPCK arises from a comparison of *B. rapa* and *B. nigra* genome structures with respect to the M/E association (Panjabi et al., 2008). Based on chromosome synteny, the block association R/W/M/O/P/E on B2 in *B. nigra* is syntenic to $R/W/E_{(MF1)}/O/P$ on chromosome A02 in *B. rapa* (**Suppl. Figure S2**).

Although the M/E association was interrupted by blocks O and P on chromosome B2, the colocalization of M/E with blocks O, P, R, and W on the same chromosome supports the idea that it originated from the ancestral associations O/P/W/R and M/E.

The larger number of GB associations shared between *B. rapa* and PCK/tPCK compared with those between *B. rapa* and ACK, as well as the tPCK-specific GB associations and chromosome breakpoints found in *B. rapa*, indicate that the three *B. rapa* subgenomes bear a close structural resemblance to the seven chromosomes of the translocation Proto-Calepineae Karyotype.

To facilitate the retrieval of gene sets of the three tPCK-like subgenomes in *B. rapa* (**Suppl. Figure S3),** with mutual homoeology and synteny to *A. thaliana* orthologs, we developed a webtool—"Search Syntenic Genes Br-tPCK"— embedded in the BRAD *Brassica* Database (BRAD; http://brassicadb.org/brad/searchSyntenytPCK.php). There was a total of 12,914, 8,905 and 7,719 genes in LF, MF1 and MF2 subgenome, respectively, syntenic to *A. thaliana* (**Suppl. Table S6**).

Paleocentromeres in the *B. rapa* Genome

Paleocentromere sequence analysis indicated that the 10 *B. rapa* centromeres were all inherited from the triplicated tPCK-like genomes. Using the *B. rapa* genome sequence, we analyzed the fate of paleocentromeres corresponding to 21 ancestral chromosomes of the three tPCK-like genomes. After aligning (peri)centromere-specific repeats, including centromeric satellite repeats CentBr (Lim et al., 2005; Koo et al., 2011), PCRBr, and TR238 (Lim et al., 2007) to the *B. rapa* genome sequence, signals were detected for 18 paleocentromere regions. The paleocentromeres of three ancestral chromosomes ($AK3_{(MF1)}$, $AK3_{(MF2)}$, and $AK6/8_{(MF1)}$) could not be located (**Figures 2B and 3, Suppl. Table S7**).

Among the 18 detected paleocentromeres, 10 were involved in the origin of the 10 *B. rapa* centromeres (**Table 2**, **Figures 2B and 3**). The centromeres on chromosomes A04, A05, A07, and A09 correspond to ancestral centromeres of AK7, AK1, AK3, and AK2/5/6/8, respectively, with two ancestrally positioned GBs facing the centromeres on either sides. The centromeres of A01, A03, and A10 were derived via a translocation or pericentric inversion event from paleocentromeres of AK2/5/6/8, AK2/5/6/8, and AK6/8, respectively, leaving only one of the two GBs associated with the ancestral centromere. The origin of centromeres on A02, A06, and A08 is directly linked to a reduction in chromosome number. For these centromeres we inferred translocation events with breakpoints within the (peri)centromeric regions of two different ancestral chromosomes resulting in the loss of one of the two paleocentromeres. The available data do not allow to decide which three of the six participating paleocentromeres remained active in *B. rapa* (**Table 2, Figure 3)**.
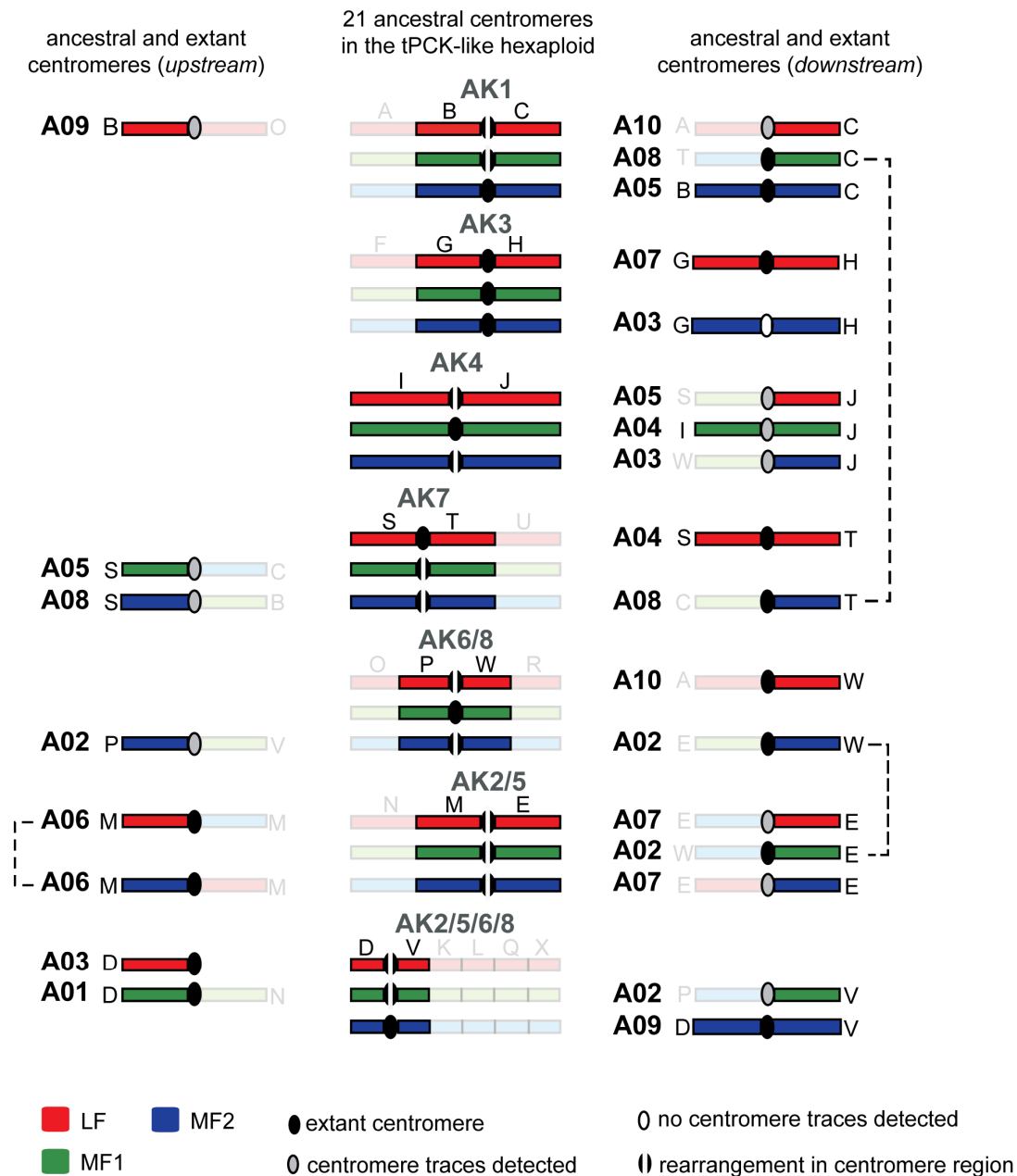
Figure 3. The fate of 21 paleocentromeres in the triplicated ancestral genome of *B. rapa*. The central column shows triplicated tPCK-like chromosomes with paleocentromeres and two adjacent genomic blocks highlighted. The left and right columns display the extant or purported location of the ancestral centromeres on *B. rapa* chromosomes. All blocks that are not ancestrally adjacent to paleocentromeres are plotted in semitransparent colors. Three instances of extant centromeres having an equally possible origin from any of two different paleocentromeres are marked by connecting dashed lines.

Of the eight remaining ancestral centromeres, six were disrupted by translocation events, with sequence traces detectable at four paleocentromere regions (**Figure 3**). Chromosomes A03 and A04 contained two entirely conserved ancestral chromosomes with inactivated/deleted centromeres. AK3 centromere left no sequence remnants between blocks G_{MF2} and H_{MF2} within A03, whereas centromere sequence traces (PCRBr and TR238 in the boundary of J/I_{MF1}) corresponding to AK4 (**Suppl. Table S7**) were detected on A04 (**Figures 2B** and **3**, **Figure 3**).

Comparison of *B. rapa* Ancestral Genomes with Other Brassicaceae Genomes

To obtain further insights into the origin and phylogenetic context of the *B. rapa* ancestral genomes, we investigated whether the tPCK-specific GB associations D/V/K/L/Q/X, O/P/W/R, M/E, and D/V are present in genomes of other *Brassica* and Brassiceae species, as well as other lineage II species (i.e., *B. oleracea, B. nigra, Raphanus sativus, Sinapis alba, Caulanthus amplexicaulis, S. parvula,* and *T. salsuginea*; **Figure 1B**). Three copies of the first two tPCK-specific block associations were identified in *B. oleracea* assembled scaffolds (**Suppl. Table S8**). In the *B. nigra* genome analyzed within the allotetraploid *B. juncea*, two copies of V/K/L/Q/X were located on chromosomes B1 and B6, whereas the third copy was split between B4 and B8. Similarly, two copies of O/P/W/R were located on B2 and B3, while the third copy was rearranged on chromosome B8 (Panjabi et al., 2008) (**Suppl. Figure S2**). Furthermore, the D/V association was found on one scaffold of *B. oleracea* (**Suppl. Table S8**) as well as on chromosomes B1 and B6 in *B. nigra* (Panjabi et al., 2008) (**Suppl. Figure S2**). Similar to *B. rapa*, the M/E association was not found in *B. oleracea*; however, one copy of M/E was rearranged with O/P/W/R on B2 in *B. nigra* (Panjabi et al., 2008) (**Suppl. Figure S2**). These analyses suggest that the three *Brassica* genomes (A, B, and C) descended from a common hexaploid ancestor that originated from the merger of three tPCK-like ancestral genomes.

For radish (*R. sativus*), we used an EST linkage map (Shirasawa et al., 2011) to align EST sequences to *A. thaliana* genes. These results were then used to determine the position of 24 GBs on the nine *R. sativus* chromosomes (**Suppl. Table S9**). Most of the 24 GBs (75%) were present in three copies in the radish genome. The two tPCK-specific block associations (V/K/L/Q/X and O/P/W/R) were detected in triplicate on chromosomes LG2, LG4, and LG6, and LG5, LG6, and LG8, respectively. In addition, two copies of V/K/L/Q/X blocks were associated with D on chromosomes LG2 and LG6, and one copy of O/P/W/R was rearranged with M/E on LG5 (**Suppl. Table S9**). This suggests that at least two tPCK-like genomes were involved in the origin of the radish genome.

In the diploid genome of *S. parvula*, tPCK-specific GB associations D/V/K/L/Q/X, O/P/W/R, and M/E were located on chromosomes SpChr2, SpChr6, and SpChr5, respectively (Dassanayake et al., 2011). tPCK-specific block associations were found to be conserved also in the *T. salsuginea* genome (Wu et al., 2012).

Comparative genome mapping in white mustard (*S. alba*; Nelson et al., 2011) and *C. amplexicaulis* (Burrell et al., 2011) also revealed tPCK-specific arrangements of genomic blocks. Because of extensive genome rearrangements and the limited resolution of the linkage maps, however, no clear conclusions could be drawn for the two species. In *S. alba*, V/K/L/V/Q on chromosomes S10 and S11 and O/P/W/R on S09 were rearranged and not readily identifiable. The M/E association was identified on S02 (Nelson et al., 2011), but the D/V association was not found. In *Caulanthus*, the block association D/V/K/L/Q/X on LG3 was fragmented to such extent that it could not be clearly identified (i.e., blocks V, L, and Q were disrupted, while K remained undetected). The O/P/W/R association was observed on LG2 (with blocks P and R fragmented), while the M/E association was not present (Burrell et al., 2011).

Analysis of Synonymous Substitution Rates

Using the syntenic orthologs identified for *B. rapa, A. thaliana, A. lyrata, S. parvula*, and *T. salsuginea*, we calculated K_s values (the number of substitutions per synonymous site) for *B. rapa* relative to the other crucifer species. K_s for *B. rapa* relative to *S. parvula* was smaller (~0.29–0.31, ~10 million years of divergence; Koch et al., 2000) than that relative to *T. salsuginea* (~0.34–0.36, ~11.7 million years), *A. lyrata* (~0.41–0.45, ~14.3 million years), and *A. thaliana* (~0.42–0.45, ~14.5 million years) (**Figure 4A**). Based on these calculations, *B. rapa* is thus phylogenetically closer to *S. parvula* than to *T. salsuginea* or to the genus *Arabidopsis*. K_s values for each *B. rapa* subgenome relative to *S. parvula* ranged from 0.29 to 0.31 (**Figure 4B**), while those for the subgenomes relative to one another ranged from 0.29 to 0.33 (i.e., ~10.3 million years) (**Figure 4C**).

To investigate the evolutionary relationship of *A. thaliana, A. lyrata, S. parvula, T. salsuginea*, and the three *B. rapa* subgenomes, we carried out a phylogenetic analysis using shared synonymous SNPs for these five genomes and the two outgroup species *Carica papaya* and *Vitis vinifera*. From the 591 syntenic orthologs shared by the three *B. rapa* subgenomes and the six other species, 63,239 K_s loci (i.e., synonymous SNPs; see Methods) were identified and concatenated into a dataset that was subjected to phylogenetic analysis. Based on the resulting tree, Brassiceae and *Arabidopsis* ancestors diverged prior to the Brassiceae-*Thellungiella* split. The Brassiceae-specific whole-genome triplication occurred near the time of the split between the progenitors of the Brassiceae and *Schrenkiella* (**Figure 4D**).
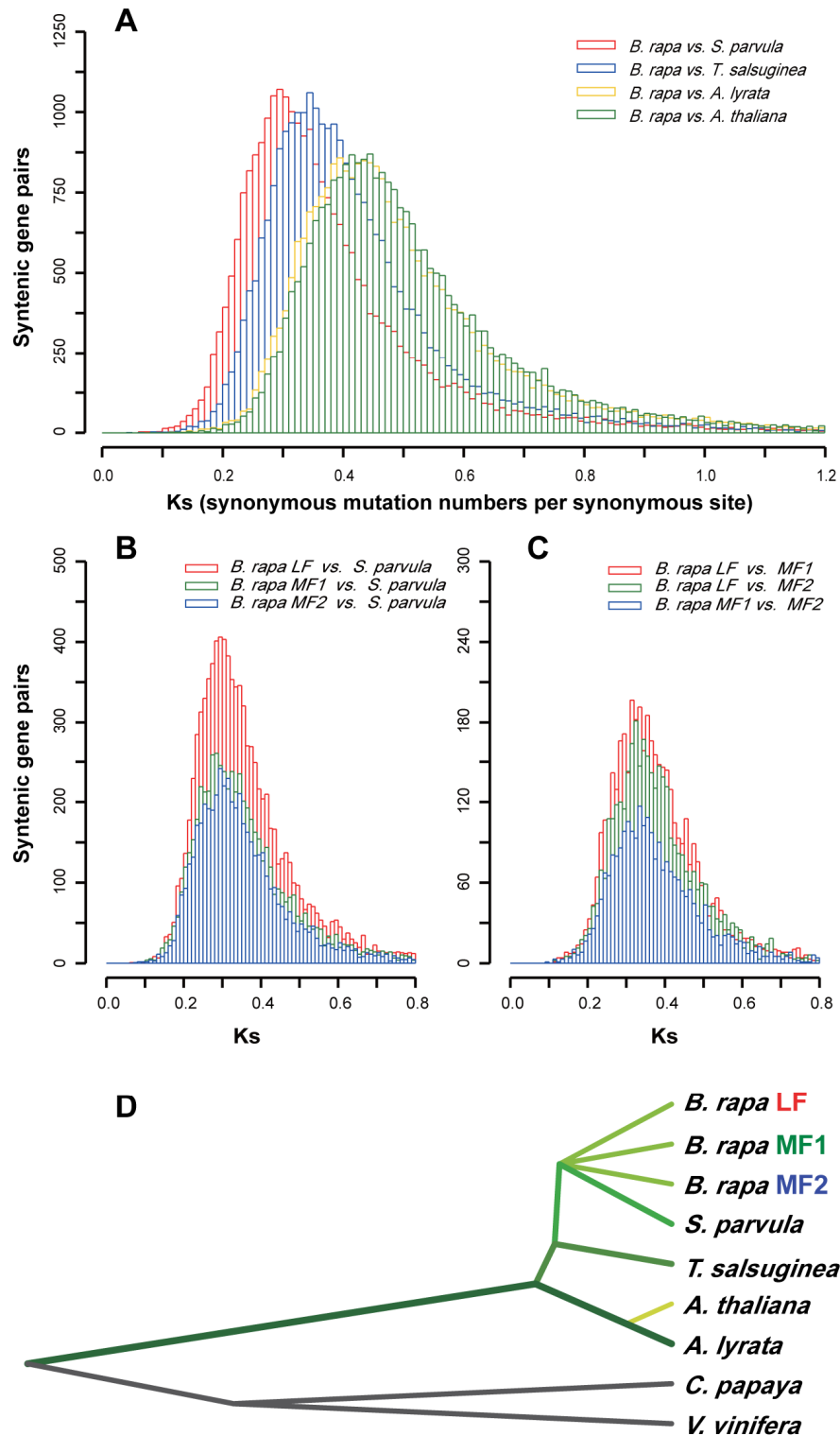
Figure 4. Pairwise comparison of K_s values and a phylogenetic tree based on K_s loci. **(A)** The distribution of Ks values between *B. rapa* and crucifer species *T. salsuginea, S. parvula, A. thaliana*, and *A. lyrata*. **(B)** The distribution of K_s values between each of the three *B. rapa* subgenomes and *S. parvula*. **(C)** The distribution of K_s values between pairs of the three *B. rapa* subgenomes. **(D)** Phylogenetic tree for the four crucifer species and the three *B. rapa* subgenomes, with non-Brassicaceae species *Carica papaya* and *Vitis vinifera* used as outgroups. The tree was constructed based on K_s loci for all syntenic orthologs shared among the seven species.

Reevaluation of Block Association V/K/L/Wa/Q/X and Interval Refinement of Seven Genomic Blocks
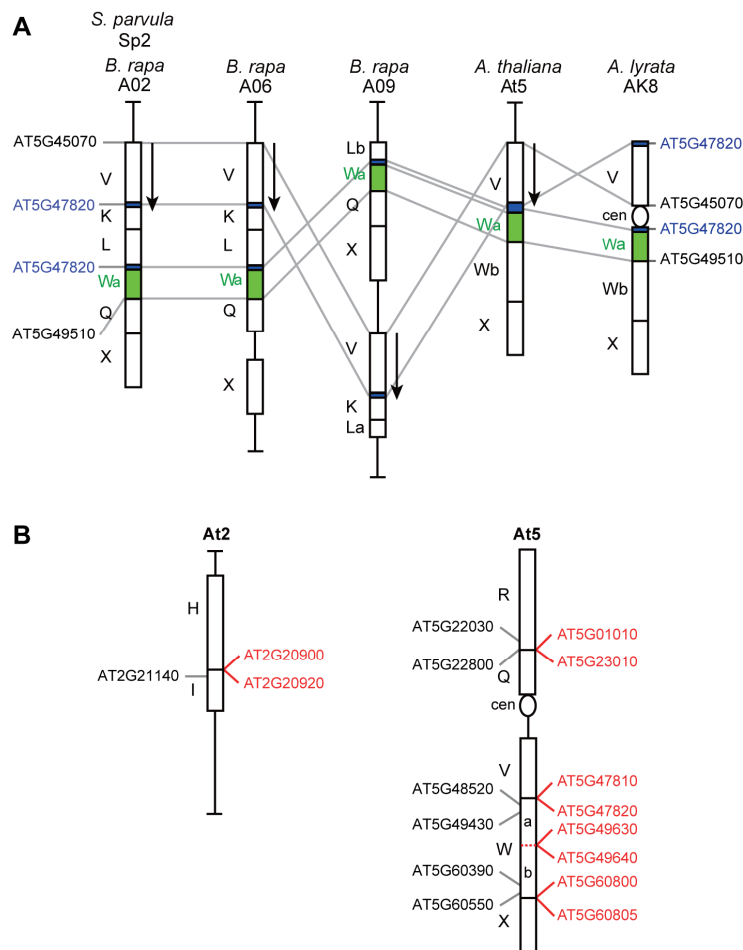
Figure 5. Redefinition of intervals for genomic blocks (GBs) H, I, Q, R, V, W, and X, and the interspecies collinearity of the diagnostic block association V/K/L/Wa/Q/X. Block intervals were redefined based on interspecies comparison among genomes of *A. lyrata* (Scaffold_8 on chromosome AK8), *A. thaliana* (At5), *S. parvula* (Sp2), and *B. rapa* (A02, A06, A09). **(A)** The diagnostic block association V/K/L/Wa/Q/X. The split of block W into Wa and Wb preceded the formation of the block association V/K/L/Wa/Q/X in *B. rapa*, *S. parvula*, and probably all species of the expanded lineage II. Block V is inverted in *A. thaliana, S. parvula,* and *B. rapa* compared with the ancestral ACK-like orientation in *A. lyrata*. **(B)** Refined intervals of seven ancestral GBs shown as corresponding *A. thaliana* genes located on chromosomes At2 and At5. GB intervals as defined by Schranz et al. (2006) appear in black on the left, whereas newly refined intervals are listed in red on the right. The boundary between Wa and Wb is located between AT5G49630 and AT5G49640.

The results of the multispecies comparison prompted a reevaluation of the V/K/L/Q/X block association considered specific for the $x = 7$ tribes of expanded lineage II (Mandakova and Lysak, 2008). This association was identified on chromosome Sp2 in *S. parvula*, albeit with a fragment of block V between L and Q (V/K/L/V'/Q/X) (Dassanayake et al., 2011). The same rearrangement was found on *B. rapa* chromosomes

A02, A06, and A09 (**Figure 5A**), *B. oleracea* (**Suppl. Table S8**) and *R. sativus* (Li et al., 2011; Shirasawa et al., 2011). We analyzed this chromosome region by performing an interspecies comparison between *B. rapa*, *B. oleracea*, *S. parvula, A. thaliana* and *A. lyrata,* with the *A. lyrata* genome (Hu et al., 2011) serving as a proxy for the ACK genome ($n = 8$). The presumed V' region turned out to be part of block W, located on the bottom arm of the AK8 (Al8) chromosome and adjacent to the centromere (**Figure 5A**). The origin of the V/K/L/Wa/Q/X rearrangement, which is specific to the ancestral PCK genome, is consistent with the previously proposed scenario (Mandakova and Lysak, 2008), except that one of the breakpoints is actually positioned within block W (**Figure 5B and Suppl. Figure S4A**). An alternative scenario involves a paracentric inversion of blocks V/K/L followed by inactivation/removal of the AK8 paleocentromere (**Suppl. Figure S4B**). These results, together with multispecies comparisons of syntenic chromosomes, allowed us to redefine the boundaries of blocks V and W (**Table 3** and **Figure 5A**).

By comparing the boundaries of blocks H, I, Q, R, V, W, and X reported by Schranz et al. (2006) with corresponding regions and/or breakpoints within genomes of *B. rapa*, *A. lyrata*, *S. parvula*, and *A. thaliana* (**Suppl. Table S10**), we revised intervals of the seven blocks of ACK (**Table 3** and **Figure 5B)**.

Discussion

We compared the whole-genome sequence of *B. rapa* (Wang et al., 2011) with genome sequences or genetic maps of other crucifer species, including *A. thaliana, A. lyrata, B. oleracea, B. nigra, C. amplexicaulis, R. sativus, S. alba, S. parvula*, and *T. salsuginea*, as well as with previously-proposed Brassicaceae ancestral genomes (Schranz et al., 2006; Mandakova and Lysak, 2008; Panjabi et al., 2008; Burrell et al., 2011; Dassanayake et al., 2011; Hu et al., 2011; Li et al., 2011; Nelson et al., 2011; Shirasawa et al., 2011; Wu et al., 2012). We have provided conclusive evidence that the *B. rapa* genome arose via a whole-genome triplication event involving three *n* = 7 genomes structurally resembling the ancestral translocation Proto-Calepineae Karyotype (tPCK).

tPCK-like Ancestral Genomes and the Origin of the Mesohexaploid *Brassica* Genome

Since the hypothesis was proposed that diploid *Brassica* species represent "balanced secondary polyploids" (e.g., Catcheside, 1934; Röbbelen, 1960), numerous lines of evidence have supported the occurrence of an ancestral whole-genome triplication prior to species radiation of the tribe Brassiceae (e.g., Lagercrantz, 1998; Lysak et al., 2005; Parkin et al., 2005; Wang et al., 2011). In this study, we inferred the tPCK-like genome with seven chromosomes to be the basic diploid genome of the hexaploid ancestor of Brassiceae. The hexaploid genome likely evolved through hybridization between a tetraploid and diploid progenitor genome. Evidence for this is based on genetic mapping (Lagercrantz, 1998; Babula et al., 2003), comparative cytogenetic analyses (Lysak et al., 2005; Ziolkowski et al., 2006), and, most convincingly, whole-genome sequence analysis. Wang et al. (2011) found that the three subgenomes in *B. rapa* have experienced different levels of gene loss. The more extensive fractionation of two subgenomes (MF1 and MF2) compared with the third subgenome (LF) suggest a two-step origin for the hexaploid genome. Several recent investigations (Cheng et al., 2012b; Tang and Lyons, 2012; Tang et al., 2012) convincingly support the origin of the Brassiceae ancestral genome according to the formula $4x \times 2x \rightarrow 3x \rightarrow 6x$. Here, we provide evidence that all three participating ancestral genomes had seven chromosomes and that the hexaploid ancestor probably possessed 42 chromosomes ($2n = 6x = 42$).

tPCK in a Phylogenetic Context

Because species of the genus *Brassica* and closely related genera, such as *Raphanus* and *Sinapis,* can be intercrossed, and generic circumscriptions based on morphological characters are unreliable, congeneric species have frequently been assigned to two different sister phylogenetic clades within Brassiceae: Rapa/Oleracea and Nigra (e.g., see Warwick and Black, 1991). To better understand the structure of *Brassica* progenitor genomes, we compared tPCK-specific block associations with chromosome structures of Rapa/Oleracea (*B. oleracea, B. rapa,* and *R. sativus*) and Nigra (*B. nigra*) species. We found that all three Rapa/Oleracea species (*B. oleracea*, *B. rapa*, and *R. sativus*) retained three copies of the V/K/L/Wa/Q/X and O/P/W/R associations. D/V or M/E associations

were also found in the three species. In addition, chromosome synteny among *B. nigra*, *B. oleracea*, and *B. rapa*, and the existence of the tPCK-specific associations D/V and M/E in *B. nigra* and D/V in *B. oleracea* and *R. sativus* indicate that *B. nigra*, *B. oleracea*, and *R. sativus* had the same origin as *B. rapa*. Furthermore, some tPCK-specific associations observed within the comparative linkage map of *S. alba* (Nelson et al., 2011), a species from the Nigra clade, suggest that the genome of this species is also derived from a triplicated tPCK-like ancestral genome.

The tribes Brassiceae, Isatideae, Thelypodieae, and Sisymbrieae were designated as lineage II of the Brassicaceae (Al-Shehbaz et al., 2006; Beilstein et al., 2006; Beilstein et al., 2008). The relationship, however, of these tribes to the two other Brassicaceae lineages and to the remaining 45 tribes remains uncertain. Franzke et al. (2011) grouped lineage II with 18 other tribes to form an expanded lineage II. Using comparative chromosome painting, Mandakova and Lysak (2008) provided valuable insight into the phylogenetics of expanded lineage II. They demonstrated that tribes Calepineae, Coluteocarpeae, Conringieae, Eutremeae, Isatideae, the genus *Ochthodium* (formerly Sisymbrieae), and probably the Brassiceae descended from an ancestral PCK genome. In Eutremeae, Isatideae and *Ochthodium*, PCK was altered by a whole-arm translocation giving rise to tPCK. The tPCK genome structure was recently confirmed by draft genome assemblies in *T. salsuginea* (from Eutremeae) (Wu et al., 2012) and in *S. parvula* (Dassanayake et al., 2011). Although the tribal placement of *Schrenkiella* is still uncertain (Al-Shehbaz, 2012), our results with respect to K_s values place Brassiceae closer to *Schrenkiella* than to Eutremeae (*Thellungiella*) and indicate that the Brassiceae triplication event occurred near the time of the divergence between Brassiceae and *Schrenkiella* (~10 million years ago). Collectively, our data suggest that the ancestor of Brassiceae, Isatideae, *Schrenkiella*, *Ochthodium*, and Eutremeae, and probably also Thelypodieae (*Caulanthus*), descended from the same diploid tPCK-like ancestral genome. While the triplicated tPCK genome has undergone extensive repatterning in Brassiceae, the ancestral structure of the tPCK genome has been conserved in diploid *x* = 7 taxa, such as *Schrenkiella* or *Thellungiella*, where no diploidization of a polyploid genome could occur.

Descending Dysploidy in the Hexaploid Ancestor of Brassiceae
The chromosome number of $2n = 42$ of the hexaploid ancestor has been reduced by a factor of two during the evolution of Brassiceae (Warwick and Al-Shehbaz, 2006). Such reductions are associated with the origin of composite chromosomes and loss of chromosome segments such as centromeres, telomeres, and nucleolar organizing regions. In *B. rapa*, descending dysploidy led to the origin of 10 "fusion" chromosomes and the elimination of 11 paleocentromeres. Some paleocentromeres were eliminated through symmetric translocations, thus resulting in products of unequal size: a large "fusion" chromosome, and minichromosome which contains mainly a centromere and telomeres and usually gets lost during meiosis (Lysak et al., 2006; Schranz et al., 2006; Schubert and Lysak, 2011). Other ancestral chromosomes were combined by asymmetric end-to-end translocations (e.g., between AK6/8 and AK3; in A03; **Figure 2B**). End-to-end translocation events must be accompanied by centromere loss or

inactivation on one of the participating chromosomes to prevent unstable dicentric chromosomes. Centromeres were apparently lost from the collinear GBs F/G/H (AK3) of A03 and I/J (AK4) of A03 and A04 (**Figure 2B**). Nested chromosome fusions (i.e., insertion of one chromosome into the centromere of a recipient chromosome), which occurred frequently in the evolution of grass species (Luo et al., 2009; Salse et al., 2009; Murat et al., 2010), have not been detected in the karyotype of *B. rapa*. In Brassicaceae, single such events apparently occurred in *Pachycladon* species (Mandakova et al., 2010a) and in *Hornungia alpina* (Lysak et al., 2006).

We also analyzed whether recombination between homoeologous chromosome segments of the three *B. rapa* subgenomes exceeded recombination between non-homoeologous segments. If recombination between homoeologous segments was prevalent, we would expect to find the same GBs located adjacent to one another (e.g., association $A_{(MF1)}/A_{(MF2)}$). In *B. rapa*, only four such associations were detected ($S_{(LF)}/S_{(MF2)}$ on A04, $M_{(LF)}/M_{(MF2)}$ on A06, $E_{(MF2)}/E_{(LF)}$ on A07, and $D_{(MF1)}/D_{(MF2)}$ on A09). These data suggest that interchromosomal rearrangements were not mediated preferentially by recombination between homoeologous GBs. Nevertheless, the true picture may have been blurred by numerous rearrangements (typically inversions) that may have occurred later and reshuffled originally adjacent homoeologous genomic blocks.

Whole-Genome Sequence Assembly Improved the Resolution of Comparative Paleogenomics in Brassicaceae

The *B. rapa* genome sequence (Wang et al., 2011) provided much more detailed information on genome structure than previously constructed linkage maps. Although most genomic blocks and parts thereof (68%) were detected by genetic mapping (Parkin et al., 2005; Panjabi et al., 2008; Kim et al., 2009; Trick et al., 2009), some blocks and block associations were missed or reported as being present in less than three copies. A comparison of the *B. rapa* genome sequence with previously reported linkage maps shows that block D on chromosome A01; blocks W, Q, X, M, and D on A03; S on A05; M, H, and U on A6; F, X, and S on A07; S on A08; K, L, and M on A09; and C on A10 were not detected by genetic mapping (Parkin et al., 2005; Panjabi et al., 2008; Kim et al., 2009; Trick et al., 2009). Although the assembly of *B. rapa* chromosome A03 by Mun et al. (2010) identified 14 GBs or block segments, three additional blocks on the same chromosome were detected within the *B. rapa* reference genome (Wang et al., 2011). Similarly, the 2.4-Mb fragment of block W within the V/K/L/Wa/Q/X association was not detected by genetic mapping. Moreover, two adjacent copies of the same GB belonging to different subgenomes have often been regarded as a single block in previous genetic maps, including blocks $S_{(LF)}/S_{(MF2)}$, $E_{(LF)}/E_{(MF2)}$, and $D_{(MF1)}/D_{(MF2)}$ on chromosomes A04, A07, and A09, respectively (see Panjabi et al., 2008 and **Figure 2B**). Furthermore, we found that earlier reports on GBs represented by more than three copies were mostly due to misinterpretation of block segments as entire blocks. For instance, Panjabi et. al (2008) had observed four copies of block A located on chromosomes A06, A08, A09, and A10. Analysis of *B. rapa* genome sequences revealed, however, that the two copies of block A on chromosomes A06 and A10 were in fact two

segments of this block (**Figure 2B**). These examples demonstrate that whole-genome sequence assemblies largely eliminate the risk of misinterpreting low-resolution genetic data.

Refining Intervals of Ancestral Genomic Blocks

The original concept of 24 GBs (Schranz et al., 2006) had resulted from the synthesis of comparative genetic and cytogenetic data for *A. thaliana*, *A. lyrata*, *C. rubella*, the A and C genomes of *B. napus*, and other lineage I species (Boivin et al., 2004; Kuittinen et al., 2004; Parkin et al., 2005; Lysak et al., 2006). Because *A. thaliana* had been the only crucifer species with a sequenced genome at that time, chromosome collinearity between all Brassicaceae species had been inferred by mapping homoeologous genetic markers from *A. thaliana* and by chromosomal location of *A. thaliana* BAC clones and contigs. This approach was largely influenced by the density of genetic markers and the selection of chromosome-specific BAC contigs. Without sufficiently accurate information on cross-species collinearity, only approximate or minimal intervals could be defined for some GBs (Schranz et al., 2006) and GB associations. This is exemplified by the PCK/tPCK-specific chromosome rearrangement initially described as V/K/L/Q/X (Mandakova and Lysak, 2008) and V/K/L/V'/Q/X (Dassanayake et al., 2011), and eventually refined as V/K/L/Wa/Q/X in this study. The extensive repatterning of ancestral GBs within the *B. rapa* genome, along with newly available sequence information for Brassicaceae genomes, allowed us to refine the intervals of seven genomic blocks. This was particularly feasible when individual GBs were found to be relocated and forming a non-ancestral GB association with corresponding breakpoints at syntenic positions in two or three *B. rapa* subgenomes. The interval refinement of seven GBs will improve the resolution of interspecies comparisons of genome collinearity across Brassicaceae.

Methods

Syntenic Ortholog Determination

Syntenic orthologs between *B. rapa, A. thaliana, A. lyrata, S. parvula, T. salsuginea, C. papaya*, and *V. vinifera* were identified using the tool SynOrths (Cheng et al., 2012a), with *B. rapa* as the query genome and each of the other species as the subject genome. SynOrths was run with default parameters (m = 20, n = 100, r = 0.2). Under these settings, SynOrths first identifies homoeologous gene pairs between the two genomes using BLASTP, defining gene pairs as those with an *E*-value more significant than 1×10^{-20} or representing the best hits through the entire genome. The 20 closest genes (m = 20) flanking either side of the gene in the query *B. rapa* genome are then compared with the 100 closest genes (n = 100) flanking either side of the gene in the subject genome. If at least 20% (r = 0.2) of the best hits for the 40 genes (20 × 2 for both sides) in the query *B. rapa* genome are found within the 200 genes (100 × 2 for both sides) in the subject genome, the original pair of homoeologous genes are designated as a syntenic ortholog candidate. The increase in the number of flanking genes searched in the subject genomes was necessary to compensate for the fractionation that has reduced the gene content of each *B. rapa* subgenome.

Comparison of Arrangements of GB Associations

We searched the distribution of predefined GBs in the three genomes of the ACK, PCK, and tPCK for adjacent pairs of GBs (GB associations). These GB associations were compared with the extant arrangement of GBs within *B. rapa*. GB associations in ACK, PCK, and tPCK that were also immediately adjacent to each other in the *B. rapa* genome were counted and recorded.

Reconstruction of Subgenomes Based on the tPCK

As previously reported (Wang et al., 2011; Cheng et al., 2012b), the *B. rapa* genome is composed of three subgenomes each of the 24 GBs. We annotated these blocks along the seven chromosomes of the tPCK and reconstructed three tPCK-like subgenomes. At first we identified all continuous chromosome segments corresponding to genomic blocks and belonging to the same chromosome of the tPCK. Then, chromosome segments were concatenated according to the following rules: (i) reconstructed chromosome could not contain redundant copies of the same syntenic segment, and (ii) the concatenated chromosome segments had to have a comparable gene density corresponding to the three *B. rapa* subgenomes (LF, MF1, and MF2).

Ancestral Centromere Detection

Centromere-specific sequences, including CentBr, PCRBr, and TR238, were obtained from previous studies (Koo et al., 2004; Lim et al., 2005; Lim et al., 2007; Koo et al., 2011). These sequences were aligned to the *B. rapa* genome using Nucmer (Kurtz et al., 2004) with parameters "-maxmatch -g 500 -c 16 -l 16 " (i.e., find all matches with minimal exact match size of 16 bp regardless of uniqueness, merge exact matches

distributed less than 500 bp apart into clusters, and report all clusters >16 bp). The nearest genes to the best Nucmer hits were determined and used as labels of genomic coordinates for the detected centromere traces (**Suppl. Table S7**).

Identification of tPCK-specific Block Associations in *B. oleracea*

Because a pseudomolecule-level assembly for the *B. oleracea* genome is not yet available, we used a different method to determine local syntenic relationships between *B. rapa* and *B. oleracea*. Thirty *B. rapa* genes located on either side of the boundaries between GBs within the GB associations O/P/W/R and D/V/K/L/Q/X were selected. These genes were compared using BLAST against *B. oleracea* genes via the web interface provided by BRAD (http://brassicadb.org/brad/blastPage.php) (Cheng et al., 2011). Two blocks were classified as being syntenic in *B. oleracea* when more than 60% of the *B. rapa* block boundary genes from each GB had BLAST hits (identity and coverage) on the same *B. oleracea* scaffold. For example, if 1) 60% of 30 *B. rapa* genes at the boundary of block O (near P) were homoeologous to genes on Scaffold000203 of *B. oleracea,* 2) 30 *B. rapa* genes at the block P boundary (near O) were also homoeologous to *B. oleracea* genes in Scaffold000203, and 3) these genes were adjacent to each other on Scaffold000203, then the O/P association was considered to be present also in the *B. oleracea* genome.

K_s Analysis

Protein sequences from *B. rapa* were aligned with those of *A. thaliana*, *A. lyrata*, *S. parvula*, and *T. salsuginea* using MUSCLE (Edgar, 2004). Protein alignments were translated into coding sequence alignments using an in-house Perl script. K_s values were calculated based on the coding sequence alignments using the method of Nei and Gojobori as implemented in KaKs_calculator (Zhang et al., 2006). K_s values of all syntenic orthologs between *B. rapa* and each of the four species were then plotted as histograms.

Phylogenetic Analysis

For phylogenetic analysis, the same MUSCLE alignments as for K_s analysis were used. The genotypes of synonymous loci in 591 syntenic orthologs among *B. rapa*, *S. parvula*, *T. salsuginea*, *A. thalaliana*, *A. lyrata*, *C. papaya*, and *V. vinifera* were extracted from multiple alignments and concatenated into a single sequence for each species. Phylogenies were constructed, based on the character states of these 63,239 synonymous loci using the neighbor-joining method, and plotted under a linearized tree model as implemented in MEGA (Tamura et al., 2011).

Accession Numbers

Gene and protein sequences for *B. rapa* were obtained from BRAD (V1.2; http://brassicadb.org) (Cheng et al., 2011). Gene and genome datasets for *A. thaliana* were downloaded from TAIR (TAIR9; http://www.arabidopsis.org/index.jsp). Datasets for *A. lyrata* were downloaded from the JGI database (Gene models 6; http://genome.jgi-psf.org/Araly1/Araly1.home.html) (Hu et al., 2011). *S. parvula*

datasets (TpV7) were obtained from Dassanayake et al. (2011). *T. salsuginea* data were downloaded from NCBI (Accession no. AHIU00000000) (Wu et al., 2012).

Supplemental Data

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

Supplemental Figure S1. Purported translocation events reshuffling tPCK-specific associations D/V and M/E in the *B. rapa* genome. **(A)** Translocation between block U_{MF1} and block association D/V_{LF} resulted in block associations Ua/D and V/Ub on *B. rapa* chromosomes A03 and A06, respectively. **(B)** Translocation between two M/E associations from subgenomes LF and MF2 resulted in associations M_{LF}/M_{MF2} and E_{LF}/E_{MF2} on chromosomes A06 and A07, respectively.

Supplemental Figure S2. Chromosome collinearity comparison among A, B, and C genomes, revised from Panjabi et al. (2008). PCK-specific GB associations O/P/W/R and V/K/L/Q/X are colored to distinguish subgenomes: red = LF, green = MF1, blue = MF2. Each copy of block D associated with block V is indicated in yellow. The entire chromosomes A1 and C1, and A2 and C2 have a similar block orders between A and C genomes; the structures of A4 and B4, A5 and B5, and A6 and B6 are shared between genomes A and B; A3, B3, and C3 share similar chromosome structures between A, B, and C genomes.

Supplemental Figure S3. Reconstruction of three tPCK-like ancestral subgenomes in *B. rapa*. tPCK-like genomes each comprising 24 genomic blocks (A–X) were reconstructed for the three *B. rapa* subgenomes (least fractionated [LF], medium fractionated [MF1], and most fractionated [MF2]). Ancestral centromeres still active in *B. rapa*, traces of ancestral centromeres, and theoretically purported but undetected ancestral centromeres are represented by ovals of different colors. Red arrows highlight breakpoints shared by two or three subgenomes. The downward pointing arrow refers to the genomic block inverted relative to the position within the ACK.

Supplemental Figure S4. Two alternative origins of translocation chromosomes AK5/6/8 (PCK) and AK2/5/6/8 (tPCK). Genomic blocks are indicated by capital letters and colored according to their positions on ancestral chromosomes (see Figure 2A). Downward pointing arrows indicate the opposite orientation of GBs compared to their ancestral position within the ACK. Scenario **(A)** involves a double inversion of genomic block V, whereas alternatively **(B)** is based on a paracentric inversion of V/K/L and accompanying inactivation of the AK8 paleocentromere. The position(s) of the inactivated centromere depend(s) on the position of the inversion breakpoint within the AK8 (peri)centromere.

Supplemental Table S1. Summary of genomic blocks identified in *B. rapa*.

Supplemental Table S2. Genomic blocks E, G, and T showing sequence deletions in *B. rapa* (Br) as compared with *A. thaliana* (At).

Supplemental Table S3. Number of genomic blocks with different fragmental status in *B. rapa* subgenomes.

Supplemental Table S4. Syntenic fragments comparison between *B. rapa* and the Ancestral Crucifer Karyotype (ACK).

Supplemental Table S5. Frequency of tPCK-specific block associations in each of the three *B. rapa* subgenomes.

Supplemental Table S6. Number of syntenic genes shared between each *B. rapa*

subgenome and the *A. thaliana* genome (At).

Supplemental Table S7. Centromere-specific sequences detected in the *B. rapa* genome.

Supplemental Table S8. The frequency of tPCK-specific block associations in *B. rapa* and *B. oleracea*.

Supplemental Table S9. List of genomic blocks comprising the nine linkage groups of *Raphanus sativus* (Shirasawa et al., 2011).

Supplemental Table S10. Redefined intervals of seven ancestral genomic blocks.

Acknowledgements

Tables

Table 1 Comparison of genomic block (GB) associations shared between *B. rapa* and three ancestral crucifer genomes (Ancestral Crucifer Karyotype [ACK], Proto-Calepineae Karyotype [PCK], and translocation Proto-Calepineae Karyotype [tPCK]).

ACK (n = 8)

Chr[a]	GB assoc.[b]	#[c]	*B. rapa* Chr
AK1	A/B	2	A06, A08
	B/C	2	A05, A08
AK2	D/E	0	-
AK3	F/G	2	A03, A07
	G/H	2	A03, A07
AK4	I/J	2	A03, A04
AK5	K/L	3	A02, A06, A09
	L/M	0	-
	M/N	3	A01, A03, A09
AK6	O/P	3	A02, A03, A09
	P/Q	0	-
	Q/R	0	-
AK7	S/T	2	A04, A08
	T/U	3	A01, A03, A08
AK8	V/W	0	-
	W/X	0	-

PCK (n = 7)

Chr	GB assoc.	#	*B. rapa* Chr
AK1	A/B	2	A06, A08
	B/C	2	A05, A08
AK2	D/E	0	-
AK3	F/G	2	A03, A07
	G/H	2	A03, A07
AK4	I/J	2	A03, A04
AK7	S/T	2	A04, A08
	T/U	3	A01, A03, A08
AK6/8	O/P	3	A02, A03, A09
	P/W	1	A03
	W/R	3	A02, A03, A10
AK5/6/8	N/M	3	A01, A03, A09
	M/V	0	-
	V/K	3	A02, A06, A09
	K/L	3	A02, A06, A09
	L/Q	3	A02, A06, A09
	Q/X	3	A02, A03, A09

tPCK (n = 7)

Chr	GB assoc.	#	*B. rapa* Chr
AK1	A/B	2	A06, A08
	B/C	2	A05, A08
AK2/5	N/M	3	A01, A03, A09
	M/E	0	-
AK3	F/G	2	A03, A07
	G/H	2	A03, A07
AK4	I/J	2	A03, A04
AK7	S/T	2	A04, A08
	T/U	3	A01, A03, A08
AK6/8	O/P	3	A02, A03, A09
	P/W	1	A03
	W/R	3	A02, A03, A10
AK2/5/6/8	D/V	1	A09
	V/K	3	A02, A06, A09
	K/L	3	A02, A06, A09
	L/Q	3	A02, A06, A09
	Q/X	3	A02, A03, A09

[a] *ancestral chromosome.*
[b] *genomic block association.*
[c] *copy number of GB association in the* B. rapa *genome.*

Table 2 The inheritance of 10 *B. rapa* centromeres from the triplicated ancestral tPCK-like genome ($n = 21$).

B. rapa chromosome	Ancestral chromosome	*B. rapa* subgenome
A01	AK2/5/6/8	MF1
A02	AK6/8 or AK2/5	MF2 or MF1
A03	AK2/5/6/8	LF
A04	AK7	LF
A05	AK1	MF2
A06	AK2/5 or AK2/5	LF or MF2
A07	AK3	LF
A08	AK1 or AK7	MF1 or MF2
A09	AK2/5/6/8	MF2
A10	AK6/8	LF

Table 3 Interval redefinition of seven genomic blocks (GBs) in the Ancestral Crucifer Karyotype (ACK). Newly defined GB intervals based on multi-genome comparison are underlined and shown in bold. Intervals of the 17 remaining blocks follow Schranz et al. (2006). Each GB interval is defined by two *A. thaliana* genes.

GB	AK chromosome	Interval	
A	1	AT1G01560	AT1G19330
B	1	AT1G19850	AT1G36240
C	1	AT1G43600	AT1G56120
D	2	AT1G63770	AT1G56530
E	2	AT1G65040	AT1G80420
F	3	AT3G01040	AT3G25520
G	3	AT2G05170	AT2G07690
H	3	AT2G15670	**AT2G20900**
I	4	**AT2G20920**	AT2G28910
J	4	AT2G31040	AT2G47730
K	5	AT2G01250	AT2G03750
L	5	AT3G25855	AT3G29770
M	5	AT3G43740	AT3G49970
N	5	AT3G50950	AT3G62790
O	6	AT4G00030	AT4G04955
P	6	AT4G12070	AT4G08690
Q	6	**AT5G28885**	**AT5G23010**
R	6	**AT5G23000**	**AT5G01010**
S	7	AT5G41900	AT5G33210
T	7	AT4G12750	AT4G16143
U	7	AT4G16250	AT4G38770
V	8	**AT5G42130**	**AT5G47810**
W	8	**AT5G47820**	**AT5G60800**
X	8	**AT5G60805**	AT5G67385

References

Al-Shehbaz, I.A. (2012). A generic and tribal synopsis of the Brassicaceae (Cruciferae). Taxon **61:** 931-954.

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.

Babula, D., Kaczmarek, M., Barakat, A., Delseny, M., Quiros, C.F., and Sadowski, J. (2003). Chromosomal mapping of *Brassica oleracea* based on ESTs from *Arabidopsis thaliana*: complexity of the comparative map. Mol. Genet. Genomics **268:** 656-665.

Beilstein, M.A., Al-Shehbaz, I.A., and Kellogg, E.A. (2006). Brassicaceae phylogeny and trichome evolution. Am. J. Bot. **93:** 607-619.

Beilstein, M.A., Al-Shehbaz, I.A., Mathews, S., and Kellogg, E.A. (2008). Brassicaceae phylogeny inferred from phytochrome A and ndhF sequence data: tribes and trichomes revisited. Am. J. Bot. **95:** 1307-1327.

Boivin, K., Acarkan, A., Mbulu, R.S., Clarenz, O., and Schmidt, R. (2004). The Arabidopsis genome sequence as a tool for genome analysis in Brassicaceae. A comparison of the *Arabidopsis* and *Capsella rubella* genomes. Plant Physiol. **135:** 735-744.

Burrell, A.M., Taylor, K.G., Williams, R.J., Cantrell, R.T., Menz, M.A., and Pepper, A.E. (2011). A comparative genomic map for *Caulanthus amplexicaulis* and related species (Brassicaceae). Mol. Ecol. **20:** 784-798.

Catcheside, D.G. (1934). The chromosomal relationships in the Swede and turnip groups of *Brassica*. Ann. Bot. **48:** 601-633.

Cheng, F., Wu, J., Fang, L., and Wang, X. (2012a). Syntenic gene analysis between *Brassica rapa* and other Brassicaceae species. Front. Plant Sci. **3:** 198.

Cheng, F., Wu, J., Fang, L., Sun, S., Liu, B., Lin, K., Bonnema, G., and Wang, X. (2012b). Biased gene fractionation and dominant gene expression among the subgenomes of *Brassica rapa*. PLoS One **7:** e36442.

Cheng, F., Liu, S., Wu, J., Fang, L., Sun, S., Liu, B., Li, P., Hua, W., and Wang, X. (2011). BRAD, the genetics and genomics database for *Brassica* plants. BMC Plant Biol. **11:** 136.

Dassanayake, M., Oh, D.H., Haas, J.S., Hernandez, A., Hong, H., Ali, S., Yun, D.J., Bressan, R.A., Zhu, J.K., Bohnert, H.J., and Cheeseman, J.M. (2011). The genome of the extremophile crucifer *Thellungiella parvula*. Nat. Genet. **43:** 913-918.

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

Franzke, A., Lysak, M.A., Al-Shehbaz, I.A., Koch, M.A., and Mummenhoff, K. (2011). Cabbage family affairs: the evolutionary history of Brassicaceae. Trends Plant Sci. **16:** 108-116.

Gao, Z., Fu, S., Dong, Q., Han, F., and Birchler, J.A. (2011). Inactivation of a centromere during the formation of a translocation in maize. Chromosome Res. **19:** 755-761.

Hu, T.T., Pattyn, P., Bakker, E.G., Cao, J., Cheng, J.F., Clark, R.M., Fahlgren, N., Fawcett, J.A., Grimwood, J., Gundlach, H., Haberer, G., Hollister, J.D., Ossowski, S., Ottilar, R.P., Salamov, A.A., Schneeberger, K., Spannagl, M., Wang, X., Yang, L., Nasrallah, M.E., Bergelson, J., Carrington, J.C., Gaut, B.S., Schmutz, J., Mayer, K.F., Van de Peer, Y., Grigoriev, I.V., Nordborg, M., Weigel, D., and Guo, Y.L. (2011). The *Arabidopsis lyrata* genome sequence and the basis of rapid genome size change. Nat. Genet. **43:** 476-481.

Kim, H., Choi, S.R., Bae, J., Hong, C.P., Lee, S.Y., Hossain, M.J., Van Nguyen, D., Jin, M., Park,

B.S., Bang, J.W., Bancroft, I., and Lim, Y.P. (2009). Sequenced BAC anchored reference genetic map that reconciles the ten individual chromosomes of *Brassica rapa*. BMC Genomics **10:** 432.

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.

Koo, D.H., Plaha, P., Lim, Y.P., Hur, Y., and Bang, J.W. (2004). A high-resolution karyotype of *Brassica rapa* ssp. *pekinensis* revealed by pachytene analysis and multicolor fluorescence in situ hybridization. Theor. Appl. Genet. **109:** 1346-1352.

Koo, D.H., Hong, C.P., Batley, J., Chung, Y.S., Edwards, D., Bang, J.W., Hur, Y., and Lim, Y.P. (2011). Rapid divergence of repetitive DNAs in *Brassica* relatives. Genomics **97:** 173-185.

Kuittinen, H., de Haan, A.A., Vogl, C., Oikarinen, S., Leppala, J., Koch, M., Mitchell-Olds, T., Langley, C.H., and Savolainen, O. (2004). Comparing the linkage maps of the close relatives *Arabidopsis lyrata* and *A. thaliana*. Genetics **168:** 1575-1584.

Kurtz, S., Phillippy, A., Delcher, A.L., Smoot, M., Shumway, M., Antonescu, C., and Salzberg, S.L. (2004). Versatile and open software for comparing large genomes. Genome Biol. **5:** R12.

Lagercrantz, U. (1998). 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:** 1217-1228.

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

Li, F., Hasegawa, Y., Saito, M., Shirasawa, S., Fukushima, A., Ito, T., Fujii, H., Kishitani, S., Kitashiba, H., and Nishio, T. (2011). Extensive chromosome homoeology among Brassiceae species were revealed by comparative genetic mapping with high-density EST-based SNP markers in radish (*Raphanus sativus* L.). DNA Res. **18:** 401-411.

Lim, K.B., de Jong, H., Yang, T.J., Park, J.Y., Kwon, S.J., Kim, J.S., Lim, M.H., Kim, J.A., Jin, M., Jin, Y.M., Kim, S.H., Lim, Y.P., Bang, J.W., Kim, H.I., and Park, B.S. (2005). Characterization of rDNAs and tandem repeats in the heterochromatin of *Brassica rapa*. Mol. Cells **19:** 436-444.

Lim, K.B., Yang, T.J., Hwang, Y.J., Kim, J.S., Park, J.Y., Kwon, S.J., Kim, J., Choi, B.S., Lim, M.H., Jin, M., Kim, H.I., de Jong, H., Bancroft, I., Lim, Y., and Park, B.S. (2007). Characterization of the centromere and peri-centromere retrotransposons in *Brassica rapa* and their distribution in related *Brassica* species. Plant J. **49:** 173-183.

Luo, M.C., Deal, K.R., Akhunov, E.D., Akhunova, A.R., Anderson, O.D., Anderson, J.A., Blake, N., Clegg, M.T., Coleman-Derr, D., Conley, E.J., Crossman, C.C., Dubcovsky, J., Gill, B.S., Gu, Y.Q., Hadam, J., Heo, H.Y., Huo, N., Lazo, G., Ma, Y., Matthews, D.E., McGuire, P.E., Morrell, P.L., Qualset, C.O., Renfro, J., Tabanao, D., Talbert, L.E., Tian, C., Toleno, D.M., Warburton, M.L., You, F.M., Zhang, W., and Dvorak, J. (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., Koch, M.A., Pecinka, A., and Schubert, I. (2005). Chromosome triplication found across the tribe Brassiceae. Genome Res. **15:** 516-525.

Lysak, M.A., Cheung, K., Kitschke, M., and Bures, 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., 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.

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

Mandakova, T., Heenan, P.B., and Lysak, M.A. (2010a). Island species radiation and karyotypic stasis in *Pachycladon* allopolyploids. BMC Evol Biol **10:** 367.

Mandakova, T., Joly, S., Krzywinski, M., Mummenhoff, K., and Lysak, M.A. (2010b). Fast diploidization in close mesopolyploid relatives of Arabidopsis. Plant Cell **22:** 2277-2290.

Mun, J.H., Kwon, S.J., Seol, Y.J., Kim, J.A., Jin, M., Kim, J.S., Lim, M.H., Lee, S.I., Hong, J.K., Park, T.H., Lee, S.C., Kim, B.J., Seo, M.S., Baek, S., Lee, M.J., Shin, J.Y., Hahn, J.H., Hwang, Y.J., Lim, K.B., Park, J.Y., Lee, J., Yang, T.J., Yu, H.J., Choi, I.Y., Choi, B.S., Choi, S.R., Ramchiary, N., Lim, Y.P., Fraser, F., Drou, N., Soumpourou, E., Trick, M., Bancroft, I., Sharpe, A.G., Parkin, I.A., Batley, J., Edwards, D., and Park, B.S. (2010). Sequence and structure of *Brassica rapa* chromosome A3. Genome Biol. **11:** R94.

Murat, F., Xu, J.H., Tannier, E., Abrouk, M., Guilhot, N., Pont, C., Messing, J., and Salse, J. (2010). Ancestral grass karyotype reconstruction unravels new mechanisms of genome shuffling as a source of plant evolution. Genome Res. **20:** 1545-1557.

Nagaharu, U. (1935). Genome analysis in *Brassica* with special reference to the experimental formation of *B. napus* and peculiar mode of fertilization. Jap. J. Bot. **7:** 389-452.

Nelson, M.N., Parkin, I.A., and Lydiate, D.J. (2011). The mosaic of ancestral karyotype blocks in the *Sinapis alba* L. genome. Genome **54:** 33-41.

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., 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.

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

Röbbelen, G. (1960). Beitra ge zur Analyse des *Brassica*-Genoms. Chromosoma **11:** 205–228.

Salse, J., Abrouk, M., Bolot, S., Guilhot, N., Courcelle, E., Faraut, T., Waugh, R., Close, T.J., Messing, J., and Feuillet, C. (2009). Reconstruction of monocotelydoneous proto-chromosomes reveals faster evolution in plants than in animals. Proc. Natl. Acad. Sci. USA **106:** 14908-14913.

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.

Schubert, I., and Lysak, M.A. (2011). Interpretation of karyotype evolution should consider chromosome structural constraints. Trends Genet. **27:** 207-216.

Shirasawa, K., Oyama, M., Hirakawa, H., Sato, S., Tabata, S., Fujioka, T., Kimizuka-Takagi, C., Sasamoto, S., Watanabe, A., Kato, M., Kishida, Y., Kohara, M., Takahashi, C., Tsuruoka, H., Wada, T., Sakai, T., and Isobe, S. (2011). An EST-SSR linkage map of *Raphanus sativus* and comparative genomics of the Brassicaceae. DNA Res. **18:** 221-232.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and

maximum parsimony methods. Mol. Biol. Evol. **28:** 2731-2739.

Tang, H., and Lyons, E. (2012). Unleashing the genome of *Brassica rapa*. Front. Plant Sci. **3:** 172.

Tang, H., Woodhouse, M.R., Cheng, F., Schnable, J.C., Pedersen, B.S., Conant, G., Wang, X., Freeling, M., and Pires, J.C. (2012). Altered patterns of fractionation and exon deletions in *Brassica rapa* support a two-step model of paleohexaploidy. Genetics **190:** 1563-1574.

Trick, M., Kwon, S.J., Choi, S.R., Fraser, F., Soumpourou, E., Drou, N., Wang, Z., Lee, S.Y., Yang, T.J., Mun, J.H., Paterson, A.H., Town, C.D., Pires, J.C., Pyo Lim, Y., Park, B.S., and Bancroft, I. (2009). Complexity of genome evolution by segmental rearrangement in *Brassica rapa* revealed by sequence-level analysis. BMC Genomics **10:** 539.

Truco, M.J., Hu, J., Sadowski, J., and Quiros, C.F. (1996). Inter- and infra-genomic homology of the *Brassica* genomes: Implications for their origin and evolution. Theor. Appl. Genet. **93:** 1225–1233.

Wang, X., Wang, H., Wang, J., Sun, R., Wu, J., Liu, S., Bai, Y., Mun, J.H., Bancroft, I., Cheng, F., Huang, S., Li, X., Hua, W., Freeling, M., Pires, J.C., Paterson, A.H., Chalhoub, B., Wang, B., Hayward, A., Sharpe, A.G., Park, B.S., Weisshaar, B., Liu, B., Li, B., Tong, C., Song, C., Duran, C., Peng, C., Geng, C., Koh, C., Lin, C., Edwards, D., Mu, D., Shen, D., Soumpourou, E., Li, F., Fraser, F., Conant, G., Lassalle, G., King, G.J., Bonnema, G., Tang, H., Belcram, H., Zhou, H., Hirakawa, H., Abe, H., Guo, H., Jin, H., Parkin, I.A., Batley, J., Kim, J.S., Just, J., Li, J., Xu, J., Deng, J., Kim, J.A., Yu, J., Meng, J., Min, J., Poulain, J., Hatakeyama, K., Wu, K., Wang, L., Fang, L., Trick, M., Links, M.G., Zhao, M., Jin, M., Ramchiary, N., Drou, N., Berkman, P.J., Cai, Q., Huang, Q., Li, R., Tabata, S., Cheng, S., Zhang, S., Sato, S., Sun, S., Kwon, S.J., Choi, S.R., Lee, T.H., Fan, W., Zhao, X., Tan, X., Xu, X., Wang, Y., Qiu, Y., Yin, Y., Li, Y., Du, Y., Liao, Y., Lim, Y., Narusaka, Y., Wang, Z., Li, Z., Xiong, Z., and Zhang, Z. (2011). The genome of the mesopolyploid crop species Brassica rapa. Nat. Genet. **43:** 1035-1039.

Warwick, S.I., and Black, L.D. (1991). Molecular systematics of *Brassica* and allied genera (subtribe Brassicinae, Brassiceae) - chloroplast genome and cytodeme congruence. Theor. Appl. Genet. **82:** 81-92.

Warwick, S.I., and Al-Shehbaz, I.A. (2006). Brassicaceae: chromosome number index and database on CD-Rom. Plant Syst. Evol. **259:** 237-248.

Wu, H.J., Zhang, Z., Wang, J.Y., Oh, D.H., Dassanayake, M., Liu, B., Huang, Q., Sun, H.X., Xia, R., Wu, Y., Wang, Y.N., Yang, Z., Liu, Y., Zhang, W., Zhang, H., Chu, J., Yan, C., Fang, S., Zhang, J., Wang, Y., Zhang, F., Wang, G., Lee, S.Y., Cheeseman, J.M., Yang, B., Li, B., Min, J., Yang, L., Wang, J., Chu, C., Chen, S.Y., Bohnert, H.J., Zhu, J.K., Wang, X.J., and Xie, Q. (2012). Insights into salt tolerance from the genome of *Thellungiella salsuginea*. Proc. Natl. Acad. Sci. USA **109:** 12219-12224.

Zhang, Z., Li, J., Zhao, X.Q., Wang, J., Wong, G.K., and Yu, J. (2006). KaKs_Calculator: calculating Ka and Ks through model selection and model averaging. Genomics Proteomics Bioinformatics **4:** 259-263.

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.

VIII. **Mandáková T.**, Heenan P.B., Lysak M.A. 2010. Island species radiation and karyotypic stasis in *Pachycladon* allopolyploids. BMC Evolutionary Biology 10: 367.

BMC
Evolutionary Biology

Island species radiation and karyotypic stasis in *Pachycladon* allopolyploids

Terezie Mandáková[1], Peter B Heenan[2], Martin A Lysak[1*]

Abstract

Background: *Pachycladon* (Brassicaceae, tribe Camelineae) is a monophyletic genus of ten morphologically and ecogeographically differentiated, and presumably allopolyploid species occurring in the South Island of New Zealand and in Tasmania. All *Pachycladon* species possess ten chromosome pairs (2n = 20). The feasibility of comparative chromosome painting (CCP) in crucifer species allows the origin and genome evolution in this genus to be elucidated. We focus on the origin and genome evolution of *Pachycladon* as well as on its genomic relationship to other crucifer species, particularly to the allopolyploid Australian Camelineae taxa. As species radiation on islands is usually characterized by chromosomal stasis, i.e. uniformity of chromosome numbers/ploidy levels, the role of major karyotypic reshuffling during the island adaptive and species radiation in *Pachycladon* is investigated through whole-genome CCP analysis.

Results: The four analyzed *Pachycladon* species possess an identical karyotype structure. The consensual ancestral karyotype is most likely common to all *Pachycladon* species and corroborates the monophyletic origin of the genus evidenced by previous phylogenetic analyses. The ancestral *Pachycladon* karyotype (n = 10) originated through an allopolyploidization event between two genomes structurally resembling the Ancestral Crucifer Karyotype (ACK, n = 8). The primary allopolyploid (apparently with n = 16) has undergone genome reshuffling by descending dysploidy toward n = 10. Chromosome "fusions" were mediated by inversions, translocations and centromere inactivation/loss. *Pachycladon* chromosome 3 (PC3) resulted from insertional fusion, described in grasses. The allopolyploid ancestor originated in Australia, from the same or closely related ACK-like parental species as the Australian Camelineae allopolyploids. However, the two whole-genome duplication (WGD) events were independent, with the *Pachycladon* WGD being significantly younger. The long-distance dispersal of the diploidized *Pachycladon* ancestor to New Zealand was followed by the Pleistocene species radiation in alpine habitats and characterized by karyotypic stasis.

Conclusions: Karyotypic stasis in *Pachycladon* suggests that the insular species radiation in this genus proceeded through homoploid divergence rather than through species-specific gross chromosomal repatterning. The ancestral *Pachycladon* genome originated in Australia through an allopolyploidization event involving two closely related parental genomes, and spread to New Zealand by a long-distance dispersal. We argue that the chromosome number decrease mediated by inter-genomic reshuffling (diploidization) could provide the *Pachycladon* allopolyploid founder with an adaptive advantage to colonize montane/alpine habitats. The ancestral *Pachycladon* karyotype remained stable during the Pleistocene adaptive radiation into ten different species.

Background

Multiple rounds of WGD events (both allopolyploidy and autopolyploidy) cyclically increase the genetic diversity in vascular plants, and subsequently this is steadily eroded by genomic fractionation toward diploid-like genomes. Lineage-specific WGD events followed by genome repatterning and descending dysploidy toward diploid-like genomes have been revealed in several angiosperm groups and are probably best characterized in grasses [1-5] and Brassicales [6-11]. Despite still scant knowledge of the number and genealogical context of ancient WGD episodes [12], c. 15% of speciation

* Correspondence: lysak@sci.muni.cz
[1]Department of Functional Genomics and Proteomics, Masaryk University, and CEITEC, Masaryk University, Brno, Czech Republic
Full list of author information is available at the end of the article

events among extant angiosperms are associated with polyploidy [13].

Polyploidy has also played a significant role in colonization and species radiation on islands. Multiple examples of long-distance dispersals of diploid progenitors or polyploid founders followed by adaptive radiation are documented on well-studied archipelagos (Canary Islands, New Zealand, Hawaiian Islands) [14-17]. Remarkably, species radiation on islands is usually characterized by chromosomal stasis, i.e. uniformity of chromosome numbers/ploidy levels [15-17]. This means that adaptive or species radiations proceed through homoploid divergence, rather than by changing the number of linkage groups by dysploidy and/or polyploidy. The reasons for insular chromosomal stasis are most likely complex and lineage-specific, albeit the young age of radiating polyploid lines and the adaptive advantage of successful polyploid founders and their descendants are suspected as crucial factors. Although chromosomal stasis does not necessarily imply karyotypic stasis [17], only a handful of reports deal with the evolution of whole chromosome complements in island endemics. With the exception of the Hawaiian silverswords (Asteraceae), analyzed through inter-species crossing experiments and meiotic chromosome pairing configurations [18], none of the homoploid species complexes on islands has been analyzed for whole-genome collinearity.

The genus *Pachycladon* (Brassicaceae) comprises nine morphologically and ecologically diverse species in mainly alpine habitats of the South Island of New Zealand, and a single species occurs in alpine habitats in Tasmania [19,20]. The morphology of *Pachycladon* is so diverse that prior to the genus being recircumscribed by [21], species were also placed in *Cheesemania* and *Ischnocarpus*. *Pachycladon* is monophyletic [21-23], characterized by little genetic variation amongst species at a variety of genetic loci [24], and the species are interfertile [25,26]. Furthermore, six *Pachycladon* species analyzed karyologically all have the same chromosome number of 2n = 20 [27,28] and comparable genome sizes (430 to 550 Mb [26,28]). *Pachycladon* is related to *Arabidopsis* [21,23], with both these genera belonging to the polyphyletic tribe Camelineae [29,30]. The close relationship between these genera is underlined by the generation of a sexually derived intergeneric hybrid between *A. thaliana* and *P. cheesemanii* [31].

Based on chromosome counts and preliminary cytogenetic data, *Pachycladon* species were thought to have a polyploid origin (M. Lysak and P. Heenan, unpublished results). Indeed, an allopolyploid origin of the genus during the Pleistocene between ~0.8 and 1.6 mya (million years ago) has been confirmed through identification of two paralogous copies of five single copy nuclear genes [23]. Phylogenetic data of Joly et al. [23] suggested

that one of the purported parents comes from the polyphyletic Camelineae or the genus *Boechera* (i.e. from crucifer lineage I [29]), whereas the *Brassica* copy from the crucifer lineage II. Our recent comparative phylogenomic study of some allopolyploid Australian Camelineae species (*Ballantinia antipoda*, 2n = 12; *Stenopetalum nutans*, 2n = 8 and *S. lineare*, 2n = 10) revealed their close phylogenetic affinity to *Pachycladon* and other Camelineae taxa [10]. The ~6 to 9 million old allopolyploid event in the ancestry of Australian genera was found to be obscured by extensive chromosome repaterrning leading to the extant diploid-like karyotypes (n = 4-6). Such concealed WGD episodes still detectable by comparative genetic and cytogenetic analysis were classified as mesopolyploid [10]. Although both recent studies [10,23] argued for an allopolyploid origin of the New Zealand and Australian Camelineae genera, the unknown genome structure of *Pachycladon* species did not yet allow to elucidate the relationship between the two polyploid Camelineae groups.

In the present paper comparative chromosome painting (CCP) has been applied to four *Pachycladon* species (*Pachycladon cheesemanii*, *P. enysii*, *P. exile*, and *P. novae-zelandiae*) that represent the morphological, ecological and phylogenetic diversity of the genus (Figure 1 and [21]), and for which genetic maps are not available. *Pachycladon enysii* is a monocarpic, lanceolate and serrate-leaved, stout terminal inflorescence species of high altitude (975-2492 m) alpine greywacke rock; *P. novae-zelandiae* is a polycarpic, lobed-leaved, lateral inflorescence species of mid-altitude (1080-2031 m)
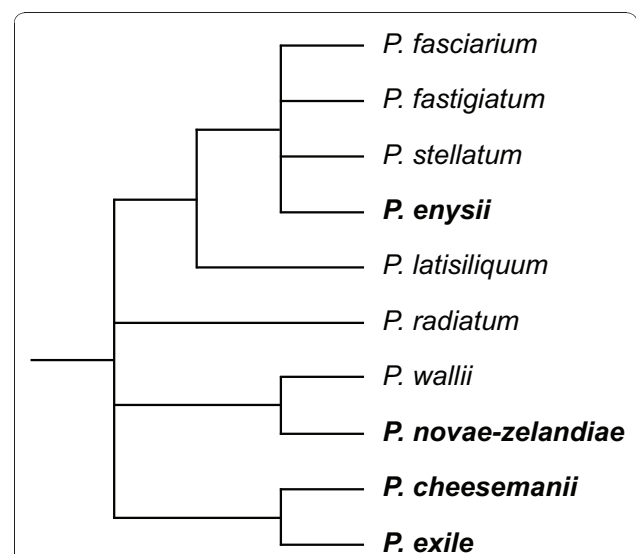
Figure 1 Phylogenetic relationships in *Pachycladon*. Strict consensus tree of the six most parsimonious trees based on the internal transcribed spacer (ITS) region of 18S-25S ribosomal DNA. Species analyzed herein are in bold. Adapted from [[21], Figure 2].

alpine schist rock; and *P. chessemanii* and *P. exile* are polycarpic, heterophyllous, slender terminal inflorescence and generalist species of high fertility rock such as limestone, schist, and volcanics and occur from near sea-level to the alpine zone (10-1600 m altitude) [19]. We used CCP to study the extent of chromosome collinearity between the ten chromosomes of *Pachycladon* species and the eight chromosomes of the theoretical Ancestral Crucifer Karyotype (ACK [32,33]). Combining comparative cytogenetic data with already published accounts on phylogenetics, biogeography, and ecology of the genus we addressed (i) genome structure and evolution of *Pachycladon* species, (ii) the genome relationship to other crucifer species, particularly to the endemic Australian Camelineae taxa, and (iii) the role of major karyotypic reshuffling in the species radiation in the island setting.

Results

Comparative structure of *Pachycladon* karyotypes

The karyotype structure of four *Pachycladon* species (*P. cheesemanii*, *P. enysii*, *P. exile*, and *P. novae-zelandiae*) has been reconstructed by comparative chromosome painting (CCP) (Figure 2). Considering the close phylogenetic relationship between *Pachycladon* and *Arabidopsis* [21,23], we assumed that both genera descended from the Ancestral Crucifer Karyotype (ACK) with eight ancestral chromosomes AK1 to AK8 [32,33]. Hence *A. thaliana* BAC clones and contigs corresponding to the 24 conserved genomic blocks (GBs) of the ACK were used as painting probes to identify collinear chromosome regions in *Pachycladon* species. The four reconstructed karyotypes showed overall similarity, comprising seven (sub)metacentric (PC1, PC3, PC4, PC6 - PC8, and PC10) and three acrocentric (PC2, PC5, and PC9) chromosomes with the identical arrangement of ancestral GBs (Figure 3). The structural uniformity of all reconstructed karyotypes suggests that this structure is the ancestral *Pachycladon* karyotype.

All 24 GBs were found to be duplicated within the analyzed pachytene complements displaying regular meiotic pairing (Figure 2 and 3). The *Pachycladon* karyotype comprises one AK chromosome (PC7), seven AK-like chromosomes discernible within the composite *Pachycladon* chromosomes (four chromosomes modified by inversions), and 14 AK-like chromosome arms (Figure 3 and Table 1). Thus, in total forty-three ancestral GBs (90%) remained intact and only five blocks were split within one chromosome arm (block L on PC5), between two arms of the same chromosome (W on PC6), or between two different chromosomes (D to PC1 and PC2, J to PC9 and PC10, R to PC4 and PC5). Except chromosome PC7 resembling chromosome AK7, all *Pachycladon* chromosomes originated through "fusion" of two or three AK chromosomes (Figure 3).

Evolution of the ten *Pachycladon* chromosomes

We have reconstructed the origin of the nine "fusion" chromosomes of the ancestral *Pachycladon* karyotype using the minimal number of rearrangements and assuming that the ten PC chromosomes originated from the duplicated ACK (i.e. from 16 AK chromosomes).

PC1 and PC2 chromosomes

(Figure 4A). PC1 originated via a reciprocal translocation between chromosomes AK1 and AK2 with breakpoints in the (peri)centromeric region of AK1 (close to block B) and in the block D of AK2. The second translocation product harbouring the AK1 centromere has been involved in a subsequent reciprocal end-to-end translocation with AK5, resulting in chromosome PC2. As the four GBs (K-N) of AK5 have the ancestral position within PC2 chromosome, we infer an inactivation and/or loss of the AK5 centromere.
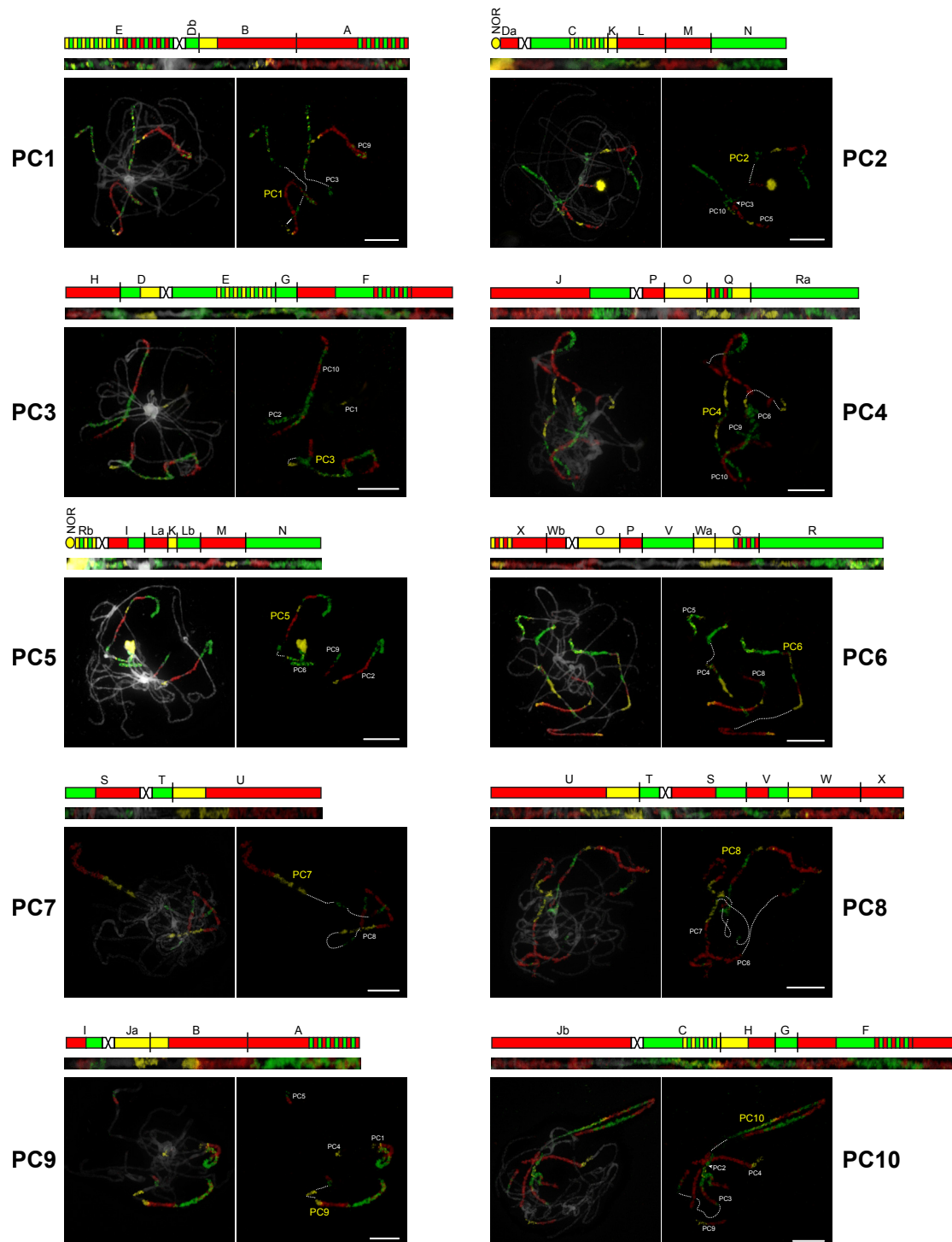
PC3 chromosome

(Figure 4B). The origin of PC3 can be reconstructed as a paracentric inversion of the block D on AK2 followed by nested "fusion" of this chromosome into the (peri)centromere of AK3. The nested "fusion" required three or four breakpoints: two at the chromosome termini of AK2 and one or two at the centromere of AK3. One breakpoint would presumably disrupt the AK3 centromere, whereas two breaks at pericentromeric regions of the opposite arms would yield a dispensable minichromosome as a second translocation product.

PC4 and PC5 chromosomes

(Figure 4C). PC4 and PC5 were generated through the reshuffling of ancestral chromosomes AK4 and AK6, and AK4, AK5 and AK6, respectively. A pericentric inversion (GBs O and P) transforming AK6 into a telocentric chromosome was followed by a reciprocal translocation between this chromosome and AK4. This translocation joined the long arm of AK4 (block J) with the AK6 telocentric (= PC4). The AK4-derived telocentric chromosome comprising only the centromere and block I has undergone a reciprocal end-to-end translocation with AK5. As the GB collinearity around the AK5 centromere between blocks L and M remained conserved, we inferred an inactivation and/or loss of this centromere on PC5. A small reciprocal translocation between the bottom arm of PC4 (block R) and the upper arm of PC5 occurred after the major reshuffling steps. A paracentric inversion between GBs K and L on PC5 could have occured before the origin of both PC chromosomes or it is a later event.
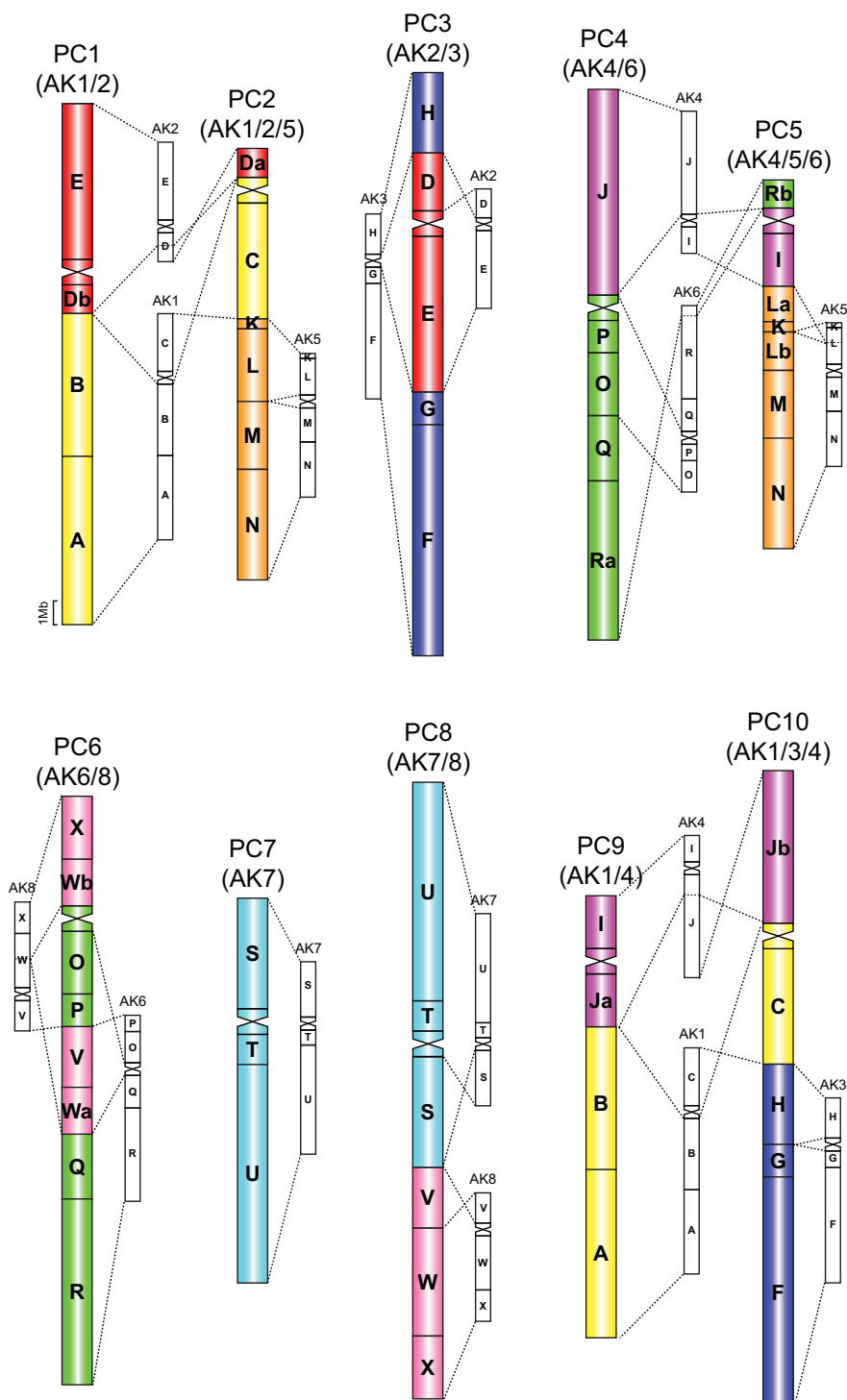
PC6 chromosome

(Figure 4D). This chromosome most likely originated via a reciprocal end-to-end translocation between AK6 and AK8 and was probably followed by a concurrent or subsequent inactivation and/or loss of the AK8 centromere, reflected by the ancestral position of blocks V and Wa

Figure 2 Comparative chromosome painting (CCP) in *Pachycladon cheesemanii*. Labelling scheme, *in situ* localization within a pachytene complement and straightened pachytene bivalent for each of the ten chromosomes (PC1-PC10) are shown. Chromosomes were identified by CCP with *Arabidopsis* BAC clones and contigs labelled by biotin-dUTP (red), digoxigenin-dUTP (green), and Cy3-dUTP (yellow). Due to the duplicated nature of *Pachycladon* genomes, each painting probe labels two homeologous chromosome regions on different chromosomes (white and yellow acronyms). Chromosomes counterstained by DAPI. NOR: nucleolar organizing region. Scale, 10 μm.

Figure 3 Comparative cytogenetic map of the ancestral *Pachycladon* karyotype based on CCP data. Collinearity relationship of the ten *Pachycladon* chromosomes (PC1 - PC10) to the duplicated Ancestral Crucifer Karyotype (ACK) comprising two sets of eight ancestral chromosomes (AK1- AK8). Dashed lines connect collinear regions shared by the two genomes. Duplicated 24 conserved genomic blocks (A-X) of the ACK are colored according to their position on chromosomes AK1 to AK8 [33]. Blocks split into two parts are labeled as „a" and „b". Centromeres of *Pachycladon* chromosomes are depicted as sandglass-like symbols colored according to their presumed origin from AK chromosomes.

Table 1 Comparison of ancestral genomic features between *Pachycladon* and Australian Camelineae species

| | PK | BA | SL | SN |
	n = 10	n = 6	n = 5	n = 4
entirely conserved AK chromosomes	**3**	4	4	3
AK chromosomes modified by inversions	**4**	2	1	4
AK chromosome arms	**14**	12	15	11
GBs not forming any AK-like structure	**2**	4	6	7
split GBs	**5**	6	6	9
lost GBs	**0**	4	2	0
non-ancestral associations of GBs	**18**	29	30	38

This table shows the extent of conservation of the eight ancestral chromosomes (AK1-8) and 24 genomic blocks (GBs) of the duplicated Ancestral Crucifer Karyotype (ACK) in *Pachycladon* and Australian Camelineae species [10]. PK- *Pachycladon* karyotype (n = 10), BA - *Ballantinia antipoda* (n = 6), SL - *Stenopetalum lineare* (n = 5), and SN - *S. nutans* (n = 4).

on PC6. This event was followed by a pericentric inversion with breakpoints in the (peri)centromeric region (close to block Q) and within block W.

PC8 chromosome

(Figure 4E). PC8 originated via a reciprocal translocation between AK7 and AK8, yielding the fusion PC8 chromosome and a meiotically unstable minichromosome containing the centromere of AK8. The translocation was preceded by a paracentric inversion on AK7 (block S) and pericentric inversion on AK8 (block V).

PC9 and PC10 chromosomes

(Figure 4F). Chromosome PC9 originated through a reciprocal translocation between AK1 and AK4 with breakpoints in the AK1 pericentromere (close to block B) and the proximal part of the bottom arm of AK4 (block J). The second translocation product (GBs C and Jb, and the AK1 centromere) participated in a reciprocal end-to-end translocation with AK3 which resulted in the origin of PC10 and small acentric fragment. Ancestral arrangement of AK3-derived GBs suggests that the AK3 centromere has been lost or inactivated.

The reconstructed chromosome origins are congruent with the reduction of 16 ancestral chromosomes (centromeres) to only 10 in *Pachycladon*. Centromeres of both homeologues of AK1, AK2, AK4, AK6 and AK7 remained functional, whereas six centromeres were lost (Figure 3). The centromere of one AK3 homeologue was eliminated due to the nested chromosome fusion, one AK8 centromere was eliminated via symmetric translocation, and both AK5 centromeres and centromeres of second homeologues of AK3 and AK8 were inactivated/lost (Figure 4).

Out of the 33 breakpoints inferred for the origin of ten *Pachycladon* chromosomes, 12 (36%) map to pericentromeric regions, 16 (49%) to telomeric regions, whereas only five (15%) occurred within GBs (Figure 4A to 4F).

### Chromosome landmarks (heterochromatin, telomeres and rDNA)

We have analyzed mitotic and pachytene chromosome complements of the four *Pachycladon* species for the distribution of heterochromatin domains, localization of ribosomal RNA genes (rDNA) and the *Arabidopsis*-type telomere repeat (Figure 5). Except for prominent heterochromatin of pericentromeres and terminal nucleolus organizing regions (NORs) (Figure 2), a single heterochromatic knob occurs in *P. enysii* and two knobs were found in *P. exile*. Whereas two of the three knobs reside within genomic blocks (B on PC1 in *P. enysii*, U on PC7 in *P. exile*), the knob on the bottom arm of PC10 in *P. exile* is localized between blocks G and H, i.e. at the site of presumably inactivated centromere of AK3 (Figure 4F). No heterochromatic domains were observed at the sites of other presumably inactivated and/or lost paleocentromeres. The telomere (TTTAGGG)n repeat hybridized only to chromosome ends and no interstitial telomeric signals were observed (data not shown). Whereas the four species have a single 5S rDNA locus at the same position, the number of terminal 45S rDNA loci varies. *P. novae-zelandiae* has one, *P. cheesemanii* and *P. exile* possess two, and *P. enysii* has three 45S rDNA loci, with 45S locus on the upper arm of PC2 being common to all species (Figure 5). Thus, the cross-species karyotypic stasis does not apply to the number of terminal 45S rDNA loci.
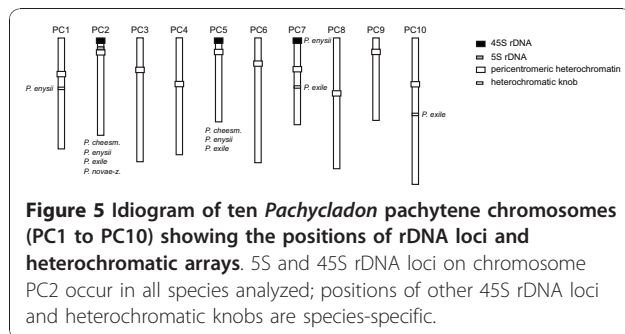
Discussion

We have used comparative chromosome painting to reconstruct karyotype structure and evolution in the genus *Pachycladon*. Interestingly, our analysis showed that the four analyzed species representing the phylogenetic, ecological and morphological diversity of the genus possess an identical karyotype, which is also most likely to be the ancestral karyotype of the genus *Pachycladon*.

Chromosomal and karyotypic stasis in *Pachycladon*

The present study of four *Pachycladon* species is the first whole-genome analysis of an island species radiation. *Pachycladon* species have uniformly ten chromosomes [27,28] and this infrageneric chromosomal stasis has been now extended for karyotypic stasis. Overall similar genome structures supported the monophyletic origin of the genus [21-23] and allowed inference of the ancestral *Pachycladon* karyotype whose structure remained conserved in the extant species. Karyotypic stasis revealed in *Pachycladon* clearly indicates that the Pleistocene species radiation on the South Island of New Zealand [19] was not associated with major chromosome rearrangements. The four karyotypes differ only by the number of heterochromatic knobs and
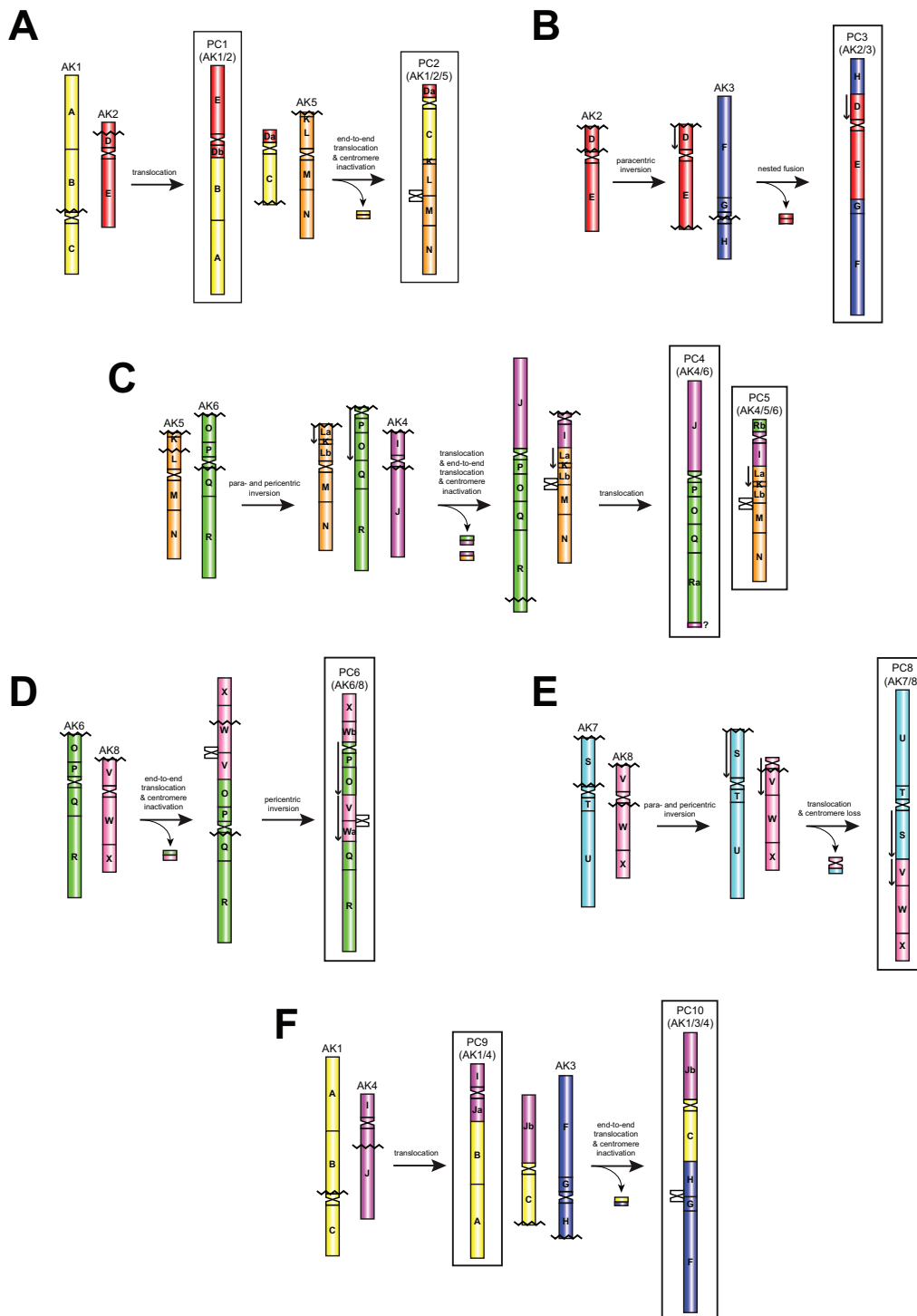
Figure 4 Tentative scenarios of the origin of nine *Pachycladon* chromosomes (PC1- PC6, PC8-PC10) from the duplicated Ancestral Crucifer Karyotype (ACK). Duplicated 24 conserved genomic blocks (A-X) of the ACK are colored according to their position on chromosomes AK1 to AK8 [33]. Blocks split into two parts are labeled as „a" and „b". Centromeres of *Pachycladon* chromosomes are depicted as sandglass-like structures colored according to their presumed origin from AK chromosomes. Inactivated and/or lost ancestral centromeres are shown outside the modern *Pachycladon* chromosomes. Downward-pointing arrows indicate the opposite orientation of genomic blocks as compared to their position within the ACK [33]. Jagged lines mark purported breakpoints of inferred chromosome rearrangements.

Figure 5 Idiogram of ten *Pachycladon* pachytene chromosomes (PC1 to PC10) showing the positions of rDNA loci and heterochromatic arrays. 5S and 45S rDNA loci on chromosome PC2 occur in all species analyzed; positions of other 45S rDNA loci and heterochromatic knobs are species-specific.

NORs, without an apparent link to infrageneric phylogenetic relationships (Figure 1). Hence, the speciation proceeded through homoploid divergence from the ancestral allopolyploid genome.

Perhaps with the exception of meiotic studies in the Hawaiian silversword alliance [18] there is virtually no data on karyotype evolution during island angiosperm speciation. Hence, only the variation in chromosome number/ploidy level can be discussed more extensively. Several surveys of angiosperm chromosome numbers showed the trend of chromosomal stasis during species radiation on islands (see reviews by [15-17]). This tendency might appear paradoxical considering geographical isolation and a wealth of diverse insular environments potentially promoting the origin of novel chromosomal races and karyotypes. However, genomes diverging on islands are under multiple constraints determining chromosomal stasis or chromosomal variation. As self-evident factors influencing the insular species radiation and genomic stability are the age of islands and their distance from the mainland, the number of colonization events, the incidence of polyploidy and phylogenetic constraints. Colonizations followed by adaptive radiation on (volcanic) islands represent often relatively young evolutionary events and therefore many island endemics represent monophyletic lineages comprising closely related species with uniform chromosome numbers. Furthermore, it was concluded that chromosomal stasis vs. lability is under a strong phylogenetic constraint as some lineages (e.g. Asteraceae, *Sideritis*) seem to be more prone to genome reshuffling than others [15,16,34].

Generally the low incidence of polyploidy has been claimed for island floras [16]. These estimates collated prior to the era of indepth whole-genome analyses revealing multiple whole-genome duplications of a different age (e.g., [7,12,10]) had to be, by definition, too conservative. Recent studies suggest that colonization of islands has been frequently associated with hybridization and allopolyploidy (see [35,36] for examples). Allopolyploid ancestors originated either on continents and spread to islands (e.g. the allopolyploid ancestor of

Hawaiian mints [37]) or diploid ancestors hybridize *in situ* after long-distance dispersal (e.g. the New Zealand and Australian *Lepidium* species [38]). The allopolyploidy-driven speciation on islands is frequently associated with chromosomal stasis as shown for the Hawaiian flora with the high incidence of polyploidy (> 80% [15]).

Polyploidy is also a pronounced feature of the New Zealand flora, with 72% of the species being polyploid in families with 25 or more species [39]. Chromosomal features of New Zealand plants indicative of polyploidy are the high number of species with even haploid numbers and/or haploid numbers n = > 10-14 [40]. Many of the polyploid genera that are like *Pachycladon* exhibiting chromosomal stasis are species-rich and generally considered to be recent species radiations, often into mainly alpine-montane habitats. They include, for example, *Aciphylla* (42 species, 2n = 22), *Brachyglottis* (30 species, 2n = 60), *Chionochloa* (22 species, 2n = 42), *Gentianella* (40 species, 2n = 36), *Epilobium* (38 species, 2n = 36), and *Ourisia* (20 species, 2n = 48) (data from [27]).

Chromosomal stasis is also observed in the few crucifer genera that have species radiations on islands. All seven *Parolinia* species endemic to the Canary Islands probably have 2n = 22 (4 species counted [41]), and seven shrubby species of *Descurainia* endemic to the Canary Islands share 2n = 14 [42]. Similarly, of nine *Diplotaxis* species in the Cape Verde Islands, five have 2n = 26 [43]. Unfortunately, insufficient chromosomal data are available for c. 40 *Cardamine* species endemic to New Zealand (P. Heenan, unpublished data) as well as for most crucifer genera endemic to Australia [41,44].

Pachycladon karyotype is derived from the duplicated Ancestral Crucifer Karyotype

Our data suggest that the ancestral *Pachycladon* karyotype (n = 10) was derived from the duplicated Ancestral Crucifer Karyotype (n = 8) through allopolyploidy. The ACK was expected to be inferred as an ancestral genome of *Pachycladon*, as all of the Camelineae genomes analyzed thus far have descended from the ACK (for instance, *Arabidopsis*, *Capsella*, *Turritis*, and *Neslia* [32,45,46], including the analyzed Australian Camelineae species [10]). Furthermore, the karyotypes of *Crucihimalaya* and *Transberingia*, two genera often found as being the closest relatives of *Pachycladon* [21,23], resemble the ACK structure [10]. Similarly, the ACK was proposed as an ancestral karyotype for tribes Boechereae and Cardamineae [[47], Mandáková and Lysak, unpublished data]. It is likely, therefore, that the ancestral *Pachycladon* genome has been derived from the hybridization between two ACK-like genomes. The primary allopolyploid had either the structure of duplicated ACK with n = 16 or the participating genome(s) were reduced (n = 8

→ n = 7-5?) prior to the hybridization event and the allopolyploid had less than 16 chromosome pairs. The fact that paralogous genomic blocks do not lay on the same chromosome suggests that the modern *Pachycladon* chromosomes were reshuffled prior to the hybridization event, rather than due to homeologous recombination between two ACK-like genomes within the allopolyploid ancestor.

The ten composite *Pachycladon* chromosomes originated through inversions, reciprocal translocations and centromere inactivation/loss events within the duplicated ACK complement (Figure 4). Chromosome "fusions" were mediated by reciprocal translocations with or without preceding para- and pericentric inversions. These translocations yielded a "fusion" chromosome and (a)centric fragment as the second translocation product. Small acentric fragments and the minichromosome harbouring one AK8 centromere were meiotically unstable and eliminated. Whereas Robertsonian-like translocations eliminating one AK centromere together with the minichromosome is a common mechanism of the karyotype evolution in Brassicaceae [32,48,49], asymmetric translocation events yielding miniature acentric fragments and dicentric chromosomes with one AK centromere apparently inactivated or removed by recombination were proposed for the origin of composite chromosomes in the Australian Camelineae species [10]. Centromere inactivation and/or loss has been inferred on bottom (long) arms of four *Pachycladon* chromosomes (PC2, PC5, PC6, and PC10) based on the absence of ancestral centromeres and conserved organization of adjacent genomic blocks (Figure 3 and 4). The incidence of centromere inactivation in Australian and New Zealand Camelineae species might be tentatively related to the common ancestry of both lineages and/or to the duplicated character of the allopolyploid ancestral genomes. Centromere inactivation of AK4 can also be suggested for the origin of chromosome At2 in *A. thaliana* [32], and centromere inactivation of AK5 for the origin of Bst5 in *Boechera stricta* [47] and chromosome AK4/5 in *Neslia paniculata* [32]. Nevertheless, an alternative mechanism of centromere removal through subsequent paracentric and pericentric inversions followed by a symmetric translocation (Figure 2C in [32]) is also plausible, though more breakpoints have to be considered. A dicentric chromosome could also be stabilized by intrachromosomal translocation, with breakpoints in pericentromeric region of one of the ancestral centromeres, followed by a loss of the resulting centric fragment.

Chromosome PC3 originated probably through a nested "fusion" of chromosome AK2 between chromosome arms of AK3. As both AK chromosomes within PC3 possess the ancestral structure of genomic blocks (except inverted block D) translocation events with breakpoints at chromosome termini of AK2 and centromere of AK3 seems to be the parsimonious scenario. In grasses (Poaceae), insertional chromosome "fusion" has been proposed as a general mechanism of descending dysploidy [4,50], whereas in crucifers it can be suggested only for the origin of chromosome AK2/5 in *Hornungia alpina* [32]. Thus, *Pachycladon* chromosome PC3 is most likely the first instance of reconstructed insertional dysploidy in Brassicaceae. An alternative mechanism of the PC3 origin via end-to-end reciprocal translocation coupled with the elimination of the AK3 centromere requires two more breakpoints.

Common origin of *Pachycladon* and Australian Camelineae species?

Based on the phylogenetic analysis of Australian Camelineae taxa and *Pachycladon* species, [10] concluded that both groups might originate from a very similar allopolyploid ancestor. Although the authors could not reject a single origin of both lineages, they considered two successive allopolyploidization events as more likely, i.e. mesopolyploid Australian Camelineae species originated and radiated in continental arid habitats before the mesopolyploid ancestor of *Pachycladon*. The present data corroborate this conclusion. Specifically, the two species groups do not share any cytogenetic signature, i.e. a taxon/lineage-specific chromosome rearrangement, such as the rearranged AK8 homeologue shared by five Australian species analyzed [10]. In the Australian species, any two paralogous GBs differ by the length and fluorescence intensity as revealed by CCP [10]. This difference was either present already in the hybridizing progenitors or was caused by preferential fractionation of paralogous regions belonging to only one subgenome [51]. In *Pachycladon*, two paralogous copies of all GBs cannot be distinguished upon CCP analysis. Furthermore, higher chromosome number in *Pachycladon* species (n = 10) than in the Australian species (n = 4-7) implies a more recent origin and less extensive diploidization in *Pachycladon*. Indeed, the significantly lower number of non-ancestral junctions of genomic blocks in *Pachycladon* compared to *Ballantinia antipoda* and the two *Stenopetalum* species (Table 1 and [10]) underlines the less extensive genome reshuffling in *Pachycladon*. Also the number of split GBs in *Pachycladon* (10%) is lower than in the Australian species (13% to 19%; [10]). Interestingly, both groups do not differ substantially by the number of preserved AK chromosomes, chromosome arms and GBs (Table 1). This comparison suggests that the most recent steps of chromosome number reduction in the Australian Camelineae species have been mediated by tandem end-to-end translocations followed by centromere inactivation/loss, not disrupting the structure of AK-like chromosomes and chromosome arms.

Altogether, the differences in genome structure between the mesopolyploid Australian and New Zealand lineage indicate two successive WGD events involving the same pool of parental species. The existence of the progenitor species in Australia for a long period of time is a credible assumption considering the remarkable stasis of the ACK and AK chromosomes across crucifer lineages I and II [32,33,49]. Further research is needed to elucidate if the ancient ACK-like karyotype could be found in some not yet analyzed Australian crucifer species. Recurrent formation of allopolyploids from the same or closely related parents has been documented, e.g. in the North American allopolyploid species of *Tragopogon* [52], in *Persicaria* [53] or *Arabidopsis kamchatica* [54], and also proven by the generating synthetic allopolyploids as *Arabidopsis suecica* [55], tobacco [56] or *Tragopogon mirus* and *T. miscellus* [57].

Although less likely, we cannot rule out that karyotypic change in the Australian Camelineae species and in *Pachycladon* had significantly different dynamics. The Australian Brassicaceae species exhibit a predominantly annual growth habit [44] in comparison to the perennial *Pachycladon* [21], and a more rapid rate of genome evolution could therefore be brought about with faster nucleotide substitution rates that occur in many annuals [58,59]. Perennials are thought to have greater chromosomal stasis than annuals [60,61]. Certainly the annuality could have accelerated genome reshuffling in the Australian lineage. However, for Brassicaceae we have insufficient data on large-scale genome evolution in relation to the life forms, reproduction systems and ecological factors, and as noted by [15] and [62] chromosomal evolution is often stochastic and does not obey the models.

Phylogeographic scenario of the origin of *Pachycladon*

Pachycladon is the only New Zealand genus from the polyphyletic tribe Camelineae (other endemic crucifer species belong to Cardamineae, Lepidieae, and Notothlaspideae), and therefore an *in situ* origin seems unlikely. The closest Camelineae relatives of *Pachycladon* occur in Australia (e.g., *Arabidella*, *Ballantinia*, and *Stenopetalum*) and Eurasia/Beringia (e.g., *Arabidopsis*, *Crucihimalaya*, *Transberingia*) [10,23]. It seems more plausible that the hybridization event giving rise to *Pachycladon* has taken place on the Australian continent.

There are strong taxonomic and biogeographic links between Australia and New Zealand and dispersal across the Tasman Sea can occur in both directions. Tasmania and New Zealand have about 200 species in common [63], and there are many genera in continental Australia and New Zealand that have species that are closely related (e.g., *Aciphylla*, *Celmisia*, *Gentianella*, *Melicytus*,

and *Ranunculus*). For these shared genera, species diversity is often highest in New Zealand and the Australian species are considered to be the result of westward dispersal from New Zealand and subsequent speciation (e.g. [64,65]). Indeed, *P. radicatum* occurs in the Tasmanian mountains and is considered to have dispersed there and diverged contemporaneously with the radiation of *Pachycladon* in New Zealand [21,22]. Other taxa are also shared between the two countries, but these are considered to have dispersed eastward from Australia to New Zealand and include, for example, *Craspedia* [66], *Montigena* [67], *Poranthera* [68], *Scleranthus* [69], and Stylidiaceae [70]. This pattern of eastward dispersal means it is plausible that *Pachycladon* could have originated in Australia and then subsequently dispersed to New Zealand.

An alternative scenario of the origin of the *Pachycladon* allopolyploid ancestor in (eastern) Eurasia followed by a later dispersal to New Zealand is unlikely and incongruent with the close phylogenetic ties of *Pachycladon* to Australian Camelineae. Also, the origin of *Pachycladon* and Australian crucifer species in New Zealand is very unlikely, considering the diversity of endemic Australian Brassicaceae taxa (15 genera and 65 species [44]).

Many of the Australian Camelineae are distributed in the arid Eremaean Zone and/or the southeastern temperate biome [44], whereas in New Zealand *Pachycladon* mainly occupies montane-alpine habitats. These three environments have expanded in both Australia and New Zealand during the Pliocene and Pleistocene and are generally considered important drivers of species radiations (e.g. [71,72]). For the Australian Camelineae their origin and diversification ~6 to 9 mya [10] is consistent with other dated molecular phylogenies of a diverse range of arid-adapted taxa [73]. These dated phylogenies show the deepest divergences of taxa are consistent with the beginning of the formation of the arid zone in the mid-Miocene and that most arid-zone species lineages date to the Pliocene or earlier. The molecular clock date of 0.8 to 1.6 mya for the origin of *Pachycladon* [23] is also consistent with its alpine distribution and habitats in the Southern Alps in the South Island of New Zealand [19]. Uplift of the Southern Alps occurred over the last 8 million years, but only reached a suitable height to permanently support alpine plants during the Pleistocene.

Reconstructed genome evolution corroborates the close relationship of *Pachycladon* to *Arabidopsis* and other Camelineae species

The phylogenetic position of *Pachycladon* has been investigated repeatedly using various nuclear, chloroplast, and mitochondrial genes [10,21,23]. All studies are

congruent in placing the genus into the crucifer lineage I, within the polyphyletic tribe Camelineae [21,29,30,74]. Although the phylogenetic relationships within Camelineae are unclear, these studies have shown *Transberingia* and *Crucihimalaya* (Camelineae), *Sphaerocardamum* (Halimolobeae), *Physaria* (Physarieae), and *Boechera* (Boechereae) to be among the closest relatives of *Pachycladon*. Based on the analysis of five single-copy nuclear genes, [23] showed that *Pachycladon* has an allopolyploid origin and that the two genomes were associated with two divergent Brassicaceae lineages (lineage I and II [29,75]). One putative parental genome was associated with Camelineae *sensu lato* (and Boechereae) and the second genome being related to Brassiceae, *Sisymbrium*, Eutremeae, Thlaspideae, and remarkably also to Cardamineae on the chalcone synthase gene tree. This pattern has been interpreted as the evidence of an inter-tribal allopolyploidization event at the origin of *Pachycladon*.

A recent study using nuclear, mitochondrial and chloroplast genes, as well as significantly increasing the sampling of Australian Camelineae (in comparison to that of [23]), has confirmed the allopolyploid origin of *Pachycladon* and provides confidence that the two gene paralogues that constitute *Pachycladon* are derived from within lineage I [10]. Most importantly, this study has disclosed the close relationship of *Pachycladon* to the Australian genera *Arabidella*, *Ballantinia*, and *Stenopetalum*, and the maternal gene paralogues of *Pachycladon* and these three genera clustered with Eurasian Camelineae (*Arabidopsis*, *Capsella*, *Crucihimalaya*, *Olimarabidopsis*, *Transberingia*) and North American Boechereae. The position of the paternal gene copy was less evident, but it was always embedded within lineage I, and therefore different from the study by [23]. Mandakova et al. [10] and the present study convincingly show that the *Pachycladon* ancestor orginated from hybridization between a Camelineae species and either another species of that tribe or a very closely related tribe of lineage I. Future phylogenomic analyses of other Australian crucifer genera are likely to further resolve the parentage and phylogenetic relationships of *Pachycladon*.

Conclusion

We have shown that the remarkable infrageneric morphological and ecological differentiation in *Pachycladon* is characterized by the genome stability manifested as chromosomal and karyotypic stasis. The monophyletic *Pachycladon* species descended from a common allopolyploid ancestor (n = 10) through a whole-genome duplication of the Ancestral Crucifer Karyotype (n = 8) and subsequent diploidization by descending dysploidy. Furthermore, the present study and the phylogenetic data of [10] clearly demonstrate the close relationship between the allopolyploid *Pachycladon* and the

allopolyploid Australian Camelineae taxa. CCP data demonstrate that both mesopolyploid groups most likely originated from two different WGD events that involved identical or very similar diploid parents. We argue that the *Pachycladon* ancestor has its origin in Australia and later dispersed to the South Island of New Zealand. The endemic Australian and New Zealand Camelineae provide an excellent framework to examine the nature and consequences of differently-aged WGD events within a complex of closely related species.

Methods
Plant material

The four species of *Pachycladon* included in this study represent the morphological and ecological diversity of the genus [19]. Plants were cultivated in a glasshouse at Landcare Research, Lincoln, New Zealand. All species have known wild origins: *P. cheesemanii* (Bobs Cove, Queenstown, Otago; 168°37'E, 47°08'S), *P. enysii* (Mount Potts, Canterbury; 170°55'E, 43°30'S), *P. exile* (Awahokomo, Otago; 170°23E, 44°42'S), and *P. novae-zelandiae* (Old Man Range, Otago; 169°12'E, 45°20'S).

Chromosome preparation

Entire inflorescences were fixed in ethanol:acetic acid (3:1) fixative overnight and stored in 70% ethanol at -20° C until use. Selected flower buds were rinsed in distilled water and citrate buffer (10 mM sodium citrate, pH 4.8) and incubated in an enzyme mix (0.3% cellulase, cytohelicase, and pectolyase; all Sigma) in citrate buffer at 37°C for 3 h. Individual flower buds were disintegrated on a microscopic slide by a needle in a drop of citrate buffer and the suspension softened by adding 20 μL of 60% acetic acid. The suspension was spread on a hot plate at 50°C for ~0.5 min. Chromosomes were fixed by adding of ethanol:acetic acid (3:1, 100 μL) and dried with a hair dryer. Suitable slides were postfixed in 4% formaldehyde in distilled water for 10 min and air-dried.

DNA probes for fluorescence *in situ* hybridization (FISH)

For CCP in *P. exile*, on average each third *Arabidopsis thaliana* BAC clone was used to establish contigs corresponding to the 24 genomic blocks of the ACK [33]. For the detail composition of the BAC contigs see [49]. After initial CCP experiments in *P. exile*, some BAC contigs were split into smaller subcontigs to pinpoint rearrangement of ancestral blocks. (Sub)conting characterizing chromosome rearrangements in *P. exile* were used as CCP probes to reconstruct karyotypes of *P. cheesemanii*, *P. enysii* and *P. novae-zelandiae*. The *A. thaliana* BAC clone T15P10 (AF167571) containing 45 S rRNA genes was used for *in situ* localization of NORs, and *A. thaliana* clone pCT4.2 (M65137), corresponding to a 500-bp 5 S rRNA repeat, was used for

localization of 5 S rDNA loci. The *Arabidopsis*-type telomere repeat (TTTAGGG)n was prepared according to [76]. All DNA probes were labeled with biotin-dUTP, digoxigenin-dUTP, or Cy3- dUTP by nick translation as described by [49].

FISH

To remove cytoplasm prior to FISH, the slides were treated with pepsin (0.1 mg/mL; Sigma) in 0.01 M HCl for 10 min, postfixed in 4% formaldehyde in 2× SSC (1× SSC is 0.15 M NaCl and 0.015 M sodium citrate) for 10 min, and dehydrated in an ethanol series (70, 80, and 96%). Selected BAC clones were pooled and ethanol precipitated. The pellet was resuspended in 20 μL of hybridization mix (50% formamide and 10% dextran sulfate in 2× SSC) per slide. The probe 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. Posthybridization washing was performed in 20% formamide in 2× SSC at 42°C. Detection of was as described by [49]. Chromosomes were counterstained with 4',6-diamidino-2-phenylindole (2 μg/mL) in Vectashield (Vector Laboratories). Fluorescence signals were analyzed with an Olympus BX-61 epifluorescence microscope and AxioCam CCD camera (Zeiss). Images were acquired separately for all four fluorochromes using appropriate excitation and emission filters (AHF Analysentechnik). The four monochromatic images were pseudocolored and merged using the Adobe Photoshop CS2 software (Adobe Systems). Pachytene chromosomes in Figure 2 were straightened using the plugin 'Straighten Curved Objects' [77] in ImageJ program (National Institutes of Health).

Author details

[1]Department of Functional Genomics and Proteomics, Masaryk University, and CEITEC, Masaryk University, Brno, Czech Republic. [2]Allan Herbarium, Landcare Research, Lincoln, New Zealand.

Authors' contributions

MAL and PBH conceived the study. TM carried out the research. TM, PBH and MAL analyzed the data and wrote the manuscript. All authors read and approved the final paper.

References

1. Devos KM: **Grass genome organization and evolution.** *Curr Opin Plant Biol* 2009, **13**:1-7.
2. Schnable PS, et al: **The B73 maize genome: complexity, diversity and dynamics.** *Science* 2009, **326**:1112-1115.
3. Paterson AH, et al: **The *Sorghum bicolor* genome and the diversification of grasses.** *Nature* 2009, **457**:551-556.
4. The International Brachypodium Initiative: **Genome sequencing and analysis of the model grass *Brachypodium distachyon*.** *Nature* 2010, **463**:763-768.
5. Salse J, Bolot S, Throude M, Jouffe V, Piegu B, Quraishi UM, Calcagno T, Cooke R, Delseny M, Feuillet C: **Identification and characterization of shared duplications between rice and wheat provide new insight into grass genome evolution.** *Plant Cell* 2008, **20**:11-24.
6. Bowers JE, Chapman BA, Rong J, Paterson AH: **Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events.** *Nature* 2003, **422**:433-438.
7. Barker MS, Kane NC, Matvienko M, Kozik A, Michelmore RW, Knap SJ, Rieseberg LH: **Multiple paleopolyploidizations during the evolution of the Compositae reveal parallel patterns of duplicate gene retention after millions of years.** *Mol Biol Evol* 2008, **25**:2445-2455.
8. Ming R, et al: **The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus).** *Nature* 2008, **452**:991-996.
9. Lysak MA, Koch MA, Pecinka A, Schubert I: **Chromosome triplication found across the tribe *Brassiceae*.** *Genome Res* 2005, **15**:516-525.
10. Mandáková T, Joly S, Krzywinski M, Mummenhoff K, Lysak MA: **Fast diploidization in close mesopolyploid relatives of *Arabidopsis*.** *Plant Cell* 2010, **22**:2277-2290.
11. Schranz ME, Mitchell-Olds T: **Independent ancient polyploidy events in the sister families Brassicaceae and Cleomaceae.** *Plant Cell* 2006, **18**:1152-1165.
12. Soltis DE, Albert VA, Leebens-Mack J, Bell CD, Paterson AH, Zheng C, Sankoff D, dePamphilis CW, Wall PK, Soltis PS: **Polyploidy and angiosperm diversification.** *Am J Bot* 2009, **96**:336-348.
13. Wood TE, Takebayashi N, Barker MS, Mayrose I, Greenspoond PB, Rieseberg LH: **The frequency of polyploid speciation in vascular plants.** *Proc Natl Acad Sci USA* 2009, **106**:13875-13879.
14. Moore DM: **Chromosome numbers of Falkland Inslands angiosperms.** *Br Antarct Surv Bull* 1967, **14**:69-82.
15. Carr GD: **Chromosome evolution and speciation in Hawaiian flowering plants.** In *Evolution and speciation of island plants.* Edited by: Stuessy TF, Ono M. Cambridge: University Press; 1998:5-47.
16. Stuessy TF, Crawford DJ: **Chromosomal stasis during speciation in angiosperms of oceanic islands.** In *Evolution and speciation of island plants.* Edited by: Stuessy TF, Ono M. Cambridge: Cambridge University Press; 1998:307-324.
17. Weiss H, Sun B-Y, Stuessy TF, Kim CH, Kato H, Wakabayashi M: **Karyology of plant species endemic to Ullung Island (Korea) and selected relatives in peninsular Korea and Japan.** *Bot J Linn Soc* 2002, **138**:93-105.
18. Carr GD, Kyhos DW: **Adaptive radiation in the Hawaiian silversword alliance: II. Cytogenetics of artificial and natural hybrids.** *Evolution* 1986, **40**:969-976.
19. Heenan PB, Mitchell AD: **Phylogeny, biogeography, and adaptive radiation of *Pachycladon* (Brassicaceae) in the mountains of South Island, New Zealand.** *J Biogeogr* 2003, **30**:1737-1749.
20. Heenan PB: **A new species of *Pachycladon* (Brassicaceae) from limestone in eastern Marlborough, New Zealand.** *New Zeal J Bot* 2009, **47**:155-161.
21. Heenan PB, Mitchell AD, Koch M: **Molecular systematics of the New Zealand *Pachycladon* (Brassicaceae) complex: generic circumscription and relationship to *Arabidopsis* sens. lat. and *Arabis* sens. lat.** *New Zeal J Bot* 2002, **40**:543-562.
22. McBreen K, Heenan PB: **Phylogenetic relationships of *Pachycladon* (Brassicaceae) species based on three nuclear and two chloroplast DNA markers.** *New Zeal J Bot* 2006, **44**:377-386.
23. Joly S, Heenan PB, Lockhart PJ: **An inter-tribal hybridization event precedes the adaptive species radiation of *Pachycladon* (Brassicaceae) in New Zealand.** *Mol Phylogen Evol* 2009, **51**:365-372.
24. Mitchell AD, Heenan PB: **Genetic variation within the *Pachycladon* (Brassicaceae) complex based on fluorescent AFLP data.** *J Roy Soc New Zeal* 2002, **32**:427-443.
25. Heenan PB: **Artificial intergeneric hybrids between the New Zealand endemic *Ischnocarpus* and *Pachycladon* (Brassicaceae).** *New Zeal J Bot* 1999, **37**:595-601.

26. Yogeeswaran K, Voelckel C, Joly S, Heenan PB: **Pachycladon.** In *Wild Crop Relatives: Genomic and Breeding Resources Wild Relatives of Oilseeds.* Edited by: Kole C. Tokyo: Springer-Verlag; .

27. Dawson MI: **Index of chromosome numbers of indigenous New Zealand spermatophytes.** *New Zeal J Bot* 2000, **38**:47-150.

28. Lysak MA, Koch MA, Beaulieu JM, Meister A, Leitch IJ: **The dynamic ups and downs of genome size evolution in Brassicaceae.** *Mol Biol Evol* 2009, **26**:85-98.

29. Al-Shehbaz IA, Beilstein MA, Kellogg EA: **Systematics and phylogeny of the Brassicaceae (Cruciferae): an overview.** *Pl Syst Evol* 2006, **259**:89-120.

30. German DA, Friesen N, Neuffer B, Al-Shehbaz IA, Hurka H: **Contribution to ITS phylogeny of the Brassicaceae, with special reference to some Asian taxa.** *Plant Syst Evol* 2009, **283**:33-56.

31. Heenan PB, Dawson MI, Smissen RD, Bicknell RA: **An artificial intergeneric hybrid derived from sexual hybridization between the distantly related *Arabidopsis thaliana* and *Pachycladon cheesemanii* (Brassicaceae).** *Bot J Linn Soc* 2008, **157**:533-544.

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

33. Schranz ME, Lysak MA, Mitchell-Olds T: **The ABC's of comparative genomics in the Brassicaceae: building blocks of crucifer genomics.** *Trends Plant Sci* 2006, **11**:535-542.

34. Barber JC, Ortega JF, Santos-Guerra A, Marrero A, Jansen RK: **Evolution of endemic *Sideritis* (Lamiaceae) in Macaronesia: Insights from a chloroplast DNA restriction site analysis.** *Syst Bot* 2000, **25**:633-647.

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

36. Baldwin BG, Wagner WL: **Hawaiian angiosperm radiations of North American origin.** *Ann Bot* 2010, **105**:849-879.

37. Lindqvist C, Albert VA: **Origin of the Hawaiian endemic mints within North American *Stachys* (Lamiaceae).** *Am J Bot* 2002, **89**:1709-1724.

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

39. Hair JB: **Biosystematics of the New Zealand flora, 1945-1964.** *New Zeal J Bot* 1966, **4**:559-595.

40. Murray BG, de Lange PJ: **Chromosomes and evolution in New Zealand endemic angiosperms and gymnosperms.** In *Biology of island floras.* Edited by: Bramwell D. Cambridge: Cambridge University Press; .

41. Warwick SI, Al-Shehbaz IA: **Brassicaceae: chromosome number index and database on CD-Rom.** *Plant Syst Evol* 2006, **259**:237-248.

42. Goodson BE, Santos-Guerra A, Jansen RK: **Molecular systematics of *Descurainia* (Brassicaceae) in the Canary Islands: biogeographic and taxonomic implications.** *Taxon* 2006, **55**:671-682.

43. Rustan ØH: **Revision of the genus *Diplotaxis* (Brassicaceae) in the Cape Verde Islands, W Africa.** *Nord J Bot* 1996, **16**:19-50.

44. Hewson HJ: **Brassicaceae.** In *Flora of Australia. Volume 8.* Edited by: George AS. Canberra: Australian Publishing Service; 1982:231-357.

45. Boivin K, Acarkan A, Mbulu R-S, Clarenz O, Schmidt R: **The *Arabidopsis* genome sequence as a tool for genome analysis in Brassicaceae. A comparison of the *Arabidopsis* and *Capsella rubella* genomes.** *Pl Physiol* 2004, **135**:735-744.

46. Kuittinen H, de Haan AA, Vogel C, Oikarinen S, Leppälä J, Koch M, Mitchell-Olds T, Langley CH, Savolainen O: **Comparing the linkage maps of the close relatives *Arabidopsis lyrata* and *A. thaliana*.** *Genetics* 2004, **168**:1575-1584.

47. Schranz ME, Windsor AJ, Song B-H, Lawton-Rauh A, Mitchell-Olds T: **Comparative genetic mapping in *Boechera stricta*, a close relative of Arabidopsis.** *Plant Physiol* 2007, **144**:286-298.

48. Schubert I: **Chromosome evolution.** *Curr Opin Plant Biol* 2007, **10**:109-115.

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

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

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

52. Soltis DE, Soltis PS, Pires JC, Kovarik A, Tate JA, Mavrodiev E: **Recent and recurrent polyploidy in *Tragopogon* (Asteraceae): cytogenetic, genomic and genetic comparisons.** *Biol J Linn Soc* 2004, **82**:485-501.

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

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

55. Comai L, Tyagi AP, Winter K, Holmes-Davis R, Reynolds SH, Stevens Y, Byers B: **Phenotypic instability and rapid gene silencing in newly formed *Arabidopsis* allotetraploids.** *Plant Cell* 2000, **12**:1551-1567.

56. Yoong LK, Souckova-Skalicka K, Sarasan V, Clarkson JJ, Chase MW, Kovarik A, Leitch AR: **A genetic appraisal of new synthetic *Nicotiana tabacum* (Solanaceae) and the Kostoff synthetic tobacco.** *Am J Bot* 2006, **93**:875-883.

57. Tate JA, Symonds VV, Doust AN, Buggs RJA, Mavrodiev EV, Majurev LC, Soltis PS, Soltis DE: **Synthetic polyploids of *Tragopogon miscellus* and *T. mirus* (Asteraceae):60 years after Ownbey's discovery.** *Am J Bot* 2009, **96**:979-988.

58. Andreasen K, Baldwin BG: **Unequal evolutionary rates between annual and perennial lineages of checker mallows (Sidalcea, Malvaceae): evidence from 18S-26 S rDNA internal and external transcribed spacers.** *Mol Biol Evol* 2001, **18**:936-944.

59. Kay KM, Whittall JB, Hodges SA: **A survey of nuclear ribosomal internal transcribed spacer substitution rates across angiosperms: an approximate molecular clock with life history effects.** *BMC Evol Biol* 2006, **6**:36.

60. Stebbins GL: *Variation and Evolution in Plants* New York: Columbia University Press; 1950.

61. Levin DA: *The role of chromosomal change in plant evolution* Oxford: Oxford University Press; 2002.

62. Kyhos DW, Carr GD: **Chromosome stability and lability in plants.** *Evol Theory* 1994, **10**:227-248.

63. Jordan GJ: **An investigation of long-distance dispersal based on species native to both Tasmania and New Zealand.** *Aust J Bot* 2001, **49**:333-340.

64. Lockhart PJ, McLenachan PA, Havell D, Glenny D, Huson D, Jensen U: **Phylogeny, radiation, and transoceanic dispersal of New Zealand alpine buttercups: molecular evidence under split decomposition.** *Ann Missouri Bot Gard* 2001, **88**:458-477.

65. Mitchell AD, Heenan PB, Murray BG, Molloy BPJ, de Lange PJ: **Evolution of the south-west Pacific genus *Melicytus* (Violaceae): evidence from DNA sequence data, cytology, and sex expression.** *Aust Syst Bot* 2009, **22**:143-157.

66. Ford KA, Ward JM, Smissen RD, Wagstaff SJ, Breitwieser I: **Phylogeny and biogeography of *Craspedia* (Asteraceae: Gnaphalieae) based on ITS, ETS and *psbAtrnH* sequence data.** *Taxon* 2007, **56**:783-794.

67. Wagstaff SJ, Heenan PB, Sanderson MJ: **Classification, origins, and patterns of diversification in New Zealand Carmichaelinae (Fabaceae).** *Am J Bot* 1999, **86**:1346-1356.

68. Vorontsova MS, Hoffmann P, Maurin O, Chase MW: **Molecular phylogenetics of tribe Poranthereae (Phyllanthaceae; Euphorbiaceae sensu lato).** *Am J Bot* 2007, **94**:2026-2040.

69. Smissen RD, Garnock-Jones PJ, Chambers GK: **Phylogenetic analysis of ITS sequences suggests a Pliocene origin for the bipolar distribution of *Scleranthus* (Caryophyllaceae).** *Aust Syst Bot* 2003, **16**:301-315.

70. Wagstaff SJ, Wege J: **Patterns of diversification in New Zealand Stylidiaceae.** *Am J Bot* 2002, **89**:865-874.

71. Crisp M, Cook L, Steane D: **Radiation of the Australian flora: what can comparisons of molecular phylogenies across multiple taxa tell us about the evolution of diversity in present-day communities?** *Phil Trans R Soc Lond B* 2004, **359**:1551-1571.

72. McGlone MS, Duncan RP, Heenan PB: **Endemism, species selection and the origin and distribution of the vascular plant flora of New Zealand.** *J Biogeogr* 2001, **28**:199-216.

73. Byrne M, Yeates DK, Joseph L, Kearney M, Bowler J, Williams MAJ, Cooper S, Donnellan SC, Keogh JS, Leys R, Melville J, Murphy DJ, Porch N, Wyrwoll K-

H: **Birth of a biome: insights into the assembly and maintenance of the Australian arid zone biota.** *Mol Ecol* 2008, **17**:4398-4417.

74. Bailey CD, Koch MA, Mayer M, Mummenhoff K, O'Kane SL, Warwick SI, Windham MD, Al-Shehbaz IA: **Toward a global phylogeny of the Brassicaceae.** *Mol Biol Evol* 2006, **23**:2142-2160.

75. Beilstein MA, Al-Shehbaz IA, Mathews S, Kellogg E: **Brassicaceae phylogeny inferred from phytochrome A and ndhF sequence data: tribes and trichomes revisited.** *Am J Bot* 2008, **95**:1307-1327.

76. Ijdo JW, Wells RA, Baldini A, Reeders ST: **Improved telomere detection using a telomere repeat probe (TTAGGG)n generated by PCR.** *Nucleic Acids Res* 1991, **19**:4780.

77. Kocsis E, Trus BL, Steer CJ, Bisher ME, Steven AC: **Image averaging of flexible fibrous macromolecules: The clathrin triskelion has an elastic proximal segment.** *J Struct Biol* 1991, **107**:6-14.