Supplementary Information for

Post-polyploid chromosomal diploidization in plants is affected by clade divergence and constrained by shared genomic features

Yile Huang, Manuel Poretti, Terezie Mandáková, Milan Pouch, Xinyi Guo, Estela Perez Roman, Manuel B. Crespo, Stefan Grob, Alexandros Bousios, Christian Parisod, and Martin A. Lysak

Supplementary Figures S1 – S22

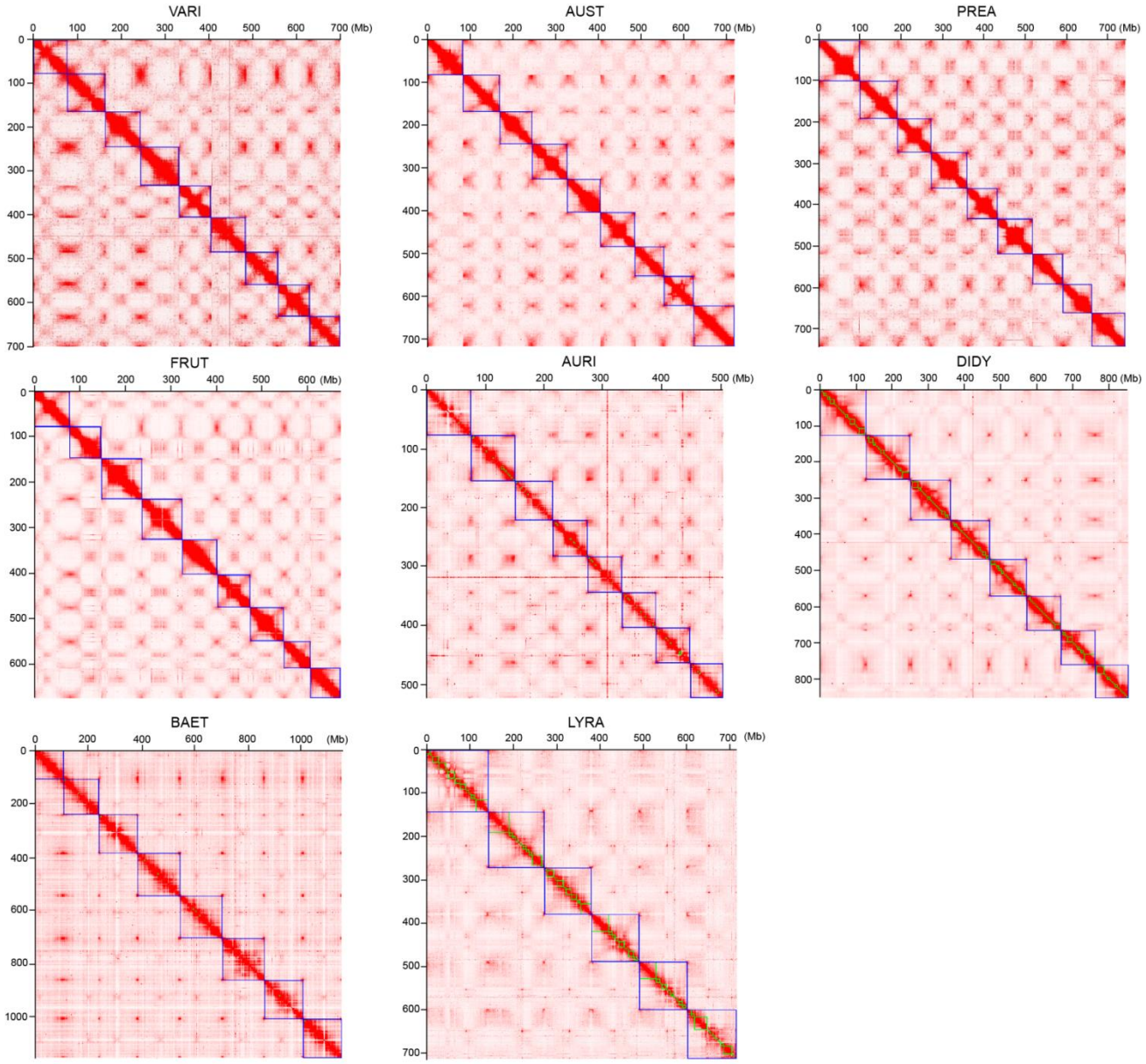


Figure S1. Genome-wide all-by-all Hi-C interactions between 6 to 9 pseudochromosomes in *Biscutella*. The intensity of the red color is proportional to the interaction frequency. The blue boxes indicate the 6 to 9 pseudochromosomes.

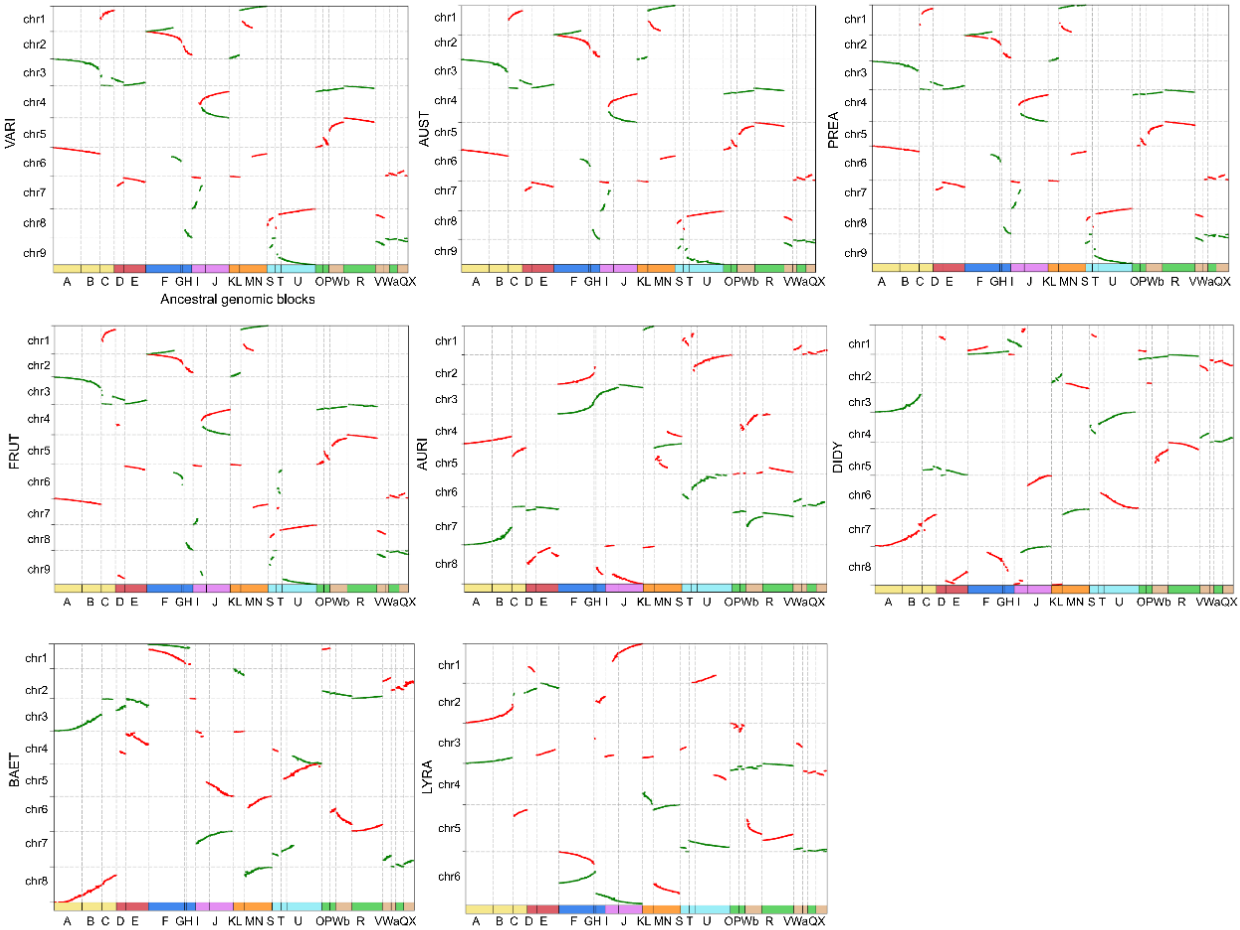


Figure S2. Syntenic relationships between 6 to 9 chromosomes of *Biscutella* genomes and 22 ancestral genomic blocks (GBs A – X; Lysak et al., 2016). Two genomic copies are indicated by red and green, respectively.

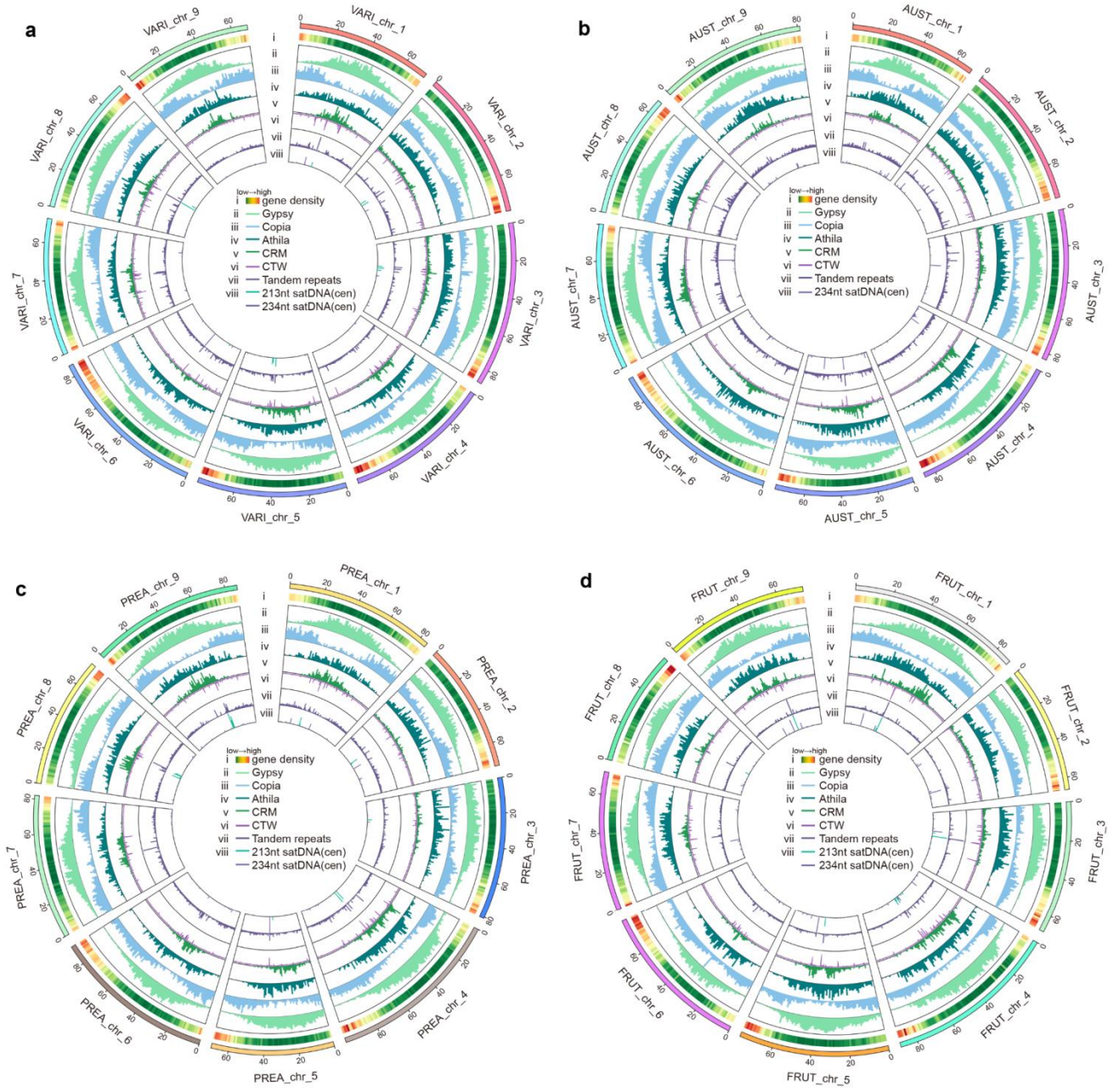


Figure S3. Distribution of genes and repeats in *B. varia* (**a**), *B. austriaca* (**b**), *B. prealpina* (**c**) and *B. frutescens* (**d**). The circles from outside to inside show (i) the gene density; (ii) the density of Gypsy retrotransposons; (iii) the density of Copia retrotransposons; (iv) the density of Athila retrotransposons; (v) the density of CRM retrotransposons; (vi) DNA compression (CTW); (vii) the density of all tandem repeats; (viii) the density of 213 and 234-bp centromeric (cen) tandem repeats.

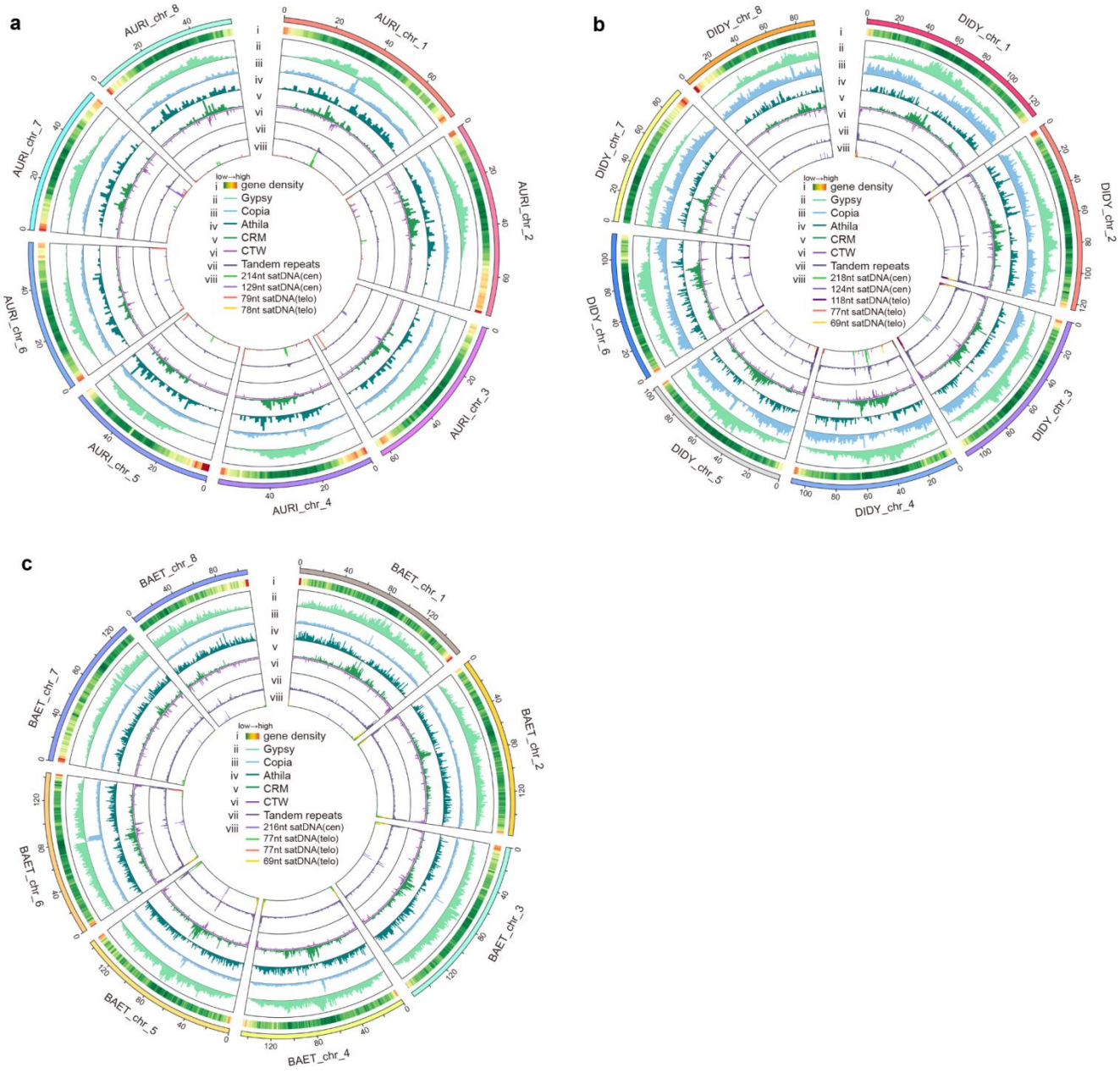


Figure S4. Distribution of genes and repeats in *B. auriculata* (**a**), *B. didyma* (**b**), and *B. baetica* (**c**). The circles from outside to inside show (i) the gene density; (ii) the density of Gypsy retrotransposons; (iii) the density of Copia retrotransposons; (iv) the density of Athila retrotransposons; (v) the density of CRM retrotransposons; (vi) DNA compression (CTW); (vii) the density of all tandem repeats; (viii) the density of the identified centromeric (cen) and subtelomeric (telo) tandem repeats.

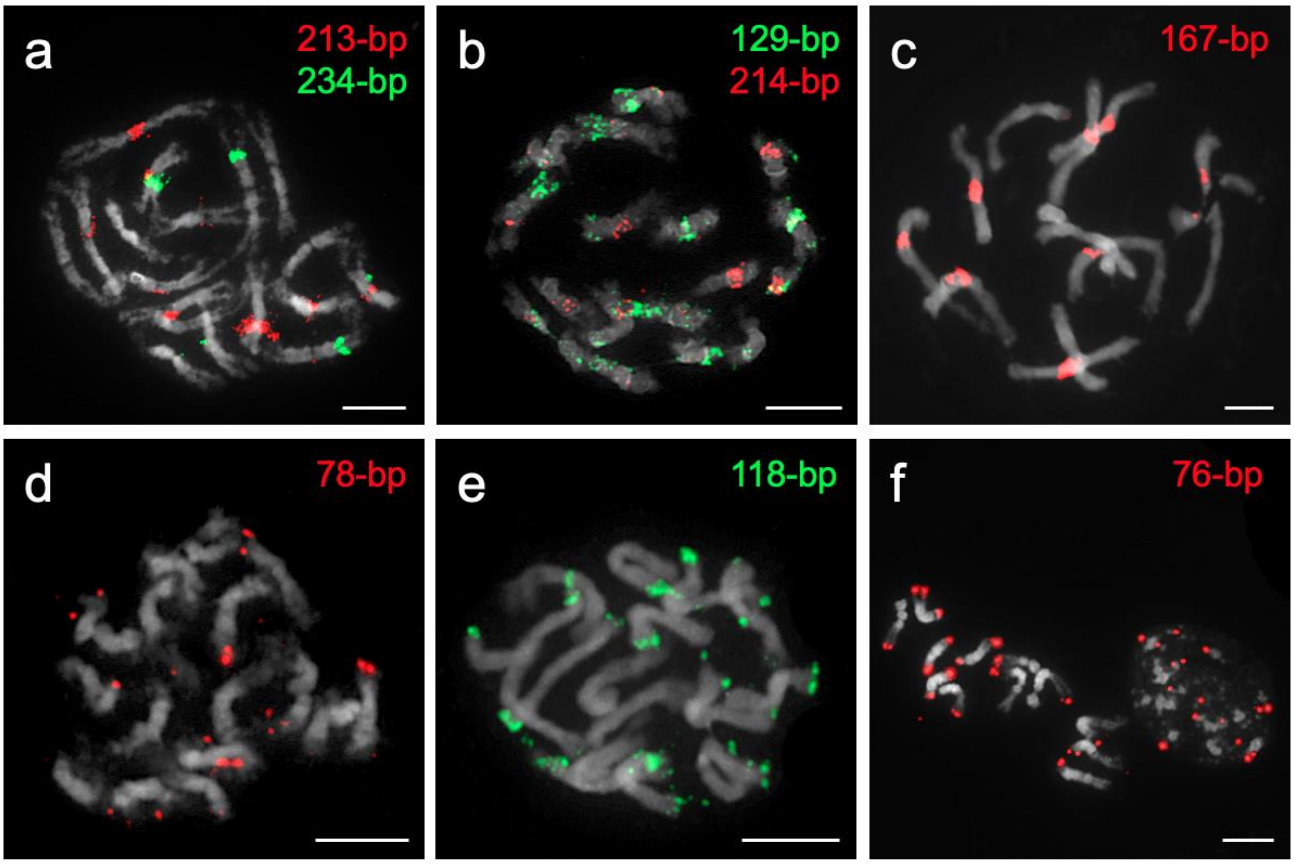


Figure S5. FISH localization of the selected tandem repeats and on mitotic metaphase chromosomes. (**a**) 213-bp and 234-bp centromeric satDNAs in *B. varia*. (**b**) 129-bp and 214-bp centromeric satDNAs in *B. auriculata*. (**c**) 167-bp centromeric satDNAs in *B. lyrata*. (**d**) 78-bp subtelomeric satDNAs in *B. auriculata*. (**e**) 118-bp subtelomeric satDNAs in *B. didyma*. (**f**) 76-bp subtelomeric satDNAs in *B. lyrata*. Chromosomes were counterstained by DAPI; FISH signals are shown in color as indicated. Scale bars, 10 μm.

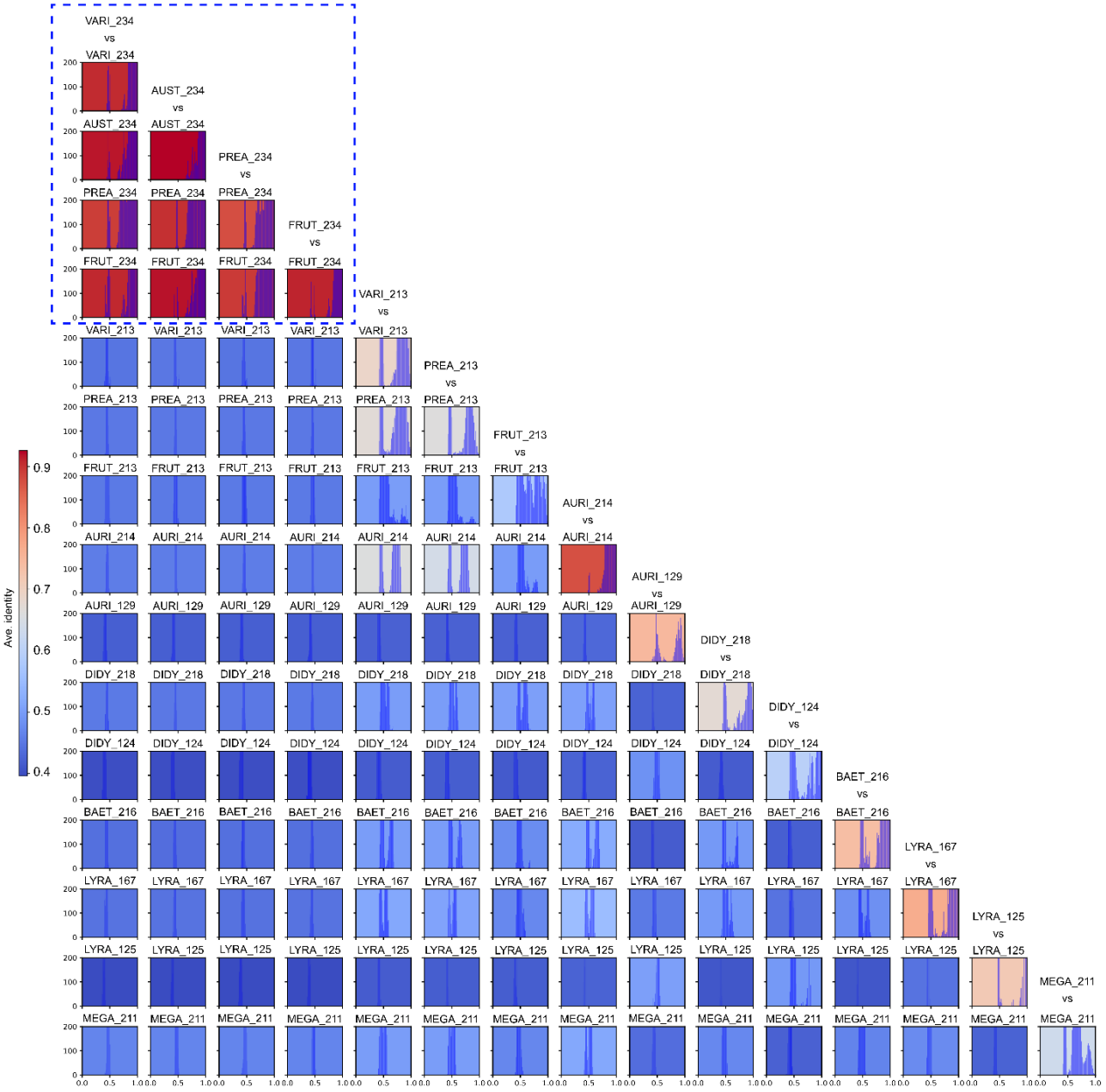


Figure S6. Pairwise comparison of satDNAs of different lengths in pericentromeric regions. Shown is a pairwise sequence identity comparison of 15 satDNA variants identified in *Biscutella* and *Megadenia pygmaea*. Frequency distribution histograms illustrate the sequence identity distribution between each pair of variants, with average identity values indicated by different background colors. The 234-bp satDNA in $n = 9$ species shows high similarity and conservation, indicated by blue dashed boxes.

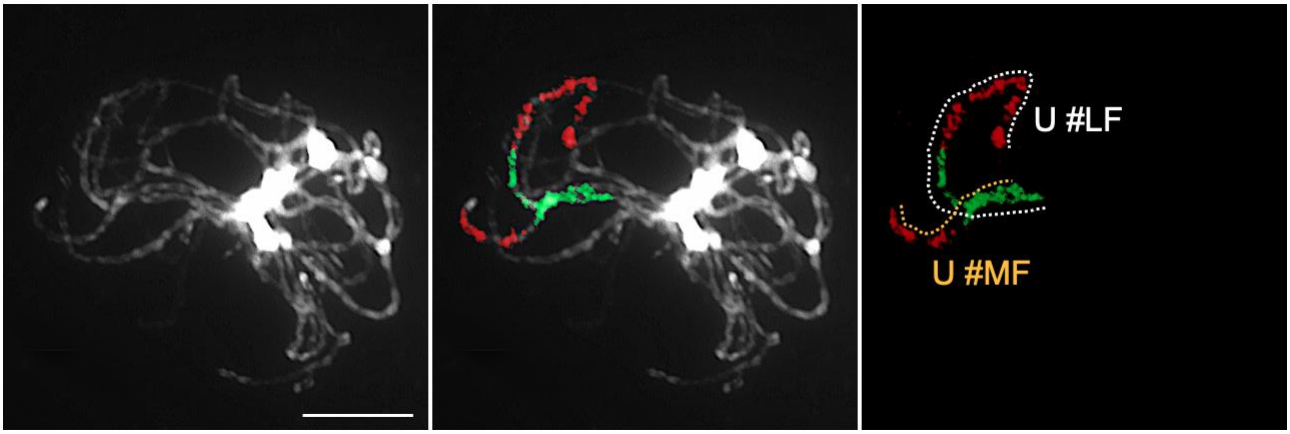


Figure S7. Cytogenetic evidence for biased subgenome fractionation during post-polyploid diploidization in *B. didyma*. Comparative chromosome painting of ancestral genomic block U using *Arabidopsis thaliana* BAC clones in pachytene complements divided into two differently labeled parts revealed two paralogous copies differing in length and fluorescence intensity. The longer and brighter copy corresponds to the U block in the LF subgenome, the shorter and weaker copy corresponds to the U block in the MF subgenome. The chromosomes were counterstained with DAPI. Scale bar, 10 μm.

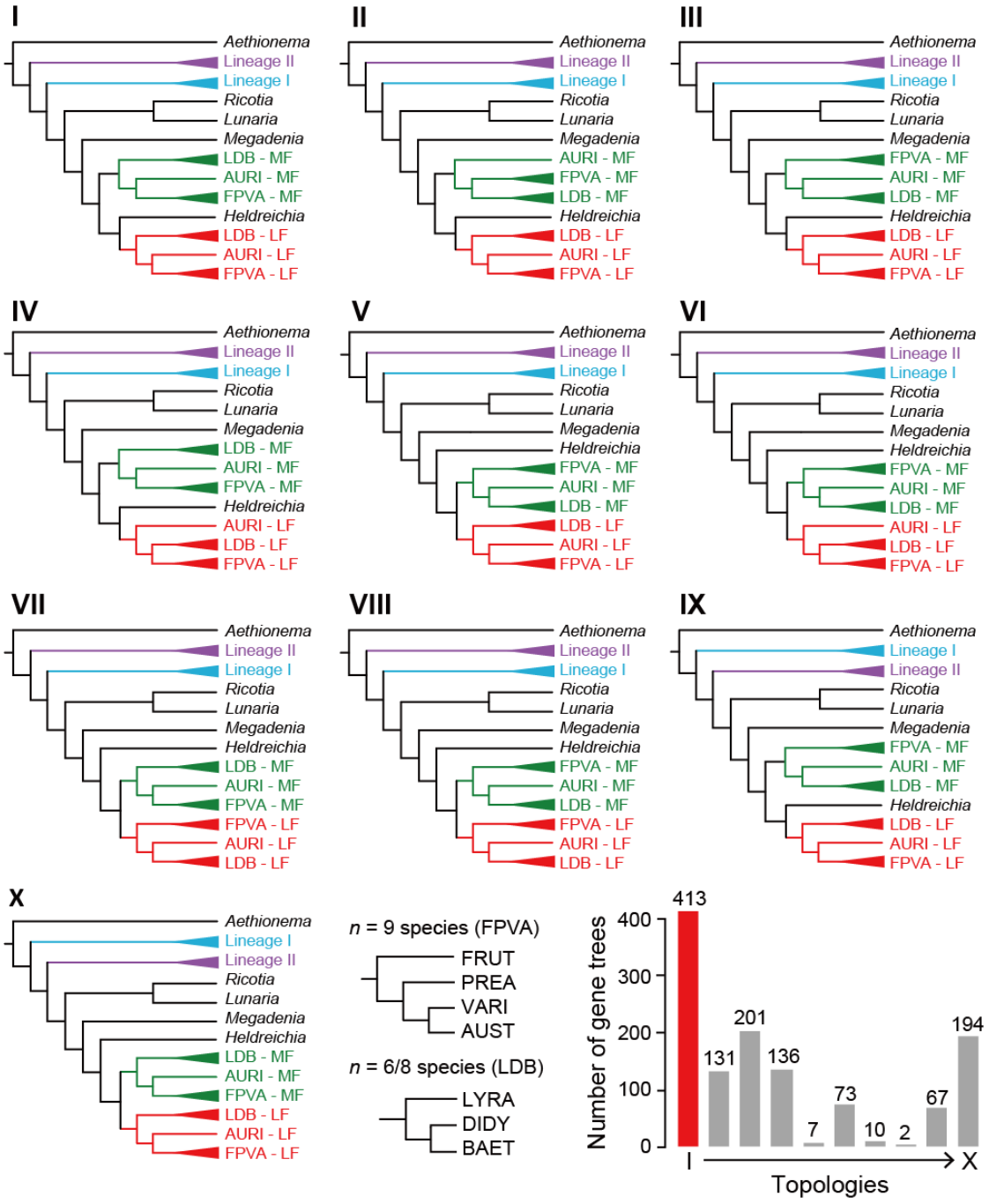


Figure S8. Alternative tree topologies (I – X) observed in 1,234 syntenic gene trees. The two subgenomes (LF and MF, marked red and green, respectively) consistently formed two distinct clades. The *n* = 9 *Biscutella* species (*B. frutescens*, *B. prealpina*, *B. varia*, and *B. austriaca*, abbreviated as FPVA) and *B. lyrata*, *B. didyma*, and *B. baetica* (abbreviated as LDB) always appeared on separate clades, whereas *B. auriculata* is either sister to FPVA or LDB, or outside of both. The graph in the bottom right corner shows the number of gene trees supporting each topology - topology I was supported by the highest number of gene trees.

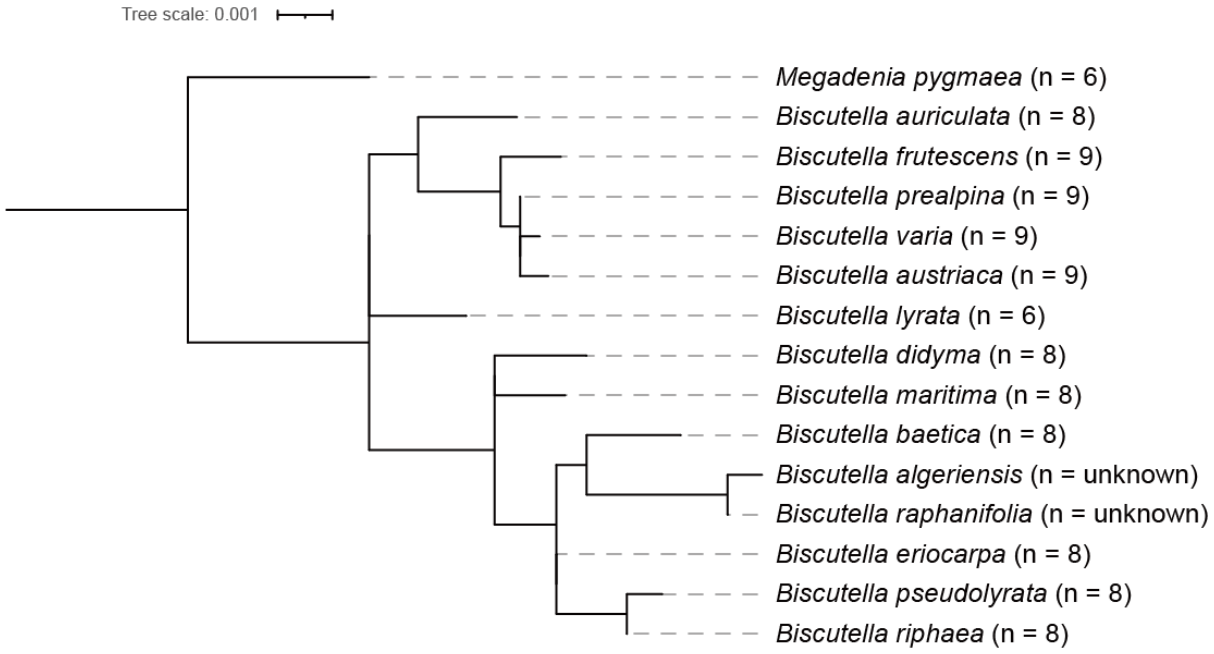


Figure S9. A chloroplast gene (rpl32-trnV) tree constructed based on 1,430 nucleotide sites. The sequences of eight species in this study were obtained from chloroplast genome assemblies (**Methods**), and the sequences of the remaining species were obtained from Vicente et al. (2020).

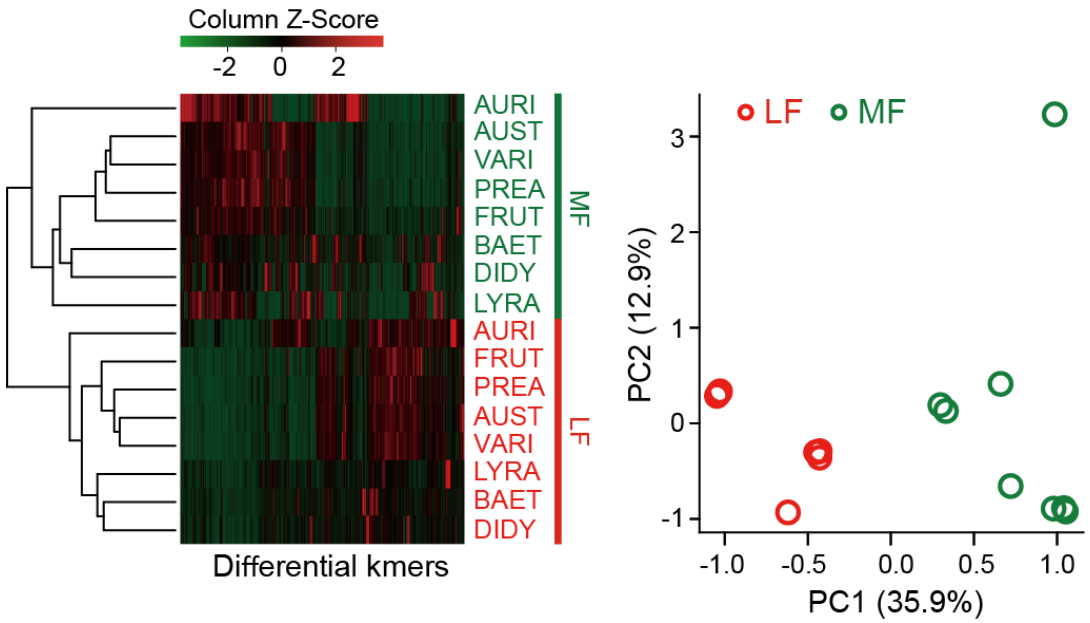
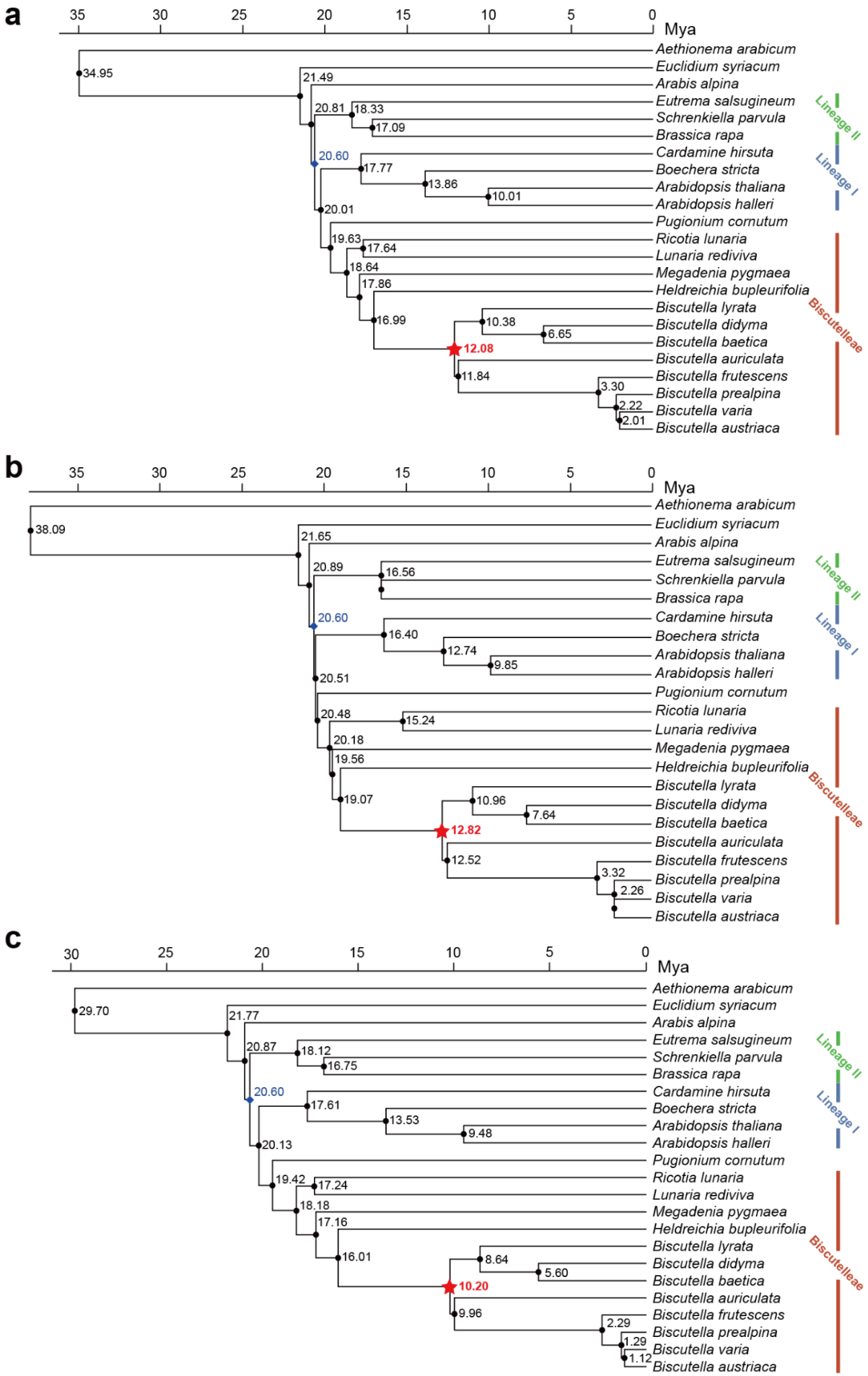
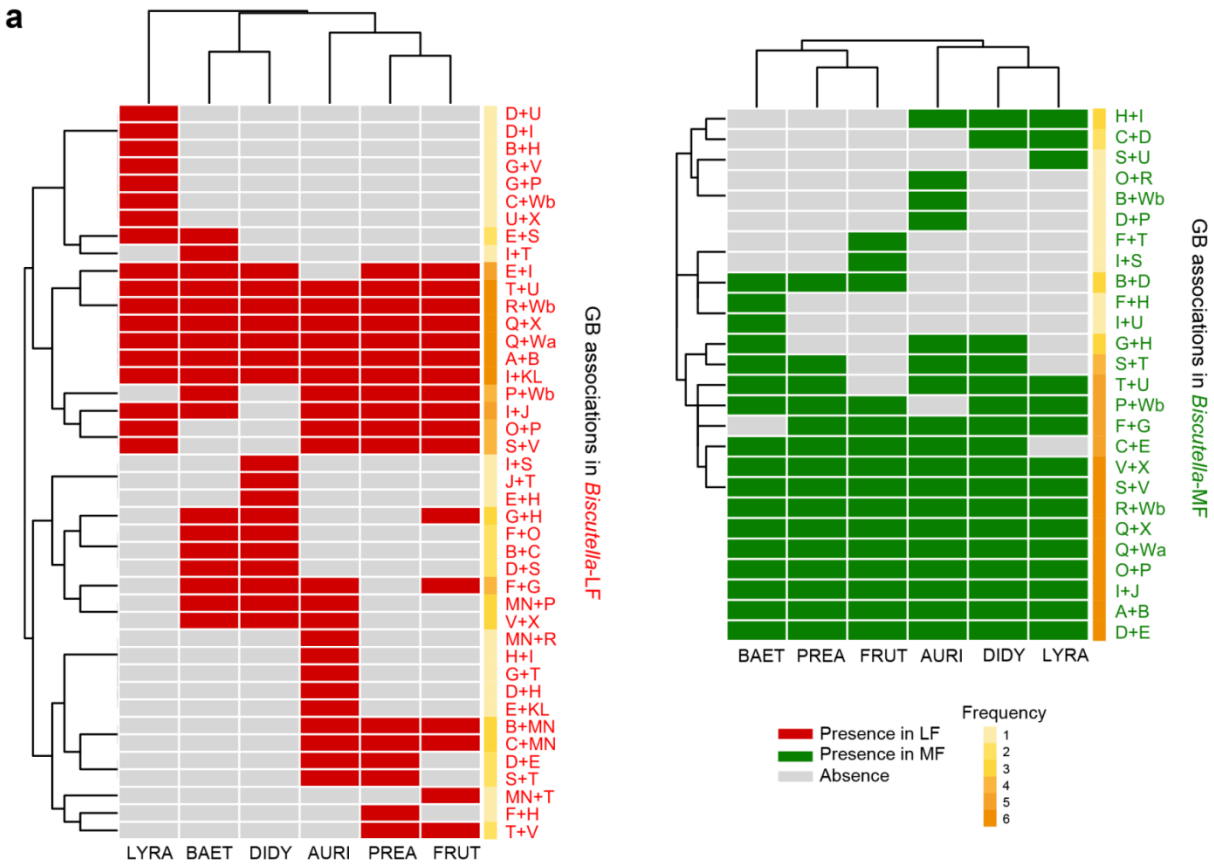


Figure S10. Hierarchical clustering of 15-mers based on the relative abundance (Z-score; left). The PCA of this data (right) shows distinct clustering of LF and MF subgenomes.


a
35
30
25
20
15
10
5
0
Mya
Aethionema arabicum
Euclidium syriacum
Arabis alpina
Eutrema salsugineum
Schrenkiella parvula
Brassica rapa
Cardamine hirsuta
Boechera stricta
Arabidopsis thaliana
Arabidopsis halleri
Pugionium cornutum
Ricotia lunaria
Lunaria rediviva
Megadenia pygmaea
Heldreichia bupleurifolia
Biscutella lyrata
Biscutella didyma
Biscutella baetica
Biscutella auriculata
Biscutella frutescens
Biscutella prealpina
Biscutella varia
Biscutella austriaca
34.95
21.49
20.81
18.33
17.09
20.60
17.77
13.86
10.01
20.01
19.63
17.64
18.64
17.86
16.99
10.38
6.65
12.08
11.84
3.30
2.22
2.01
Lineage II
Lineage I
Biscutelleae
b
35
30
25
20
15
10
5
0
Mya
Aethionema arabicum
Euclidium syriacum
Arabis alpina
Eutrema salsugineum
Schrenkiella parvula
Brassica rapa
Cardamine hirsuta
Boechera stricta
Arabidopsis thaliana
Arabidopsis halleri
Pugionium cornutum
Ricotia lunaria
Lunaria rediviva
Megadenia pygmaea
Heldreichia bupleurifolia
Biscutella lyrata
Biscutella didyma
Biscutella baetica
Biscutella auriculata
Biscutella frutescens
Biscutella prealpina
Biscutella varia
Biscutella austriaca
38.09
21.65
20.89
16.56
20.60
16.40
12.74
9.85
20.51
20.48
15.24
20.18
19.56
19.07
10.96
7.64
12.82
12.52
3.32
2.26
Lineage II
Lineage I
Biscutelleae
c
30
25
20
15
10
5
0
Mya
Aethionema arabicum
Euclidium syriacum
Arabis alpina
Eutrema salsugineum
Schrenkiella parvula
Brassica rapa
Cardamine hirsuta
Boechera stricta
Arabidopsis thaliana
Arabidopsis halleri
Pugionium cornutum
Ricotia lunaria
Lunaria rediviva
Megadenia pygmaea
Heldreichia bupleurifolia
Biscutella lyrata
Biscutella didyma
Biscutella baetica
Biscutella auriculata
Biscutella frutescens
Biscutella prealpina
Biscutella varia
Biscutella austriaca
29.70
21.77
20.87
18.12
16.75
20.60
17.61
13.53
9.48
20.13
19.42
17.24
18.18
17.16
16.01
8.64
5.60
10.20
9.96
2.29
1.29
1.12
Lineage II
Lineage I
Biscutelleae

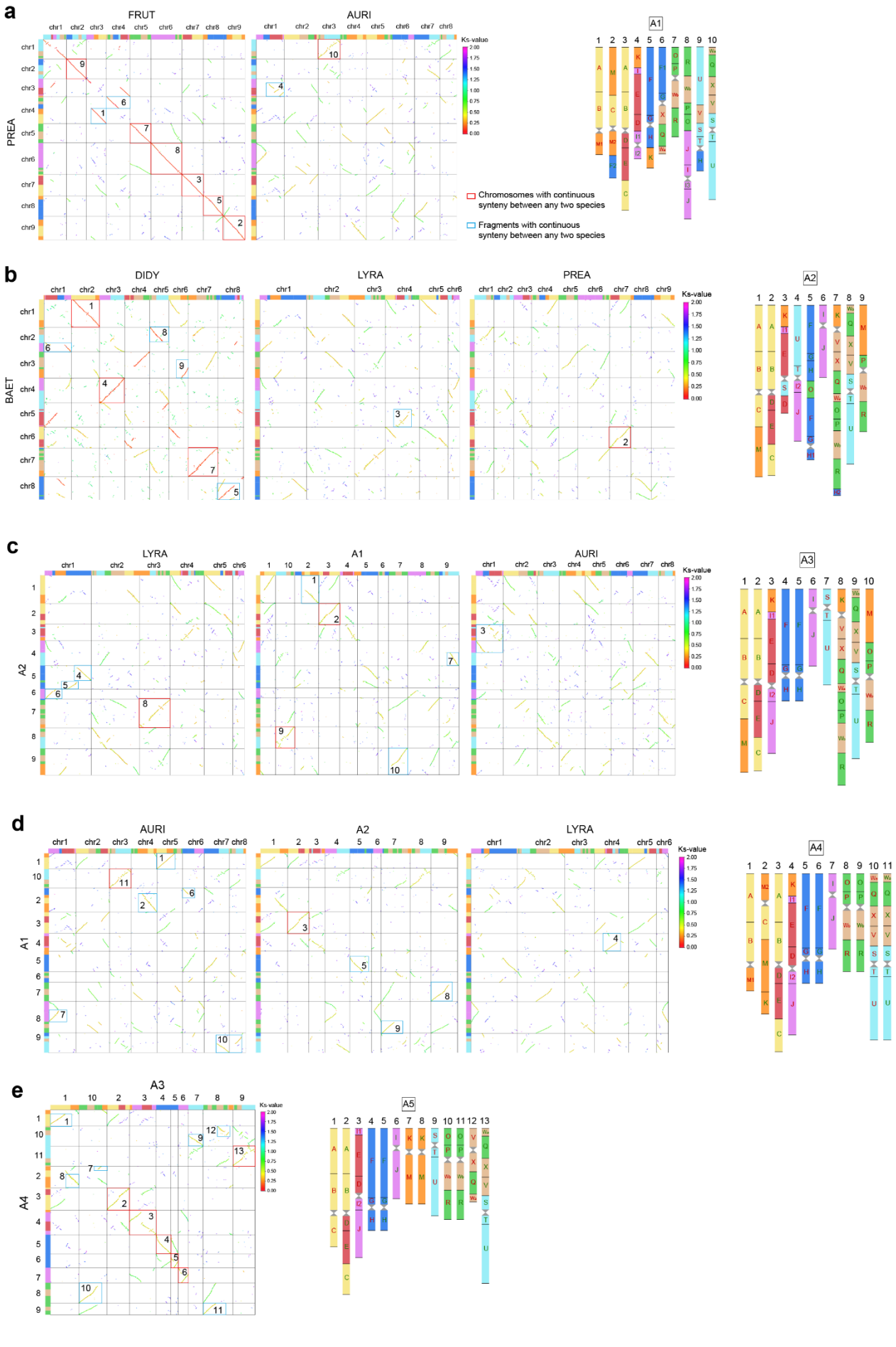
1

Figure S11. Time-calibrated trees inferred by r8s (**a**), PATHd8 (**b**), and RelTime (**c**) methods. The three different calculating strategies showed high consistency. The divergence time of Lineage I and Lineage II was used as the calibration node (blue dot at 20.6 Mya, http://www.timetree.org/). The crown node of the *Biscutella* species was labeled as red stars.


a
D+U
D+I
B+H
G+V
G+P
C+Wb
U+X
E+S
I+T
E+I
T+U
R+Wb
Q+X
Q+Wa
A+B
I+KL
P+Wb
I+J
O+P
S+V
I+S
J+T
E+H
G+H
F+O
B+C
D+S
F+G
MN+P
V+X
MN+R
H+I
G+T
D+H
E+KL
B+MN
C+MN
D+E
S+T
MN+T
F+H
T+V
LYRA BAET DIDY AURI PREA FRUT
GB associations in Biscutella-LF
H+I
C+D
S+U
O+R
B+Wb
D+P
F+T
I+S
B+D
F+H
I+U
G+H
S+T
T+U
P+Wb
F+G
C+E
V+X
S+V
R+Wb
Q+X
Q+Wa
O+P
I+J
A+B
D+E
BAET PREA FRUT AURI DIDY LYRA
GB associations in Biscutella-MF
Presence in LF
Presence in MF
Absence
Frequency
1
2
3
4
5
6


b
GB associations in
LF
33
9
(64.3%
of ACK)
ACK
5
GB associations in
LF
31
11
(68.8%
of PCK)
PCK
5
GB associations in
LF
31
11
(73.3%
of ancPCK)
ancPCK
4
GB associations in
MF
18
8
(57.1%
of ACK)
ACK
6
GB associations in
MF
16
10
(62.5%
of PCK)
PCK
6
GB associations in
MF
16
10
(66.7%
of ancPCK)
ancPCK
5

Figure S12. Analysis of GB associations in *Biscutella* karyotypes. (**a**) Presence and absence of pairwise neighboring GB associations within the LF and MF subgenomes. The present GB associations are marked in red (LF) and green (MF). The missing GB associations are marked in grey. The pheatmap R package (v1.0.12; Kolde R. & Kolde M. R., 2015) was used for hierarchical clustering. (**b**) Shared and unique GB associations in LF and MF subgenomes compared to the three known ancestral Brassicaceae karyotypes (ACK, PCK, and ancPCK; Mandáková & Lysak, 2008; Lysak et al., 2016; Geiser et al., 2016).


a
FRUT
AURI
PREA
Ks-value
Chromosomes with continuous synteny between any two species
Fragments with continuous synteny between any two species
A1
b
DIDY
LYRA
PREA
BAET
A2
c
LYRA
A1
AURI
A2
A3
d
AURI
A2
LYRA
A1
A4
e
A3
A4
A5

1

Figure S13. Bottom-up strategy for inferring ancestral karyotypes at each phylogenetic node (A1 – A5, from the youngest to the oldest). (**a**) Inference of A1, the common ancestor of *n* = 9 species. Based on synteny maps comparing *B. prealpina* (Y-axis) with *B. frutescens* and *B. auriculata* (X-axis), intact chromosomes that maintain collinear relationships between any two species are assumed to have originated from A1 (red box; chromosomes 2, 3, 5, 7, 8, 9, and 10 in A1). The GB A+B+MN, KL+I+E+D, and F+G+X+Q+Wa associations are observed in any two species among *B. prealpina*, *B. frutescens*, and *B. auriculata*, suggesting that these associations also originated from A1 (blue box; corresponding to chromosomes 1, 4, and 6 in A1). Therefore, we inferred that A1 had 10 chromosome pairs. (**b**) Inference of A2, the common ancestor of *B. baetica-B.didyma* clade. Based on synteny maps comparing *B. baetica* (Y-axis) with *B. didyma*, *B. lyrata*, and *B. prealpina* (X-axis), intact chromosomes that maintain collinear relationships between any two species are assumed to have originated from A2 (red box; chromosomes 1, 2, 4, and 7 in A2). The S+D, F+G+H+O+F+G+H, I+J, Wa+Q+X+V+S+T+U, and MN+P associations are observed in any two species among *B. baetica*, *B.didyma*, and *B. lyrata* (blue box; corresponding to chromosomes 3, 5, 6, 8, and 9 in A2). Therefore, we inferred that A2 had 9 chromosome pairs. (**c**) Inference of A3, the common ancestor of clade II in **Figure 3b**. Based on synteny maps comparing A2 (Y-axis) with *B. lyrata*, A1, and *B. auriculata* (X-axis), intact chromosomes that maintain collinear relationships between any two species are assumed to have originated from A3 (red box; chromosomes 2, 8, and 9 in A3). The A+B+C+MN, I1+E+D+I2+J (I1, I2, ..., Ix referring to the broken I GB, other GBs with the same abbreviation), F+G+H (in both LF and MF), I+J, S+T+U, and O+P+Wb+R associations are observed in any two species among A2, *B. lyrata*, A1, and *B. auriculata* (blue box; corresponding to chromosomes 1, 3, 4, 5, 6, 7, and 10 in A2). Therefore, we inferred that A3 had 10 chromosome pairs. (**d**) Inference of A4, the common ancestor of clade I in **Figure 3b**. Based on synteny maps comparing A1 (Y-axis) with *B. auriculata*, A2, and *B. lyrata* (X-axis), intact chromosomes that maintain collinear relationships between any two species are assumed to have originated from A4 (red box; chromosomes 3 and 11 in A4). The B+MN1, C+MN2, D+I+J, F+G+H (in both LF and MF), I+J, O+P+Wb+R (in both LF and MF), and Wa+Q+X+V+S+T+U associations are observed in any two species among A1, *B. auriculata*, A2, and *B. lyrata* (blue box; corresponding to chromosomes 1, 2, 4, 5, 6, 7, 8, 9, and 10 in A2). Therefore, we inferred that A4 had 11 chromosome pairs. (**e**) Inference of A5. Based on synteny maps comparing A3 (Y-axis) with A4 (X-axis), intact chromosomes that maintain collinear relationships are assumed to have originated from A5 (red box; chromosomes 2, 3, 4, 5, 6, and 13 in A5). The A+B+C, KL+MN (in both LF and MF), S+T+U, O+P+Wb+R (in both LF and MF), and V+X+Q+Wa associations are inferred as seven independent chromosomes. Therefore, we inferred that A5 had 13 chromosome pairs. The GB association I1+E+D+I2+J in chromosome 3 of A5 suggests an insertion

event (E+D inserted in the middle of the I block), but this event was not observed in *Heldreichia*, a species closely related to the LF subgenome, and thus this insertion should be specific to *Biscutella* and originate at an early stage of *Biscutella* speciation.

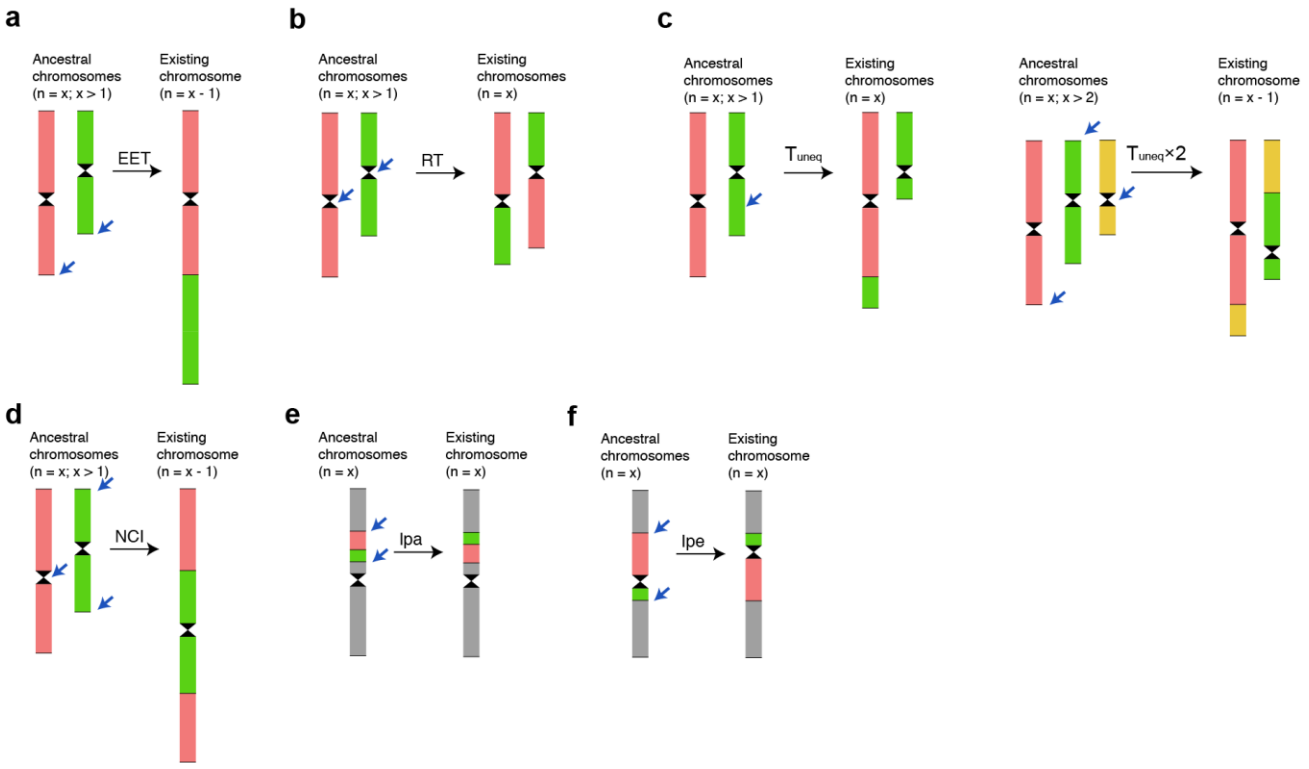


Figure S14. Six main chromosome rearrangements (CR) type scenarios, adopted from Mandáková & Lysak (2018). (**a**) End-to-end translocation (EET) between two non-homologous chromosomes is mediated by two DNA double-strand breaks (blue arrow) at chromosome ends and results in the formation of a dicentric chromosome. One of the two centromeres has to become inactive and/or be removed. (**b**) Reciprocal translocation (RT) involves the breakage of two non-homologous chromosomes. The broken segments are exchanged, forming derivative chromosomes. (**c**) Unequal reciprocal translocation (T_{uneq}) happens when chromosomes break and the pieces attach to other chromosomes (left). If all pieces are attached to other chromosomes (right) and the centromere to become inactive or be removed, the number of chromosomes will be reduced. (**d**) Nested chromosome insertion (NCI) involves the translocation of an "insertion" chromosome into a "recipient" chromosome. (**e**) Paracentric inversion (Ipa) occurs within a chromosome arm, without involving the centromere. (**f**) Pericentric inversion (Ipe) involves the inversion of a chromosome segment that includes the centromere.

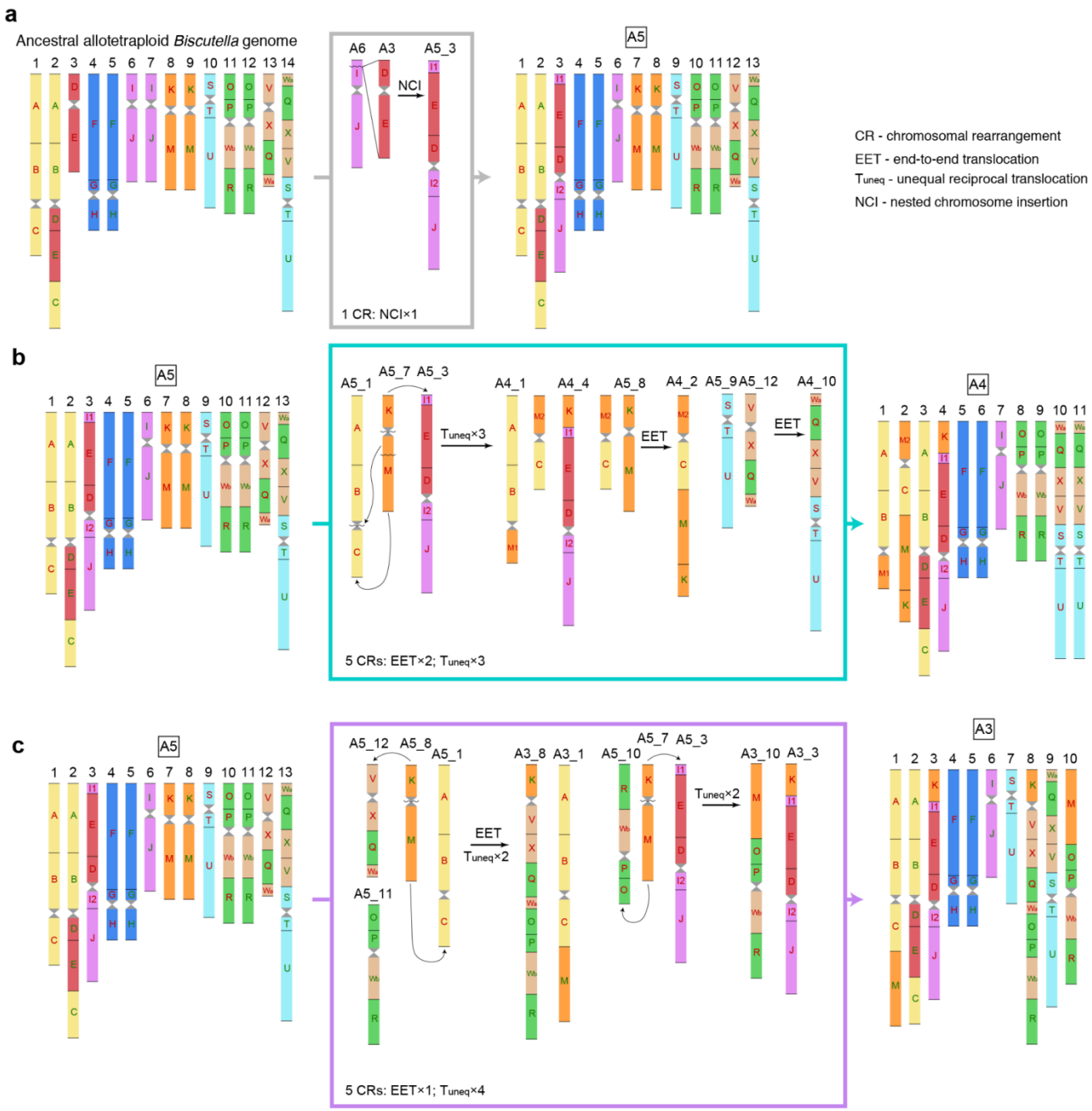

a
Ancestral allotetraploid *Biscutella* genome
A6 A3 A5_3
NCI
1 CR: NCI×1
A5
CR - chromosomal rearrangement
EET - end-to-end translocation
Tuneq - unequal reciprocal translocation
NCI - nested chromosome insertion
b
A5
A5_1 A5_7 A5_3
Tuneq×3
A4_1 A4_4 A5_8 A4_2 A5_9 A5_12 A4_10
EET
EET
5 CRs: EET×2; Tuneq×3
A4
c
A5
A5_12 A5_8 A5_1
A5_11
EET
Tuneq×2
A3_8 A3_1 A5_10 A5_7 A5_3
Tuneq×2
A3_10 A3_3
5 CRs: EET×1; Tuneq×4
A3

Figure S15. Evolutionary pathways from ancestral allotetraploid *Biscutella* genome (n = 14) to A4 and A3 nodes. (**a**) Evolutionary pathway from ancestral *Biscutella* genome to A5 (n = 14 → n = 13). The ancestral chromosomes 3 (GB D+E; LF) and 6 (I+J; LF) underwent a nested chromosome insertion (NCI) event to form the chromosome 3 (I1+E+D+I2+J; I1, I2, ..., Ix referring to the broken I GB, other GBs with the same abbreviation) of A5. (**b**) Evolutionary pathway from A5 to A4 (n = 13 → n = 11). The chromosomes 1 (A+B+MN1; LF) and 2 (MN2+C+MN+KL; LF+MF), 4 (KL+I1+E+D+I2+J; LF) in A4 originated from chromosomes 1 (A+B+C; LF), 3 (I1+E+D+I2+J; LF), 7 (KL+MN; LF), and 8 (KL+MN; MF) of A5 genome after three unequal reciprocal translocations (Ts_{uneq}) and one end-to-end translocation (EET). The chromosome 10 (Wa+Q+X+V+S+T+U; LF) originated from chromosomes 9 (S+T+U; LF) and 12 (V+X+Q+Wa; LF) of A5 genome after an EET event. Therefore, the A5 genome underwent five CRs to form A4. (**c**) Evolutionary pathway from A5 to A3 (n = 13 → n = 10). The chromosomes 1 (A+B+C+MN; LF+MF) and 8 (KL+V+X+Q+Wa+O+P+Wb+R; LF+MF) in A3 originated from chromosomes 1 (A+B+C; LF), 8 (KL+MN; MF), 11 (O+P+Wb+R; MF), and 12 (V+X+Q+Wa; LF) of A5 genome after two Ts_{uneq} and one EET. The chromosomes 3 (KL+I1+E+D+I2+J; LF) and 10 (MN+O+P+Wb+R; LF) in A3 originated from chromosomes 3 (I1+E+D+I2+J; LF), 7 (KL+MF; LF), and 10 (R+Wb+P+O; LF) after two Ts_{uneq}. Therefore, the A5 genome also underwent 5 CRs to form A3, while being distinct from the A5 → A4 pathway in (**b**).

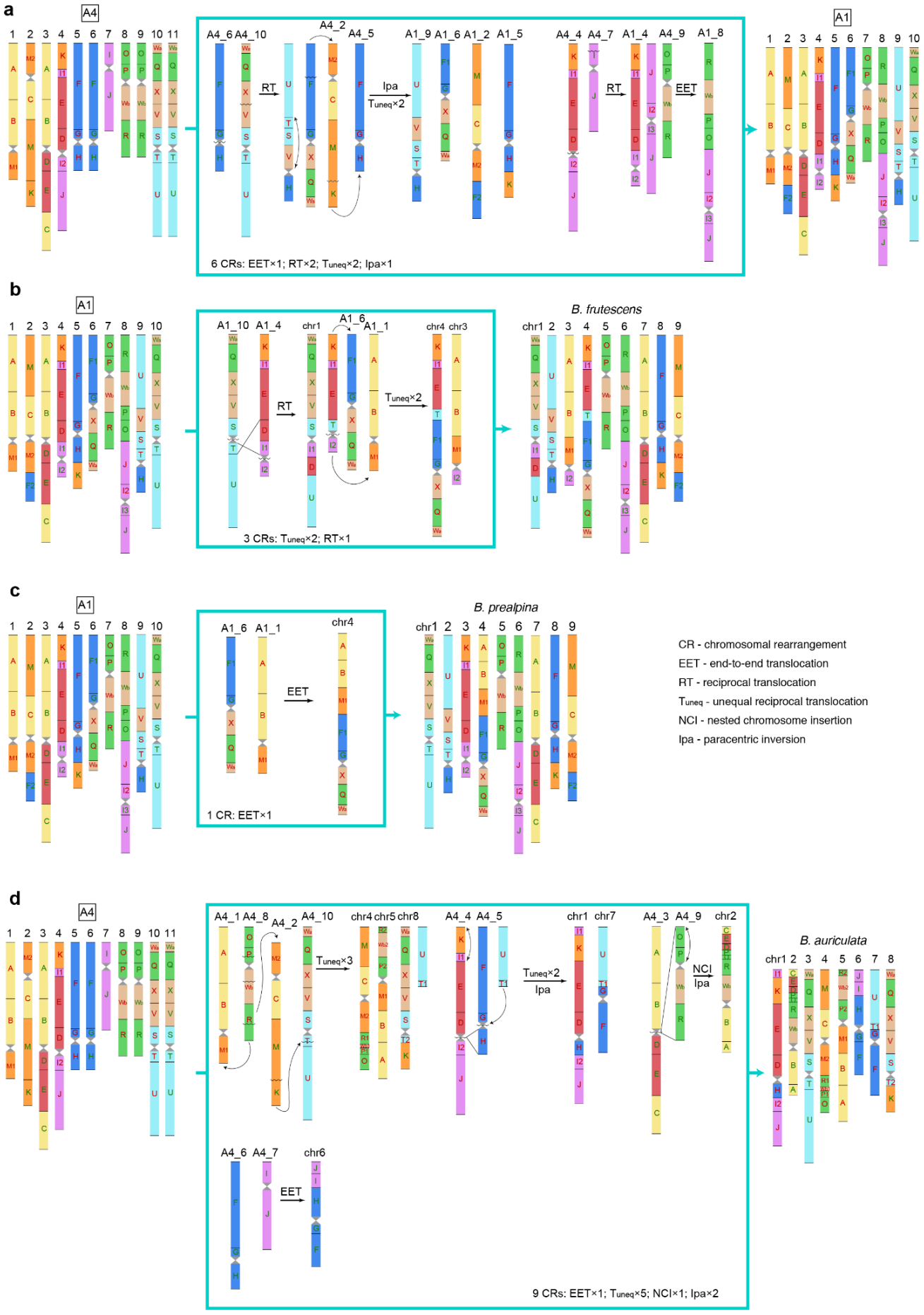

a
A4
A4_6 A4_10
A4_2
A4_5
RT
Ipa
Tuneq×2
A1_9 A1_6 A1_2 A1_5
A4_4 A4_7 A1_4 A4_9 A1_8
RT
EET
A1
6 CRs: EET×1; RT×2; Tuneq×2; Ipa×1
b
A1
A1_10 A1_4
chr1
A1_6
A1_1
chr4 chr3
RT
Tuneq×2
3 CRs: Tuneq×2; RT×1
B. frutescens
chr1
c
A1
A1_6 A1_1
chr4
EET
1 CR: EET×1
B. prealpina
chr1
CR - chromosomal rearrangement
EET - end-to-end translocation
RT - reciprocal translocation
Tuneq - unequal reciprocal translocation
NCI - nested chromosome insertion
Ipa - paracentric inversion
d
A4
A4_1 A4_8
A4_2
A4_10
Tuneq×3
chr4 chr5 chr8
A4_4 A4_5
Tuneq×2
Ipa
chr1 chr7
A4_3 A4_9
chr2
NCI
Ipa
A4_6 A4_7
chr6
EET
9 CRs: EET×1; Tuneq×5; NCI×1; Ipa×2
B. auriculata
chr1

Figure S16. Evolutionary pathways from A4 (n = 11) to n = 9 species and *B. auriculata*. (**a**) Evolutionary pathway from A4 to A1 (n = 11 → n = 10). The chromosomes 2 (MN+C+MN2+F2; LF+MF), 5 (F+G+H+KL; LF+MF), 6 (F1+G+X+Q+Wa; LF+MF) and 9 (U+V+S+T+H; LF+MF) in A1 originated from chromosomes 2 (MN2+C+MN+KL; LF+MF), 5 (F+G+H; LF), 6 (F+G+H; MF), and 10 (Wa+Q+X+V+S+T+U; LF) of A4 genome after two Ts_{uneq}, one reciprocal translocation (RT) and one paracentric inversion (Ipa). The chromosomes 4 (KL+I+E+D+I1+I2; LF+MF) and 8 (R+Wb+P+O+J+I2+J3+J; LF+MN) originated from chromosomes 4 (KL+I+E+D+I1+I2+J; LF), 7 (I+J; MF) and 9 (O+P+Wb+R; MF) of A4 genome after an EET and RT event. Therefore, the A4 genome underwent six CRs to form A1. (**b**) Evolutionary pathway from A1 to *B. frutescens* (n = 10 → n = 9). The chromosomes 1 (Wa+Q+X+V+S+I1+D+U; LF+MF), 3 (A+B+MN1+I2; LF+MF), and 4 (KL+I1+E+T+F1+G+X+Q+Wa; LF+MF) in *B. frutescens* originated from chromosomes 1 (A+B+MN1; LF), 4 (KL+I1+E+D+I1+I2; LF+MF), 6 (F1+G+X+Q+Wa; LF+MF), and 10 (Wa+Q+X+V+S+T+U; MF) of A1 genome after two Ts_{uneq} and one RT event. Therefore, the A1 genome underwent three CRs to form *B. frutescens.* (**c**) Evolutionary pathway from A1 to *B. prealpina* (n = 10 → n = 9). The chromosome 4 (A+B+MN1+F1+G+X+G+Wa; LF+MF) in *B. prealpina* originated from chromosomes 1 (A+B+MN1; LF) and 6 (F1+G+X+Q+Wa; LF+MF) of A1 genome after an EET event. Therefore, the A1 genome underwent one CR to form *B. prealpina.* (**d**) Evolutionary pathway from A4 to *B. auriculata* (n = 11 → n = 8). The chromosomes 1 (Wa+Q+X+V+S+I1+D+U; LF+MF), 4 (MN+C+MN2+R1+Wb1+P1+O; LF+MF), 5 (R2+Wb2+P2+MN1+B+A; LF), 7 (U+T1+G+F; LF), and 8 (Wa+Q+X+V+S+T2+KL; LF+MF) in *B. auriculata* originated from chromosomes 1 (A+B+MN1; LF), 2 (MN2+C+MN+KL; LF+MF), 4 (KL+I1+E+D+I2+J; LF), 5 (F+G+H; LF), 8 (O+P+Wb+R; LF), and 10 (Wa+Q+X+V+S+T+U; LF) of A4 genome after five Ts_{uneq} and one Ipa event. The chromosome 2 (C+E+D+P+O+R+Wb+B+A; MF) in *B. auriculata* originated from chromosomes 3 (A+B+D+E+C; MF) and 9 (O+P+Wb+R) of A4 genome after an NCI and Ipa event. The chromosome 6 (J+I+H+G+F; MF) in *B. auriculata* originated from chromosomes 6 (F+G+H; MF) and 7 (I+J; MF) of A4 genome after an EET event. Therefore, the A4 genome underwent nine CRs to form *B. auriculata.*

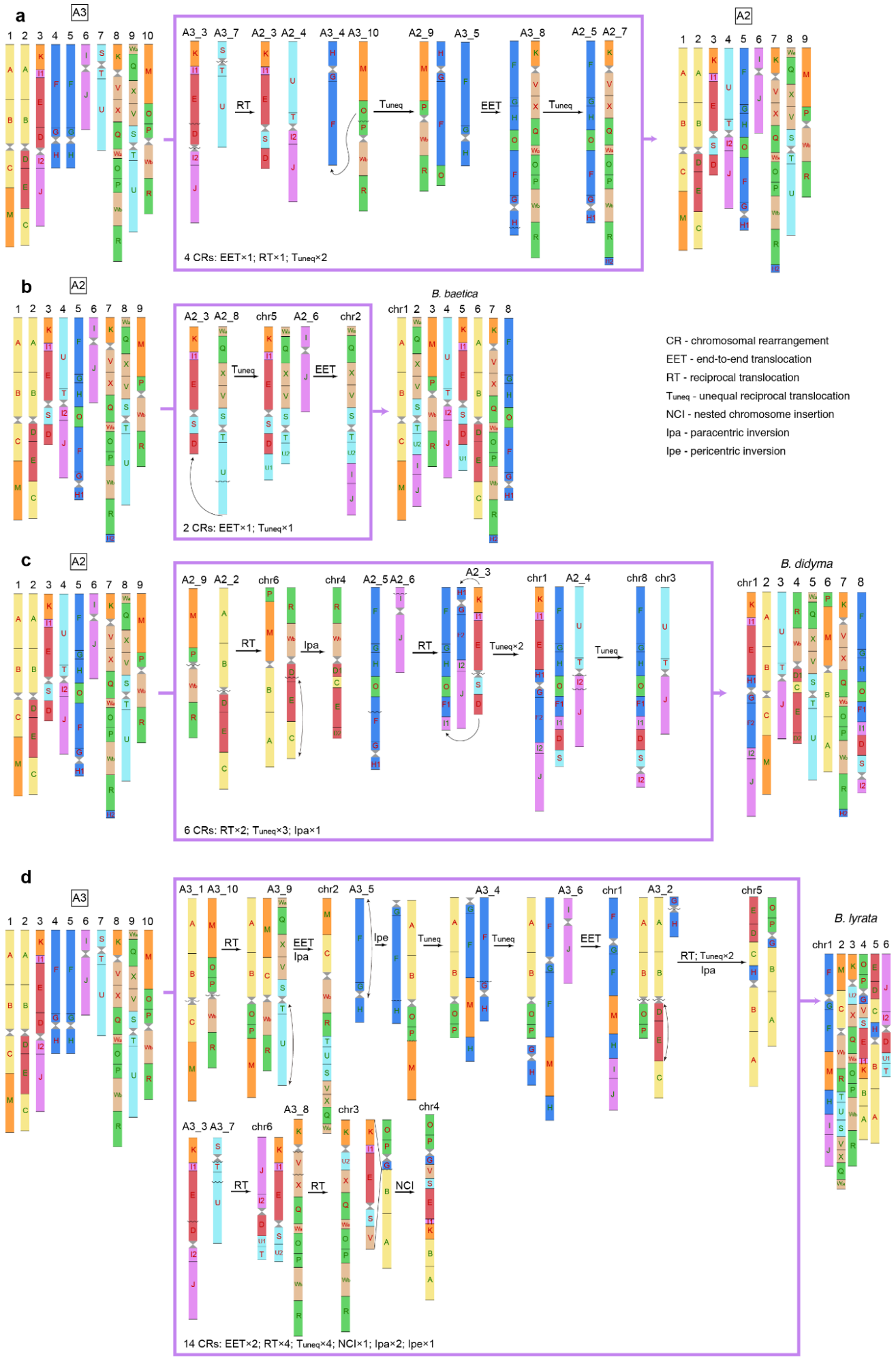

a
A3
A3_3 A3_7 A2_3 A2_4 A3_4 A3_10 A2_9 A3_5 A3_8 A2_5 A2_7
RT
Tuneq
EET
Tuneq
4 CRs: EET×1; RT×1; Tuneq×2
A2
b
A2
A2_3 A2_8 chr5 A2_6 chr2
Tuneq
EET
2 CRs: EET×1; Tuneq×1
B. baetica
chr1 2 3 4 5 6 7 8
CR - chromosomal rearrangement
EET - end-to-end translocation
RT - reciprocal translocation
Tuneq - unequal reciprocal translocation
NCI - nested chromosome insertion
Ipa - paracentric inversion
Ipe - pericentric inversion
c
A2
A2_9 A2_2 chr6 chr4 A2_5 A2_6 A2_3 chr1 A2_4 chr8 chr3
RT
Ipa
RT
Tuneq×2
Tuneq
6 CRs: RT×2; Tuneq×3; Ipa×1
B. didyma
chr1 2 3 4 5 6 7 8
d
A3
A3_1 A3_10 A3_9 chr2 A3_5 A3_4 A3_6 chr1 A3_2 chr5
RT
EET
Ipa
Ipe
Tuneq
Tuneq
EET
RT; Tuneq×2
Ipa
A3_3 A3_7 chr6 A3_8 chr3 chr4
RT
RT
NCI
14 CRs: EET×2; RT×4; Tuneq×4; NCI×1; Ipa×2; Ipe×1
B. lyrata
chr1 2 3 4 5 6

Figure S17. Evolutionary pathways from A3 (n = 10) to *B. baetica*, *B. didyma*, and *B. lyrata*. (**a**) Evolutionary pathway from A3 to A2 (n = 10 → n = 9). The chromosomes 3 (KL+I1+E+S+D; LF) and 4 (U+T+I2+J; LF) in A2 originated from chromosomes 3 (KL+I1+E+D+I2+J; LF) and 7 (S+T+U; LF) of A3 genome after an RT event. The chromosomes 5 (F+G+H+O+F+G+H1; LF+MF), 7 (KL+V+X+Q+Wa+O+P+Wb+R+H2; LF+MF), and 9 (MN+P+Wb+R; LF) originated from chromosomes 4 (F+G+H; LF), 5 (F+G+H; MF), 8 (KL+V+X+Q+Wa+O+P+Wb+R; LF+MF), and 10 (MN+O+P+Wb+R; LF) of A3 genome after two Ts_{uneq} and an EET event. Therefore, the A3 genome underwent four CRs to form A2. (**b**) Evolutionary pathway from A2 to *B. baetica* (n = 9 → n = 8). The chromosomes 2 (Wa+Q+X+V+S+T+U2+I+J; MF) and 5 (KL+I1+E+S+D+U1; LF+MF) in *B. baetica* originated from chromosomes 3 (KL+I1+E+S+D; LF), 6 (I+J; MF), and 8 (Wa+Q+X+V+S+T+U; MF) of A2 genome after a T_{uneq} and an EET event. Therefore, the A2 genome underwent two CRs to form *B. baetica*. (**c**) Evolutionary pathway from A2 to *B. didyma* (n = 9 → n = 8). The chromosomes 4 (R+Wb+D1+C+E+D2; LF+MF) and 6 (P+MN+A+B; LF+MF) in *B. didyma* originated from chromosomes 2 (A+B+D+E+C; MF) and 9 (MN+P+Wb+R; LF) of A2 genome after an RT and Ipa event. The chromosomes 1 (KL+I1+E+H1+G+F2+I2+J; LF+MF), 3 (U+T+J; LF), and 8 (F+G+H+O+F1+I1+D+S+I2; LF+MF) in *B. didyma* originated from chromosomes 3 (KL+I1+E+S+D; LF), 4 (U+T+I2+J; LF), 5 (F+G+H+O+F+G+H; LF+MF), 6 (I+J; MF) of A2 genome after three Ts_{uneq} and an RT event. Therefore, the A2 genome underwent six CRs to form *B.didyma*. (**d**) Evolutionary pathway from A3 to *B. lyrata* (n = 10 → n = 6). The chromosomes 1 (F+G+F+MN+H+I+J; LF+MF) and 2 (MN+C+Wb+R+T+U+S+V+X+Q+Wa; LF+MF) in *B. lyrata* originated from chromosomes 1 (A+B+C+MN; LF+MF), 4 (F+G+H; LF), 5 (F+G+H; MF), 9 (Wa+Q+X+V+S+T+U; MF), and 10 (MN+O+P+Wb+R; LF) after two Ts_{uneq}, two EETs, one Ipa, one Ipe, and an RT event. The two intermediate fragments (A+B+O+P; LF, and G+H; LF) produced in this process underwent two Ts_{uneq}, one RT, and one Ipa with chromosome 2 (A+B+D+E+C; MF) in A3 to form chromosome 5 (E+D+C+H+B+A; LF+MF) in *B. lyrata* and another fragment (O+P+G+B+A; LF+MF). The chromosomes 3 (KL+U2+X+Q+Wa+O+P+Wb+R; LF+MF) and 6 (J+I2+D+U1+T; LF) in *B. lyrata* originated from chromosomes 3 (KL+I1+E+D+I2+J; LF), 7 (S+T+U; LF), and 8 (KL+V+X+Q+Wa+O+P+Wb+R; LF+MF) after two RT events. Two intermediate fragments (KL+I1+E+S+V; LF, and O+P+G+B+A; LF+MF) then underwent an NCI event to form chromosome 4 (O+P+G+V+S+E+I1+KL+B+A; LF+MF) in *B. lyrata*. Therefore, the A3 genome underwent 14 CRs to form *B. lyrata*.

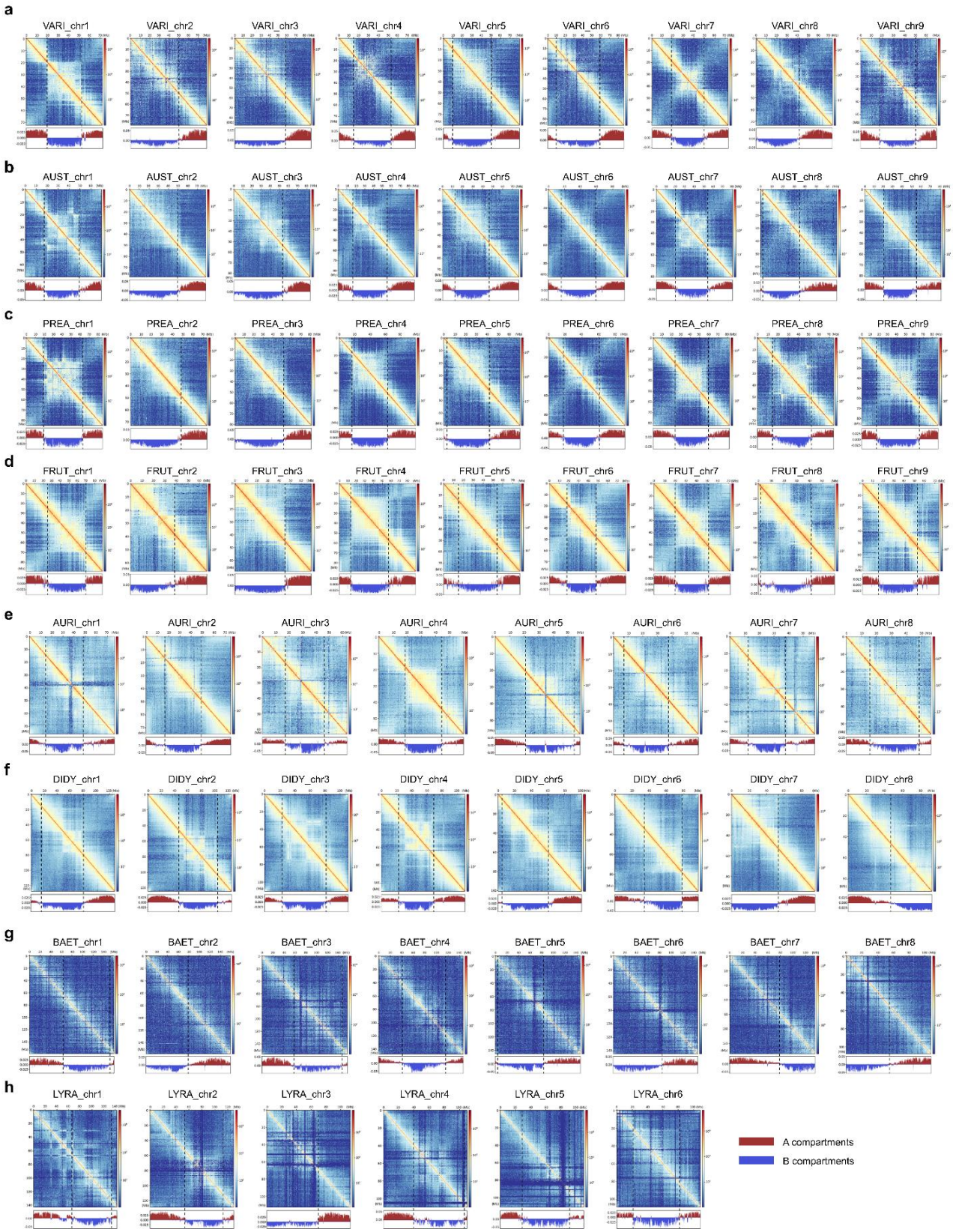

a
VARI_chr1
VARI_chr2
VARI_chr3
VARI_chr4
VARI_chr5
VARI_chr6
VARI_chr7
VARI_chr8
VARI_chr9
b
AUST_chr1
AUST_chr2
AUST_chr3
AUST_chr4
AUST_chr5
AUST_chr6
AUST_chr7
AUST_chr8
AUST_chr9
c
PREA_chr1
PREA_chr2
PREA_chr3
PREA_chr4
PREA_chr5
PREA_chr6
PREA_chr7
PREA_chr8
PREA_chr9
d
FRUT_chr1
FRUT_chr2
FRUT_chr3
FRUT_chr4
FRUT_chr5
FRUT_chr6
FRUT_chr7
FRUT_chr8
FRUT_chr9
e
AURI_chr1
AURI_chr2
AURI_chr3
AURI_chr4
AURI_chr5
AURI_chr6
AURI_chr7
AURI_chr8
f
DIDY_chr1
DIDY_chr2
DIDY_chr3
DIDY_chr4
DIDY_chr5
DIDY_chr6
DIDY_chr7
DIDY_chr8
g
BAET_chr1
BAET_chr2
BAET_chr3
BAET_chr4
BAET_chr5
BAET_chr6
BAET_chr7
BAET_chr8
h
LYRA_chr1
LYRA_chr2
LYRA_chr3
LYRA_chr4
LYRA_chr5
LYRA_chr6
A compartments
B compartments

Figure S18. Hi-C contact maps and PC1 value tracks of chromosomes in *B. varia* (**a**), *B. austriaca* (**b**), *B. prealpina* (**c**), *B. frutescens* (**d**), *B. auriculata* (**e**), *B. didyma* (**f**), *B. baetica* (**g**), and *B. lyrata* (**h**). The red and blue bars of PC1 value tracks represent the A/B compartments, respectively. The dashed lines show the main boundaries of the A/B compartments.

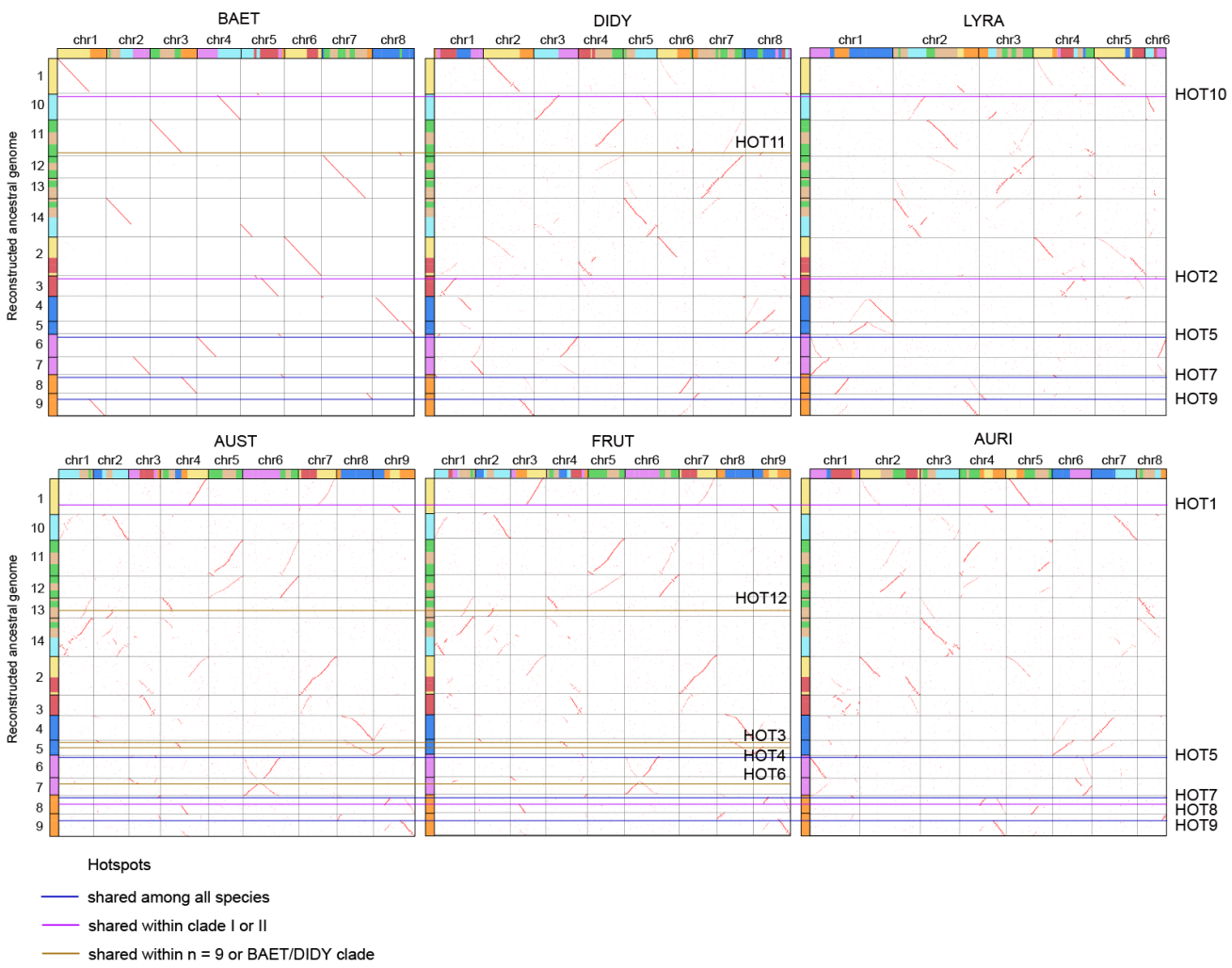


Figure S19. Homology comparisons between the reconstructed ancestral genome (Y-axis) and the six *Biscutella* genomes (X-axis). The positions of ancestral chromosome breakage hotspots are labelled, with varying colors representing three situations in which breakpoints are shared among species, i.e., three hotspots shared by all species (dark blue), four hotspots shared by either clade I (*B. auriculata* and the *n* = 9 species) or clade II (*B. baetica*, *B. didyma*, and *B. lyrata*; purple) species, and five hotspots shared by either *n* = 9 species or by *B. baetica* and *B. didyma* (gold).

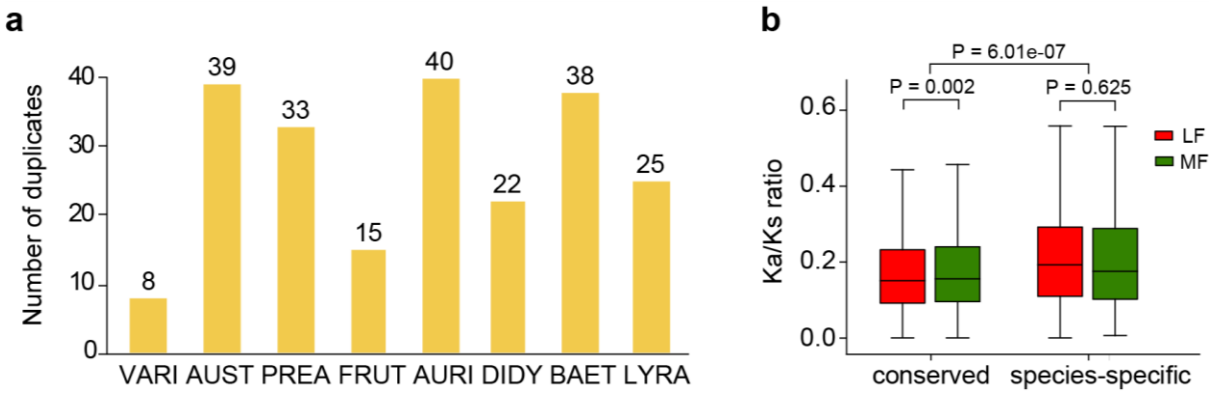


Figure S20. Sub/neo-functionalization in 1,049 CD gene families. (**a**) Number of duplicates with species-specific sub/neo-functionalization. (**b**) *K*a/*K*s ratios of genes with conserved protein domains and species-specific sub/neo-functionalization in the CD family, with syntelogs in *M. pygmaea* used as reference (Wilcoxon rank-sum test).

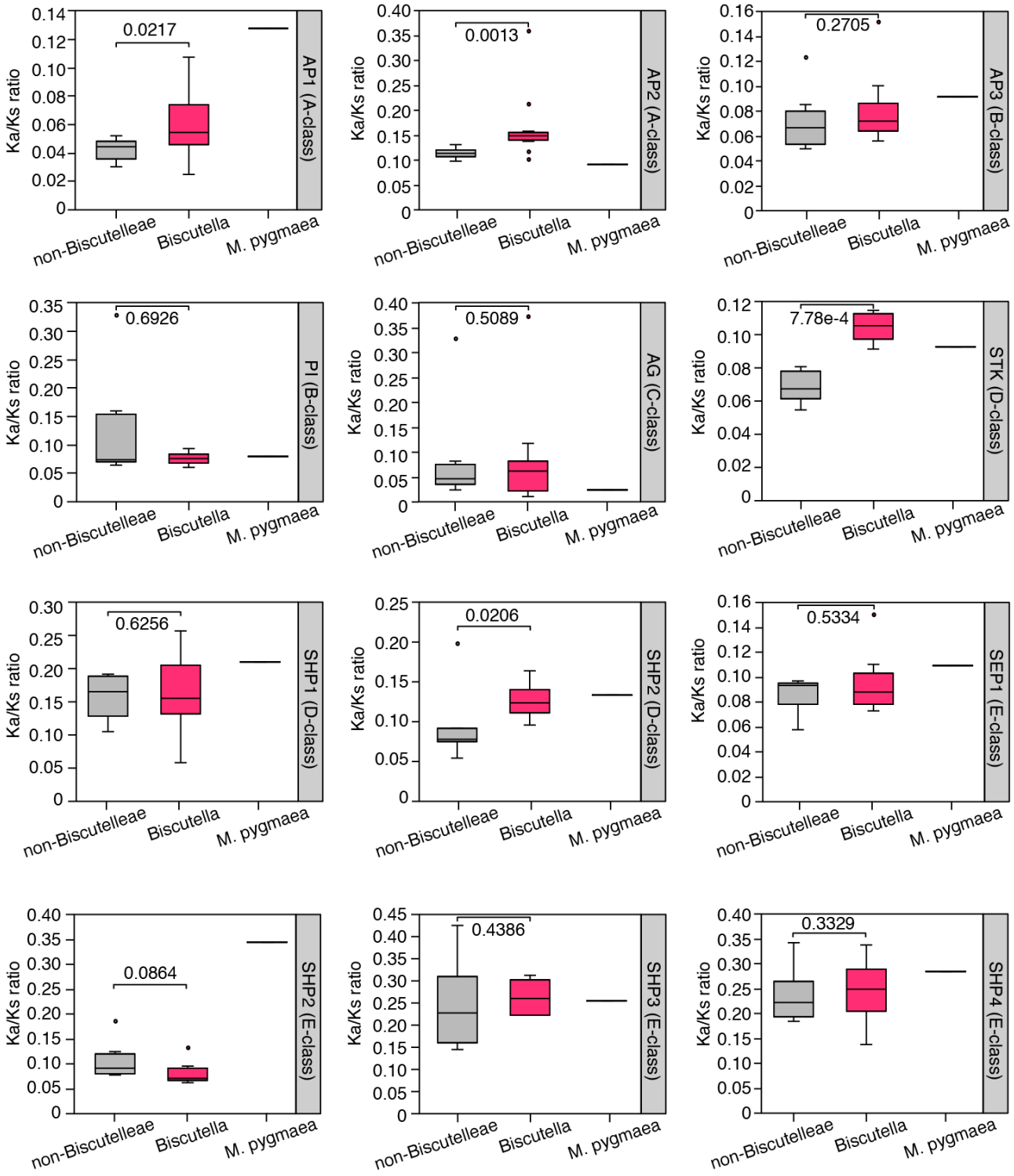
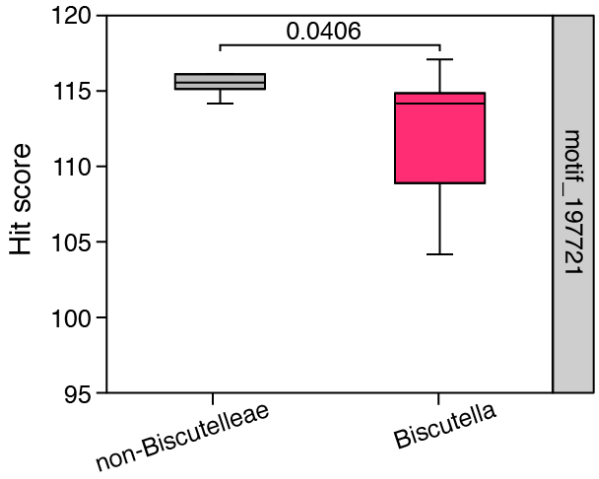
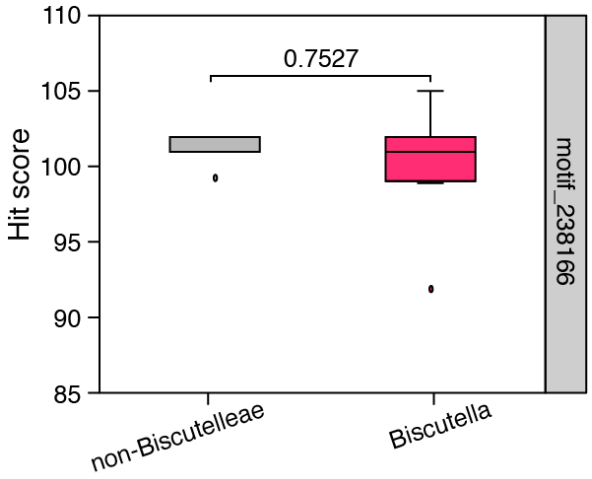
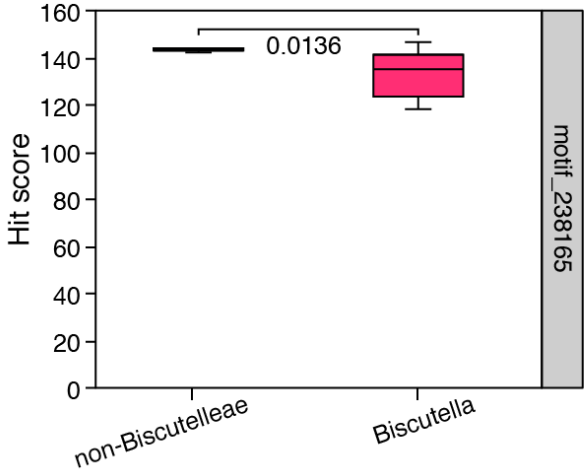
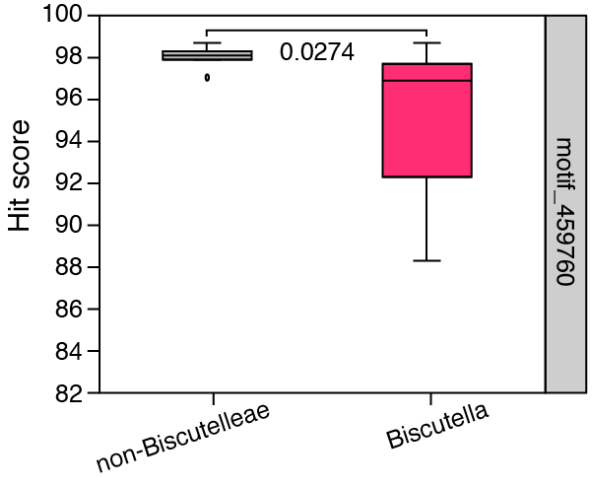


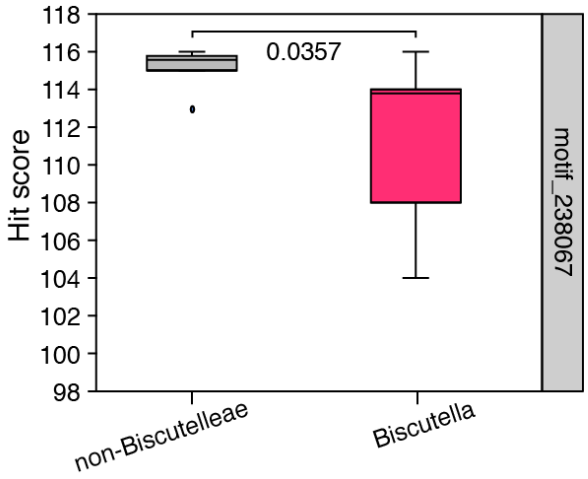
Figure S21. Box plots of *K*a/*K*s ratios for gene members in the ABCDE flowering model across seven non-Biscutelleae species (including *A. thaliana*, *Bo. stricta*, *Br. rapa*, *Car. hirsuta*, *Cap. rubella*, *P. cornutum*, and *S. parvula*), eight *Biscutella* species, and *M. pygmaea*, with syntelogs in *Aethionema arabicum* used as reference (Wilcoxon rank-sum test).


0.0406
Hit score
motif_197721
non-Biscutelleae
Biscutella


0.7527
Hit score
motif_238166
non-Biscutelleae
Biscutella


0.0136
Hit score
motif_238165
non-Biscutelleae
Biscutella


0.0274
Hit score
motif_459760
non-Biscutelleae
Biscutella


0.0357
Hit score
motif_238067
non-Biscutelleae
Biscutella

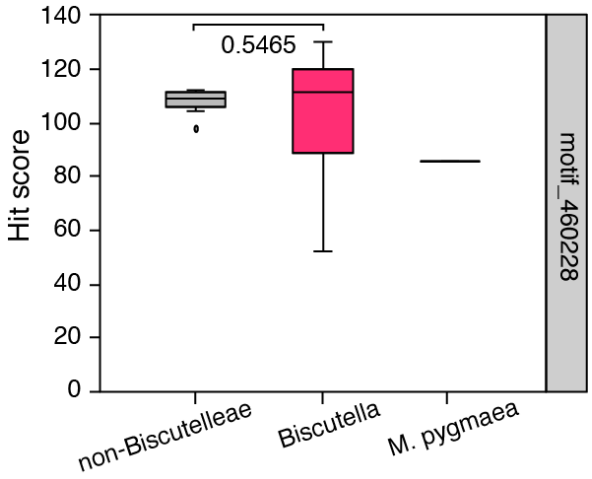

0.5465
Hit score
motif_460228
non-Biscutelleae
Biscutella
M. pygmaea

Figure S22. Hit scores of motifs in the SHP1 gene compared to the NCBI-CDD database. Six major motifs of SHP1 were found based on the MOTIFSearch tool (https://www.genome.jp/tools/motif/), and four of six motifs had significantly lower hit scores in *Biscutella* species than in other Brassicaceae species (including *A. thaliana*, *Bo. stricta*, *Br. rapa*, *Car. hirsuta, Cap. rubella, P. cornutum,* and *S. parvula*). The SHP1 gene in *M. pygmaea* had only one motif (motif_460228) identified and obviously had a low hit score. The Wilcoxon rank-sum tests were performed.

Supplementary Tables S1 – S13 (see MS Excel spreadsheet)

Table S1. Sequencing reads used for assemblies and gene annotations of the *Biscutella* genomes
Table S2. Statistics of the eight *Biscutella* genome assemblies
Table S3. The BUSCO scores of eight *Biscutella* genomes using the genome and protein models
Table S4. Statistics of various features related to genes based on annotated gene models
Table S5. Syntenic fragments and pericentromeric regions identified in *Biscutella*
Table S6. Annotated elements of each TE superfamily in *Biscutella* genomes
Table S7. Statistics of tandem repeats in eight *Biscutella* genomes
Table S8. Tandem repeat types abundant in centromeric and subtelomeric regions
Table S9. Chromosomal rearrangement rates in *Biscutella* species
Table S10. Number of solo- and intact-LTRs in the whole-genome, LF, MF, and pericentromeric regions
Table S11. GO functional enrichment of the CD gene family
Table S12. GO functional enrichment of the CS gene family
Table S13. The ABCDE orthologous gene families in 16 Brassicaceae species

Supplementary References

Geiser, C., Mandáková, T., Arrigo, N., Lysak, M.A., & Parisod, C. (2016). Repeated whole-genome duplication, karyotype reshuffling, and biased retention of stress-responding genes in Buckler mustard. *Plant Cell*, 28, 17-27.
Kolde, R., & Kolde, M.R. (2015). Package 'pheatmap'. *R package*, *1*, 790.
Lysak, M.A., Mandáková, T. & Schranz, M.E. (2016) Comparative paleogenomics of crucifers: ancestral genomic blocks revisited. *Current Opinion in Plant Biology*, 30, 108–115.
Mandáková, T., & Lysak, M.A. (2008). Chromosomal phylogeny and karyotype evolution in x = 7 crucifer species (Brassicaceae). *Plant Cell*, 20, 2559-2570.
Mandáková, T., & Lysak, M.A. (2018). Post-polyploid diploidization and diversification through dysploid changes. *Current Opinion in Plant Biology*, 42, 55-65.
Vicente, A., Alonso, M.Á., & Crespo, M.B. (2020). Born in the Mediterranean: Comprehensive taxonomic revision of *Biscutella* ser. *Biscutella* (Brassicaceae) based on morphological and phylogenetic data. *Annals of the Missouri Botanical Garden*, 105, 195-231.