Reconstruction of ancestral diploid karyotype and evolutionary trajectories leading to the formation of \textit{Camelina sativa} chromosomes

\textbf{CURRENT STATUS:} UNDER REVIEW

\textbf{BMC Genomics} \hspace{1cm} \textbf{BMC Series}

Zhikang Zhang  
North China University of Science and Technology

Fanbo Meng  
North China University of Science and Technology

Pengchuan Sun  
North China University of Science and Technology

Jiaqing Yuan  
North China University of Science and Technology

Ke Ge  
North China University of Science and Technology

Chao Liu  
North China University of Science and Technology

Weijie Wang  
North China University of Science and Technology

Xiyin Wang  
wang.xiyin@gmail.com  
North China University of Science and Technology

Corresponding Author  
ORCiD: 0000-0003-3454-0374

\textbf{DOI:} 10.21203/rs.2.20143/v1

\textbf{SUBJECT AREAS}  \hspace{1cm} Epigenetics & Genomics
KEYWORDS

Brassicaceae, C. sativa, Chromosome, Karyotype, Polyploid
Abstract

Background: Belonging to lineage I of Brassicaceae, Camelina sativa is formed by two hybridizations of three species (three sub-genomes). The three sub-genomes were diverged from a common ancestor, likely derived from lineage I (Ancestral Crucifer karyotype, ACK). The karyotype evolutionary trajectories of the C. sativa chromosomes are currently unknown. Here, we managed to adopt a telomere-centric theory proposed previously to explain the karyotype evolution in C. sativa.

Results: By characterizing the homology between A. lyrate and C. sativa chromosomes, we inferred ancestral diploid karyotype of C. sativa (ADK), including 7 ancestral chromosomes, and reconstructed the karyotype evolutionary trajectories leading to the formation of C. sativa genome. The process involved 2 chromosome fusions. We found that sub-genomes Cs-G1 and Cs-G2 may share a closer common ancestor than Cs-G3. Together with other lines of evidence from Arabidopsis, we propose that the Brassicaceae plants, even the eudicots, follow a chromosome fusion mechanism favoring end-end joining of different chromosomes, rather than a mechanism favoring the formation circular chromosomes and nested chromosome fusion preferred by the monocots.

Conclusions: The present work will contribute to understanding the structural and functional innovation of C. sativa chromosomes, providing insight into Brassicaceae karyotype evolution.

Background

Brassicaceae (mustard family) is one of the largest groups in plants, being composed of an approximate 3700 species, classified into 338 genera [1]. It
includes several species of prominent scientific and economic importance, such as model plants (e.g., Arabidopsis thaliana), vegetable producing crops (Brassica rapa, AA genome), and oil producing crops (Brassica napus, AACC genome). According to phylogenetic relationship, Camelineae species (A. thaliana, Arabidopsis lyrata, and Capsella rubella) and Brassiceae species (B. rapa, Brassica nigra, BB genome and Brassica oleracea, CC genome) respectively represent lineage I and lineage II, two of three well-supported lineages among the Brassicaceae [2, 3].

With the rapid increase of Brassicaceae genome assemblies, reconstructing ancestral genome can help understand the evolutionary history of the extant Brassicaceae families and species. Parkin et al. identified 21 common genomic blocks (GBs) between A. thaliana and B. napus by using restriction fragment length polymorphism (RFLP) mapping approach [4]. With genetic maps of A. lyrata and C. rubella, Schranz et al. defined 24 conversed GBs (labelled as A-X) related to ancestral karyotype (AK, n = 8), by referring to the extant A. lyrata and C. rubella karyotypes [5]. Ancestral crucifer karyotype (ACK, n = 8), improved from AK, is recognized as ancestral state of lineage I (Figs. 1 and 4a), based on the fact that most base common number of chromosomes is eight in the family and comparative chromosome painting (CCP) analysis of A. thaliana, A. lyrata, and C. rubella [6]. Besides, they reconstructed Proto-Calepineae karyotype (PCK, n = 7) as ancestral karyotype of 6 Brassicaceae tribes (Calepineae, Conringieae, Noccaeeae, Eutremeae, Isatideae, and Sisymbrieae) (Fig. 1), which involved three translocations and two inversions comparing to AK. While PCK is inherited in three of the six tribes (Calepineae, Conringieae, and Noccaeeae), which belongs to lineage II, the rest of three of the six tribes (Eutremeae, Isatideae, and Sisymbrieae) is characterized by an additional translocation comparing to PCK. The derivative of PCK within the three
tribes is referred as translocation Proto-Calepineae Karyotype (tPCK, n = 7). Cheng et al. provided conclusive evidence to support that tPCK was adapted to be ancestral karyotype of the mesohexaploid B. rapa, the genus Brassica, and the tribe Brassiceae, by comparing three ancestral sub-genomes of Chinese cabbage (B. rapa) with PCK and tPCK [7].

Running through the evolutionary history of plant kingdom, polyploidization continually led to genome doubling/tripling, genome repatterning, and gene loss, characterizing genome instability and fractionation [8–10]. Interestingly, chromosome numbers could be much reduced to a kind of normal range after rounds of polyploidization. For example, it was inferred that the eudicot proto-chromosome number before the major eudicot-common whole-genome triplication is seven [11]. After two extra Brassicaceae-common duplications (BCD) [12], A. thaliana has only five base chromosomes. It was proposed that chromosome number reduction (CNR) was often the result of reciprocal translocations, which combined two chromosomes into a larger one and a smaller one, and the smaller chromosome got lost during meiosis [13]. For example, ACK and PCK shared the same karyotype of five chromosomes (AK1-4,7), AK6/8 and AK5/6/8 in PCK formed by reciprocal translocations between AK5, AK6 and AK8, resulting in chromosome number reduction from eight to seven [6]. However, an alternative explanation considers that the removal of telomeres caused chromosome fusion and chromosome number reduction during the karyotype evolution, and explained the molecular dynamics of karyotype evolution [14]. This is a telomere-centric theory. A chromosome may form a circular form and cross-over may occur near its two telomeres, and the resolution of the cross-over may produce a telomere-free chromosome and a satellite chromosome of two telomeres and little DNA; the telomere-free chromosome may
invade another chromosome and eventually result in the merge of the invading one into the invaded one, referred to nested chromosome fusion (NCF). Alternatively, two chromosomes may cross over near one telomere of each chromosome, resulting in chromosome end-end joining (EEJ) and formation of a satellite chromosome. Besides, reciprocal translocation of arms (RTA) may although occur. The loss of satellite chromosome explains chromosome number reduction. Based on the telomere-centric theory, ancestral karyotype and evolutionary trajectories of chromosomes were reconstructed for Arabidopsis, grasses and legumes [14-16]. Belonging to lineage I of Brassicaceae [3] (Fig. 1), C. sativa (false flax) is a high-quality oilseed crop with several advantages of high production and resistance to drought and diseases for industrial production of biodiesel [17]. It was proposed that C. sativa represented a whole-genome triplication event relative to A. thaliana, and three sub-genomes were defined (Cs-G1, Cs-G2 and Cs-G3) [18, 19] (Fig. 2). The sub-genomes Cs-G1 and Cs-G2 are more closely related to each other than any of the diploids assayed based on phylogenetic relationship, Cs-G3 shows a clear expression level advantage over the other two sub-genomes, and the three sub-genomes have an almost identical Ks distribution of synteny genes with A. thaliana. The three sub-genomes were likely diverged from a common ancestor and the extant C. sativa hexaploidy genome result from a two-stage allopolyploid pathway. The ancestral diploid karyotype of C. sativa (derivative of ACK, dACK) and the three sub-genomes has been reconstructed based on synteny and collinearity between C. sativa and Arabidopsis species [19]. Here, using the theory of telomere-centric genome repatterning, we inferred a different ancestral diploid karyotype of C. sativa (ADK) comparing to previous study and reconstructed the karyotype evolutionary trajectories of the extant C. sativa genome. The present work will contribute to
understanding the structural and functional innovation of chromosomes in C. sativa and Brassicaceae.

Results
Inference of ancestral diploid karyotype of C. sativa
Parsimony-based phylogenomic analysis can help understand karyotype evolution.
To understand the evolutionary trajectories of ADK before divergence of three C. sativa sub-genomes, we analyzed the syntenic conservation and chromosome repatterning between the genomes of the ancestor of lineage I and C. sativa. Here, we take the A. lyrata genome as the reference of ancestral genome of lineage I for the sake of the high similarity between their karyotype. By searching homologous genes between them, we drew homologous gene dot-plots (Figs. 2 and 3), which showed orthologous correspondence between ancestral genomes of lineage I and C. sativa genomes.
In the homologous gene dot-plots of the two genomes, produced directly by using BLASTP hits and further highlighted by using inferred collinear genes, every chromosome in the ancestral genome has three homoeologous chromosomes or groups of homoeologous chromosome regions in C. sativa genome. We found that 5 ACK chromosomes had nearly perfect orthologous correspondence with at least one or more complete chromosomes in C. sativa (Fig. 3a, b, c, d and e), showing that the integrity of each of these 5 chromosomes in ADK (correspondingly defined as ADK chromosomes 1, 2, 5, 6, 7) inherited the chromosome structure of ACK (AK chromosomes 1, 3, 6, 7, 8) without prominent DNA rearrangements. This means that during the formation process of ADK, 5 ADK chromosomes nearly perfectly retained the chromosome structures of the corresponding ACK ones.
Notably, orthologous correspondence between AK2, 4, 5 and Cs4, 16 (Cs-G1) is nearly the same as that between AK2, 4, 5 and Cs6, 7 (Cs-G2) (Fig. 3g and h), indicating that Cs-G1 and Cs-G2 shared two ancestral chromosomes, which majorly formed through RTA and EEJ between AK2, 4, 5. By searching synteny chains between A. lyrate and C. sativa genomes, we further inferred the crossing-over positions between chromosomes (AK4, 5) are respectively between gene AL482377 (Corresponding C. sativa ortholog: Csa16g006880.1) and AL321151 (Csa04g046610.1) in AK4, and that between gene AL486375 (Csa04g046590.1) and AL486377 (Csa16g006870.1) in AK5. Actually, two trajectories could explain the evolution of these chromosomes. One of the complex evolutionary processes could occur as follows: AK2 and AK4 crossed over near one telomere of each of them, resulting in EEJ to produce AK2/4 and formation of a satellite chromosome of two telomeres (and possibly little DNA); then cross-over between AK5 and neo-AK2/4, which experienced one extra translocation and pericentric inversion, resulting in RTA between the two chromosomes to produce AK5/4 (ADK3) and ADK2/4/5 (ADK4) (Fig. 4c). An alternative trajectory could occur as follows: a cross-over between AK4 and AK5 resulted in reciprocal translocation of arms (RTA) to produce AK5/4, forming ADK3, and intermediate AK4/5. Then, AK4/5 and AK2 crossed over near one telomere of each of them, resulting in chromosome end–end joining (EEJ) to produce AK2/4/5 and likely formation of a satellite chromosome by two telomeres (and possibly little DNA). The neo-chromosome AK2/4/5 experienced one extra translocation and pericentric inversion to form ADK4 (Fig. 4d). No matter which trajectory was the actual one, the satellite chromosome likely produced was lost, eventually reducing the chromosome number from 8 in ACK to 7 in ADK.

Orthologous correspondence between AK2, 4, 5 and Cs5, 9 (Cs-G3) (Fig. 3i) is much
different from that between AK2, 4, 5 and Cs4, 16 (Cs-G1) or Cs6, 7 (Cs-G2),
showing that Cs5, 9 has particular structures not shared with the other two sets of
chromosomes (Cs4, 16 and Cs6, 7). It seems that Cs-G3 does not share the two
ancestral chromosomes (ADK3, 4) with Cs-G1 and Cs-G2. However, orthologous
correspondence between Cs4, 16 or Cs6, 7 and Cs5, 9 (Fig. 3j and k), showing that
Cs5, 9 are majorly formed by RTA between ADK3 and ADK 4 (Fig. 5). By searching
synteny chains between A. lirrate and C. sativa genomes, we further inferred the
crossing-over positions between chromosomes (ADK3, 4) are respectively between
gene AL486375 (Corresponding C. sativa ortholog: Csa04g046590.1) and AL321151
(Csa04g046610.1) in ADK3 (where chromosome arms of AK4, 5 combined), and that
between gene AL476152 (Csa09g071500.1) and AL926342 (Csa09g071510.1) in
ADK4. This finding provides a clear evidence to support that the Cs-G3 actually
inherited karyotype structures of the two ancestral chromosomes (ADK3, 4), which
are shared with Cs-G1 and Cs-G2.
Inferring evolutionary trajectories from ADK to extant C. sativa karyotype
Chromosome structure can help understand phylogenomic relationship. In
homologous gene dot-plots, orthologous correspondence between AK7 (ADK6) and
Cs10, 11, 12 (Fig. 3d) suggested that one paracentric inversion is common to Cs10
and Cs 11, from Cs-G1 and Cs-G2, respectively, but not in chromosome Cs12 from
Cs-G3. It suggested that Cs-G1 and Cs-G2 are not directly diverged from ADK, but
share a common ancestor with one paracentric inversion as compared to ADK6.
The formation process of the three sub-genomes and C. sativa genomes can occur
as follows: ancestral diploid of C. sativa differentiate into species A and B firstly,
then species A differentiate into species C and D after one paracentric inversion
occurred in ADK6 (Fig. 5). Cross-over between ADK6 and ADK7 near one telomere of
each chromosome in species C, resulting in chromosome end-end joining (EEJ) to produce ADK6/7 and formation of a satellite chromosome of two telomeres and little DNA. ADK5 in species D experienced one paracentric inversion independently (Fig. 3c and 5). Cross-over between ADK3 and ADK4 in species B, which experienced one translocation, resulting in reciprocal translocation of arms (RTA) to produce ADK3/4 and ADK4/3, which experienced one pericentric inversion (Fig. 3j, k and 5). RTA between ADK5 and ADK7 in species B occurred to produce ADK5/7 and ADK7/5 (Fig. 3j and 5). The crossing-over positions between chromosomes (ADK5, 7) are respectively between gene AL489681 (Corresponding C. sativa ortholog: Csa20g058860.1) and AL351869 (Csa02g002270.1) in ADK5 (the region where the centromere of ADK5 is located), and that between gene AL494932 (Csa20g041660.1) and AL494934 (Csa02g033470.1) in ADK7 (the region where the centromere of ADK7 is located). An initial hybridization event between species C (Cs-G1) and D (Cs-G2), resulting in a tetraploid genome, followed by an additional hybridization event between the tetraploid genome and species B (Cs-G3), eventually forming the extant hexaploid genome of C. sativa [19] (Fig. 5).

During the formation of the karyotype of C. sativa, 14 chromosomes of C. sativa inherited the chromosome structures of ADK ones. While one paracentric inversion occurred in Cs-G2 to produce one new chromosome, two RTAs occurred in Cs-G3 with one translocation and pericentric inversion to produce four new chromosomes. EEJ occurred in Cs-G1 to produce one new chromosome and one satellite chromosome. The loss of the satellite chromosomes resulted in the chromosome number reduction from 21 to 20.

Discussion
The ancestral diploid karyotype of C. sativa, inferred by previous study, only involves chromosome repatterning between AK2 and AK4, but ignores that AK5 also took part in the formation of the ancestral diploid karyotype of C. sativa, which is strongly suggested in homologous gene dot-plots between A. lyrate and C. sativa genomes. Besides, we found clear evidences to support that Cs-G1 and Cs-G2 shared a closer common ancestor, having evolved from an ADK chromosome, than Cs-G3. This explains that Cs-G1 and Cs-G2 were shown to be closely related to each other than Cs-G3 based on phylogenetic relationship from previous study.

Comparing to previous study [19], we further inferred the evolutionary trajectories from ADK to extant C. sativa karyotype, which involved one EEJ and two RTAs. While homologous gene dot-plots are always used to observe the occurrence and scale of polyploidization [12, 20], it is always overlooked that homologous gene dot-plots can also intuitively show the information of chromosome structure and trace of genome repatterning. It reminds us that homologous gene dot-plot is a powerful tool to help understand karyotype evolution and even phylogenetic relationship.

Chromosome number reduction (CNR) in Brassicaceae plants took always the end-end-joining or EEJ mechanism rather than the nested-chromosome fusion or NCF mechanism. NCF and EEJ, which can generate satellite chromosome(s), the loss of which resulted in the CNR. Interestingly, the occurrence of the two mechanisms of CNR always showed an obvious plant family preference. The number of occurrences of the two mechanisms in grass family is summerized as follows: from the common 12 ancestral chromosomes, 7 NCFs and 0 EEJ occurred to produce 5 extant Brachypodium chromosomes, 5 NCFs and 1 EEJ to form wheat chromosomes, 1 NCF and 0 EEJ to form foxtail millet chromosomes, 13 NCFs and 4 EEJs to form maize chromosomes [14]. In summary, there are 23 NCFs and 5 EEJs occurring
independently to form extant grass chromosomes, showing NCFs were significantly more preferred than EEJ (Chisq-test P-value = 0.02395). In contrast, the CNR during the formation of A. thaliana chromosomes from eight ancestral chromosomes involved only three EEJs but not NCF [14]. Similar to A. thaliana, the formation of ADK, and the formation of the extant hexaploid genome of C. sativa, EEJ is the only mechanism that causes CNR. This shows an exclusive preference of EEJ in Brassicaceae. A significant preference of EEJ over NCF was also observed in legumes. Though the sampled families are still too limited, it seems that eudicots prefer EEJ, and monocots prefer NCF, resulting in CNR.

Conclusions

By using the telomere-centric model, we inferred ancestral diploid karyotype of C. sativa (ADK), including 7 ancestral chromosomes, and reconstructed the karyotype evolutionary trajectories leading to the formation of C. sativa genome. The process involved 2 chromosome fusions. Notably, we found that sub-genomes Cs-G1 and Cs-G2 may share a closer common ancestor than Cs-G3. The present work will contribute to understanding the structural and functional innovation of C. sativa chromosomes, providing insight into Brassicaceae karyotype evolution.

Methods

Plant genome data sets
The genome of A. lyrata [21] and C. sativa [19] were downloaded from Brassica database (http://brassicadb.org/brad/).

Dot-plot generation
We used BLASTP [22] to search for homologous pairs (E-value < 1 × 10^-5) between every possible pair of chromosomes in two genomes. The best, second best, and
other matches with E-value > 1e-5 were displayed in different colors, to help distinguish orthology from paralogy, or layers of paralogy as a result of recursive WGD events. Dot-plots were produced using home-made Python scripts.

Inferring positions of breakpoints
Homologous pairs detected by BLASTP were used as input for ColinearScan 1.01 [23], searching syntenic regions between A. lyrata and C. sativa genomes to infer positions of breakpoints caused by RTA. The maximum gap length (mg) was set to be 50 intervening genes between neighboring genes in collinearity on both chromosomes.

Abbreviations
RFLP
restriction fragment length polymorphism
CCP
comparative chromosome painting
AK
ancestral karyotype
ACK
Ancestral crucifer karyotype
PCK
Proto-Calepineae karyotype
tPCK
translocation Proto-Calepineae Karyotype
BCD
Brassicaceae-common duplications
CNR
chromosome number reduction
NCF
nested chromosome fusion
EEJ
end–end joining
RTA
reciprocal translocation of arms
ADK
ancestral diploid karyotype of C. sativa

Declarations

Acknowledgments
We thank the center for genomics and computational biology lab team for
discussion and support

Funding
This work was supported by the China National Science Foundation (3117022 to
XW), the Ministry of Science and Technology of the People’s Republic of China
(2016YFD0101001), Hebei 100 Talented Scholars project, and Tangshan Key
Laboratory Project to XW.

Availability of data and materials
All data generated or analyzed during this study were included in this published
article

Authors’ contributions
XW and WW conceived and led the research. ZZ and FM performed the analysis. PS,
KG and CL contributed to data collection and joined constructive discussions. XW
and ZZ wrote the paper.

Ethics approval and consent to
participate

Not applicable.

Consent for publication

Not applicable

Competing interests

The authors declare no conflict of interest.

Author details

1 School of Life Sciences, North China University of Science and Technology, Tangshan, Hebei 063210, China.

2 Institute for Genomics and Bio-Big-Data, Chengdu University of Traditional Chinese Medicine, Chengdu 610075, China.

3 College Pharmacy, Chengdu University of Traditional Chinese Medicine, Chengdu 610075, China.

References

1. Al-Shehbaz IA, Beilstein MA, Kellogg EA. Systematics and phylogeny of the Brassicaceae (Cruciferae): an overview. Plant Syst Evol. 2006;259(2-4):89-120.
2. Bailey CD, Koch MA, Mayer M, Mummenhoff K, O’Kane SL Jr, Warwick SI, Windham MD, Al-Shehbaz IA. Toward a global phylogeny of the Brassicaceae. Molecular biology evolution. 2006;23(11):2142-60.
3. Koenig D, Weigel D. Beyond the thale: comparative genomics and genetics of Arabidopsis relatives. Nature reviews Genetics. 2015;16(5):285-98.
4. Parkin IA, Gulden SM, Sharpe AG, Lukens L, Trick M, Osborn TC, Lydiate DJ. Segmental structure of the Brassica napus genome based on comparative analysis with Arabidopsis thaliana. Genetics. 2005;171(2):765-81.
5. Schranz ME, Lysak MA, Mitchell-Olds T. The ABC’s of comparative genomics in the Brassicaceae:
building blocks of crucifer genomes. Trends in plant science. 2006;11(11):535-42.

6. Mandakova T, Lysak MA. Chromosomal phylogeny and karyotype evolution in x = 7 crucifer species (Brassicaceae). Plant Cell. 2008;20(10):2559-70.

7. Cheng F, Mandakova T, Wu J, Xie Q, Lysak MA, Wang X. Deciphering the diploid ancestral genome of the Mesohexaploid Brassica rapa. Plant Cell. 2013;25(5):1541-54.

8. Charon C, Bruggeman Q, Thareau V, Henry Y. Gene duplication within the Green Lineage: the case of TEL genes. J Exp Bot. 2012;63(14):5061-77.

9. Kim C, Wang X, Lee TH, Jakob K, Lee GJ, Paterson AH. Comparative Analysis of Miscanthus and Saccharum Reveals a Shared Whole-Genome Duplication but Different Evolutionary Fates. Plant Cell. 2014;26(6):2420-9.

10. Liu S, Liu Y, Yang X, Tong C, Edwards D, Parkin IA, Zhao M, Ma J, Yu J, Huang S, et al. The Brassica oleracea genome reveals the asymmetrical evolution of polyploid genomes. Nature communications. 2014;5:3930.

11. Jaillon O, Aury JM, Noel B, Policriti A, Clepet C, Casagrande A, Choisne N, Aubourg S, Vitulo N, Jubin C, et al. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature. 2007;449(7161):463-7.

12. Bowers JE, Chapman BA, Rong J, Paterson AH. Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. 2003, 422(6930):433-438.

13. Schubert I, Lysak MA. Interpretation of karyotype evolution should consider chromosome structural constraints. Trends Genet. 2011;27(6):207-16.

14. Wang X, Jin D, Wang Z, Guo H, Zhang L, Wang L, Li J, Paterson AH. Telomere-centric genome repatterning determines recurring chromosome number reductions during the evolution of eukaryotes. New Phytol. 2015;205(1):378-89.

15. Wang Z, Wang J, Pan Y, Lei T, Ge W, Wang L, Zhang L, Li Y, Zhao K, Liu T, et al. Reconstruction of evolutionary trajectories of chromosomes unraveled independent genomic repatterning between Triticeae and Brachypodium. BMC Genomics. 2019;20(1):180.

16. Zhuang W, Chen H, Yang M, Wang J, Pandey MK, Zhang C, Chang WC, Zhang L, Zhang X, Tang R, et al. The genome of cultivated peanut provides insight into legume karyotypes, polyploid evolution and crop domestication. Nat Genet. 2019;51(5):865-76.

17. Moser BR. Biodiesel from alternative oilseed feedstocks: camelina and field pennycress. Biofuels. 2012;3(2):193-209.

18. Hutcheon C, Ditt RF, Beilstein M, Comai L, Schroeder J, Goldstein E, Shewmaker CK, Nguyen T, De Rocher J, Kiser J. Polyploid genome of Camelina sativa revealed by isolation of fatty acid synthesis genes. BMC plant biology. 2010;10:233.
19. Kagale S, Koh C, Nixon J, Bollina V, Clarke WE, Tuteja R, Spillane C, Robinson SJ, Links MG, Clarke C, et al. The emerging biofuel crop Camelina sativa retains a highly undifferentiated hexaploid genome structure. Nat Commun. 2014;5:3706.

20. Wang J, Sun P, Li Y, Liu Y, Yang N, Yu J, Ma X, Sun S, Xia R, Liu X, et al. An Overlooked Paleotetraploidization in Cucurbitaceae. Mol Biol Evol. 2018;35(1):16–26.

21. Hu TT, Pattyn P, Bakker EG, Cao J, Cheng J-F, Clark RM, Fahlgren N, Fawcett JA, Grimwood J, Gundlach H, et al. The Arabidopsis lyrata genome sequence and the basis of rapid genome size change. Nat Genet. 2011;43(5):476–81.

22. Wang H, Ooi BC, Tan KL, Ong TH, Zhou L. BLAST++: BLASTing queries in batches. Bioinformatics. 2003;19(17):2323–4.

23. Wang X, Shi X, Li Z, Zhu Q, Kong L, Tang W, Ge S, Luo J. Statistical inference of chromosomal homology based on gene colinearity and applications to Arabidopsis and rice. BMC Bioinform. 2006;7:447.

Figures
Figure 1

Phylogenetic relationship of selected tribes and species within the Brassicaceae.
Figure 2

Dot-plot of homologous genes comparing A. lyrata and C. sativa genomes. C. sativa genome is divided into yellow (Cs-G1), red (Cs-G2), and blue (Cs-G3) according to syntenic regions. Adopted and modified from Kagale et al. (2014).
Figure 3

Homologous dot-plots between selected A. lyrata or C. sativa and C. sativa chrom
Figure 4

Evolutionary trajectories from ACK to ADK. a Ancestral Crucifer Karyotype (ACK).
Figure 5

Evolutionary trajectories from ADK to extant *C. sativa* karyotype.