PLANT GENETICS

Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome

Boulos Chalhoub,[1]*† France Denoeud,[2,3,4]* Shengyi Liu,[5]* Isobel A. P. Parkin,[6]† Haibao Tang,[7,8] Xiyin Wang,[9,10] Julien Chiquet,[11] Harry Belcram,[1] Chaobo Tong,[5] Birgit Samans,[12] Margot Corréa,[2] Corinne Da Silva,[2] Jérémy Just,[1] Cyril Falentin,[13] Chu Shin Koh,[14] Isabelle Le Clainche,[1] Maria Bernard,[2] Pascal Bento,[2] Benjamin Noel,[2] Karine Labadie,[2] Adriana Alberti,[2] Mathieu Charles,[15] Dominique Arnaud,[1] Hui Guo,[9] Christian Daviaud,[16] Salman Alamery,[17] Kamel Jabbari,[1,18] Meixia Zhao,[19] Patrick P. Edger,[20] Houda Chelaifa,[1] David Tack,[21] Gilles Lassalle,[13] Imen Mestiri,[1] Nicolas Schnel,[13] Marie-Christine Le Paslier,[15] Guangyi Fan,[22] Victor Renault,[23] Philipp E. Bayer,[17] Agnieszka A. Golicz,[17] Sahana Manoli,[17] Tae-Ho Lee,[9] Vinh Ha Dinh Thi,[1] Smahane Chalabi,[1] Qiong Hu,[5] Chuchuan Fan,[24] Reece Tollenaere,[17] Yunhai Lu,[1] Christophe Battail,[2] Jinxiong Shen,[24] Christine H. D. Sidebottom,[14] Xinfa Wang,[5] Aurélie Canaguier,[1] Aurélie Chauveau,[15] Aurélie Bérard,[15] Gwenaëlle Deniot,[13] Mei Guan,[25] Zhongsong Liu,[25] Fengming Sun,[22] Yong Pyo Lim,[26] Eric Lyons,[27] Christopher D. Town,[7] Ian Bancroft,[28] Xiaowu Wang,[29] Jinling Meng,[24] Jianxin Ma,[19] J. Chris Pires,[30] Graham J. King,[31] Dominique Brunel,[15] Régine Delourme,[13] Michel Renard,[13] Jean-Marc Aury,[2] Keith L. Adams,[21] Jacqueline Batley,[17,32] Rod J. Snowdon,[12] Jorg Tost,[16] David Edwards,[17,32]† Yongming Zhou,[24]† Wei Hua,[5]† Andrew G. Sharpe,[14]† Andrew H. Paterson,[9]† Chunyun Guan,[25]† Patrick Wincker[2,3,4]†

Oilseed rape (*Brassica napus* L.) was formed ~7500 years ago by hybridization between *B. rapa* and *B. oleracea*, followed by chromosome doubling, a process known as allopolyploidy. Together with more ancient polyploidizations, this conferred an aggregate 72× genome multiplication since the origin of angiosperms and high gene content. We examined the *B. napus* genome and the consequences of its recent duplication. The constituent A_n and C_n subgenomes are engaged in subtle structural, functional, and epigenetic cross-talk, with abundant homeologous exchanges. Incipient gene loss and expression divergence have begun. Selection in *B. napus* oilseed types has accelerated the loss of glucosinolate genes, while preserving expansion of oil biosynthesis genes. These processes provide insights into allopolyploid evolution and its relationship with crop domestication and improvement.

The Brassicaceae are a large eudicot family (*1*) and include the model plant *Arabidopsis thaliana*. Brassicas have a propensity for genome duplications (Fig. 1) and genome mergers (*2*). They are major contributors to the human diet and were among the earliest cultigens (*3*).

B. napus (genome $A_nA_nC_nC_n$) was formed by recent allopolyploidy between ancestors of *B. oleracea* (Mediterranean cabbage, genome C_oC_o) and *B. rapa* (Asian cabbage or turnip, genome A_rA_r) and is polyphyletic (*2*, *4*), with spontaneous formation regarded by Darwin as an example of unconscious selection (*5*). Cultivation began in Europe during the Middle Ages and spread worldwide. Diversifying selection gave rise to oilseed rape (canola), rutabaga, fodder rape, and kale morphotypes grown for oil, fodder, and food (*4*, *6*).

The homozygous *B. napus* genome of European winter oilseed cultivar 'Darmor-*bzh*' was assembled with long-read [>700 base pairs (bp)] 454 GS-FLX+ Titanium (Roche, Basel, Switzerland) and Sanger sequence (tables S1 to S5 and figs. S1 to S3) (*7*). Correction and gap filling used 79 Gb of Illumina (San Diego, CA) HiSeq sequence. A final assembly of 849.7 Mb was obtained with SOAP (*8*) and Newbler (Roche), with 89% nongapped sequence (tables S2 and S3). Unique mapping of ~5× nonassembled 454 sequences from *B. rapa* ('Chiifu') or *B. oleracea* ('TO1000') assigned most of the 20,702 *B. napus* scaffolds to either the A_n (8294) or the C_n (9984) subgenomes (tables S4 and S5 and fig. S3). The assembly covers ~79% of the 1130-Mb genome and includes 95.6% of *Brassica* expressed sequence tags (ESTs) (*7*). A single-nucleotide polymorphism (SNP) map (tables S6 to S9 and figs. S4 to S8) genetically anchored 712.3 Mb (84%) of the genome assembly, yielding pseudomolecules for the 19 chromosomes (table S10).

The assembled C_n subgenome (525.8 Mb) is larger than the A_n subgenome (314.2 Mb), consistent with the relative sizes of the assembled C_o genome of *B. oleracea* (540 Mb, 85% of the ~630-Mb genome) and the A_r genome of *B. rapa* (312 Mb, 59% of the ~530-Mb genome) (*9*–*11*). The *B. napus* assembly contains 34.8% transposable elements (TEs), less than the 40% estimated from raw reads (tables S11 to S14) (*7*), with asymmetric distribution in the A_n and C_n subgenomes (table S12) as in the progenitor genomes (*9*–*11*). A small TE fraction has proliferated since *B. napus* separated from its progenitors (*7*), at lower rates in the *B. napus* subgenomes than the corresponding progenitor genomes (table S14 and figs. S9 and S10).

The *B. napus* genome contains 101,040 gene models estimated from 35.5 Gb of RNA sequencing (RNA-seq) data (tables S15 and S16) in combination with ab initio gene prediction, protein and EST alignments, and transposon masking (*7*). Of these, 91,167 were confirmed by matches with *B. rapa* and/or *B. oleracea* predicted proteomes. Genes are abundant in distal euchromatin but sparse near centromeric heterochromatin (Fig. 2). RNA-seq data revealed alternative splicing in 48% of genes, with frequent intron retention (62%) and rare exon skipping (3%) (tables S17 and S18 and fig. S11).

The *B. napus* A_n and C_n subgenomes are largely colinear to the corresponding diploid A_r and C_o

[1]Institut National de Recherche Agronomique (INRA)/Université d'Evry Val d'Essone, Unité de Recherche en Génomique Végétale, UMR1165, Organization and Evolution of Plant Genomes, 2 rue Gaston Crémieux, 91057 Evry, France. [2]Commissariat à l'Energie Atomique (CEA), Institut de Génomique (IG), Genoscope, BP5706, 91057 Evry, France. [3]Université d'Evry Val d'Essone, UMR 8030, CP5706, Evry, France. [4]Centre National de Recherche Scientifique (CNRS), UMR 8030, CP5706, Evry, France. [5]Key Laboratory of Biology and Genetic Improvement of Oil Crops, Ministry of Agriculture of People's Republic of China, Oil Crops Research Institute, Chinese Academy of Agricultural Sciences, Wuhan 430062, China. [6]Agriculture and Agri-Food Canada, 107 Science Place, Saskatoon, SK S7N 0X2, Canada. [7]J. Craig Venter Institute, Rockville, MD 20850, USA. [8]Center for Genomics and Biotechnology, Fujian Agriculture and Forestry, University, Fuzhou 350002, Fujian Province, China. [9]Plant Genome Mapping Laboratory, University of Georgia, Athens, GA 30602, USA. [10]Center of Genomics and Computational Biology, School of Life Sciences, Hebei United University, Tangshan, Hebei 063000, China. [11]Laboratoire de Mathématiques et Modélisation d'Evry–UMR 8071 CNRS/Université d'Evry val d'Essonne–USC INRA, Evry, France. [12]Department of Plant Breeding, Research Center for Biosystems, Land Use and Nutrition, Justus Liebig University, Heinrich-Buff-Ring 26-32, 35392 Giessen, Germany. [13]INRA, Institut de Génétique, Environnement et Protection des Plantes (IGEPP) UMR1349, BP35327, 35653 Le Rheu Cedex, France. [14]National Research Council Canada, 110 Gymnasium Place, Saskatoon, SK S7N 0W9, Canada. [15]INRA, Etude du Polymorphisme des Génomes Végétaux, US1279, Centre National de Génotypage, CEA–IG, 2 rue Gaston Crémieux, 91057 Evry, France. [16]Laboratory for Epigenetics and Environment, Centre National de Génotypage, CEA–IG, 2 rue Gaston Crémieux, 91000 Evry, France. [17]Australian Centre for Plant Functional Genomics, School of Agriculture and Food Sciences, University of Queensland, St. Lucia, QLD 4072, Australia. [18]Cologne Center for Genomics, University of Cologne, Weyertal 115b, 50931 Köln, Germany. [19]Department of Agronomy, Purdue University, WSLR Building B018, West Lafayette, IN 47907, USA. [20]Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720, USA. [21]Department of Botany, University of British Columbia, Vancouver, BC, Canada. [22]Beijing Genome Institute–Shenzhen, Shenzhen 518083, China. [23]Fondation Jean Dausset–Centre d'Étude du Polymorphisme Humain, 27 rue Juliette Dodu, 75010 Paris, France. [24]National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan 430070, China. [25]College of Agronomy, Hunan Agricultural University, Changsha 410128, China. [26]Molecular Genetics and Genomics Laboratory, Department of Horticulture, Chungnam National University, Daejeon-305764, South Korea. [27]School of Plant Sciences, iPlant Collaborative, University of Arizona, Tucson, AZ, USA. [28]Department of Biology, University of York, Wentworth Way, Heslington, York YO10 5DD, UK. [29]Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences, Beijing, China. [30]Division of Biological Sciences, University of Missouri, Columbia, MO 65211, USA. [31]Southern Cross Plant Science, Southern Cross University, Lismore, NSW 2480, Australia. [32]School of Plant Biology, University of Western Australia, WA 6009, Australia.

*These authors contributed equally to this work. **†Corresponding author. E-mail: chalhoub@evry.inra.fr (B.C.); other e-mail addressees are given in the supplementary materials.**

Corrected 4 September 2014; see full text.

genomes, with asymmetric gene distribution (42,320 and 48,847, respectively) and 93% of the diploid gene space in orthologous blocks (fig. S12) (*7*). We identified 34,255 and 38,661 orthologous gene pairs between the A_n and C_n subgenomes and their respective progenitor genomes (fig. S13). Comparison of A_n-A_r and C_n-C_o orthologous gene pairs suggested a divergence 7500 to 12,500 years ago (fig. S14), indicating formation of *B. napus* after this date. Synteny with *Arabidopsis* (table S19) confirmed the triplicated mesoploid structure (*9*–*11*) of the A_n and C_n subgenomes, with the recent allopolyploidy conferring on *B. napus* an aggregate 72× genome multiplication since the origin of angiosperms (Fig. 1) (*7*).

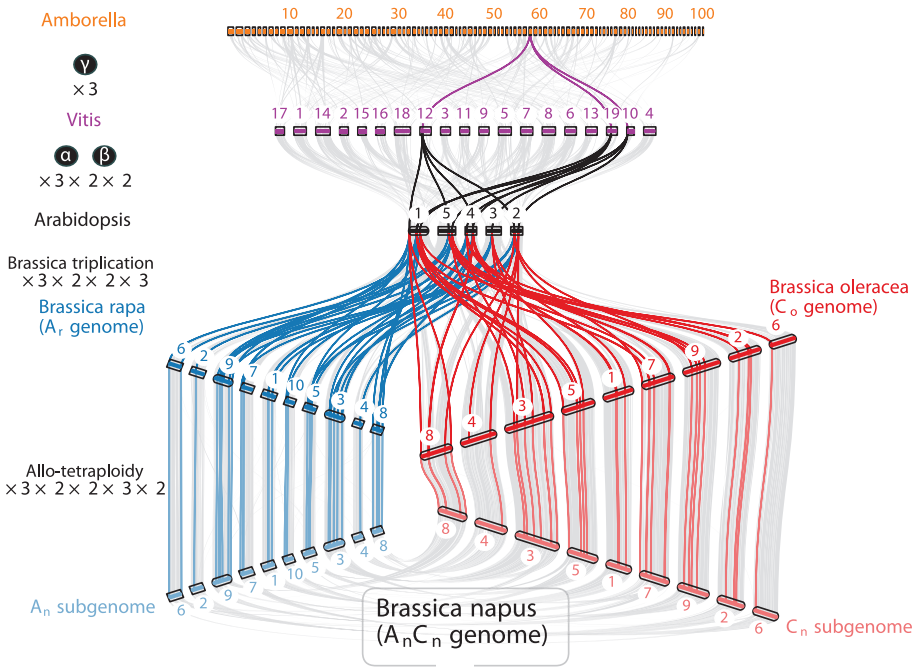

Fig. 1. Recurrent genome duplications in *B. napus*. Genomic alignments between the basal angiosperm *Amborella trichopoda* (*24*), the basal eudicot *Vitis vinifera* (*25*), and the model crucifer *A. thaliana*, as well as *B. rapa* (*9*), *B. oleracea* (*10*, *11*), and *B. napus*, are shown. A typical ancestral region in *Amborella* is expected to match up to 72 regions in *B. napus* (69 were detected for this specific region). Gray wedges in the background highlight conserved synteny blocks with more than 10 gene pairs.

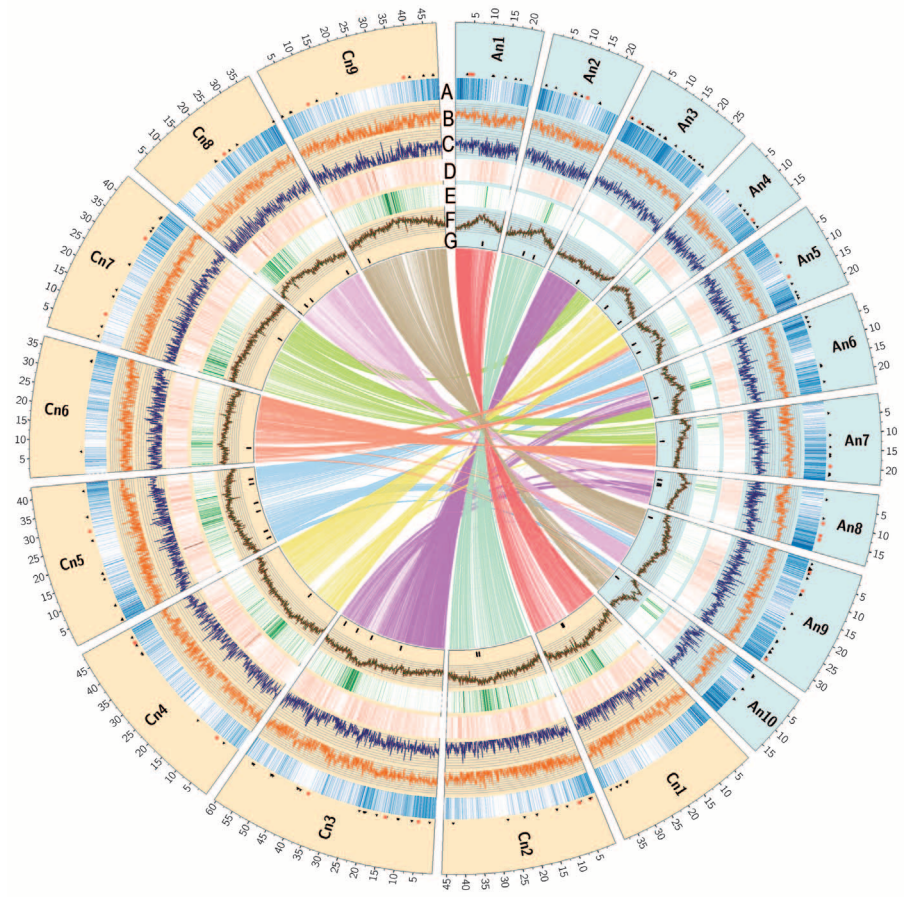

Fig. 2. The genome of the *B. napus* oilseed cultivar 'Darmor-*bzh*'. The genome comprises 9 chromosomes belonging to the C_n subgenome and 10 to the A_n subgenome, scaled on the basis of their assembled lengths. Tracks displayed are (**A**) gene density (nonoverlapping, window size = 100 kb for all tracks). Positions showing loss of one or more consecutive genes are displayed (triangles) along with homeologous exchanges, detected as missing genomic segments that have been replaced by duplicates of corresponding homeologous segments (red rectangles). (**B** and **C**) Transcription states estimated by RNA-seq in leaves (B) and roots (C) (in nonoverlapping 100-kb windows). (**D**) DNA transposon density. (**E**) Retrotransposon density. (**F**) CpG methylation in leaves (green) and roots (brown); both curves are overlapping. (**G**) Centromeric repeats (densities exaggerated for visual clarity). Homeologous relationships between A_n and C_n chromosomes are displayed with connecting lines colored according to the C_n chromosomes.

Corrected 4 September 2014; see full text.

Most orthologous gene pairs in *B. rapa* and *B. oleracea* remain as homeologous pairs in *B. napus* (tables S19 to S25 and figs. S12 to S17) (*7*). DNA sequence analysis (*7*) confirmed the loss of 112 A_n and 91 C_n genes in *B. napus* 'Darmor-*bzh*' (tables S21 to S26), ~2.6 times higher than the 41 and 37 genes lost in *B. rapa* 'Chiifu' and *B. oleracea* 'TO1000' respectively (tables S26 and S27; χ^2 test $P = 5.3 \times 10^{-14}$). Further analyses of a *Brassica* diversity set showed that ~47% of Darmor-*bzh* A_n and 31% of C_n deleted genes were also deleted in at least one additional progenitor genotype (tables S28 and S29), indicating that their deletion probably predated allopolyploidization of *B. napus* (*7*). A high proportion (27% to 54%) of the remaining Darmor-*bzh* deleted genes were also deleted from diverse *B. napus* genotypes (tables S28 and S29).

Homeologous exchanges (HEs), including crossovers and noncrossovers, are frequent between *B. napus* subgenomes and range in size from large segments to single SNPs (*7*) (Fig. 3, figs. S17 to S24, and tables S30 to S39).

At the chromosome segment level, HEs are characterized by replacement of a chromosomal region with a duplicated copy from the corresponding homeologous subgenome (*7*). We identified 17 HEs, 14 C_n to A_n and 3 A_n to C_n (Fig. 3, fig. S19, and tables S30 and S31). Sequences from seven diverse *B. napus* genotypes revealed both shared and specific segmental HEs. These are of varying sizes and are most frequent between chromosomes A_n1-C_n1, A_n2-C_n2, and A_n9-C_n9 (table S32, Fig. 3, and fig. S19). Larger HEs found in the synthetic *B. napus* H165 affect, for example, most of chromosomes A_n1-C_n1 and A_n2-C_n2 (Fig. 3 and fig. S19). Functional annotation of genes within HEs suggests some have experienced selection, contributing to the

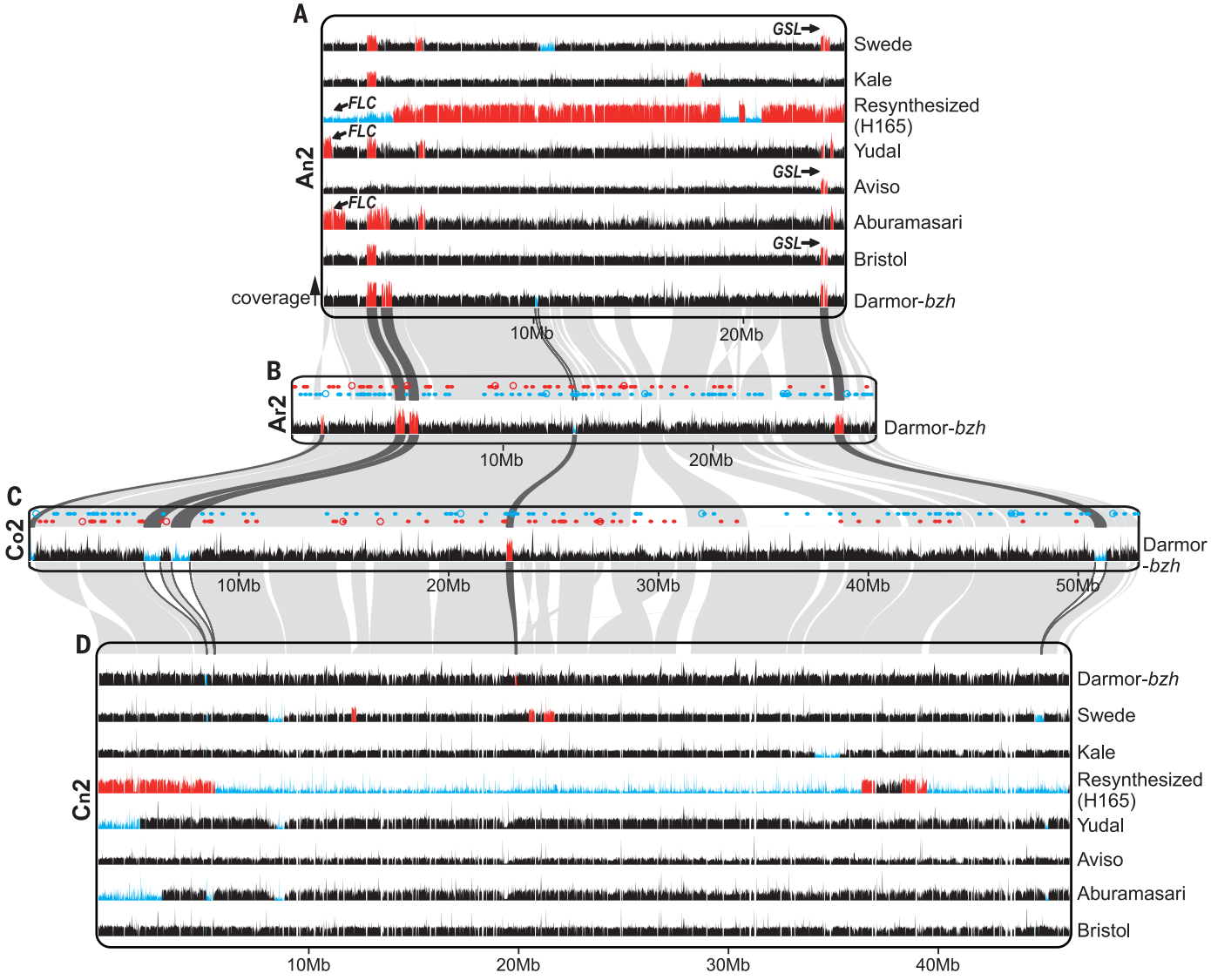

Fig. 3. HEs between *B. napus* chromosomes A_n2 and C_n2. (**A**) Coverage depth obtained along the A_n2 chromosome after mapping Illumina sequence reads from seven natural and one resynthesized *B. napus* genotypes (named on the right) to the reference genome of *B. napus* 'Darmor-*bzh*.' (**B** and **C**) Coverage depth obtained for A_r2 and C_o2 chromosomes, respectively, after mapping >21 genome-equivalents of Illumina sequence reads from *B. napus* Darmor-*bzh* on the *B. rapa* and *B. oleracea* genome assemblies concatenated together. (**D**) Similar to (A), where the C_n2 chromosome of Darmor-*bzh* is displayed. Segmental HEs are revealed based on sequence read coverage analysis, where a duplication (red) is revealed by significantly greater coverage for a given segment than the rest of the genome (black) and a deletion (blue) by little or no coverage for the corresponding homeologous segment. Sizes of chromosomes are indicated in Mb. A_n-to-C_n converted genes (at 60% or more conversion sites) are plotted as blue dots on A_r2 (B) and red dots on C_o2 (C). C_n-to-A_n converted genes are plotted as blue dots on C_o2 (C) and red dots on A_r2 (B). Open circles denote entirely converted genes using the same color code. Light gray lines connecting (A), (B), (C), and (D) indicate orthology relationships, and dark gray lines highlight segmental HEs in Darmor-*bzh* (names and descriptions detailed in table S31). Further HEs occurring between other homeologous chromosomes are shown in fig. S19. Black arrows in (A) indicate HEs involving *GSL* and *FLC* genes.

Corrected 4 September 2014; see full text.

diversification of winter, spring, and Asian types of oilseed rape, rutabaga, and kale vegetables (Fig. 3B, fig. S19, and table S33).

We also identified 37 C_n to A_n and 56 A_n to C_n whole-gene conversions (*12*) (table S34).

At the single-nucleotide scale, exchanges between homeologous subgenomes account for ~86% of allelic differences between *B. napus* and its progenitors, with nearly ~1.3 times more conversions from the A_n to the C_n subgenome than the reverse (χ^2, $P < 1.6 \times 10^{-16}$) (tables S35 and S36). A total of 16,938 A_n and 13,429 C_n genes (with 10,258 from homeologous pairs) had at least two conversion sites (table S37); 842 A_n and 579 C_n genes were highly converted with 60 to 90% conversion sites (table S37).

Transcript abundance indicated that 96% of genes are expressed in leaves, roots, or both (*7*) (figs. S25 to S29 and tables S40 to S42). Subgenome and tissue effects and tissue-by-subgenome interactions were statistically significant (χ^2, $P < 0.01$), with 45 expression patterns (fig. S26) grouping into nine clusters (table S41).

A_n and C_n homeologs contributed similarly to gene expression for 17,326 (58.3%) gene pairs (χ^2, $P < 0.01$) (fig. S27 and table S41). Both tissues showed higher expression for 4665 (15.7%) A_n homeologs and 5437 (17.3%) C_n homeologs (fig. S28 and table S41). There were 1062 gene pairs (3.7%) with higher expression of the A_n homeolog over the C_n homeolog in leaves, whereas the reverse was true in roots (fig. S28 and table S41). Conversly, for 966 gene pairs (3.3%), A_n homeologs had lower expression than C_n homeologs in leaves, with the pattern inverted in roots (fig. S28 and table S41).

Gene expression is generally inversely related to CpG, CHG, and CHH cytosine DNA methylation levels (p is phosphate, implying a C is directly followed by a G, and H is A, C, or T) (*7*). Methyl bisulfite sequencing in Darmor-*bzh* (figs. S30 to S32 and tables S43 to S45) showed 4 to 8% higher methylation in C_n genes than in their homoelogous A_n genes (table S44), possibly because of greater transposon density in the C_n subgenome (Fig. 2F). Of the ~3100 gene pairs with differential gene body and/or untranslated region methylation between A_n and C_n homeologs in both roots and leaves, 51% were equally expressed. Only ~34% showed higher expression for the less-methylated homeologs, and the remaining ~15% showed the opposite pattern (table S45).

It is interesting that partitioning of homeolog gene expression is largely established in *B. napus* with patterns of both genome dominance and genome equivalence. The absence of significant bias toward either subgenome of the recent *B. napus* allopolyploid contrasts with many old and recent polyploids (*13*–*17*) but concurs with other old polyploids (*18*).

Oilseed *B. napus* has undergone intensive breeding to optimize seed oil content and lipid composition, decrease nutritionally undesirable erucic acid and glucosinolates (GSLs), optimize flowering behavior, and improve pathogen resistance.

The expansion of *B. napus* lipid biosynthesis genes exceeds that known in other oilseed plants, with 1097 and 1132 genes annotated in the A_n and C_n subgenomes, respectively (*7*) (tables S46 to S48). Most lipid biosynthesis genes identified in the progenitor genomes are conserved in *B. napus*. For 18 acyl lipid orthologs, 3 and 2 genes appeared to be deleted from A_n and C_n subgenomes, respectively. Another 13 have been converted by HEs, nine from A_n to C_n and four from C_n to A_n (tables S47 and S48) (*7*).

Genetic variation for reduced seed GSLs also appears to be under breeding-directed selection. GSLs are sulfur-rich secondary metabolites important for plant defense and human health (*19*); however, high levels in seeds form toxic breakdown products in animal feeds (*20*). All 22 GSL catabolism genes identified in *B. rapa* and *B. oleracea* (*10*) are conserved in *B. napus* (*7*), and orthologs of only three C_o and one A_r GSL biosynthesis genes are missing (table S49). One deleted homeologous pair, corresponding to orthologs of *B. oleracea Bo2g161590* and *B. rapa Bra02931*, colocates with two quantitative trait loci (QTLs) for total aliphatic GSL content (*21*) and corresponds to a HE in which a segment of A_n2, with a missing GSL gene, has replaced the C_n2 homeolog (Fig. 3). Two additional QTLs for aliphatic GSL content (*21*) colocalize with a deletion of the *B. rapa Bra035929* ortholog on A_n9 and its nondeleted homeolog on C_n9 (*BnaC09g05300D*, fig. S17).

We identified 425 nucleotide binding site leucine-rich repeat (NBS-LRR) sequences encoding resistance gene homologs (245 on C_n and 180 on A_n). Of these, 75% (153 A_n and 224 C_n) are syntenic to A_r and C_o progenitors (*7*) (table S50 and figs. S33 and S34). We confirmed the absence of five NBS-LRR genes from the A_n subgenome, three from the C_n subgenome, and three from *B. rapa* (A_r), with none absent from *B. oleracea* C_o. This variation may reflect differential selection for resistance to diseases.

B. napus morphotypes show broad adaptation to different climatic zones and latitudes. A key adaptive gene controlling vernalization and photoperiod responses, *FLOWERING LOCUS C* (*FLC*) is expanded from one copy in *A. thaliana* to four in *B. rapa* and *B. oleracea* and nine or more in *B. napus* (*7*) (table S51). Different *FLC* homologs lie within HEs, from C_n2 to A_n2 in the Asian semiwinter oilseed forms Yudal and Aburamasari (Fig. 3) and C_n9 to A_n10 in late-flowering swedes (fig. S19 and table S51). These loci correspond to important QTLs for vernalization requirement and flowering time (*22*).

Human cultivation and breeding of *B. napus* morphotypes may have selected favorable HEs, causing subgenome restructuring of regions containing genes controlling valuable agronomic traits such as those shown here for oil biosynthesis, seed GSL content, disease resistance, and flowering. Because *B. napus* is a young allopolyploid beginning gene loss and genome reorganization, further partitioning of expression may become a key determinant for the long-term preservation of its duplicated genes (*23*). The integrative genomic resources that we report provide unique perspectives on the early evolution of a domesticated polyploid and will facilitate the manipulation of useful variation, contributing to sustainable increases in oilseed crop production to meet growing demands for both edible and biofuel oils.

REFERENCES AND NOTES

1. I. Al-Shehbaz, *Taxon* **61**, 931 (2012).
2. U. Nagaharu, *Jpn. J. Bot.* **7**, 389 (1935).
3. S. Prakash, in Brassica *Crops and Wild Allies: Biology and Breeding*, S. Tsunoda, K. Hinata, C. Gomez-Campo, Eds. (Japan Scientific Society Press, Tokyo, 1980), pp. 151–163.
4. C. J. Allender, G. J. King, *BMC Plant Biol.* **10**, 54 (2010).
5. C. Darwin, *The Variation of Animals and Plants Under Domestication* (Murray, London, 1905), vol. 2, pp. 398–403.
6. H. L. Liu, *Rapeseed Genetics and Breeding* (China Agricultural University, Beijing, 2000), pp. 1–81.
7. Supplementary materials for this article are available on *Science* Online.
8. R. Li, Y. Li, K. Kristiansen, J. Wang, *Bioinformatics* **24**, 713–714 (2008).
9. X. Wang *et al*., *Nat. Genet.* **43**, 1035–1039 (2011).
10. S. Liu *et al*., *Nat. Commun.* **5**, 3930 (2014).
11. I. A. Parkin *et al*., *Genome Biol.* **15**, R77 (2014).
12. A. H. Paterson *et al*., *Nature* **492**, 423–427 (2012).
13. J. C. Schnable, N. M. Springer, M. Freeling, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 4069–4074 (2011).
14. H. Tang *et al*., *Genetics* **190**, 1563–1574 (2012).
15. M. J. Yoo, E. Szadkowski, J. F. Wendel, *Heredity* **110**, 171–180 (2013).
16. D. C. Ilut *et al*., *Am. J. Bot.* **99**, 383–396 (2012).
17. R. J. Buggs *et al*., *Curr. Biol.* **21**, 551–556 (2011).
18. O. Garsmeur *et al*., *Mol. Biol. Evol.* **31**, 448–454 (2014).
19. I. E. Sønderby, F. Geu-Flores, B. A. Halkier, *Trends Plant Sci.* **15**, 283–290 (2010).
20. B. Wittkop, R. J. Snowdon, W. Friedt, *Euphytica* **170**, 131–140 (2009).
21. R. Delourme *et al*., *BMC Genomics* **14**, 120 (2013).
22. X. Zou *et al*., *PLOS ONE* **7**, e45751 (2012).
23. M. Lynch, A. Force, *Genetics* **154**, 459–473 (2000).
24. A. G. Sharpe, D. J. Lydiate, *Genome* **46**, 461–468 (2003).
25. N. Foisset, R. Delourme, P. Barret, M. Renard, *Theor. Appl. Genet.* **91**, 756–761 (1995).

ACKNOWLEDGMENTS

Sequence Read Archive accession numbers of *B. napus* sequencing data are ERP005275 and PRJEB6069, and those of *B. rapa* and *B. oleracea* data are PRJNA248388 and PRJNA158027. The *B. napus* assembly is available at ENA (European Nucleotide Archive), in the WGS section for contigs (accession numbers CCCW010000001 to CCCW010044187) and the CON section for scaffolds, chromosomes, and annotation (accession numbers LK031787 to LK052685). The *B. napus* genome is also available at CoGe (https://genomevolution.org/CoGe/) and at (www.genoscope.cns.fr/brassicanapus) the Genoscope Genome Database, with additional tools for comparative genomic analysis. The *B. napus* segregating populations Darmor-*bzh* × Yudal and Darmor × Bristol are available at INRA-IGEPP, Rennes, France, under a material transfer agreement. This project was funded by the French ANR (Agence Nationale de la Recherche, www.agence-nationale-recherche.fr) 2009 (ANR-09-GENM-021) to B.C., P.W., D.B., and R.D., with additional funding from Sofi-Proteol for bioinformatic personnnel (J.J.); the National Basic Research Program of China (2011CB109300) for S.L., Y.Z., C.G., and W.H.; and the Canadian Canola Genome Sequencing Initiative (http://aafc-aac.usask.ca/canseq/, Genome Alberta, and industry partners) for I.A.P.P., A.G.S., C.K., and C.S. Research leaders are B.C., S.L., I.A.P.P., X.W., I.B., R.D., J.B., D.E., Y.Z., W.H., A.G.S., A.H.P., C.G., and P.W. Additional acknowledgements and author contributions are included in the supplementary materials.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/345/6199/950/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S34
Tables S1 to S51
References (*26*–*145*)

17 March 2014; accepted 29 July 2014
10.1126/science.1253435

Corrected 4 September 2014; see full text.

Science

Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome

Boulos Chalhoub, France Denoeud, Shengyi Liu, Isobel A. P. Parkin, Haibao Tang, Xiyin Wang, Julien Chiquet, Harry Belcram, Chaobo Tong, Birgit Samans, Margot Corréa, Corinne Da Silva, Jérémy Just, Cyril Falentin, Chu Shin Koh, Isabelle Le Clainche, Maria Bernard, Pascal Bento, Benjamin Noel, Karine Labadie, Adriana Alberti, Mathieu Charles, Dominique Arnaud, Hui Guo, Christian Daviaud, Salman Alamery, Kamel Jabbari, Meixia Zhao, Patrick P. Edger, Houda Chelaifa, David Tack, Gilles Lassalle, Imen Mestiri, Nicolas Schnel, Marie-Christine Le Paslier, Guangyi Fan, Victor Renault, Philipp E. Bayer, Agnieszka A. Golicz, Sahana Manoli, Tae-Ho Lee, Vinh Ha Dinh Thi, Smahane Chalabi, Qiong Hu, Chuchuan Fan, Reece Tollenaere, Yunhai Lu, Christophe Battail, Jinxiong Shen, Christine H. D. Sidebottom, Xinfa Wang, Aurélie Canaguier, Aurélie Chauveau, Aurélie Bérard, Gwenaëlle Deniot, Mei Guan, Zhongsong Liu, Fengming Sun, Yong Pyo Lim, Eric Lyons, Christopher D. Town, Ian Bancroft, Xiaowu Wang, Jinling Meng, Jianxin Ma, J. Chris Pires, Graham J. King, Dominique Brunel, Régine Delourme, Michel Renard, Jean-Marc Aury, Keith L. Adams, Jacqueline Batley, Rod J. Snowdon, Jorg Tost, David Edwards, Yongming Zhou, Wei Hua, Andrew G. Sharpe, Andrew H. Paterson, Chunyun Guan and Patrick Wincker

Science **345** (6199), 950-953.
DOI: 10.1126/science.1253435

The genomic origins of rape oilseed

Many domesticated plants arose through the meeting of multiple genomes through hybridization and genome doubling, known as polyploidy. Chalhoub *et al.* sequenced the polyploid genome of *Brassica napus*, which originated from a recent combination of two distinct genomes approximately 7500 years ago and gave rise to the crops of rape oilseed (canola), kale, and rutabaga. *B. napus* has undergone multiple events affecting differently sized genetic regions where a gene from one progenitor species has been converted to the copy from a second progenitor species. Some of these gene conversion events appear to have been selected by humans as part of the process of domestication and crop improvement.

Science, this issue p. 950

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.