

Genome Structure of *xBrassicoraphanus*, an  
Intergeneric Hybrid between *Brassica rapa*  
and *Raphanus sativus*

Jin Hoe Huh

Department of Plant Science  
Seoul National University

# Hybrids

- Hybrids occur in many plants and some animals.
- Naturally occurring hybridization in plants is often accompanied with polyploidization events (whole genome duplication).
- Accommodation of divergent genomes in the same nucleus.

# Hybrid vigor (heterosis)

- superior levels of biomass, stature, growth rate, and/or fertility in the hybrid offspring compared to their parents



Mo17

F<sub>1</sub>

B73



Chen, 2010, Trends Plant Sci.

# Hybridization barrier between distantly related species

- Infertility and genetic instability
- **Genome shock**
  - the release of chromatin constraints of gene expression including TE activation
- **Transcriptome shock**
  - extensive changes of parental gene expression patterns

# Outbreeding depression

- Reduced fertility
- Reduced fitness
- Breakdown of biochemical and physiological compatibilities between genes

# Polyploids

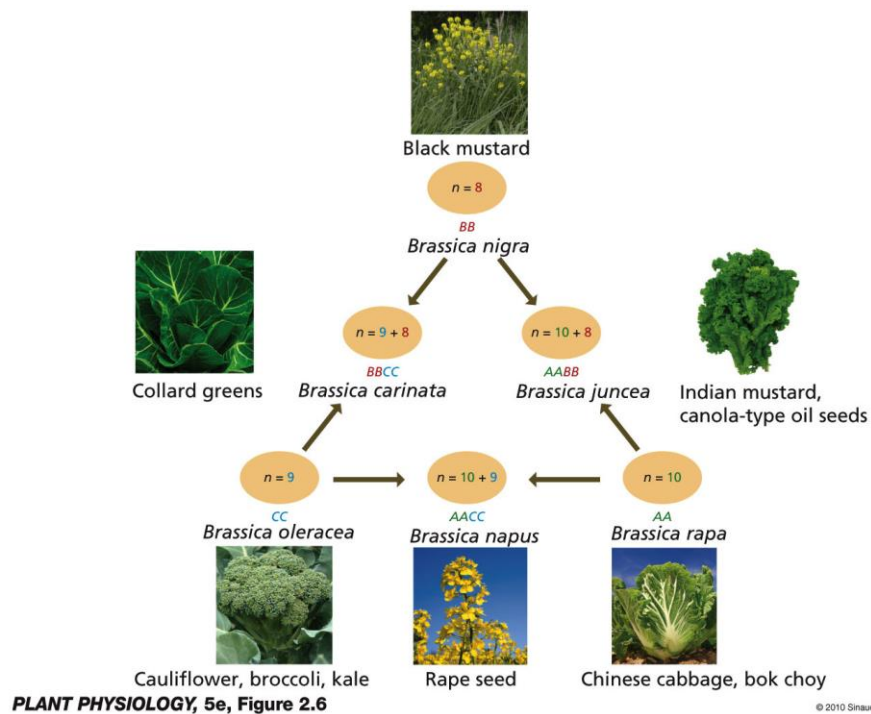
- Polyploids possess more than two complete sets of chromosomes.
- Many sequenced plant genomes display the signature of polyploidy ancestry.
- Polyploidy provides long-term evolutionary flexibility.
- **Autopolyploids**
  - chromosome doubling
- **Allopolyploids**
  - hybridization between different species followed by chromosome doubling



# “The Synthesis of Species” (1935)



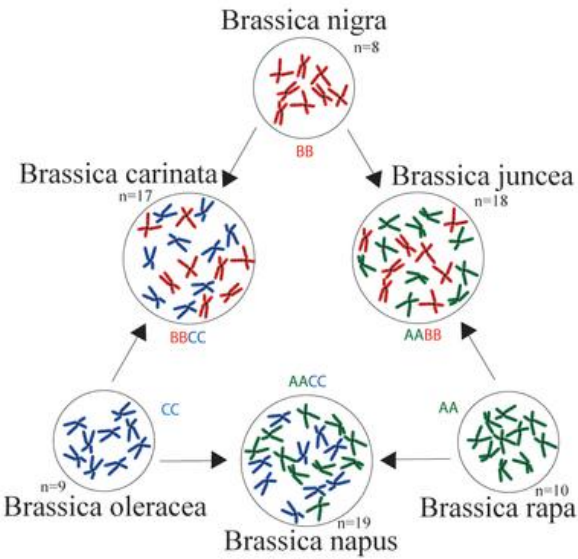
Woo Jang-Choon  
(1898~1959)



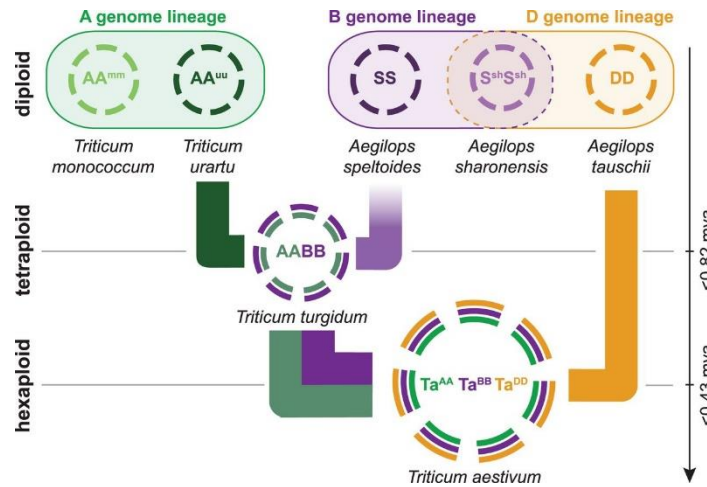
U's triangle

# Natural allopolyploid crops

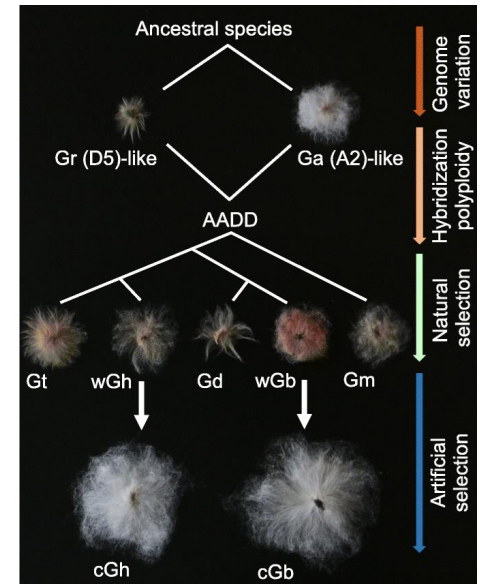
## <Brassica>



## <Triticum>



## <Gossypium>



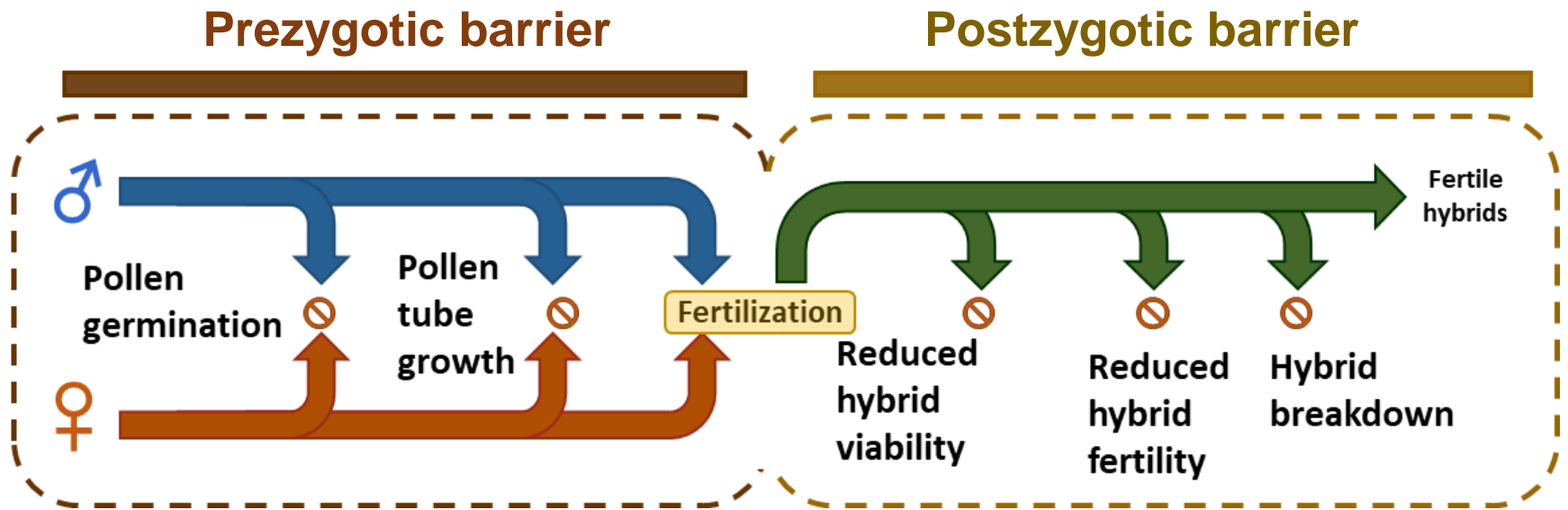
[https://en.wikipedia.org/wiki/Triangle\\_of\\_U](https://en.wikipedia.org/wiki/Triangle_of_U)

The International Wheat Genome Sequencing Consortium (2014), *Science*

Song *et al* (2017), *Genome Biol.*

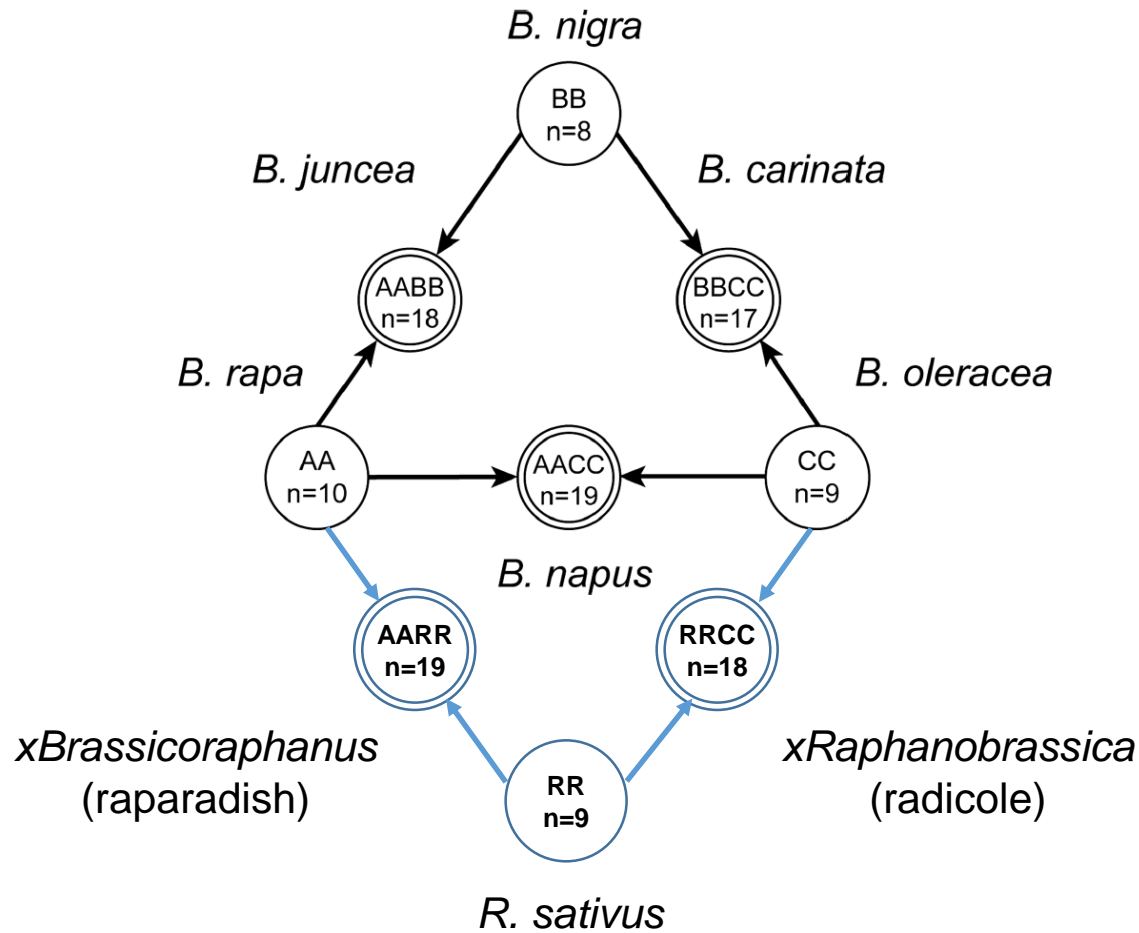


# Hybridization barrier



“Genome shock” and “Transcriptome shock” are likely responsible for a post-zygotic barrier.

# Intergeneric hybrids in the *Brassicaceae* family





*Brassica rapa* (♀)  
AA, n=10

X



*Raphanus sativus* (♂)  
RR, n=9





*Brassica rapa* (♀)  
AA, n=10

X



*Raphanus sativus* (♂)  
RR, n=9



*xBrassicoraphanus* (F1)  
aAARR, n=19

Major questions:

1. Intergeneric hybrid genome structure
2. Fates of homoeologous genes in the same nucleus
3. Possible epigenetic mechanisms for hybrid genome stabilization

# Plant materials

- *xBrassicoraphanus* 'BB1'
- *Brassica rapa* L. cv. Chiifu (2011, Nat. Genet.)
- *Raphanus sativus* L. cv. WK10039 (2016, Theor. Appl. Genet.)



# Developing stable progenies of $\times$ *Brassicoraphanus*, an intergeneric allopolyploid between *Brassica rapa* and *Raphanus sativus*, through induced mutation using microspore culture

Soo-Seong Lee · Sun-Ae Lee · Jungmin Yang · Jongkee Kim



# Phenotypes of $xBrassicoraphanus$ intermediate between *B. rapa* and *R. sativus*



# Neomorphic phenotype in petal color



*B. rapa*

*xBrassicoraphanus*

*R. sativus*

Genome structure of *xBrassicoraphanus*

# Genome assembly of *xBrassicoraphanus*

(Assembly information)	Contig	Scaffold
Total length / Number	652.44 Mb / 68,454 ea	<u>692.83 Mb / 20,299 ea</u>
Average / Median	9.53 kb / 2.40 kb	34.13 kb / 901 bp
Max / Min length	190.62 kb / 200 bp	16.46 Mb / 213 bp
N50	28,581 bp (6,854 <sup>th</sup> )	4,479,746 bp (49 <sup>th</sup> )
N90	5,982 bp (24,969 <sup>th</sup> )	166,698 bp (284 <sup>th</sup> )
GC contents	35.75%	33.68%

## <Generated genome sequences>

Insert size	Total length(Gb)	Filtered length (Gb)
200bp	39.1	26.3
400bp	42.4	31.8
3kbp	24.6	6.2
8kbp	22.7	3.5
10kbp	32.7	12.7
15kbp	33.6	2.9
Total	195.0	<u>83.4</u>

## <Size estimation by flow cytometry>

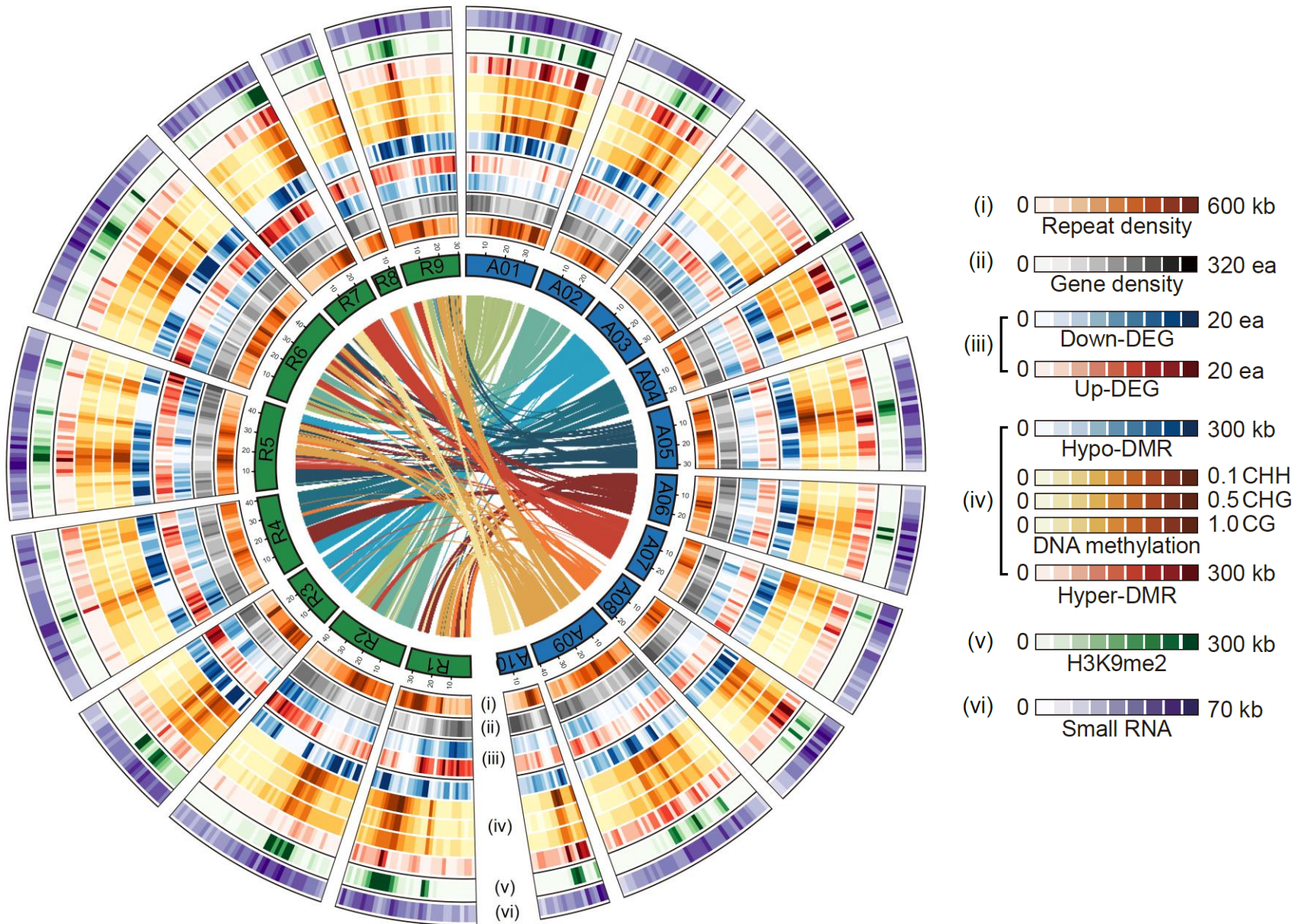
		Peak value	Genome size (Mb)
1st	xB/Br	98.29/197.90	976.5
2nd	xB/Br	92.29/190.69	1,002.1

# Br : 485Mb, Rs : 510Mb

Covering **69~71%** of the genome

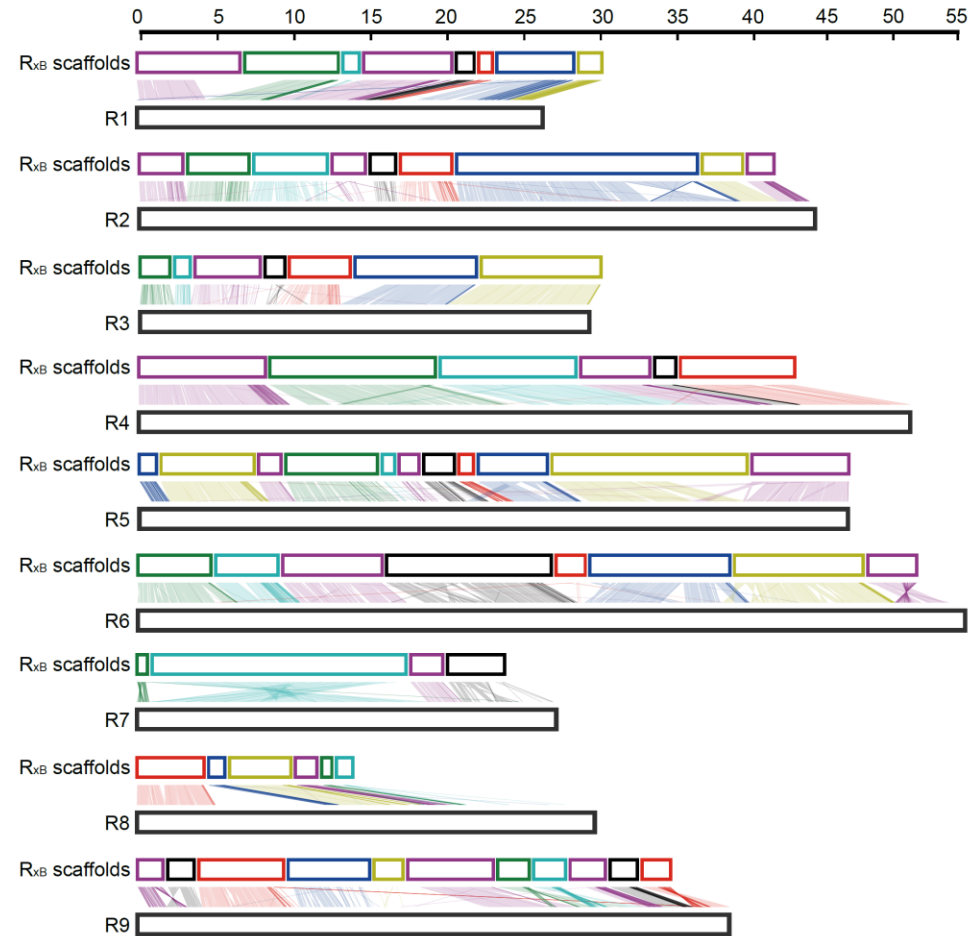
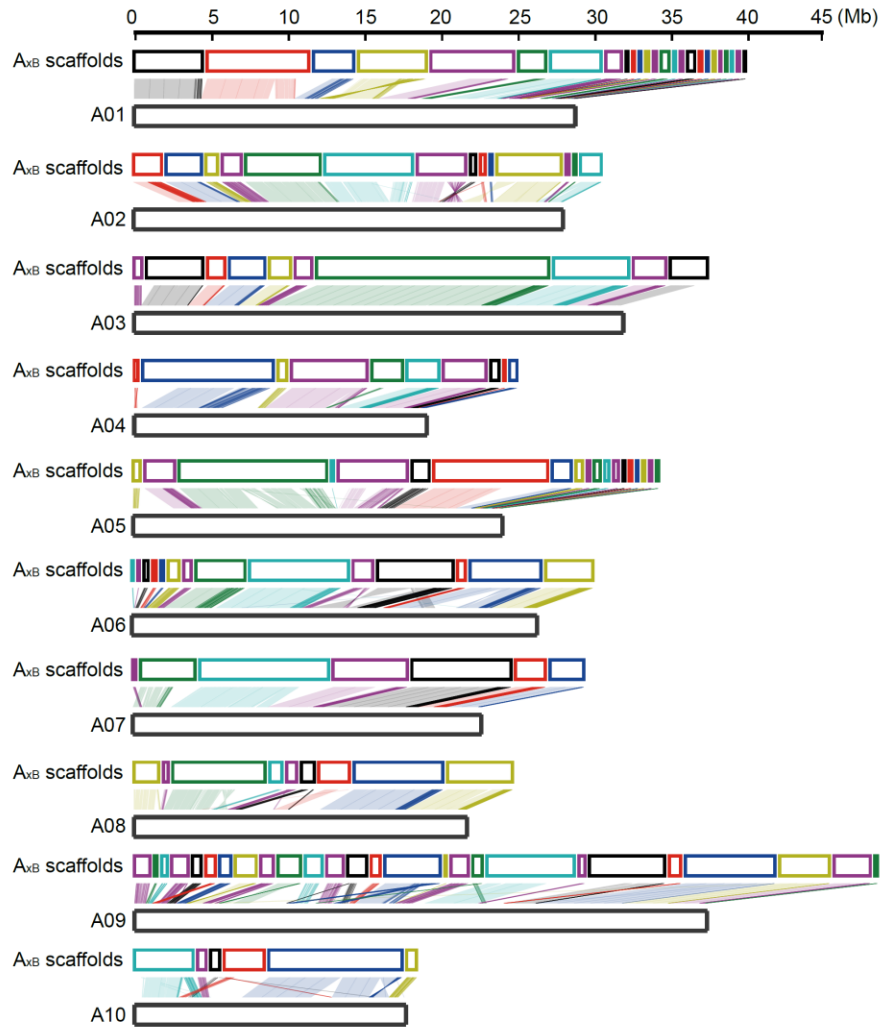


# Genome structure of *xBrassicoraphanus*



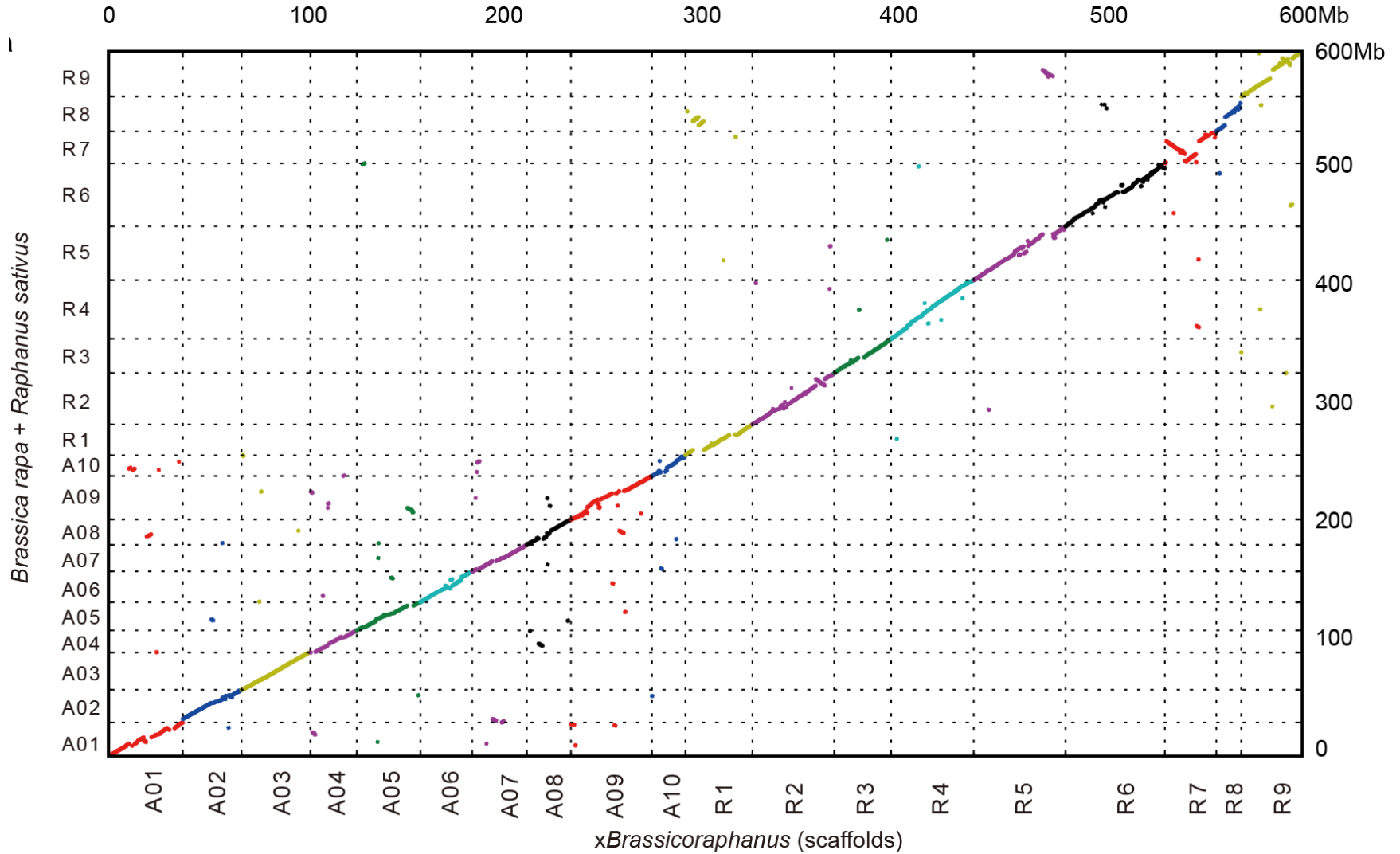


# Assignment of scaffolds



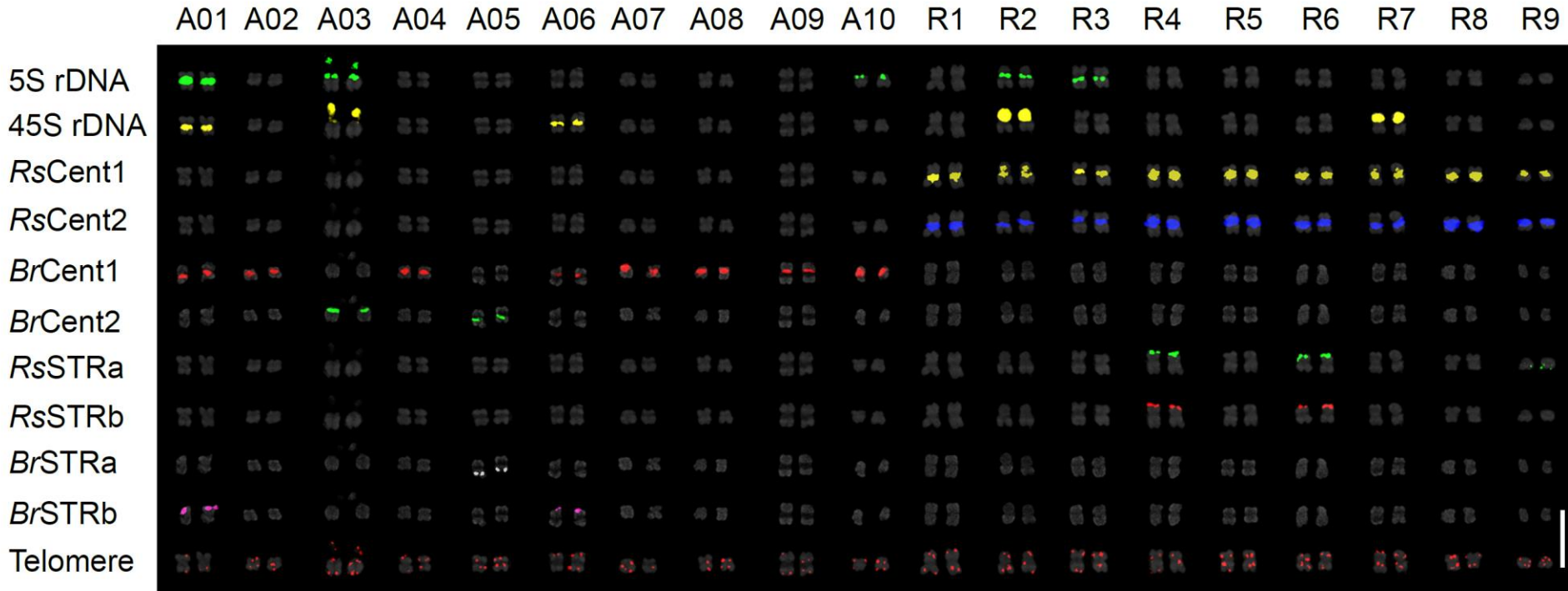
Both *Br* and *R*s genomes exist in entirety without losses in allotetraploid *x**B*.

# Assignment of scaffolds



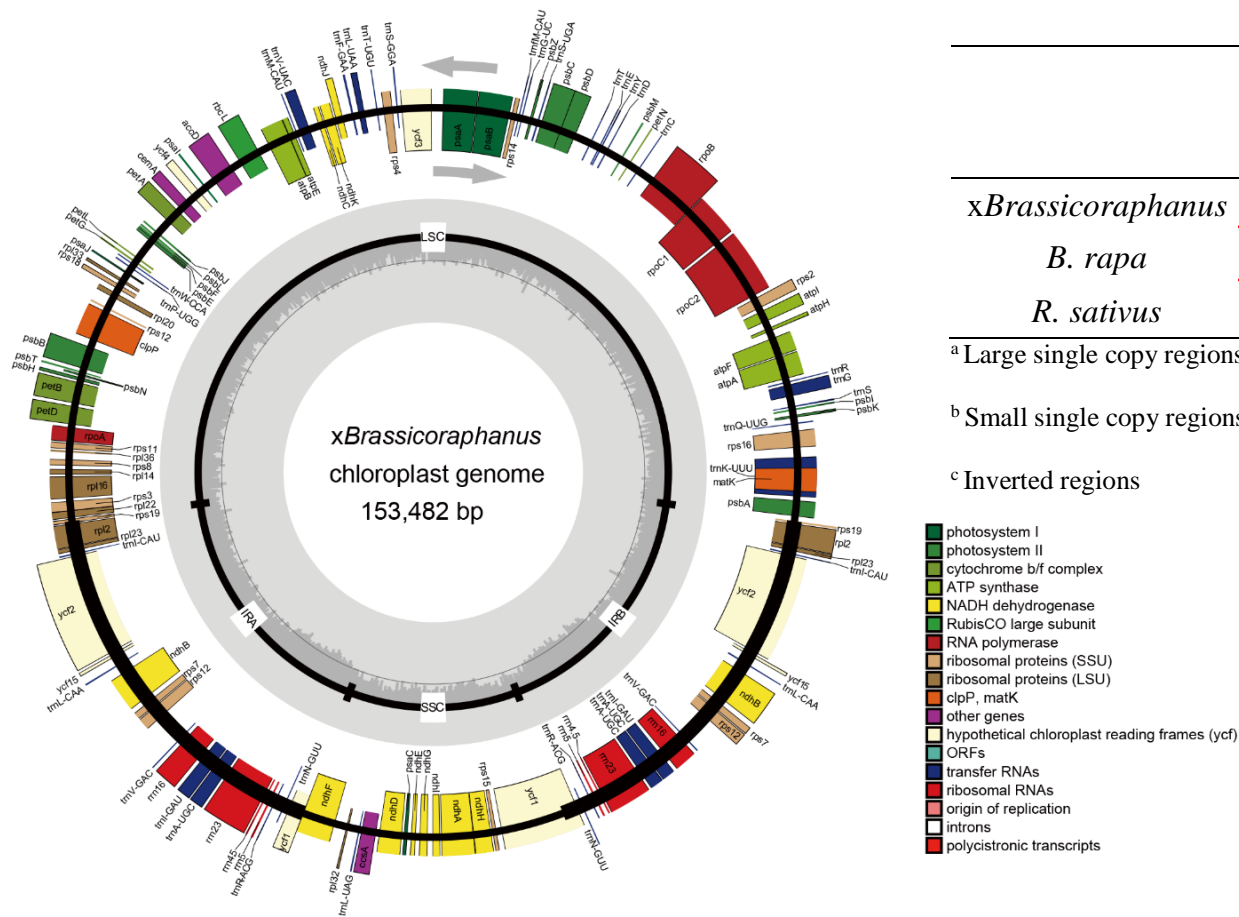
xB underwent only minor genome structural changes after allopolyploidization.

# Conservation of parental chromosome complements in *xBrassicoraphanus*



A total of 19 chromosome pairs present in *xB* without apparent chromosomal rearrangements such as translocations and deletions.

# *xBrassicoraphanus* harbors the chloroplast derived from *B. rapa*



	Structural annotation			
	Size (bp)	LSC <sup>a</sup> (bp)	SSC <sup>b</sup> (bp)	IR <sup>c</sup> (bp)
<i>xBrassicoraphanus</i>	153,482	83,281	17,775	26,213
<i>B. rapa</i>	153,482	83,281	17,775	26,213
<i>R. sativus</i>	153,368	83,171	17,765	26,215

<sup>a</sup> Large single copy regions

<sup>b</sup> Small single copy regions

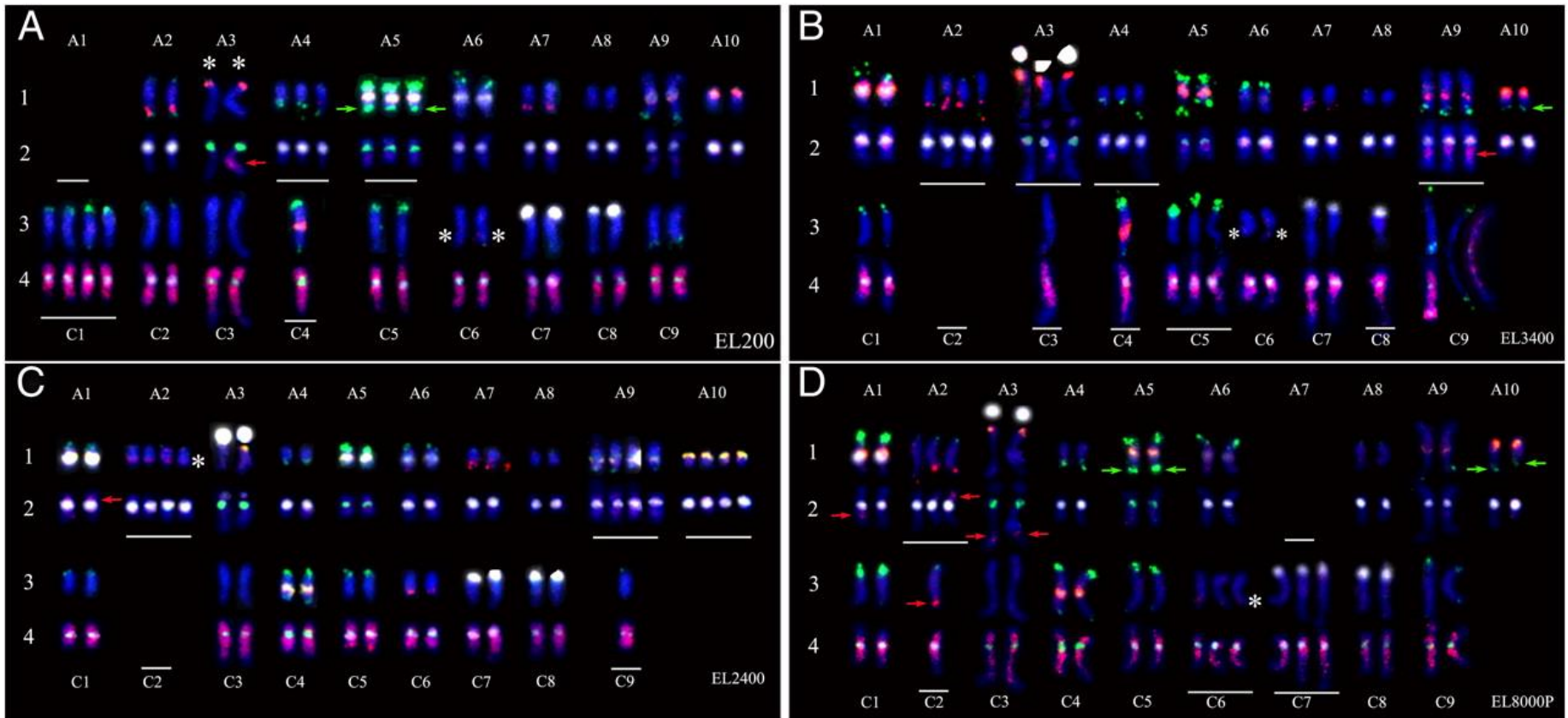
<sup>c</sup> Inverted regions

- photosystem I
- photosystem II
- cytochrome b/f complex
- ATP synthase
- NADH dehydrogenase
- RubisCO large subunit
- RNA polymerase
- ribosomal proteins (SSU)
- ribosomal proteins (LSU)
- clpP, matK
- other genes
- hypothetical chloroplast reading frames (ycf)
- ORFs
- transfer RNAs
- ribosomal RNAs
- origin of replication
- introns
- polycistronic transcripts

The chloroplast genome of *xB* has 99.9% similarity to *Br*, indicating its maternal inheritance.

Chromosome behavior in  
*xBrassicoraphanus*

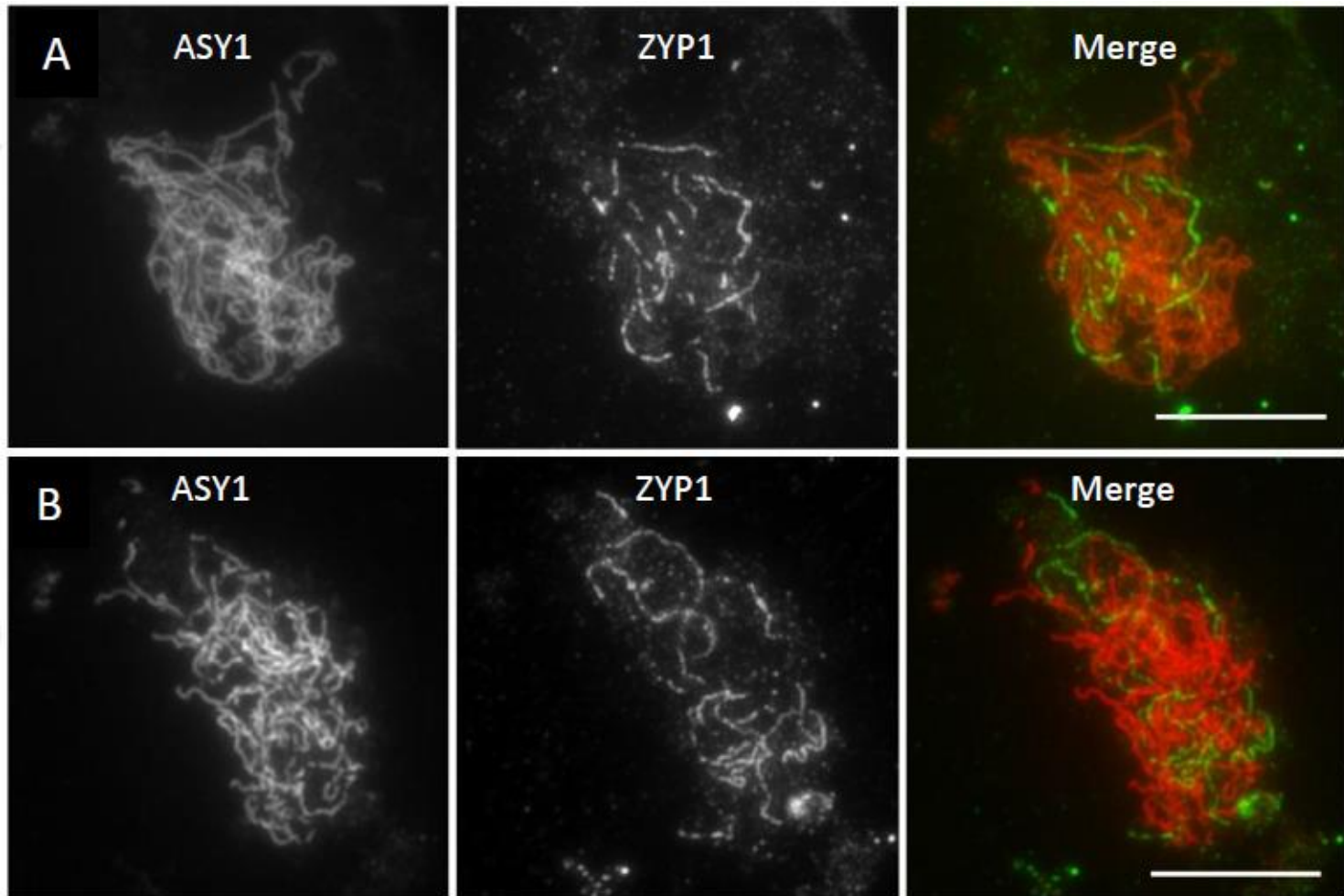
# Chromosomal instability in early generations of *B. napus*



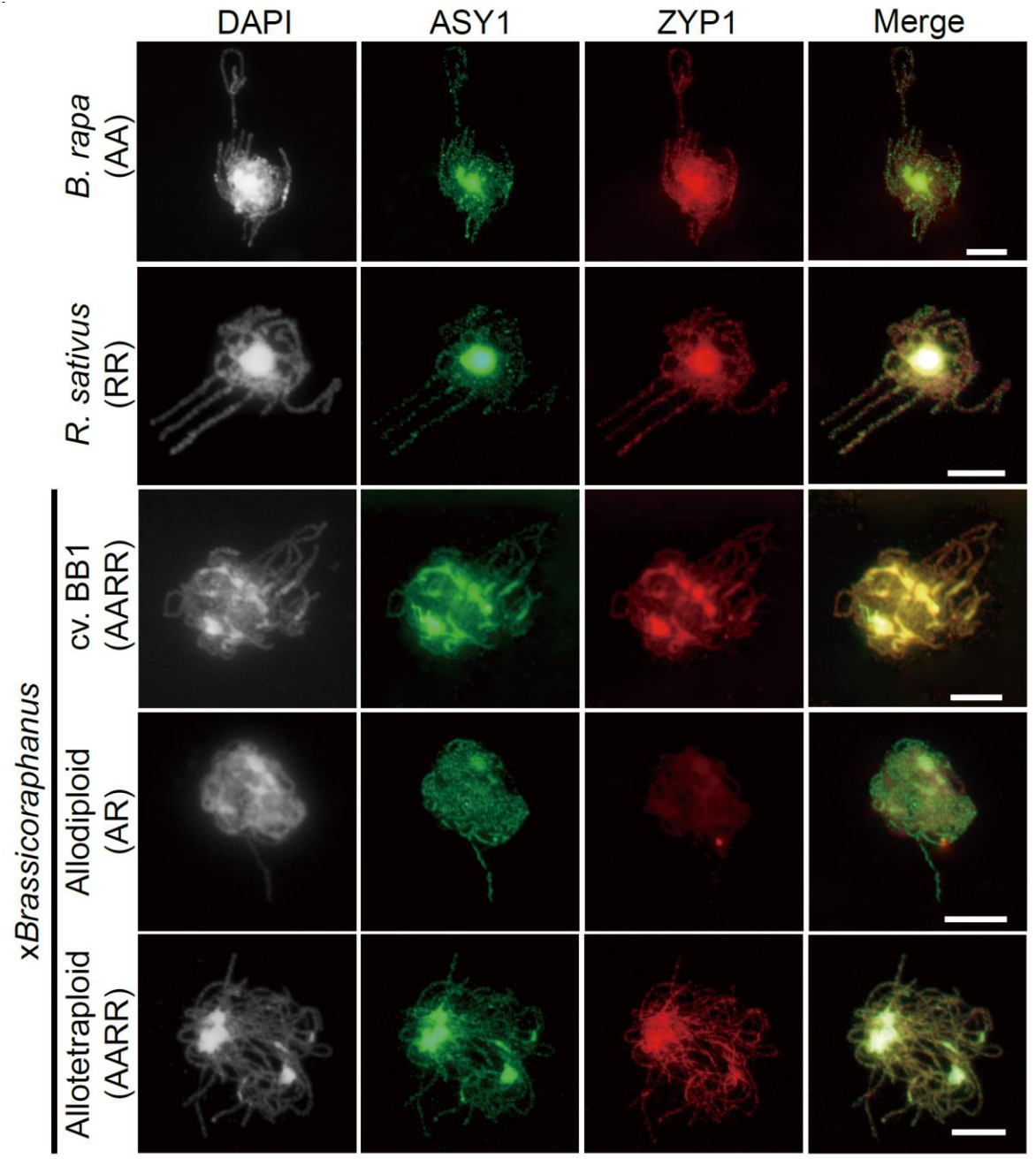
Early generations of resynthesized allopolyploids often suffer from aneuploidy and chromosomal rearrangements.



# Incomplete synapsis in allohaploids *B. napus* cv. *Darmor-bzh* and *Yudal* (AC)

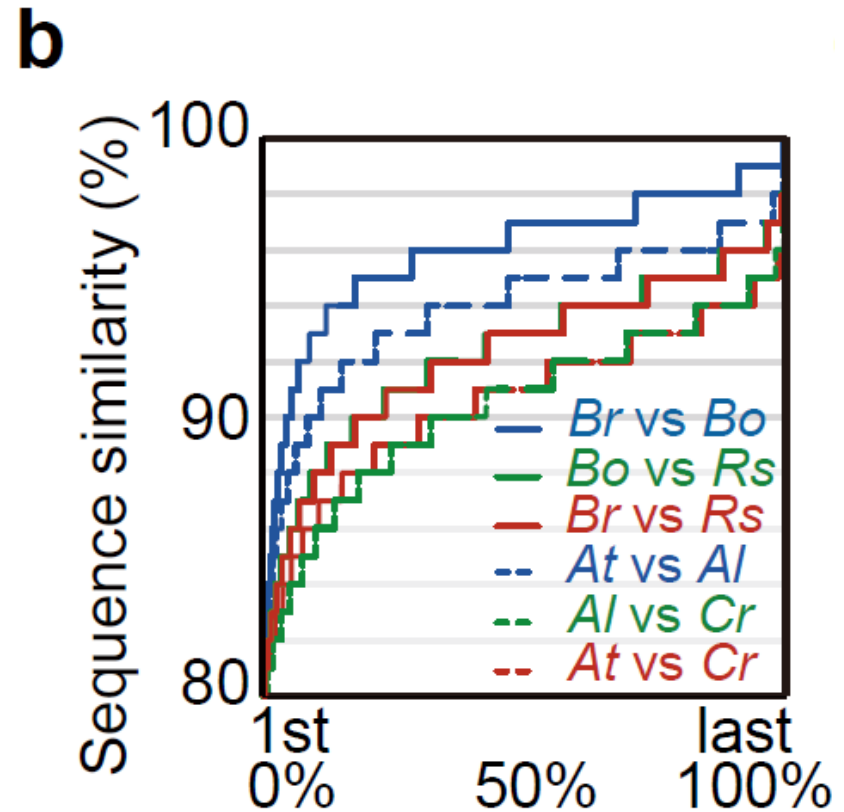
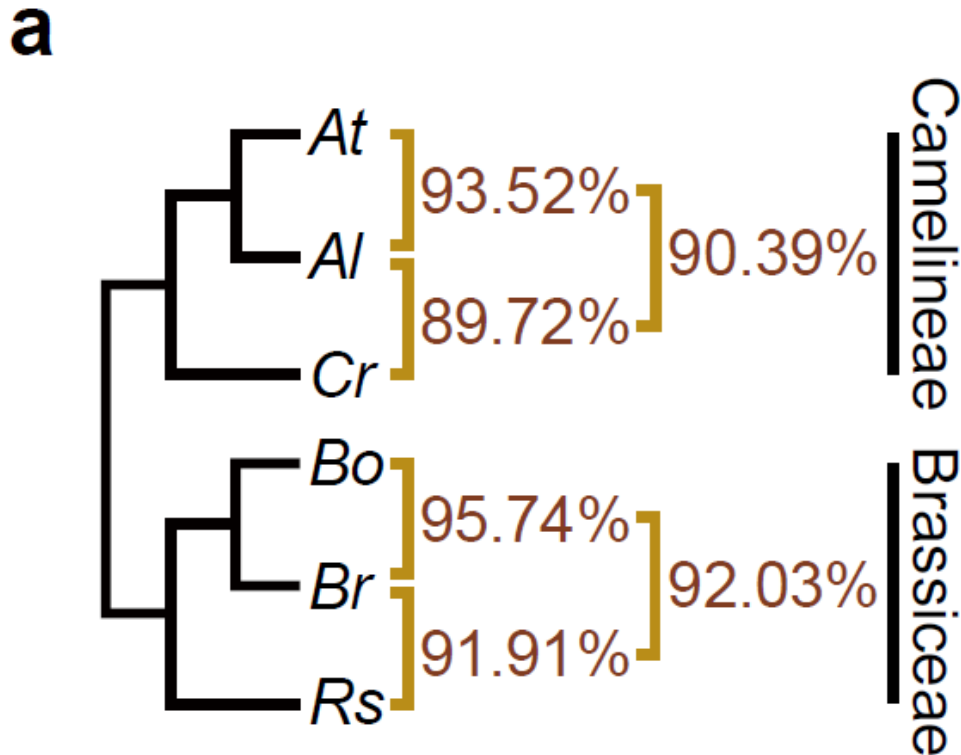


# Absence of homeologous pairing between Br and Rs chromosomes

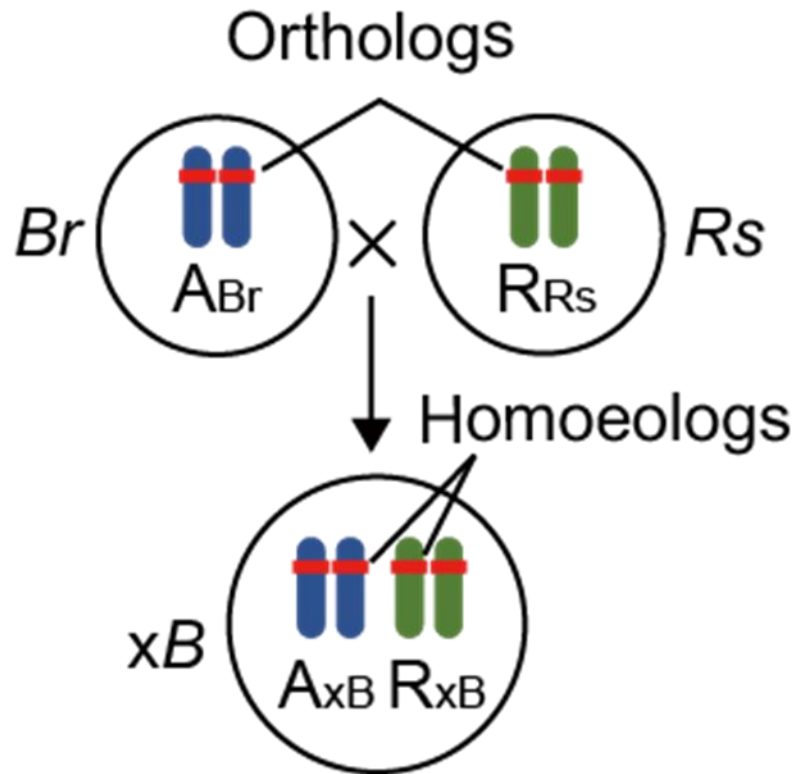


Transcriptome of *xBrassicoraphanus*

# Sequence divergence for interspecific and intergeneric relationships

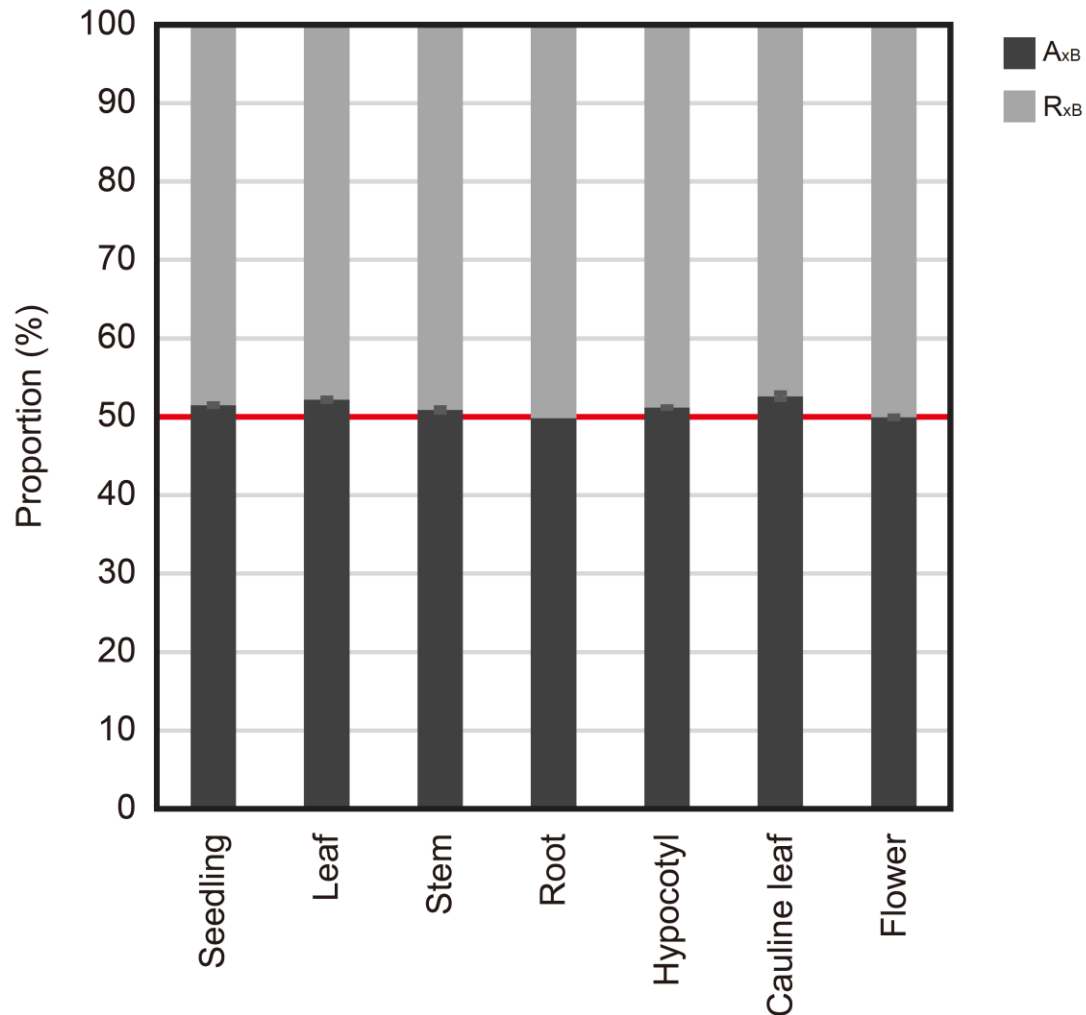


# Relationship between orthologous and homoeologous genes in $xB$ and its progenitors

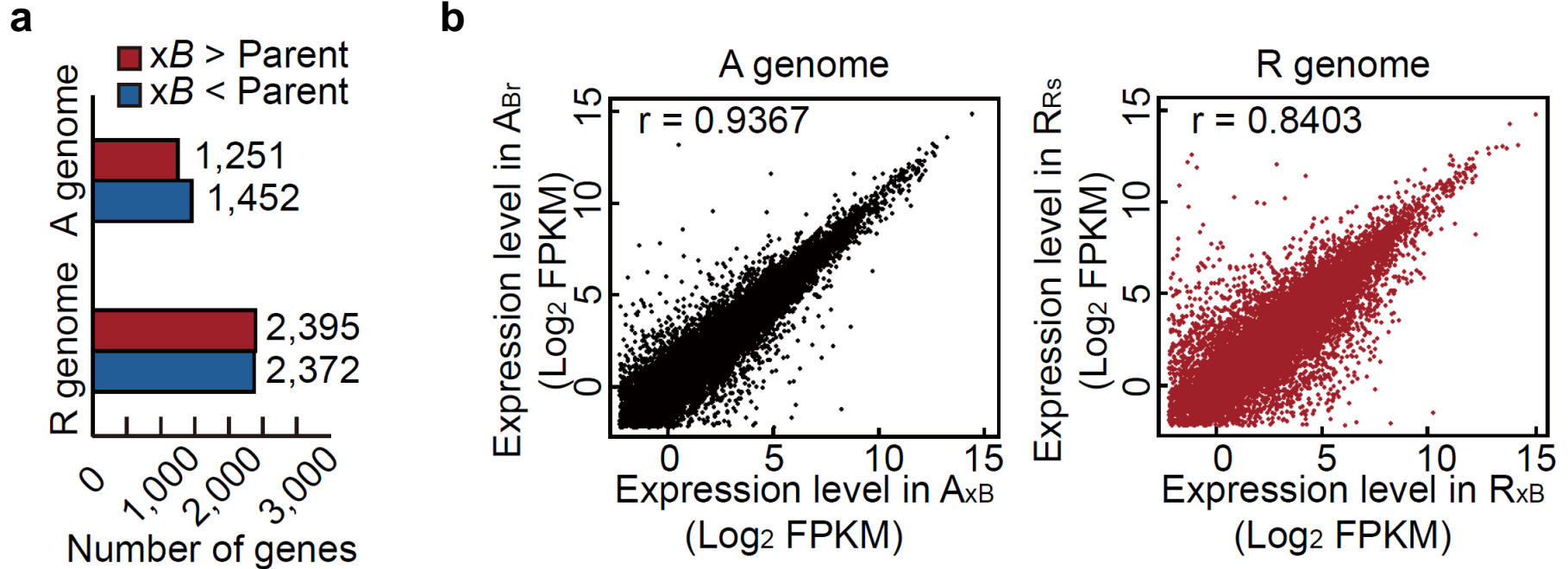


- **Orthologs** are genes in different species that originated by vertical descent from a single gene of the last common ancestor.
- **Homoeologs** are pairs of genes that originated by speciation and were brought back together in the same genome by allopolyploidization.

# Equal contribution of parental genomes to xB transcriptome



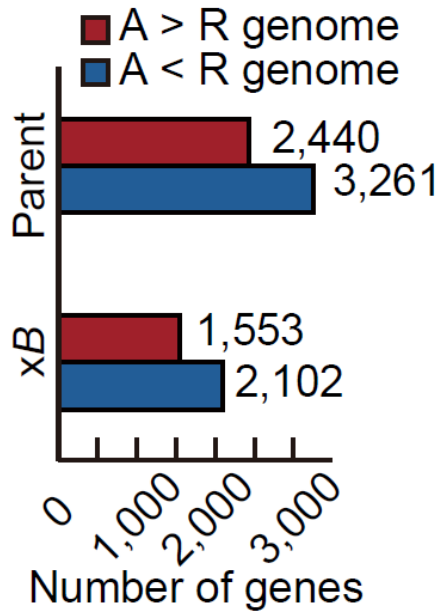
# Genome dominance of A subgenome



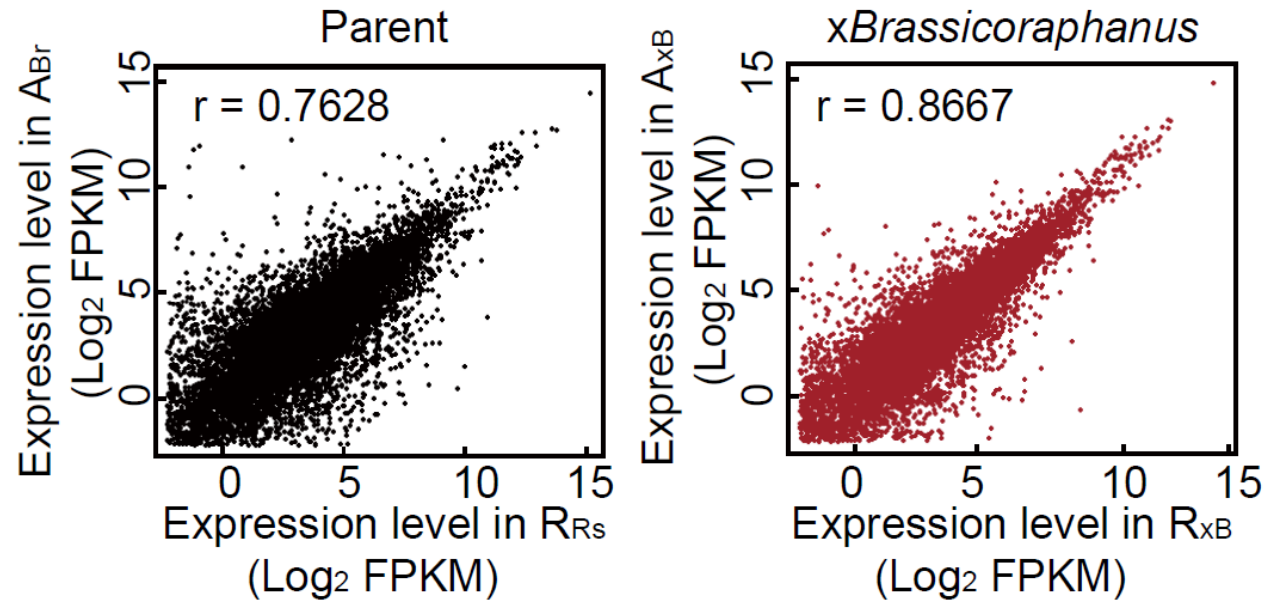
$A_{x_B}$  subgenome displays 'expression level dominance' over the  $R_{x_B}$  subgenome in  $x_B$ , in which expression of  $A_{x_B}$  subgenome is similar to its parental  $A_{B_r}$  genome.

# Expression adjustment of homeologous genes

**a**



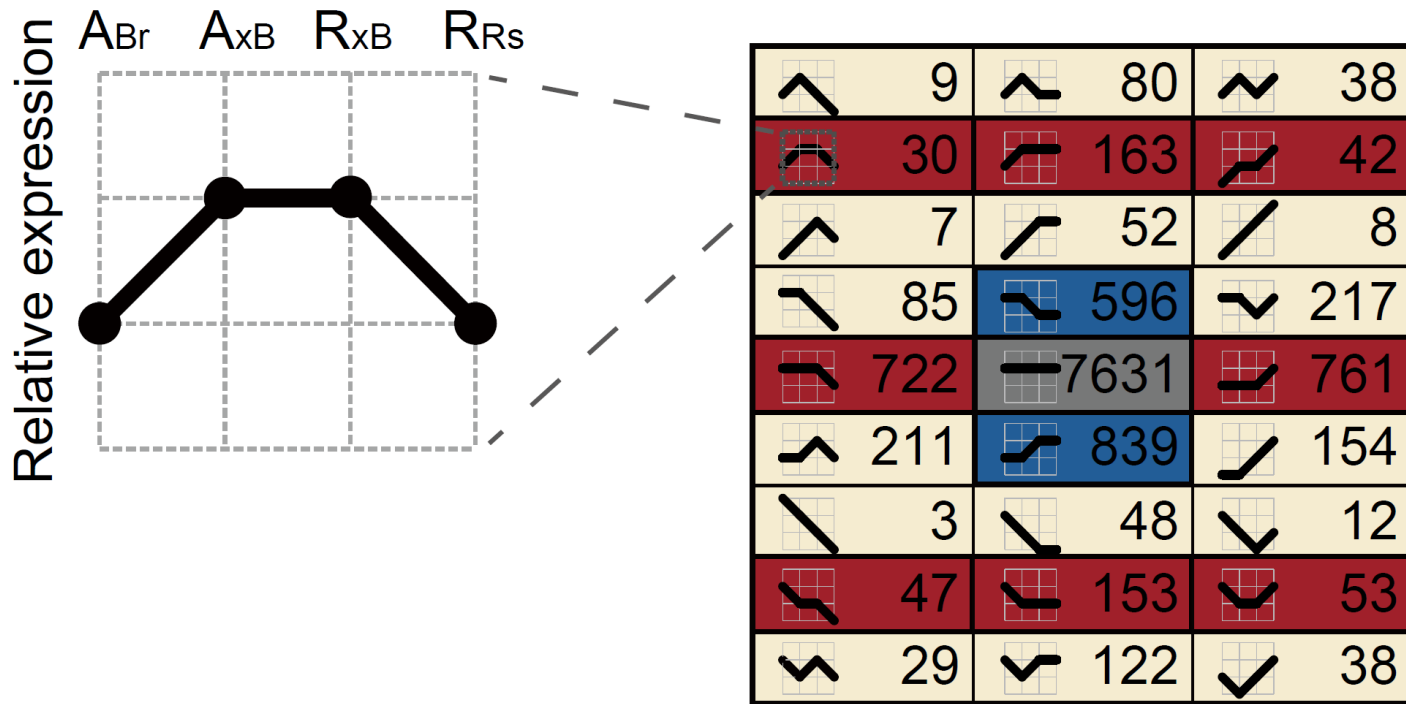
**b**



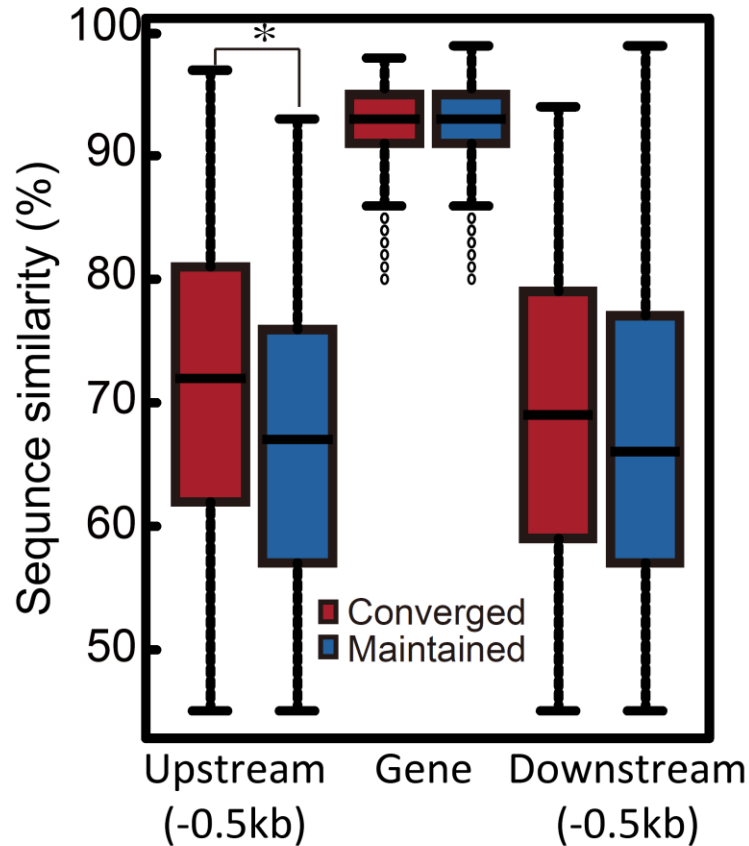
Distinct expressions of many orthologous genes are adjusted to similar levels in the context of homeologous relationship in the hybrid genome.



# Convergent expression of homoeologous genes adjusted to similar levels in the hybrids



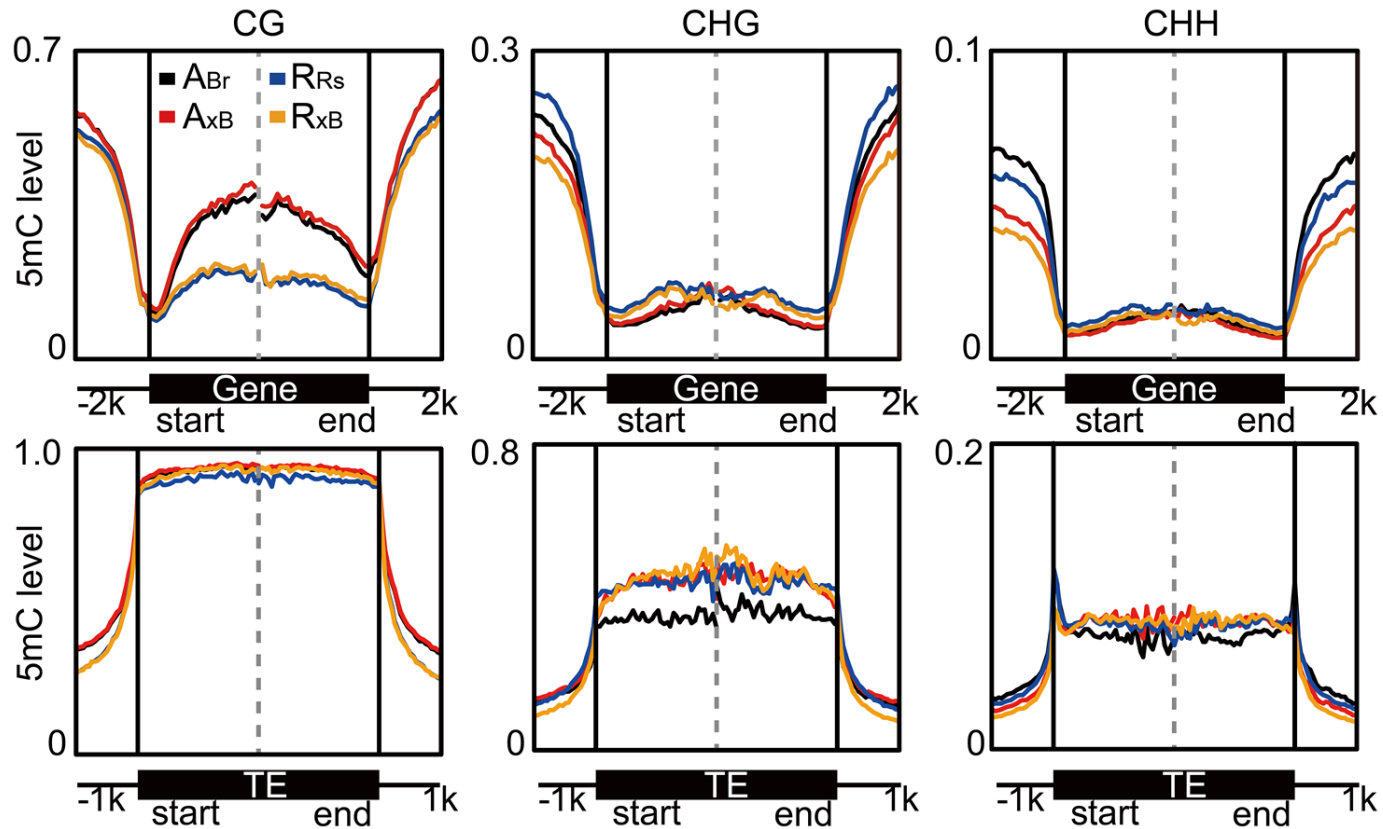
# Convergent expression of homoeologous genes due possibly to similar *cis*-elements



Convergent expression is related to sequence similarity in regulatory regions.

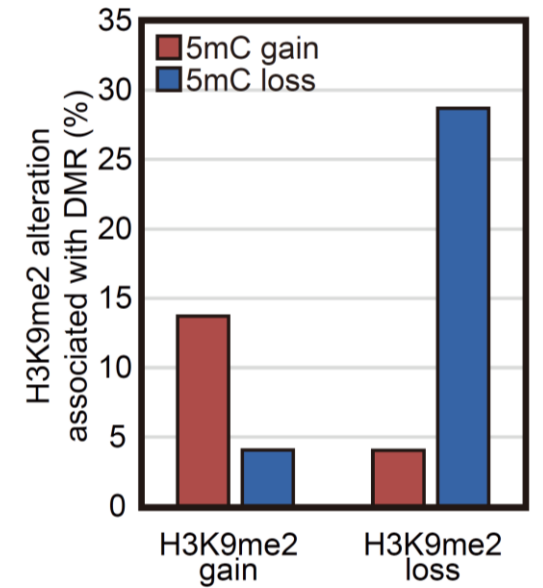
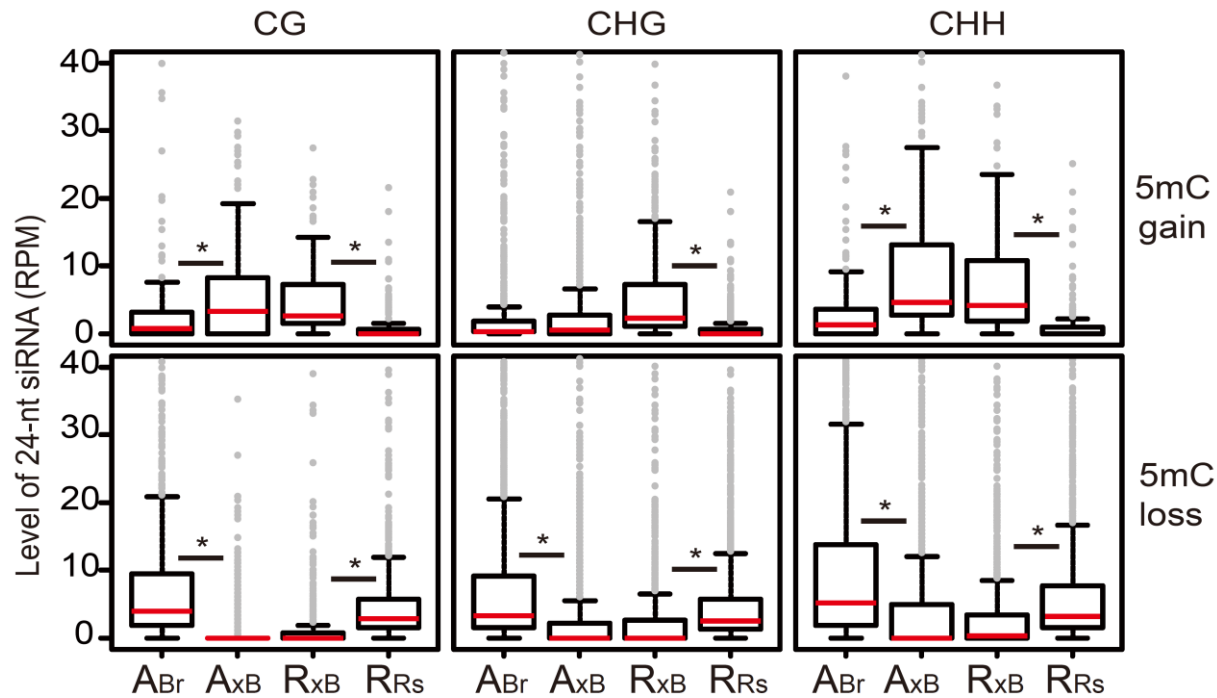
Epigenome of *xBrassicoraphanus*

# DNA methylation levels at genic and TE regions



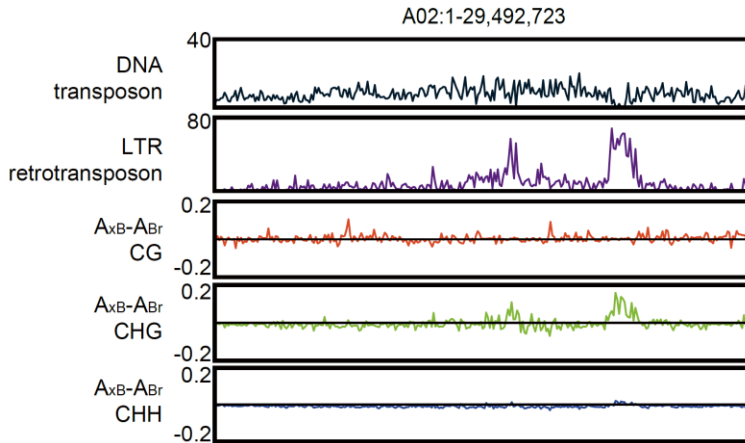
- Gene body DNA methylation is relatively conserved.
- TE methylation levels of the  $A_{xB}$  subgenome had become similar to those of the  $R_{RS}$  subgenome, especially at the CHG contexts.

# Association among DNA methylation, H3K9me2, and siRNA in xB

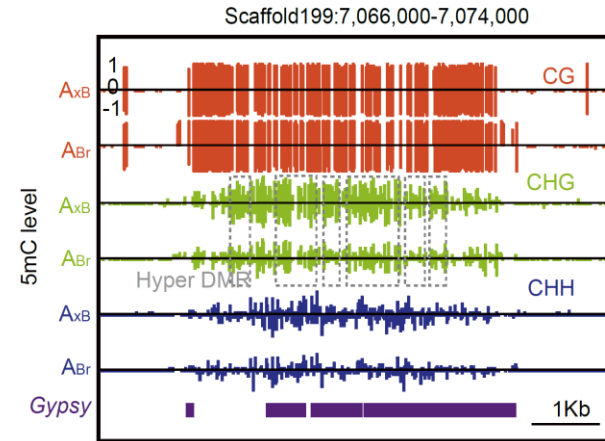


# Hyper methylation in TE regions

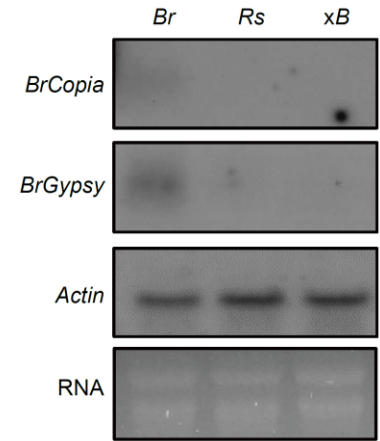
**a**



**b**



**c**



Some LTR retrotransposons become hypermethylated in the hybrid genome with reduced expression.



Intraspecific hybrids between  
closely related species

# Turnip (*Brassica rapa* subsp. *rapa*)

- Chinese cabbage and turnip are of the same species with different morphological traits.
- Turnip is morphologically similar to radish.

**Chinese cabbage**

*Brassica rapa* subsp. *pekinensis*



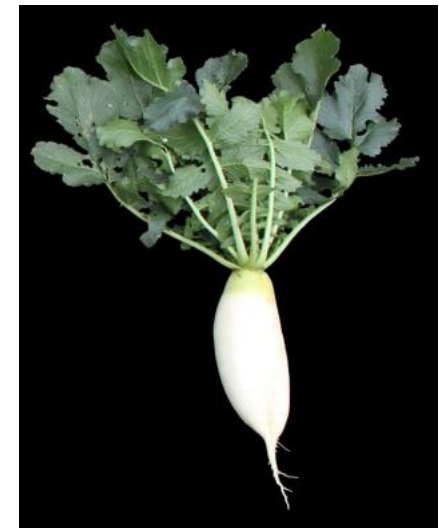
**Turnip**

*Brassica rapa* subsp. *rapa*

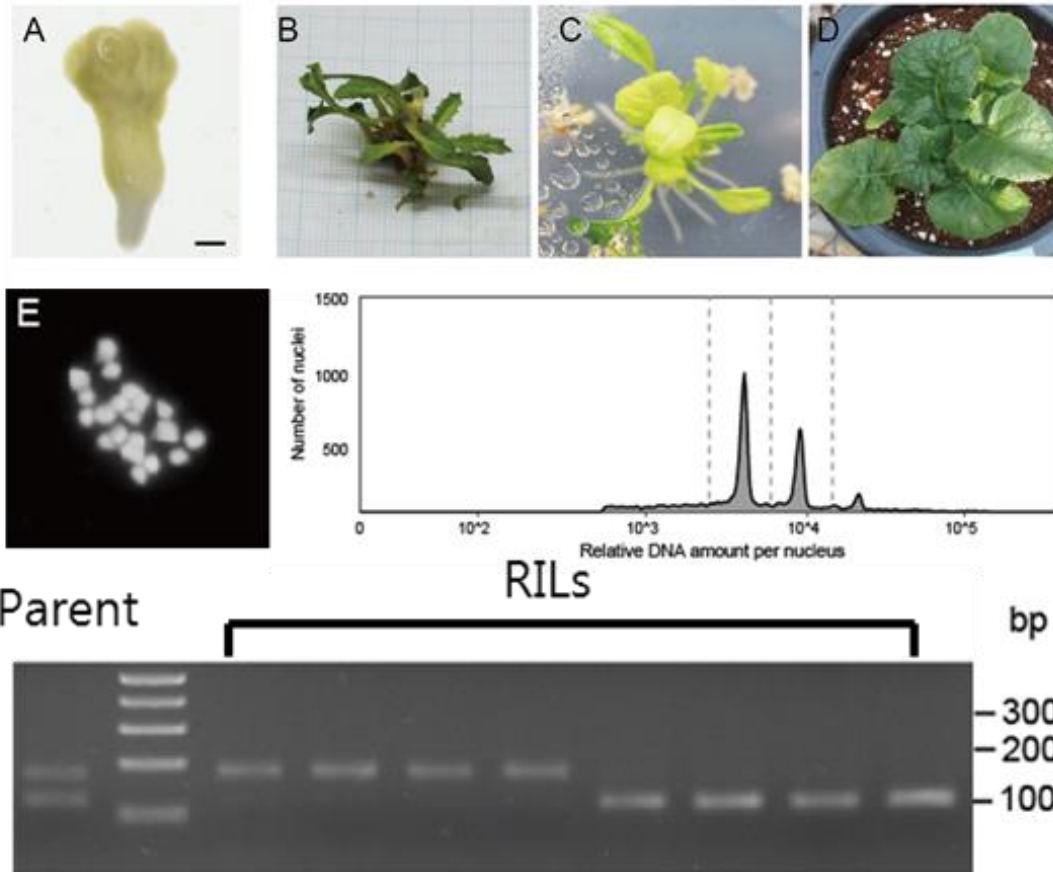


**Radish**

*Raphanus sativus*

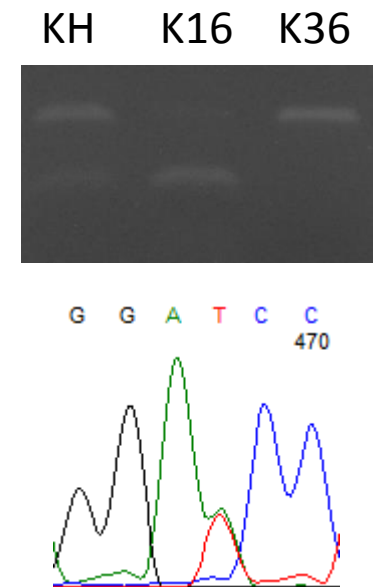
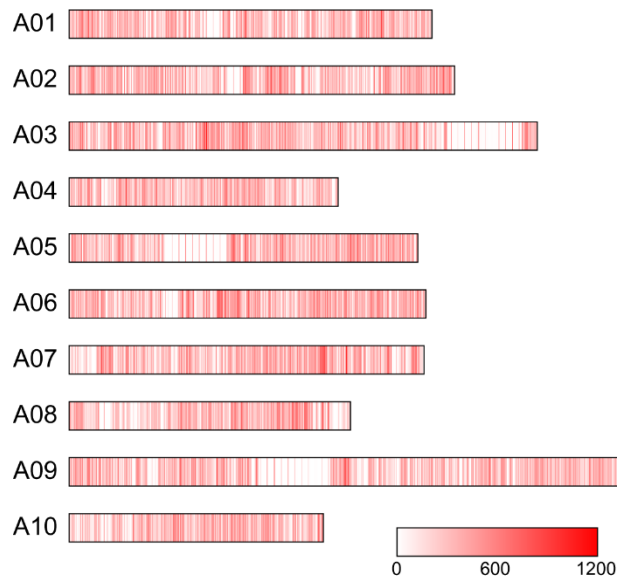


# Doubled haploid (DH) lines of Ganghwa turnip by microspore culture

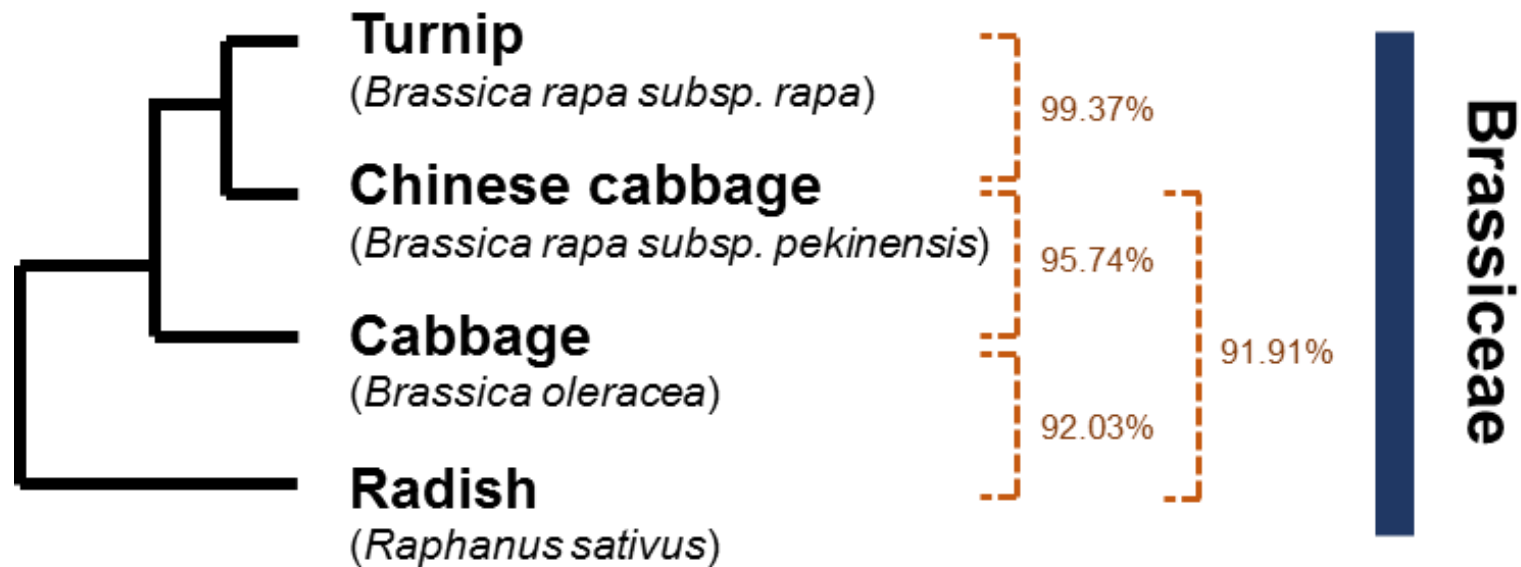


# whole genome sequencing of turnip DH lines

Chromosome	Number of SNPs
A01	179002
A02	199988
A03	223407
A04	137077
A05	170218
A06	200940
A07	185088
A08	136009
A09	246239
A10	127586



# Sequence similarity of coding genes between turnip and Chinese cabbage





Turnip  
*B. Rapa* subsp. *rapa*



Chinese cabbage  
*B. Rapa* subsp. *pekinensis*



F1



# Genome and epigenome analysis of intraspecific hybrid between turnip and Chinese cabbage

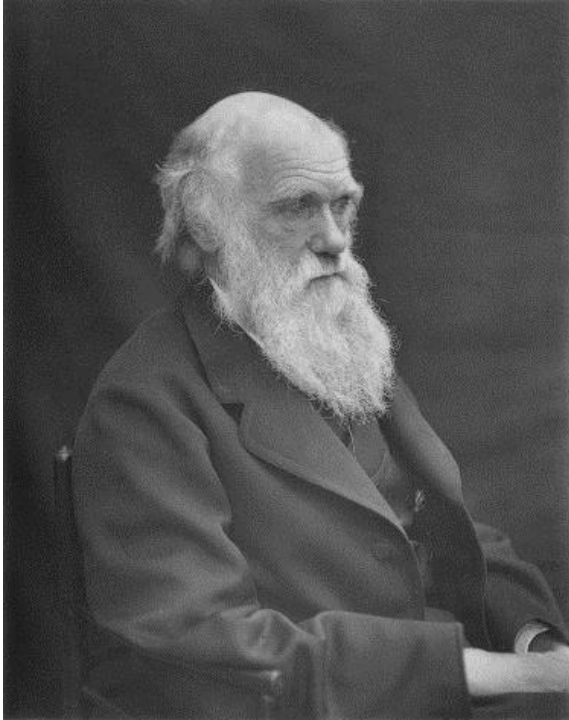
- Genetic analysis in F2 population
- DEG analysis
- DNA methylome analysis
- Small RNAs
- Identification of enhancers (GRO-seq, Hi-C analysis)
- Comparative analysis between intergeneric and intraspecific hybrids

# Summary

- *xBrassicoraphanus* has whole sets of *B. rapa* and *R. sativus* without noticeable genome reconstruction.
- Many homoeologs tend to be similarly expressed in *xBrassicoraphanus*, showing **convergent expression**, which is likely to ameliorate a transcriptome shock.
- Increase in DNA methylation via siRNA may inhibit TE activity to prevent a genomic shock.
- A certain degree of **genome divergence** has an advantage for hybrid genome stability by suppressing nonhomologous chromosome exchanges.

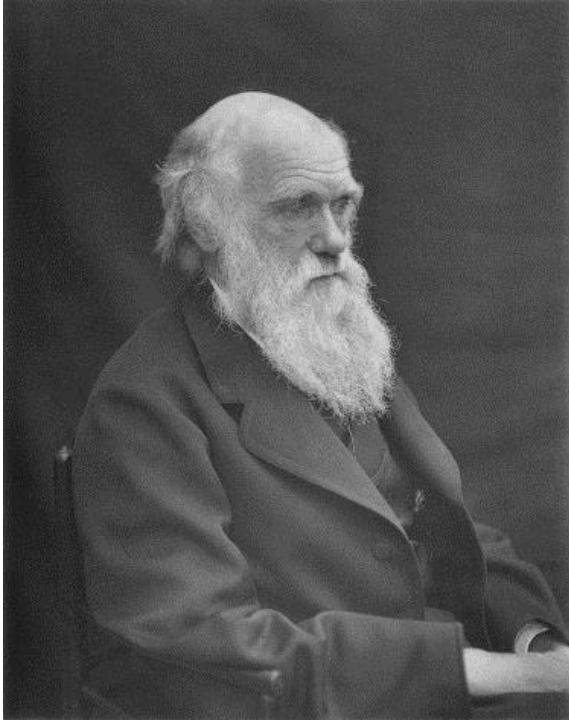
Insights into hybridization,  
polyploidization and evolution of  
angiosperms

# Darwinism and Abominable Mystery



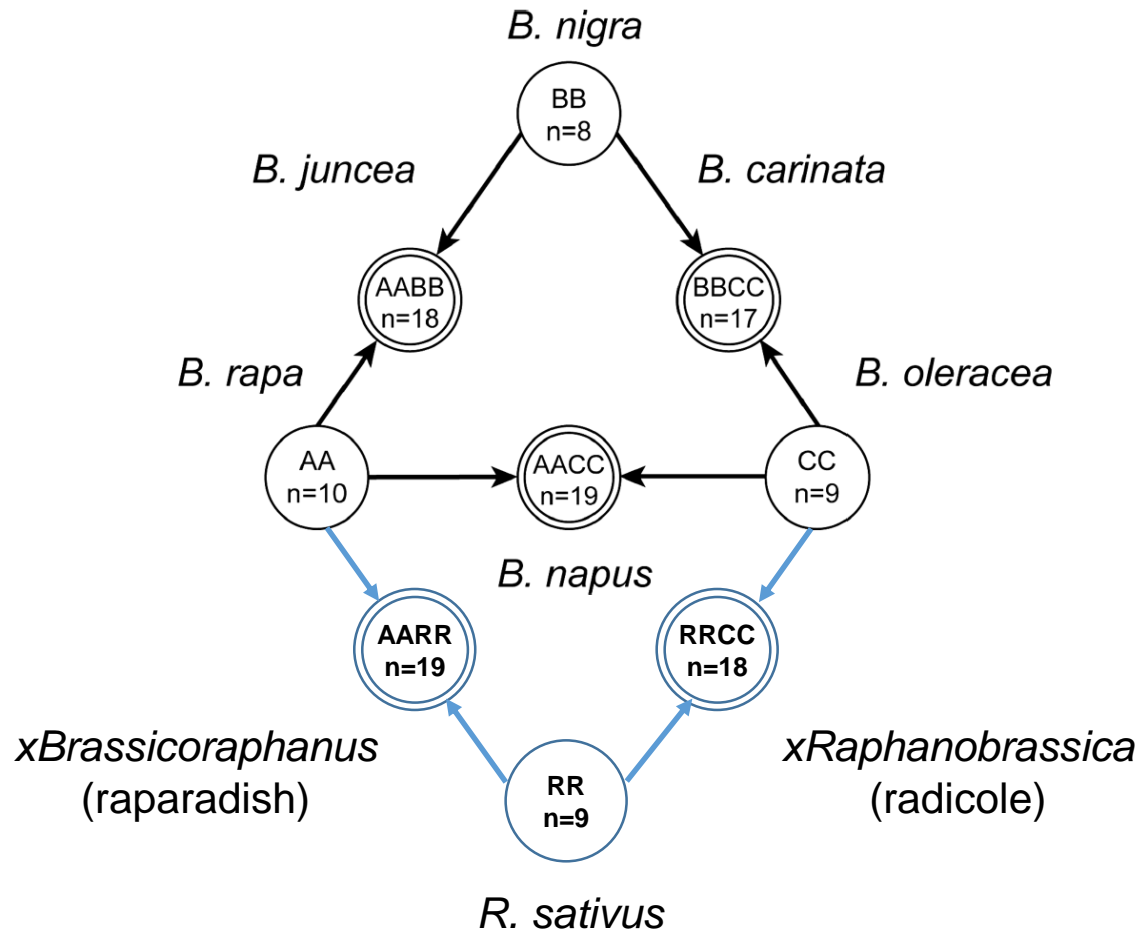
**Darwinism** is a theory of biological evolution developed by the English naturalist Charles Darwin (1809–1882) and others, stating that **all species of organisms arise and develop through the natural selection of small, inherited variations that increase the individual's ability to compete, survive, and reproduce.**

# Darwinism and Abominable Mystery



Darwin was deeply bothered by what he perceived to be an **abrupt origin and highly accelerated rate of diversification of flowering plants in the mid-Cretaceous**. This led Darwin to create speculative arguments for a long, gradual, and undiscovered pre-Cretaceous history of flowering plants on a lost island or continent.

# Intergeneric hybrids in the *Brassicaceae* family



# Acknowledgements



- **Plant Developmental Genetics Laboratory:** Hosub Shin, Gibum Yi, Jeong-Eun Park, Hyerang Park, Woo Lee Choi, Seung Hwa Yu, Joo Young Lim, Junghyo Kim, June-Sik Kim, Hye Yeon Soh, Taegu Kang, Ye Seul Kim, Hyeon Min Jeong, Wonjun Koh; **Biobreeding Institute:** Soo-Seong Lee; **Sahmyook University:** Hyun Hee Kim, Hadassah Roa Belandres; **SNU:** Doil Choi, Seung-Il Kim, Tae-Jin Yang, Kyung-Hee Kim; **KRIBB:** Yong Min Kim, Namjin Koo; **RDA:** Beom Seok Park, Soo Hyung Park, Jin-A Kim; **CNU:** Yong Pyo Lim; **MIT:** Mary Gehring; **NCSU:** Tzung-Fu Hsieh
- Supported by “Next Generation BioGreen 21 Project” and “Next Generation Multi-Ministerial Genome Project” by **RDA**, Korea, and “Agri-Bio Industry Technology Development Program” by **IPET**, MAFRA, Korea





This work is dedicated to late Dr. Woo Jang-Choon (1898–1959) for his 60th memorial anniversary.