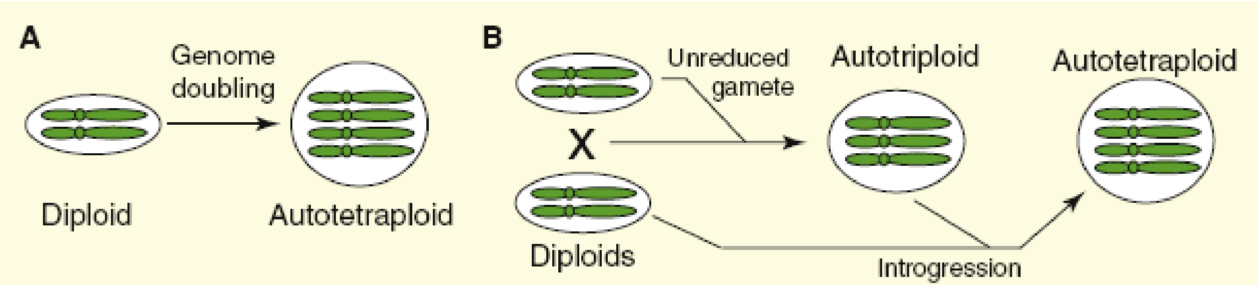
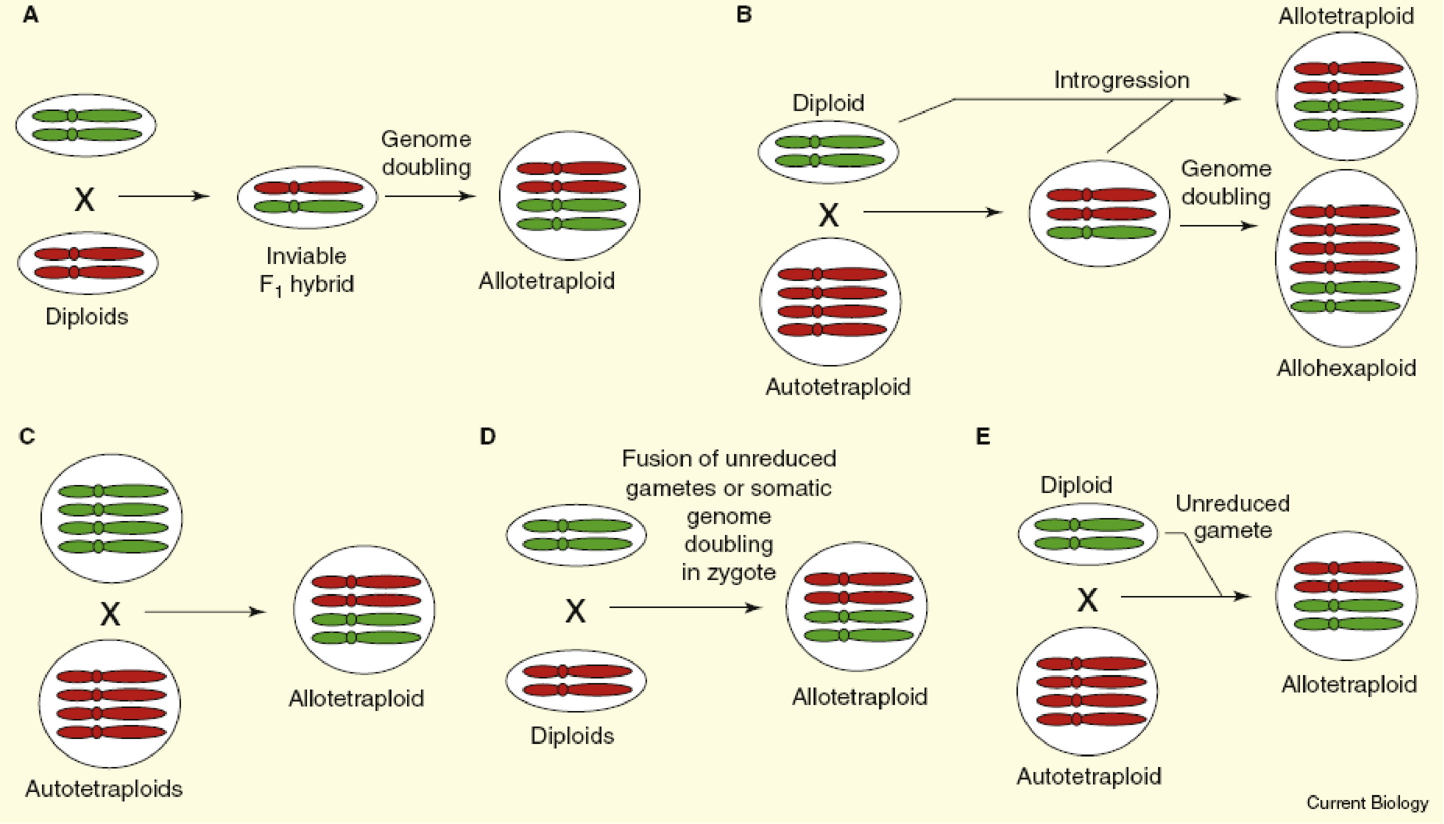
(Paleo)Polyploidy – When Things Get Bigger

Whole-genome duplications



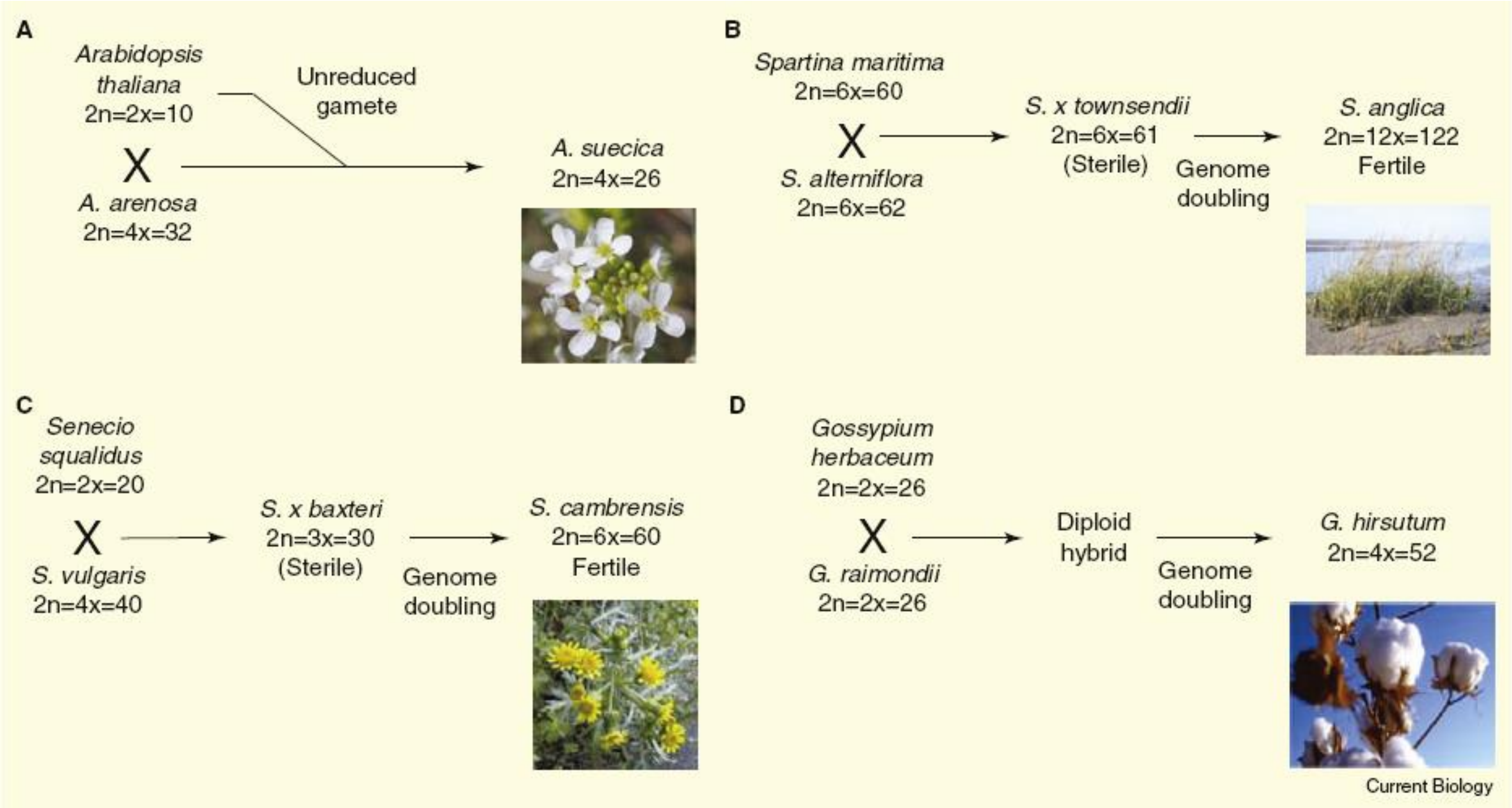
AUTOPOLYPLOIDY

ALLOPOLYPLOIDY



Hegarty and Hiscock 2008, Current Biology 18

Examples of allopolyploid speciation



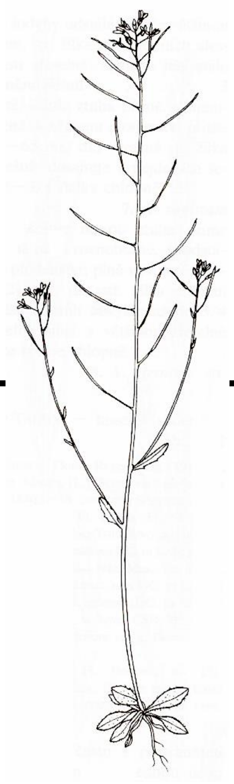
Hegarty and Hiscock 2008,
Current Biology 18

Evolutionary significance of polyploidy

Leitch & Leitch (2008) *Science* 320

Whole-genome duplications of different age

time



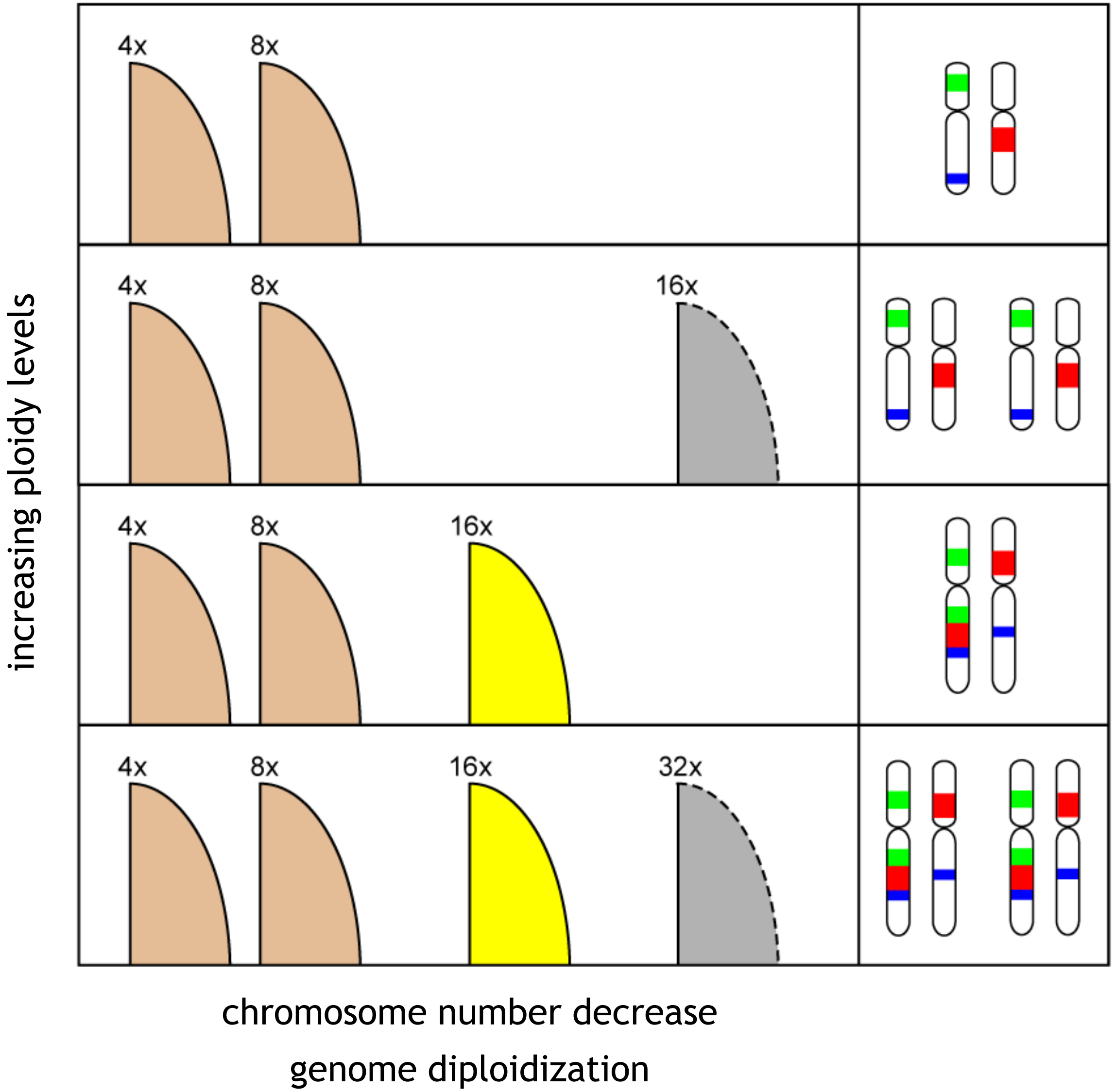
neopolyploidy

mesopolyploidy

paleopolyploidy

time

Multiple Whole-Genome Duplication Events, Diploidization and Extant Karyotype Structure

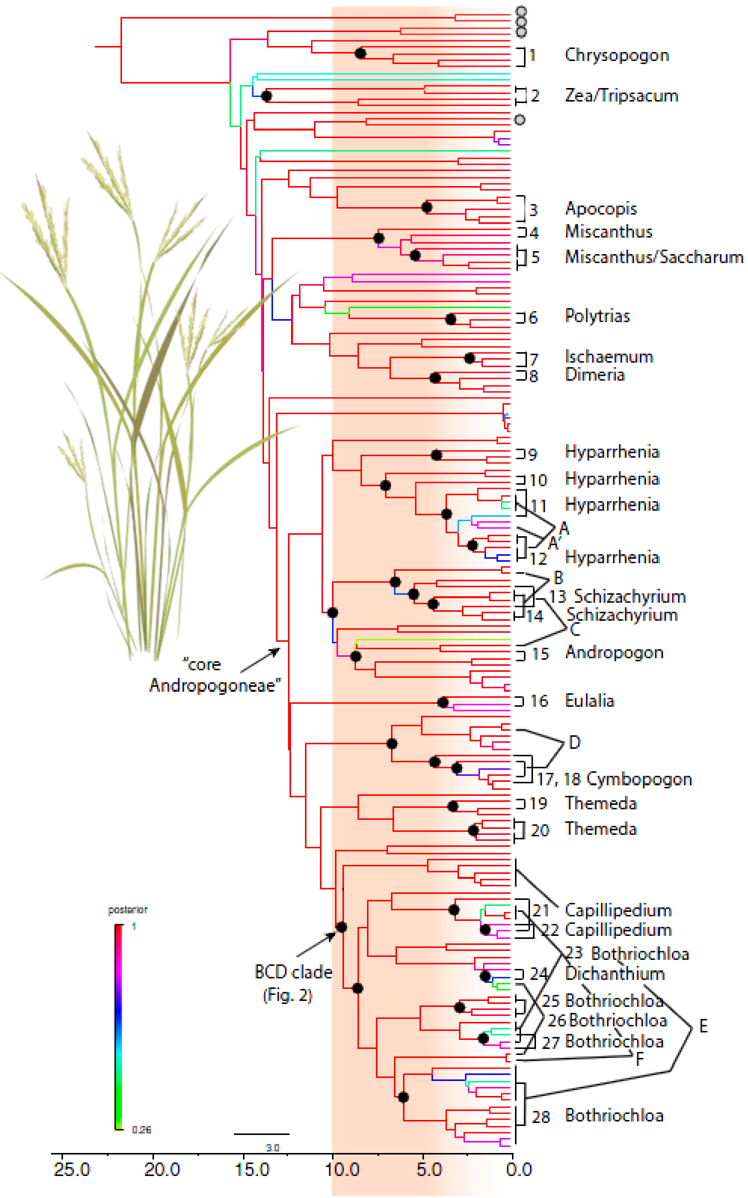


PALEOpolyploidy

NEOpolyploidy

MESOpolyploidy

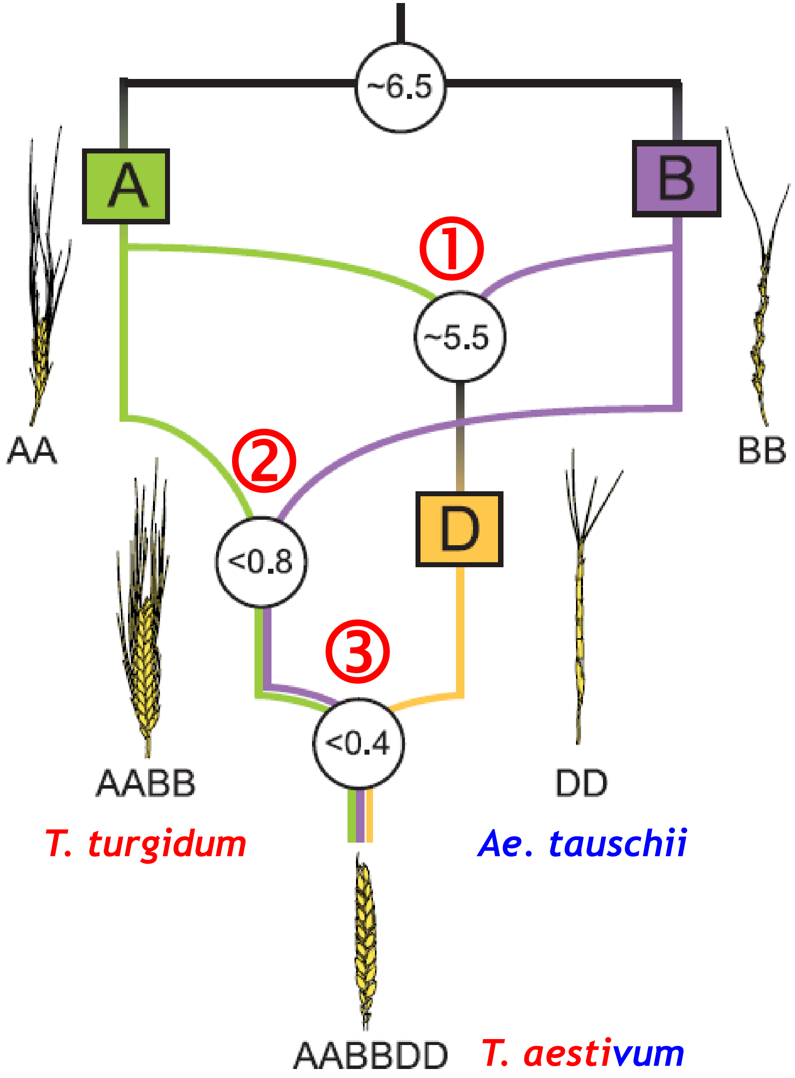
Allopolyploidy, diversification, and the Miocene grassland expansion



- Most of the allopolyploidization events identified here occurred in the Late Miocene, simultaneous with or following the well documented expansion of the C4 grasslands.
- The dominant species of modern C4 grasslands are members of Andropogoneae, and most are allopolyploid. Many of these ecological dominants whose origin is dated to about 10.5 million years ago (mya) correlates closely with the date when C4 species came to dominate grasslands in Africa and Southern Asia (Pakistan), also estimated about 10–11 mya; the expansion in North America is dated about 7 mya.
- Allopolyploidy is thus correlated with ecological success.

Estep et al., PNAS (2014)

Model of the phylogenetic history of bread wheat (*Triticum aestivum*; AABBDD). Three rounds of hybridization/polyploidy.



Marcussen et al. (2014), *Science*

Whole-genome duplications in protozoa

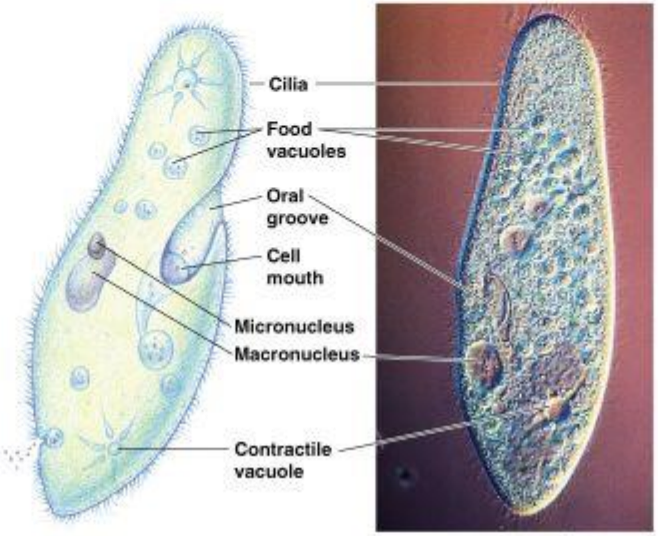
- Aury *et al.* (2006) analyzed the unicellular eukaryote *Paramecium tetraurelia*

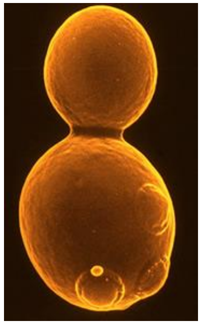
- most of 40,000 genes arose through at least 3 successive whole-genome duplications (WGDs)

- most recent duplication most likely caused an explosion of speciation events that gave rise to the *P. aurelia* complex (15 sibling species)

- some genes have been lost, some retained

- many retained (duplicated) genes do not generate functional innovations but are important because of the gene dosage effect





Whole-genome duplications in yeast

- genome comparison between two yeast species, *Saccharomyces cerevisiae* (n=16) and *Kluyveromyces waltii* (n=8)

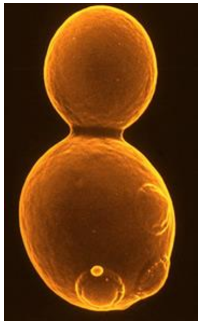
- each region of *K. waltii* corresponding to two regions of *S. cerevisiae*

- the *S. cerevisiae* genome underwent a WGD after the two yeast species diverged

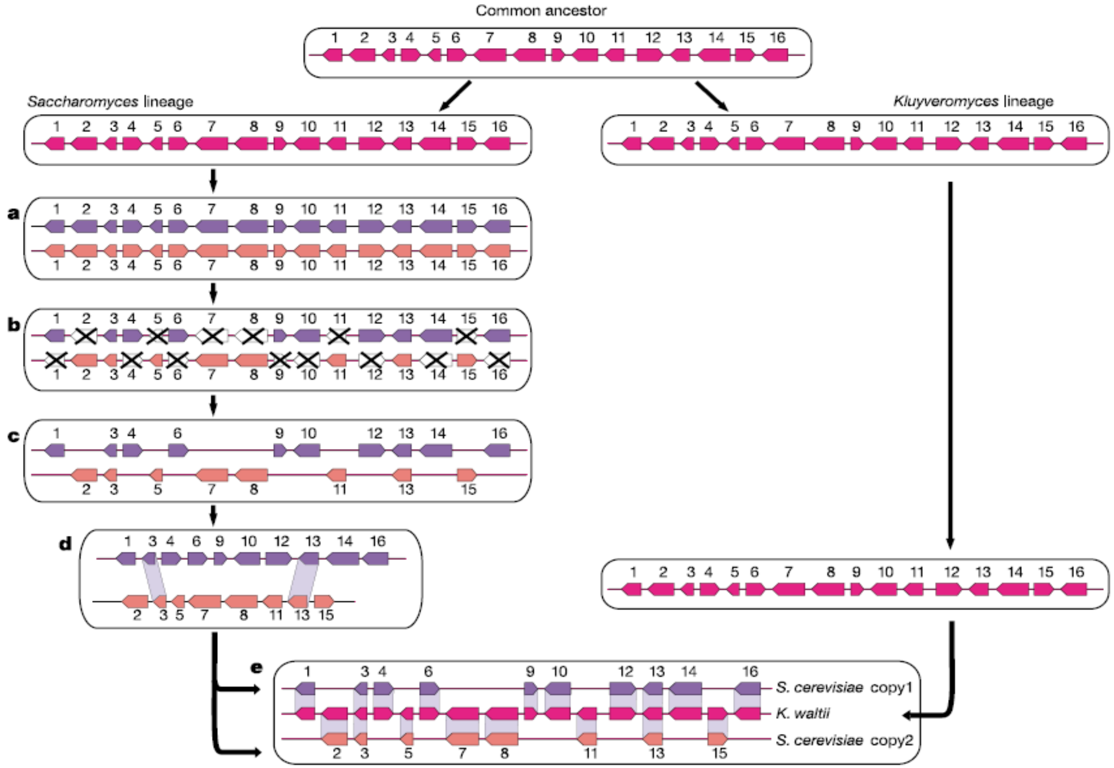
- in nearly every case (95%), accelerated evolution was confined to only one of the two paralogues (= one of the paralogues retained an ancestral function, the other was free to evolve more rapidly and acquired a derived function)

Kellis *et al.* 2004, Nature 428

Whole-genome duplications in yeast



Kellis *et al.* 2004, Nature 428



a) after divergence from *K. waltii*, the *Saccharomyces* lineage underwent a genome duplication event (2 copies of every gene and chromosome)

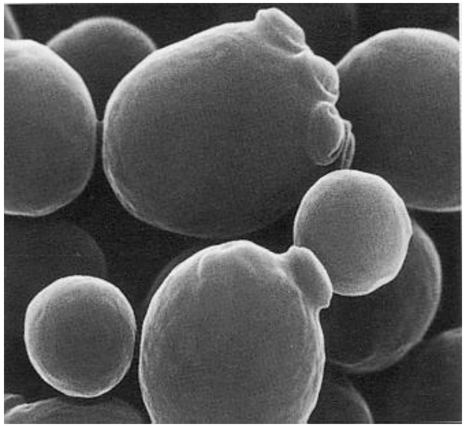
b) duplicated genes were mutated and some lost

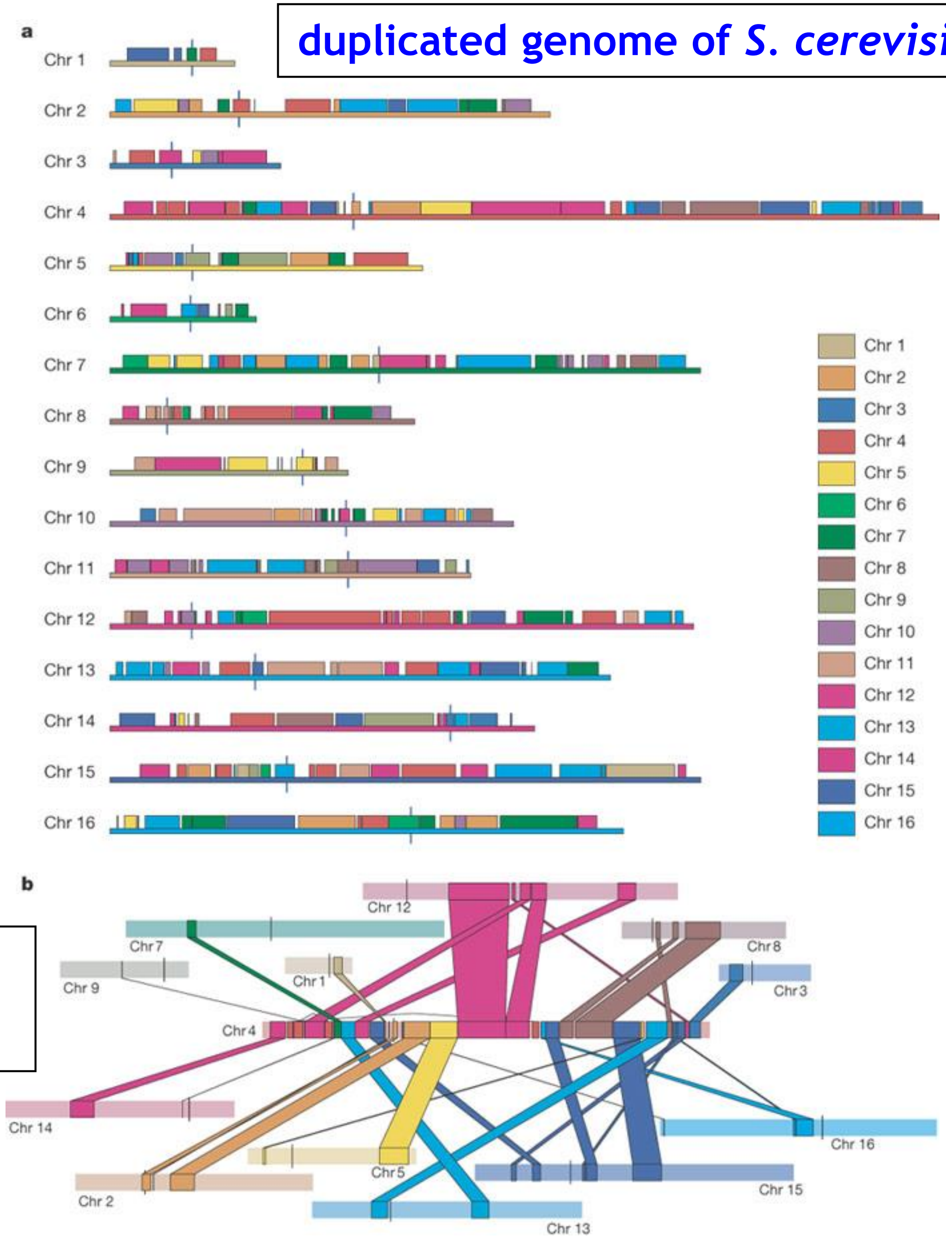
c) two copies kept for only a small minority of duplicated genes

d) the conserved order of duplicated genes (nos. 3-13) across different chromosomal segments

e) comparison between genomes of *S. cerevisiae* and *K. waltii* reveals the duplicated nature of the *S. cerevisiae* genome

Duplicated nature of the *S. cerevisiae* genome





duplicated genome of *S. cerevisiae*

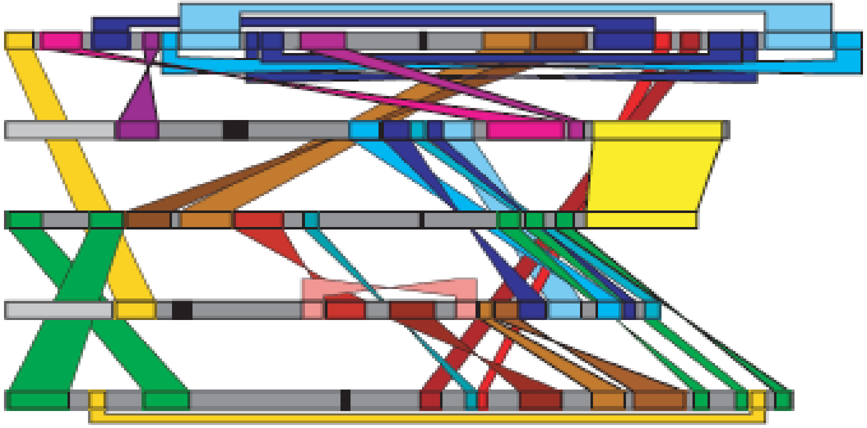
***S. cerevisiae* chromosome 4 with sister regions in other chromosomes**

Kellis *et al.* 2004, Nature 428

First evidence of a WGD in plants

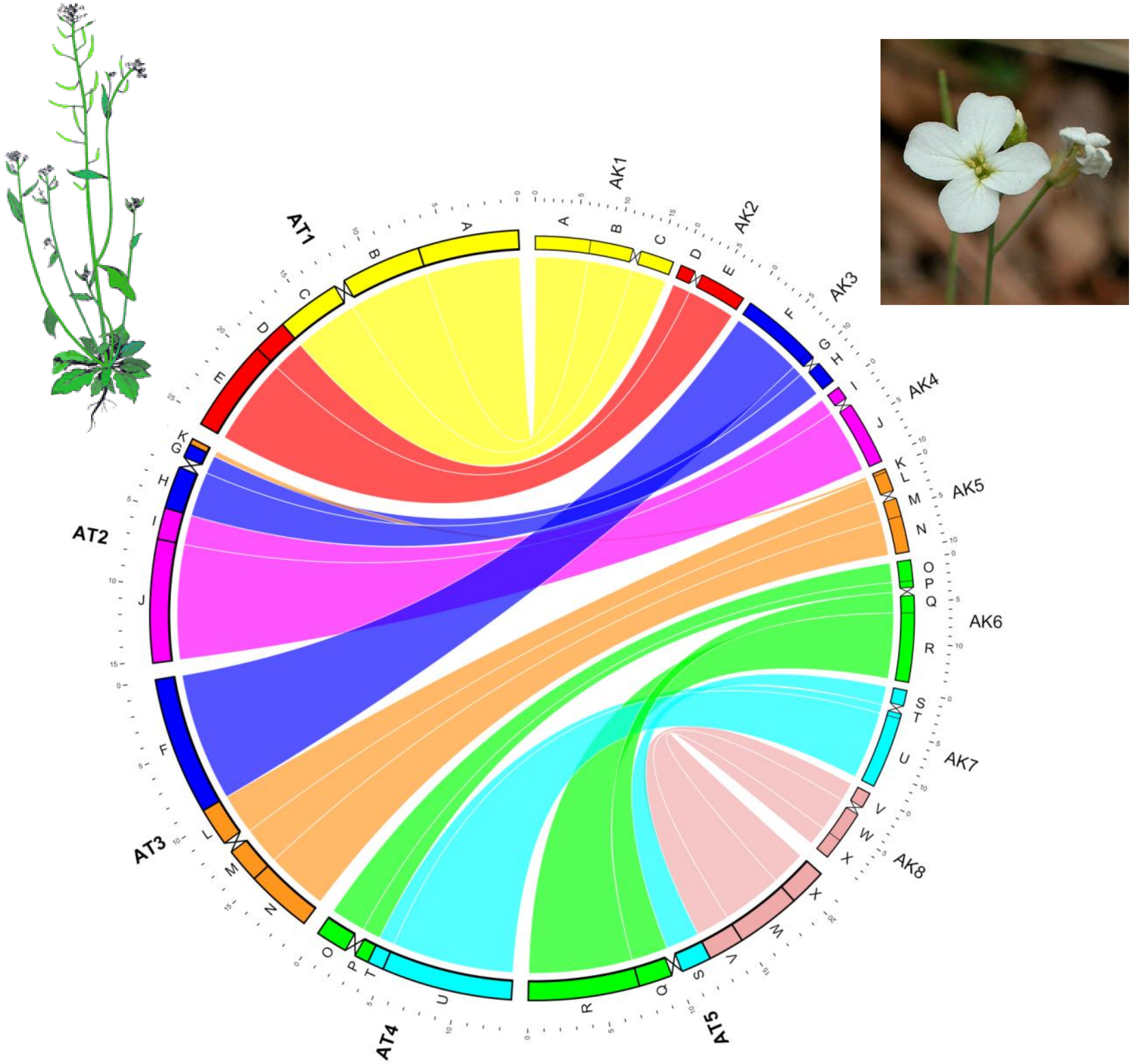
Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*

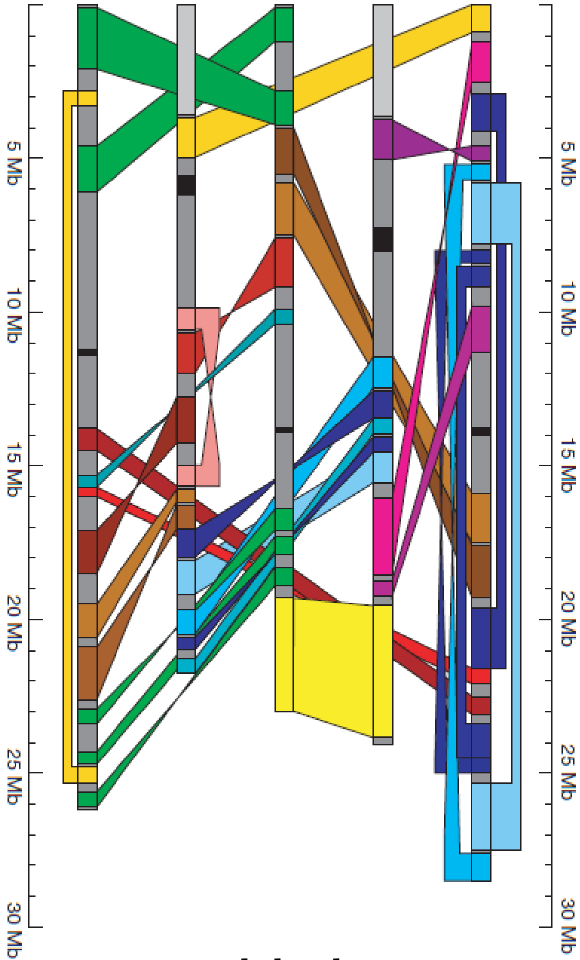
The Arabidopsis Genome Initiative* AGI (2000)



What does the duplication in the Arabidopsis genome tell us about the ancestry of the species? As the majority of the Arabidopsis genome is represented in duplicated (but not triplicated) segments, it appears most likely that Arabidopsis, like maize, had a tetraploid ancestor. …The diploid genetics of Arabidopsis and the extensive divergence of the duplicated segments have masked its evolutionary history.

Arabidopsis Species Are „Paleotetraploids“ with 8 or 5 Chromosomes





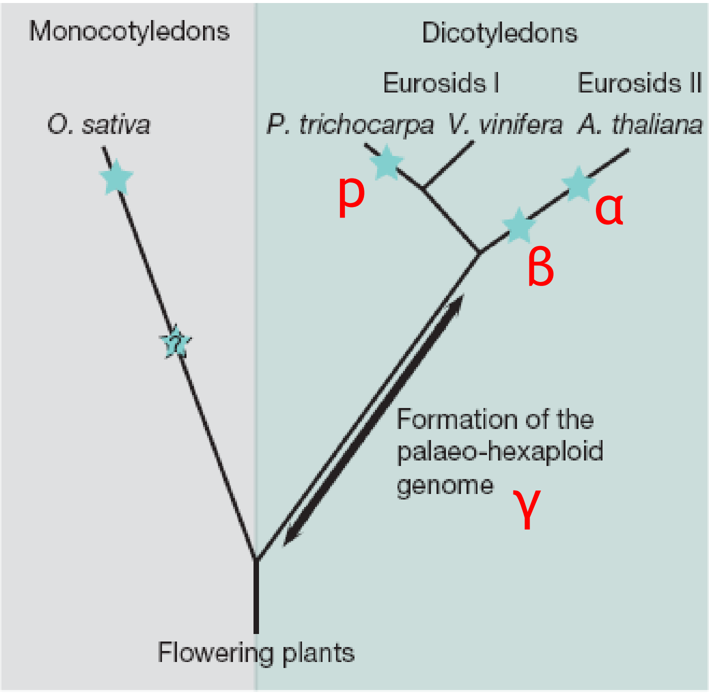
segmental duplications in the *A. thaliana* genome

AGI (2000) Nature, Hu et al. (2011) Nat Genet

The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla

The French–Italian Public Consortium for Grapevine Genome Characterization*

Nature 449, 2007

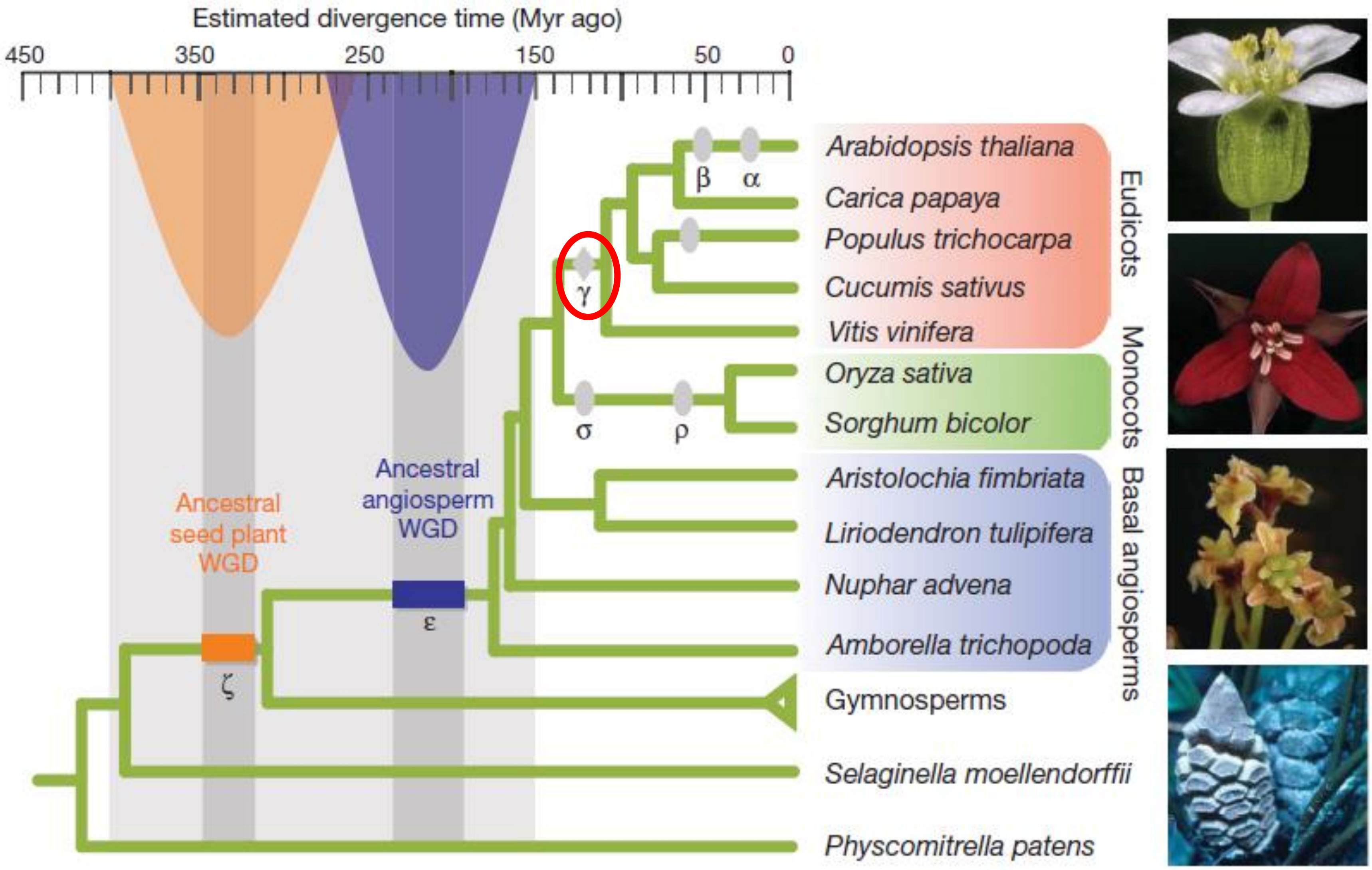


The formation of the palaeo-hexaploid ancestral genome occurred after divergence from monocots and before the radiation of the Eurosids. Star = a WGD (tetraploidization) event.

The γ **triplication** may have been an ancient **auto-hexaploidy** formed from fusions of three identical genomes, or **allo-hexaploidy** formed from fusions of three somewhat diverged genomes.

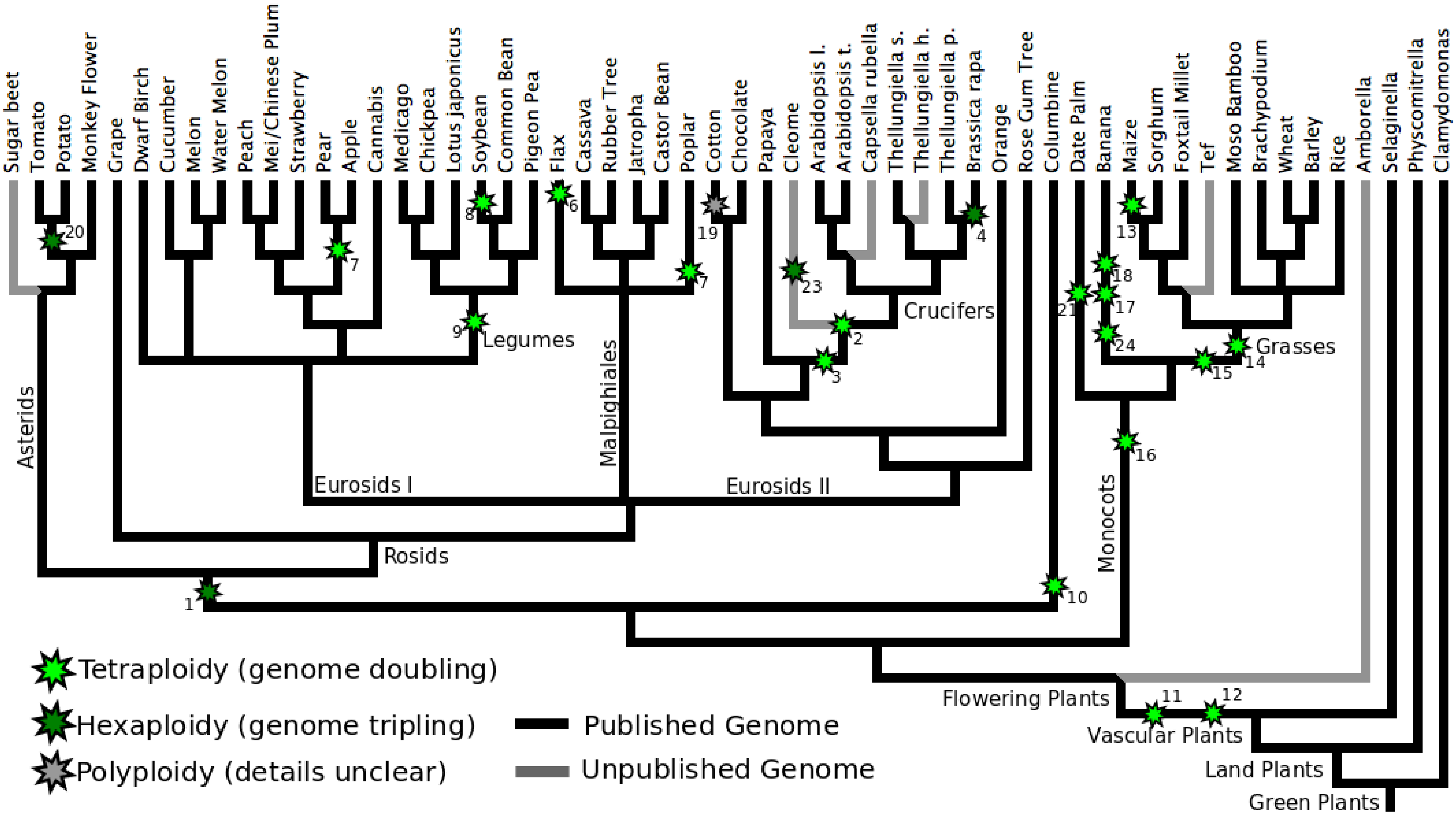
Tang *et al.* 2008, Genome Research

WGD events in seed plants and angiosperms



Jiao et al. (2011) Nature

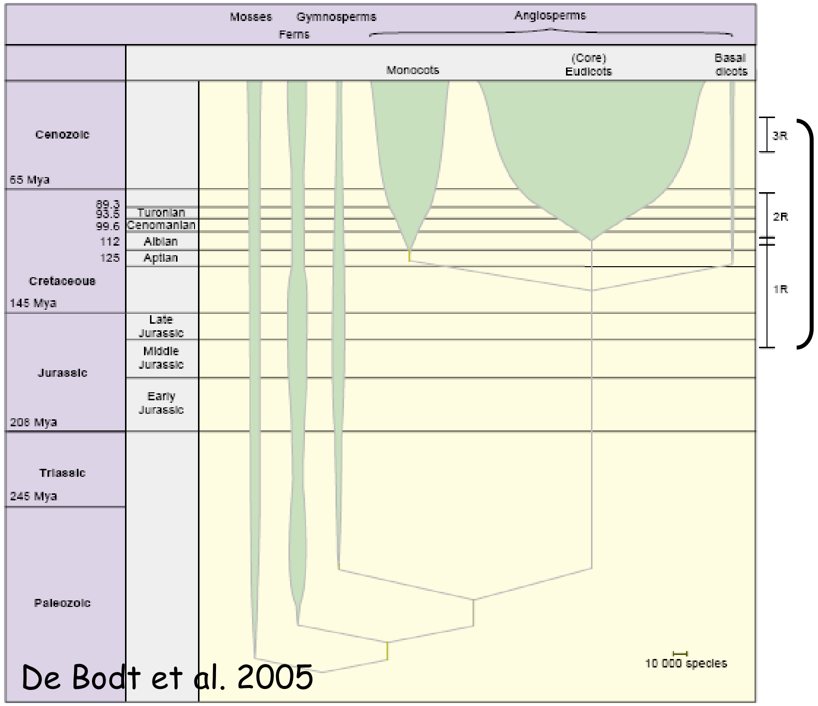
Phylogenetic Tree of Sequenced Genomes with Whole Genome Duplications Marked



CoGePedia (http://genomevolution.org/wiki/)

Charles Darwin's abominable mystery solved?

"The rapid development as far as we can judge of all the higher plants within recent geological times is an abominable mystery."
(Charles Darwin in a letter to Sir Joseph Hooker, 1879)



De Bodt et al. 2005

assumed ancient whole-genome duplication events

Archaefructus liaoningensis
(140 million year old fossil)

The leaf-like structures on the stem are pods containing the seeds, a characteristic unique to flowering plants.

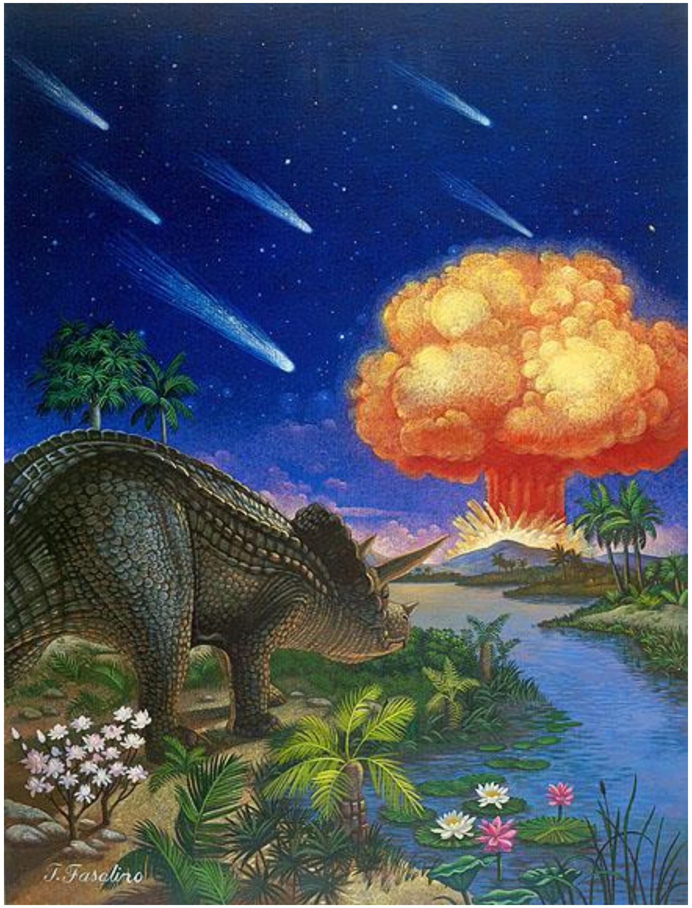
Theres is evidence of ancient polyploidy throughout the major angiosperm lineages. It means that a genome-scale duplication event probably occurred PRIOR to the rapid diversification of flowering plants

Plants with double genomes might have had a better chance to survive the Cretaceous–Tertiary extinction event

Jeffrey A. Fawcett[a,b,1], Steven Maere[a,b,1], and Yves Van de Peer[a,b,2]

[a]Department of Plant Systems Biology, Flanders Institute for Biotechnology, 9052 Gent, Belgium; and [b]Department of Plant Biotechnology and Genetics, Ghent University, 9052 Gent, Belgium

PNAS 106 (2009)



Could WGD event(s) help plants to survive the mass extinction (one or more catastrophic events such as a massive asteroid impact) at the Cretaceous-Tertiary boundary ?

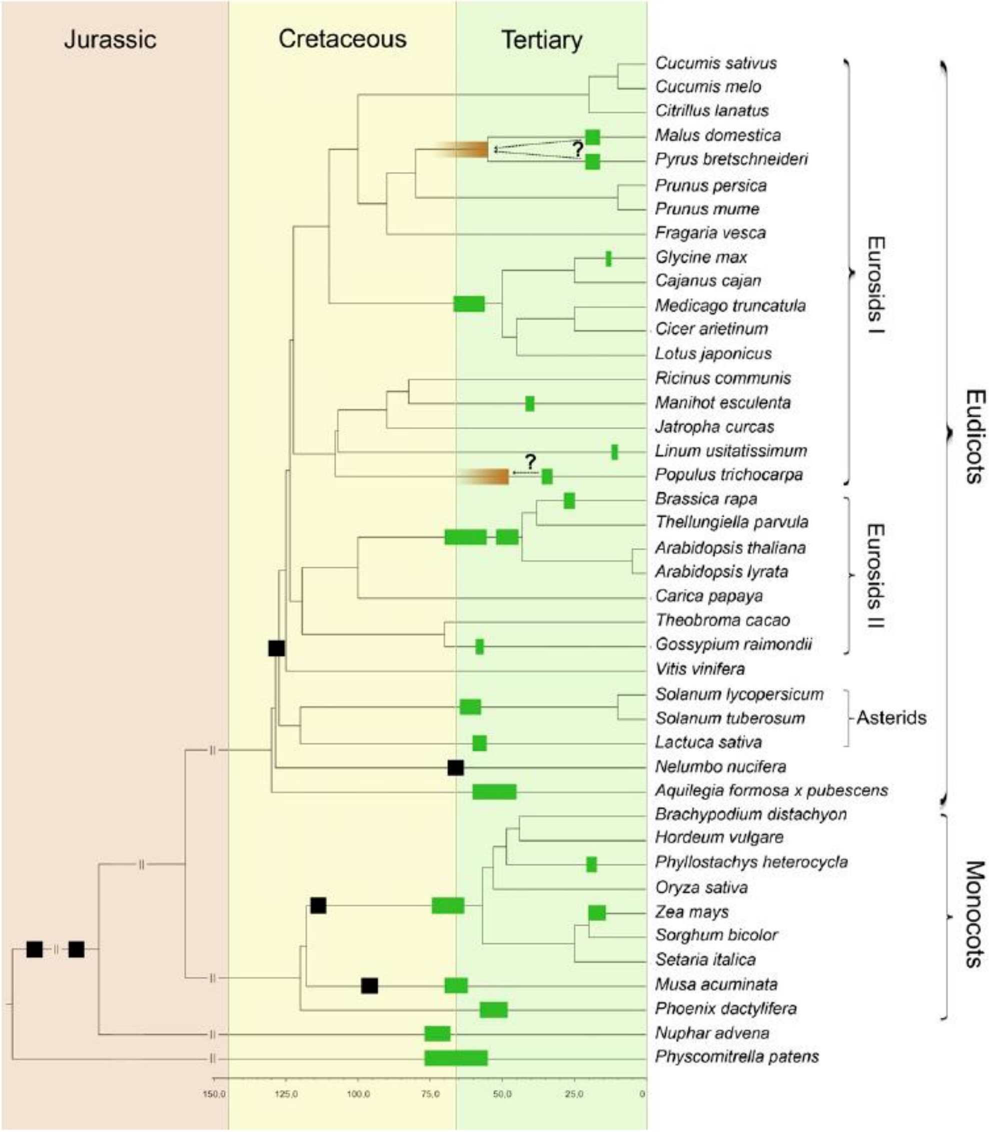
Analysis of 41 plant genomes supports a wave of successful genome duplications in association with the Cretaceous–Paleogene boundary

Kevin Vanneste,[1,2] Guy Baele,[3] Steven Maere,[1,2] and Yves Van de Peer[1,2,4]

Genome Res (2014)

Phylogenetic tree of flowering plants with assumed WGD events

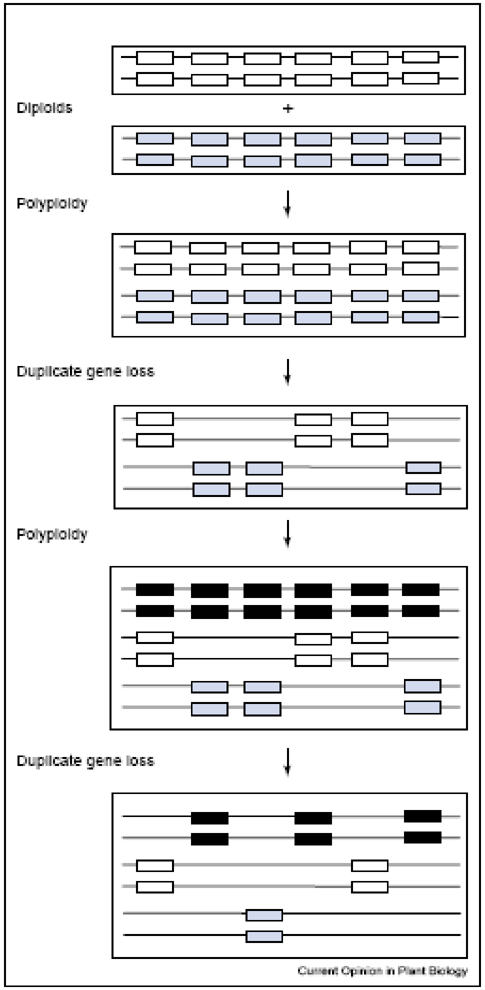
- WGDs clustered around the Cretaceous-Tertiary (KT) boundary
- the KT extinction event - the most recent mass extinction (one or more catastrophic events such as a massive asteroid impact and/or increased volcanic activity)
- the KT extinction event - extinction of 60% of plant species, as well as a majority of animals, including dinosaurs



Fawcett et al., PNAS (2009), **Vanneste et al., Genome Res (2014)**

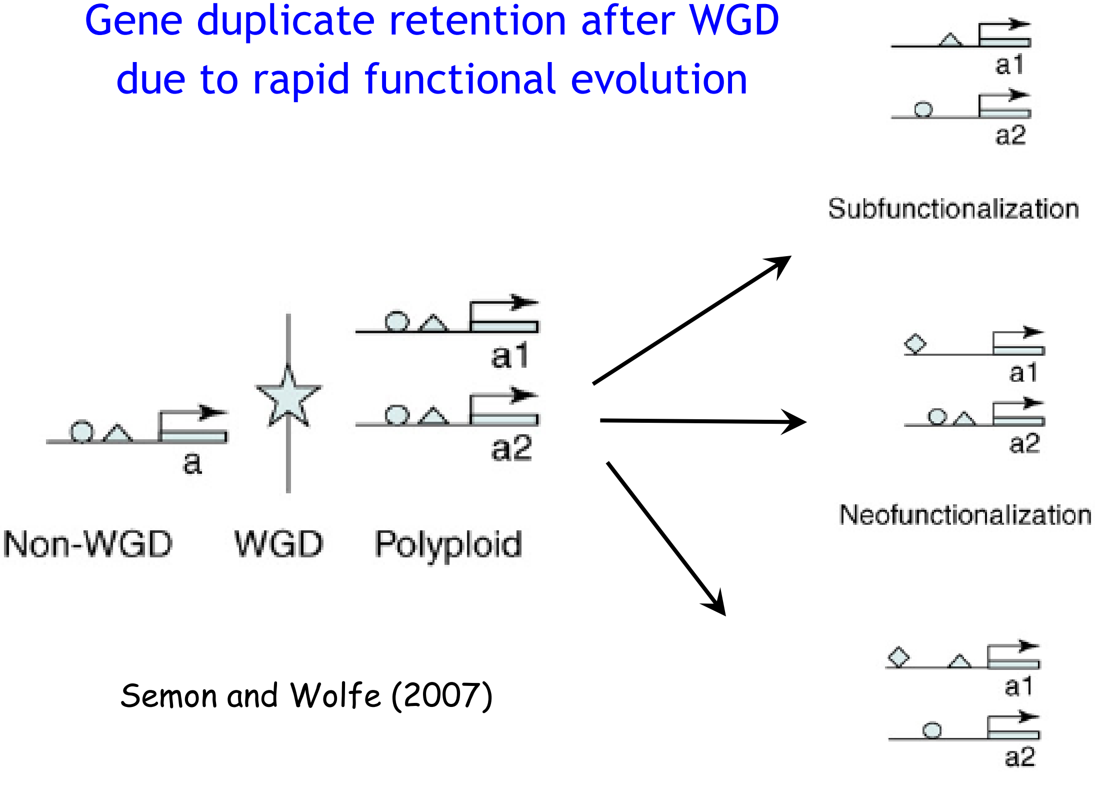
Whole-genome duplication, diploidization, and the consequences

Genome evolution through cyclic polyploidy

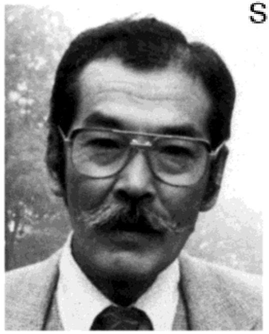


Adams and Wendel (2005)

Gene duplicate retention after WGD due to rapid functional evolution

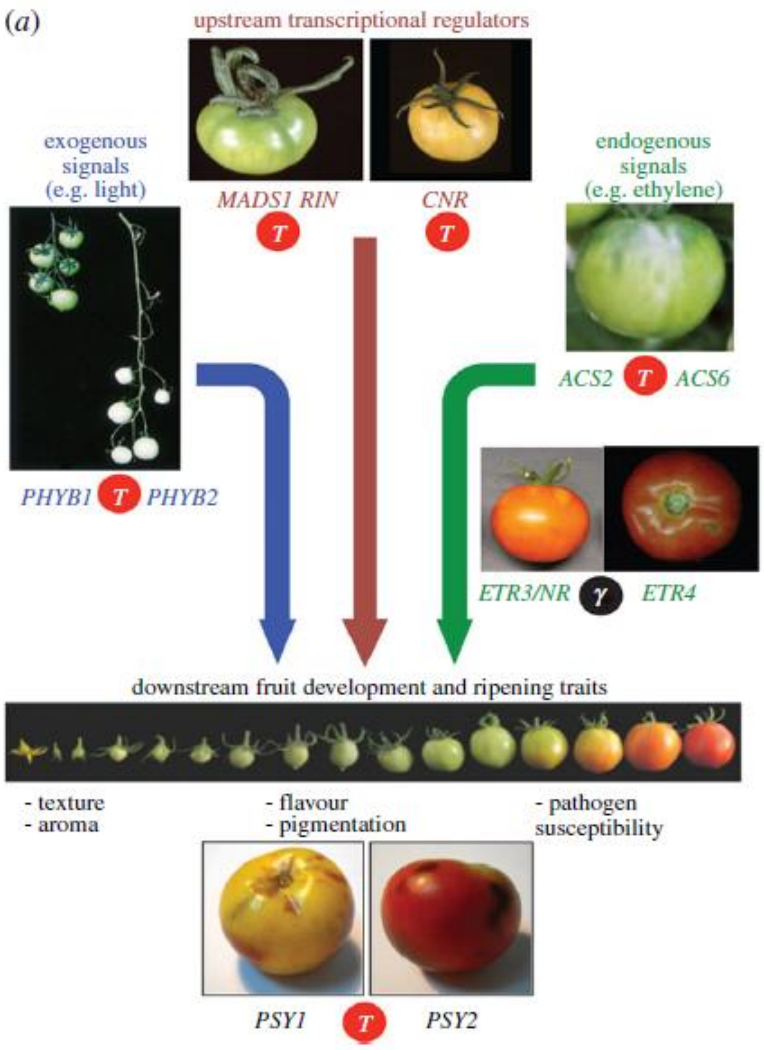


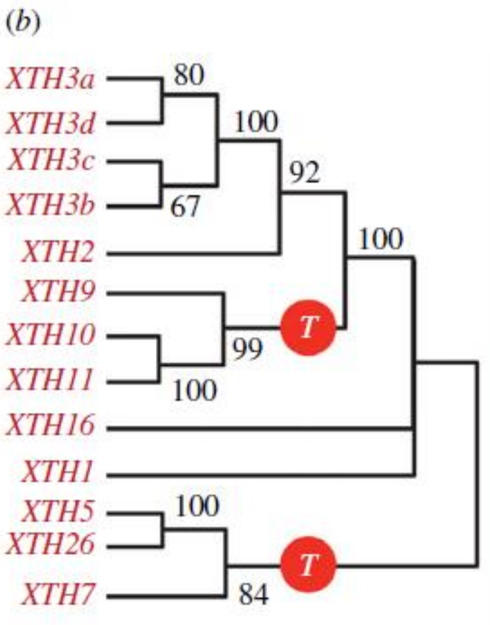
Semon and Wolfe (2007)



Susumu Ohno

Consequences of WGD events: the Solanaceae-specific genome triplication contributed to the evolution of the tomato fruit





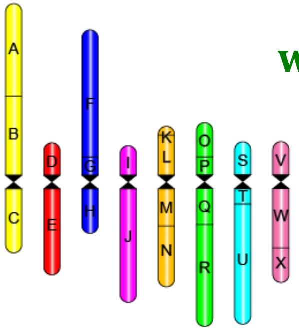
phylogeny of xyloglucan endotransglucosylase/hydrolases (XTHs)

T Solanaceae-speficific genome triplication

γ core eudicot shared hexaploidy

Evolution of the Ancestral Crucifer Genome –

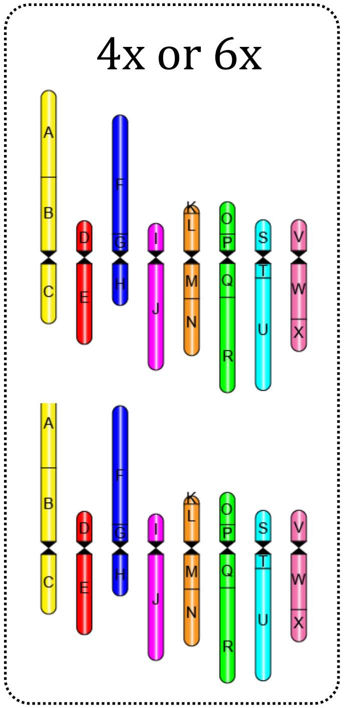
ANCIENT POLYPLOIDS



n = 8

(n = 7)

whole-genome duplication



n = (15) 16 or (21) 24

descending dysploidy

n = 8

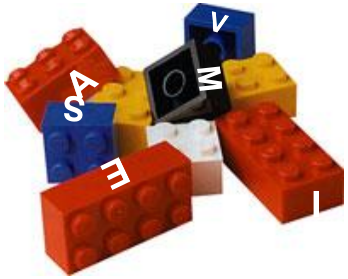
n = 7

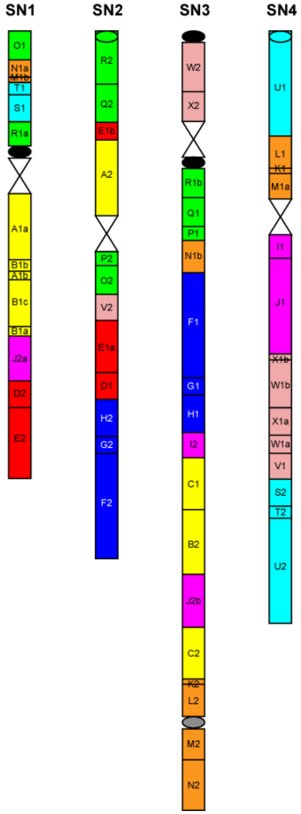
n = 6

n = 5

n = 4

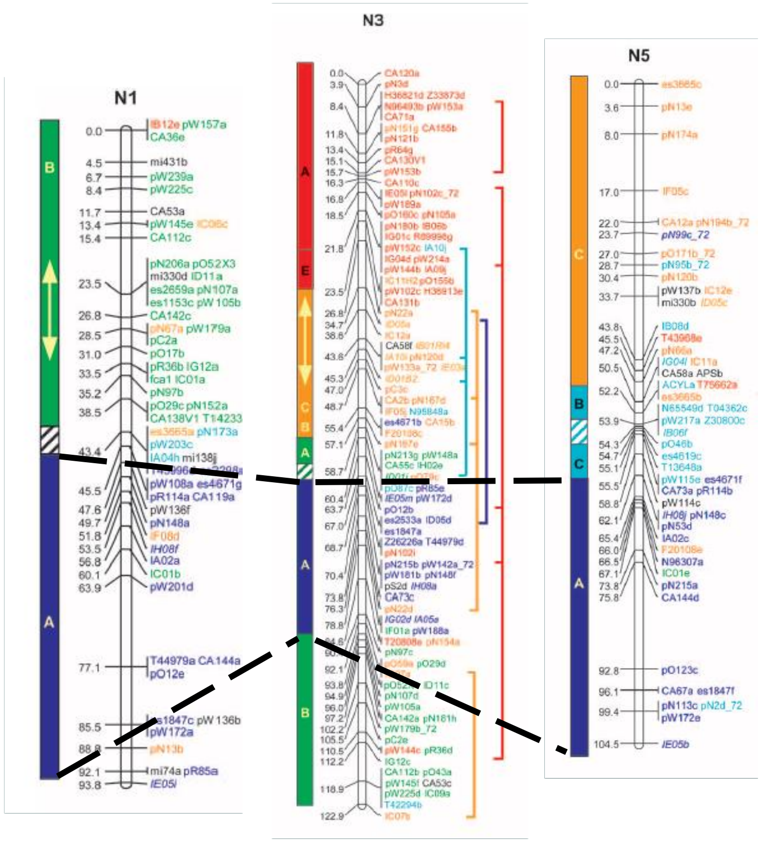
block resfuffling





Brassicas Are Ancient Hexaploids (Mesopolyploids)

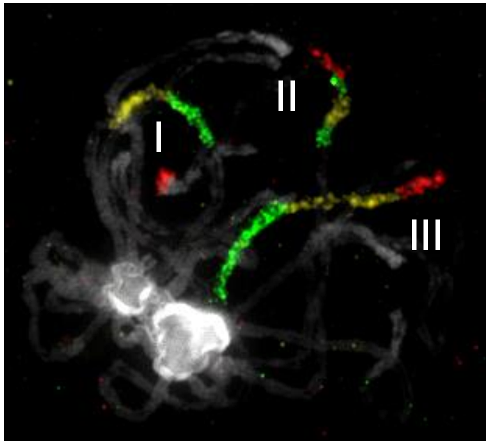
Brassica napus (AACC, n = 19), A genome (N1-N10)



Parkin et al. (2005) Genetics

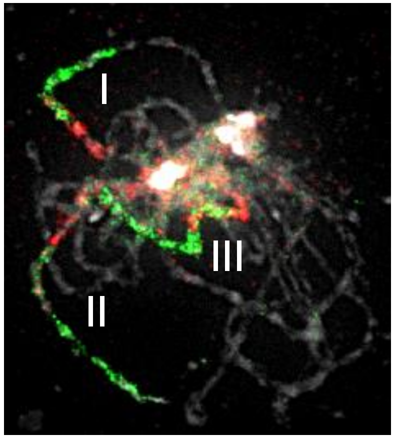
Diplotaxis erucoides

2n = 14



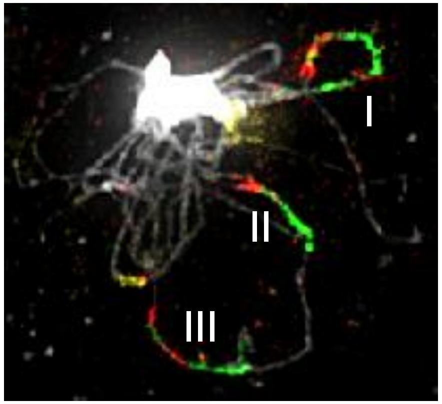
Brassica oleracea

2n = 18



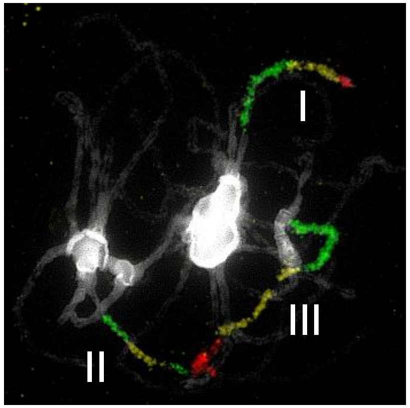
Morisia monanthos

2n = 14



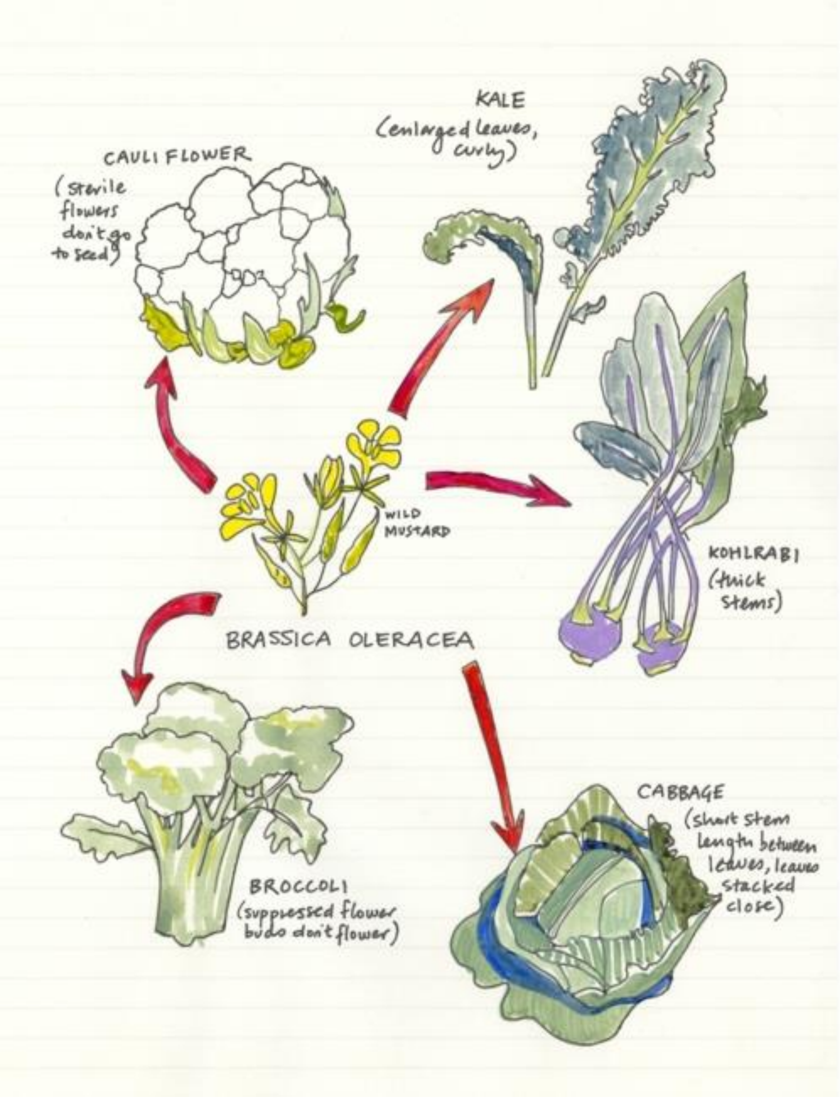
Moricandia arvensis

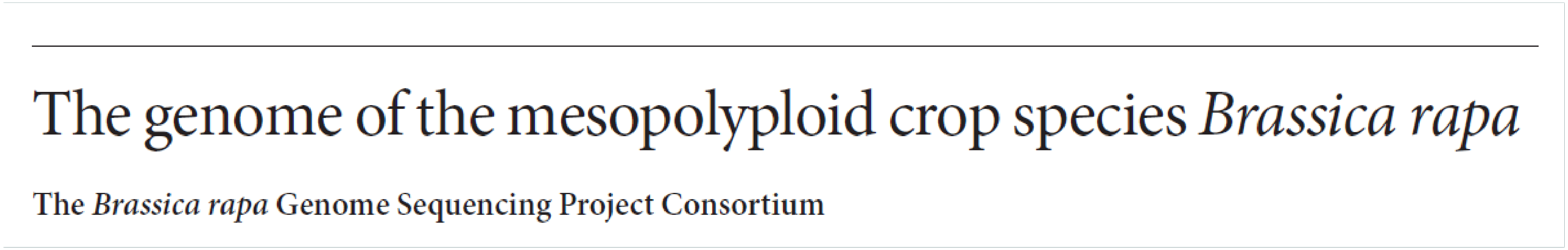
2n = 28



Lysak et al. (2005) Genome Res, (2007) Plant Physiol

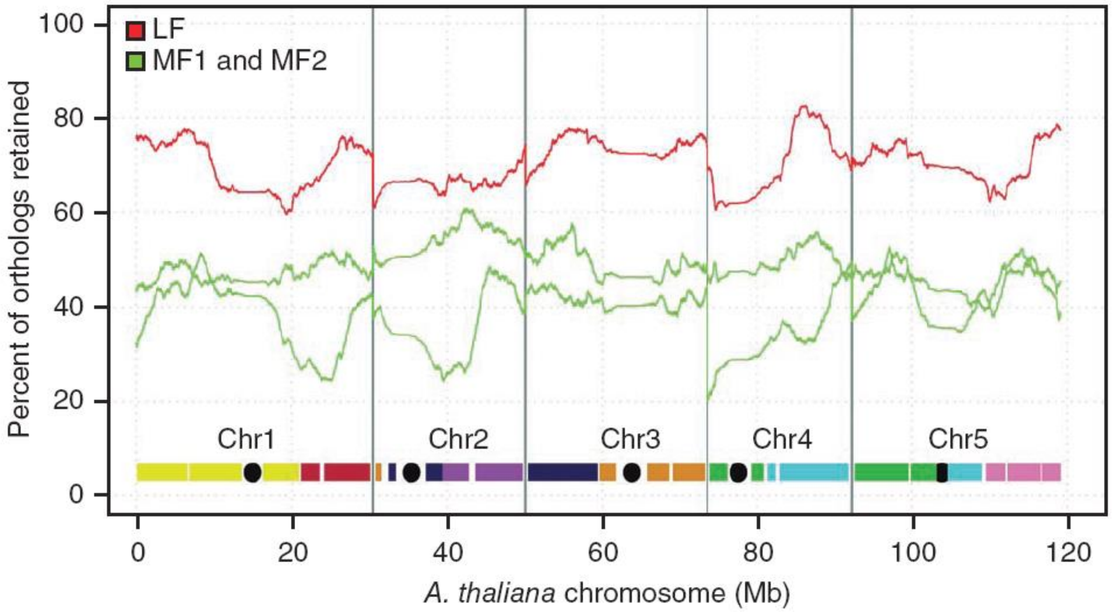
Diploidization in *Brassica* is marked by the asymmetrical evolution of polyploid genomes





The genome of the mesopolyploid crop species *Brassica rapa*

The *Brassica rapa* Genome Sequencing Project Consortium

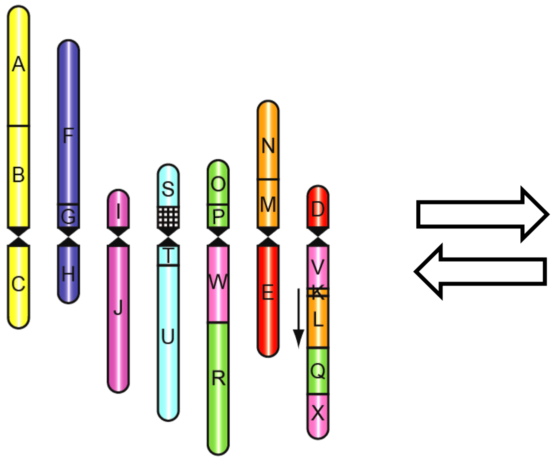


The density of orthologous genes in **three subgenomes** (LF, MF1 and MF2) of *B. rapa* compared to *A. thaliana.*

Wang et al. (2011) Nat Genet

Three *B. rapa* Subgenomes Contain Genome Block Associations Unique to the tPCK Ancestral Genome

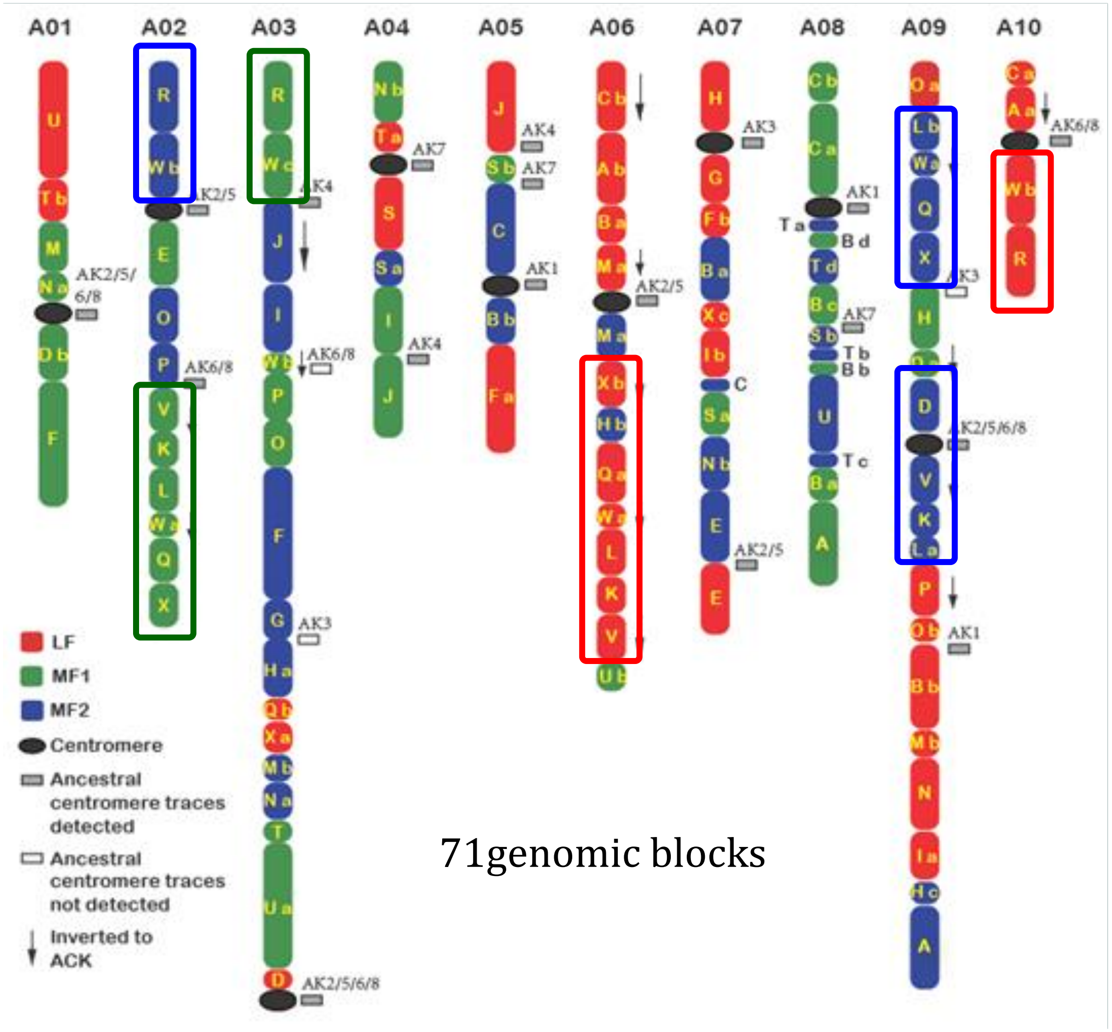
n = 7



translocation

Proto-Calepineae Karyotype

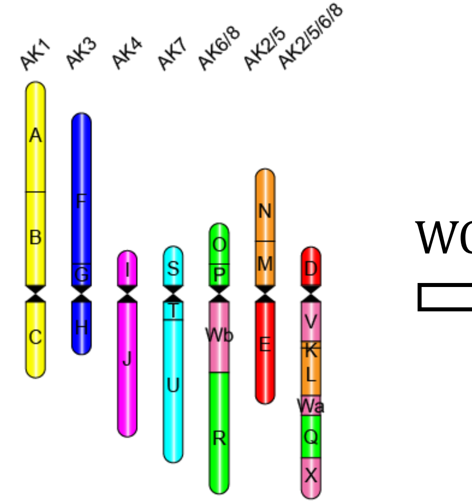
tPCK



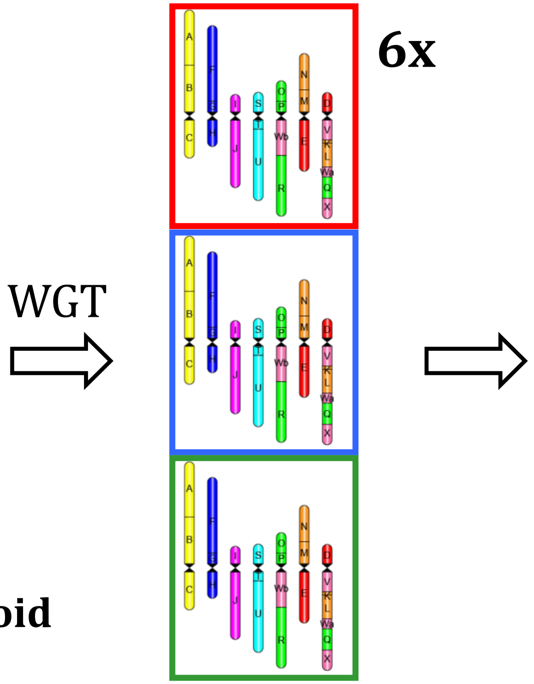
71genomic blocks

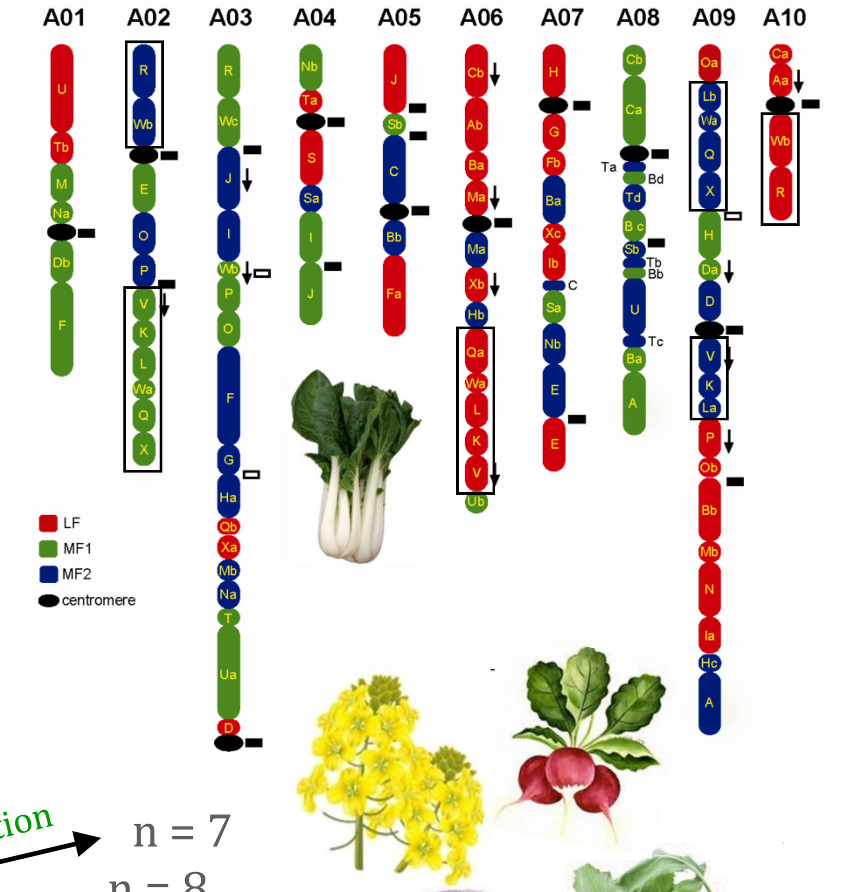
Cheng et al. (2013) Plant Cell, Mandakova and Lysak (2008) Plant Cell

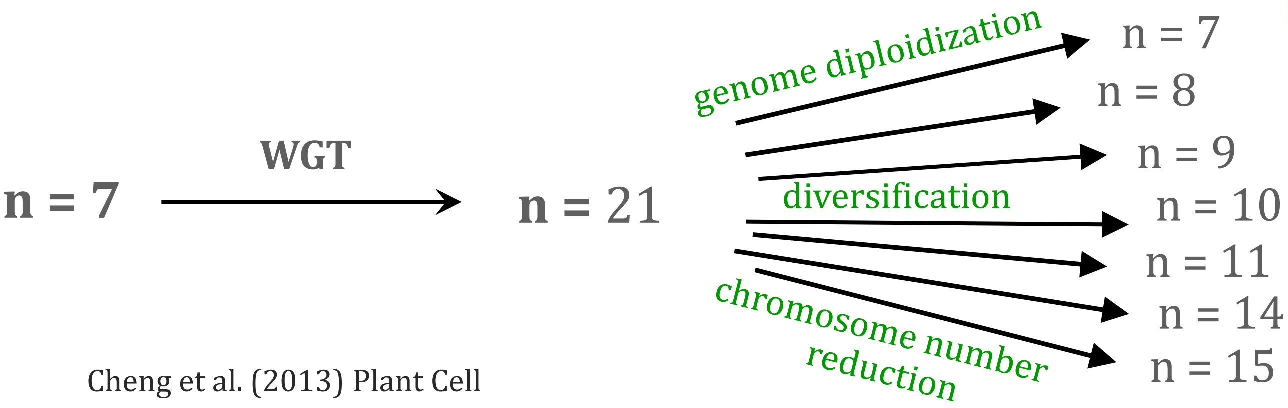
Whole-Genome Triplication Spurred Genome and Taxonomic Diversity in *Brassica* and Tribe Brassiceae



tPCK – ancestral diploid genome of *Brassica*







n = 7 —WGT→ n = 21

genome diploidization

diversification

chromosome number reduction

n = 7
n = 8
n = 9
n = 10
n = 11
n = 14
n = 15

Cheng et al. (2013) Plant Cell

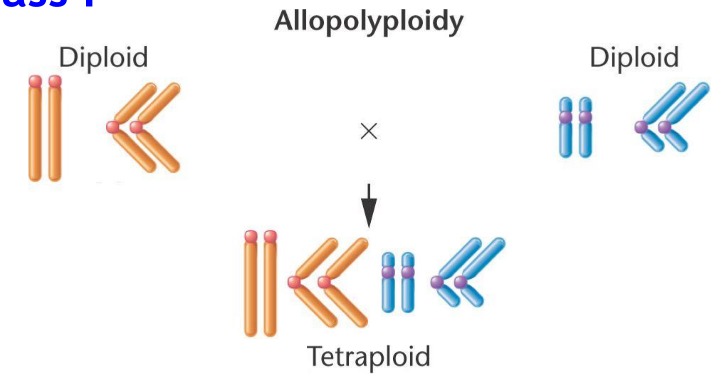
Two Evolutionarily Distinct Classes of Paleopolyploidy

Olivier Garsmeur,[†,1] James C. Schnable,[†,2] Ana Almeida,[2] Cyril Jourda,[1] Angélique D'Hont,[*,†,1] and Michael Freeling[*,†,2]

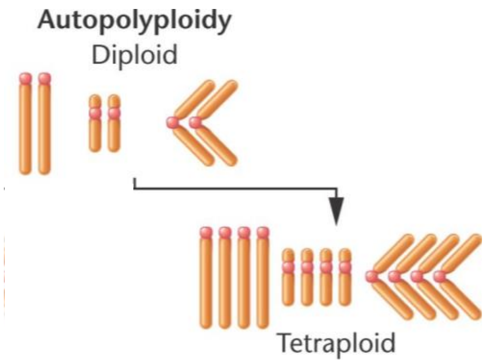
Table 2. Fractionation Pattern and Genome Dominance in Eight Species.

Species	WGD Class	Substitution Rate (Ks)	Bias Ratio between Duplicate Regions	Fractionation Pattern	Genome Dominance	Expression Data from
Medicago	I	0.87	1.23	Biased	No data	
Sorghum	I	0.95	1.24	Biased (Schnable et al. 2012)	Yes	Dugas et al. (2011)
Arabidopsis	I	0.76	1.17	Biased (Thomas et al. 2006)	Yes	Gan et al. (2011)
Brassica	I	0.34	1.47	Biased (Wang et al. 2011)	Yes (Cheng et al. 2012)	
Maize	I	0.17	1.46	Biased (Woodhouse et al. 2010)	Yes (Schnable et al. 2011)	
Poplar	II	0.23	1.05	Unbiased	No data	
Soybean	II	0.15	1.03	Unbiased	No	Schmidt et al. (2011)
Banana	II	0.39	1.06	Unbiased	No	D'Hont et al. (2012) and supplementary table S4, Supplementary Material online

Class I



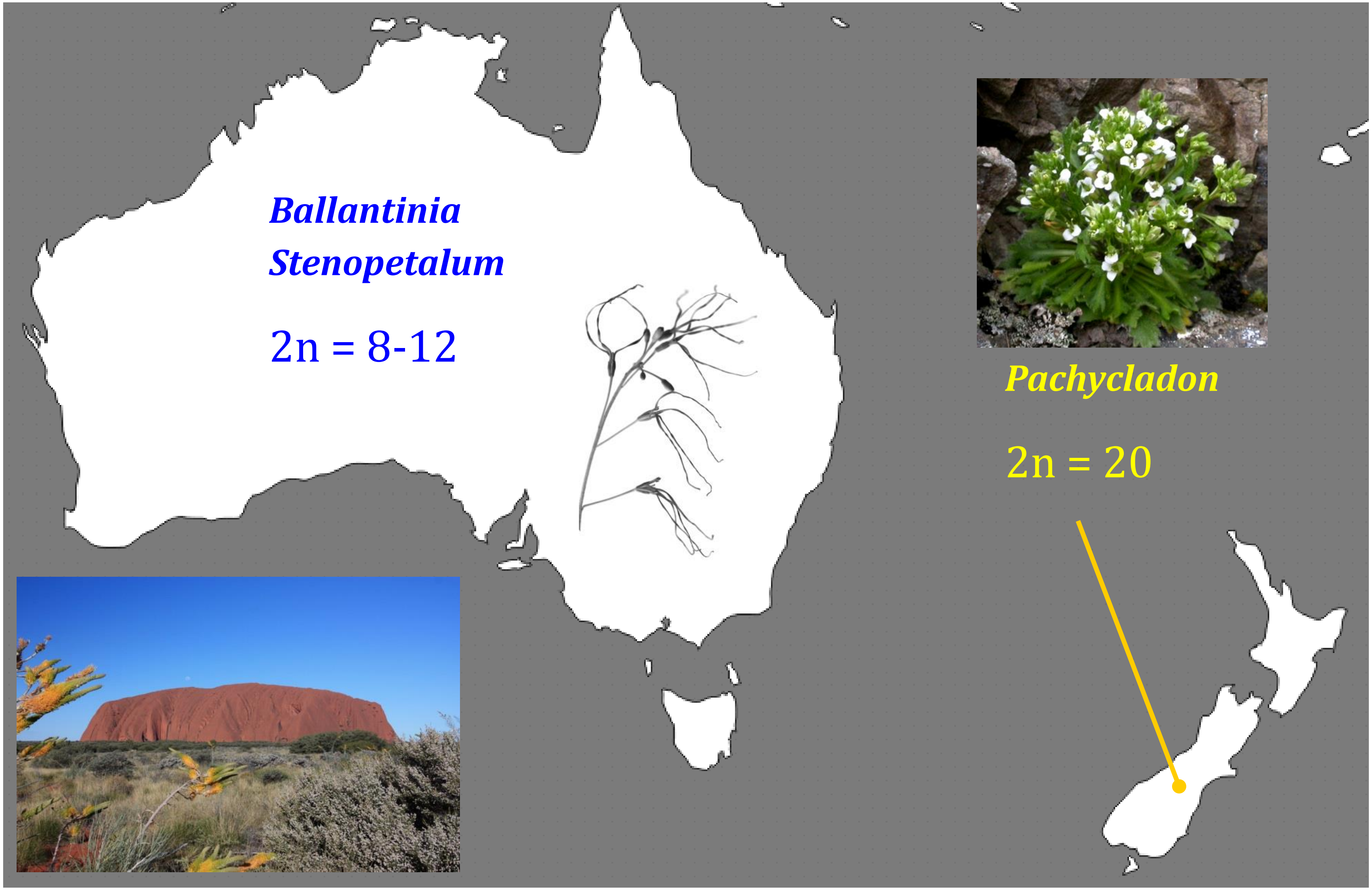
Class II



?

MBE 31 (2013)

Polyploid Evolution in Australian and New Zealand Crucifer Genera



Ballantinia

Stenopetalum

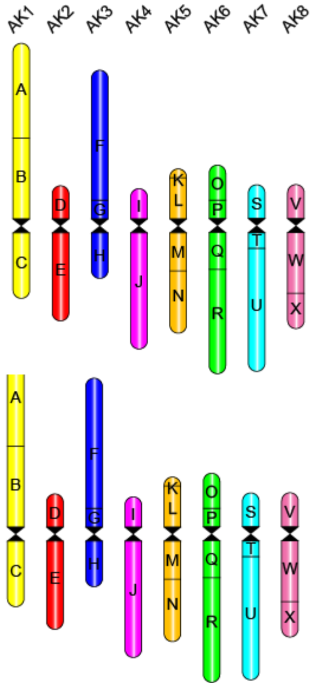
2n = 8-12

Pachycladon

2n = 20

Polyploid Origin of *Pachycladon* (n=10)

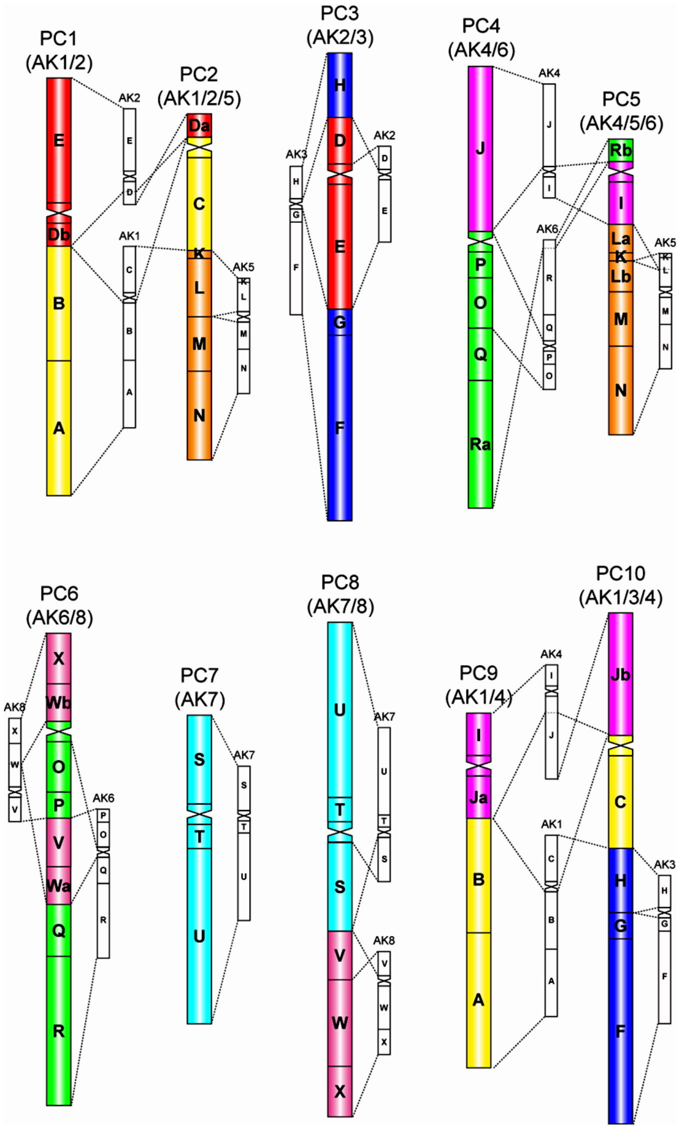
~ 1 – 2 mya



Ancestral Karyotype (4x)

16 chromosomes

48 blocks



10 chromosomes

48 blocks

P. enysii

P. novae-zelandiae

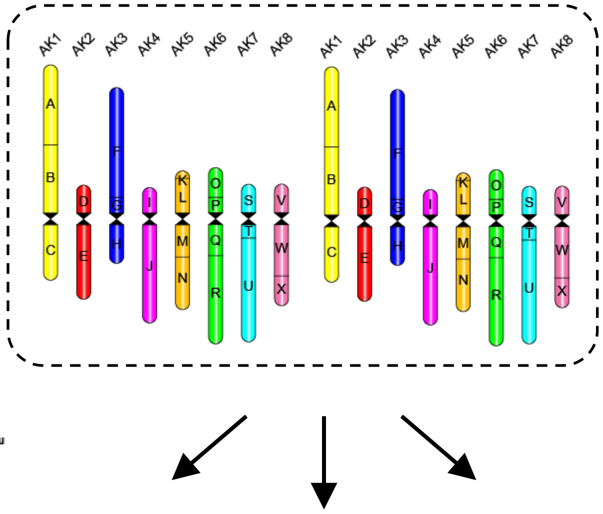
P. cheesemanii

P. exile

Mandakova, Heenan and Lysak (2010) BMC Evol Biol

16 chromosomes

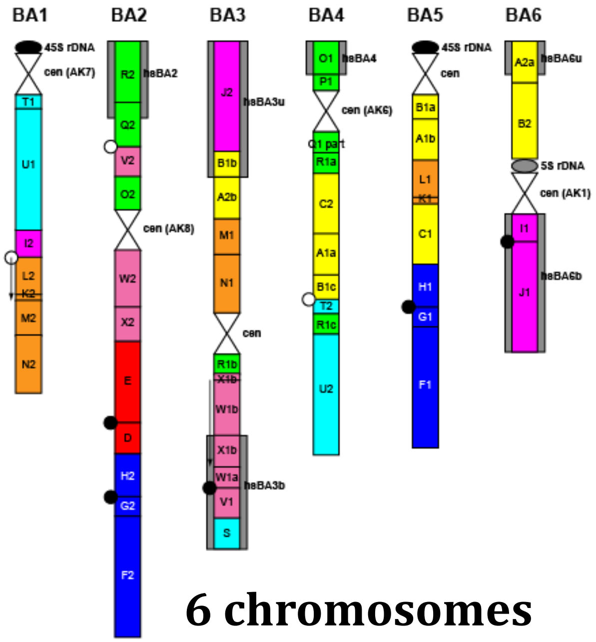
48 blocks



ACK, 4x

~ 6 – 9 mya

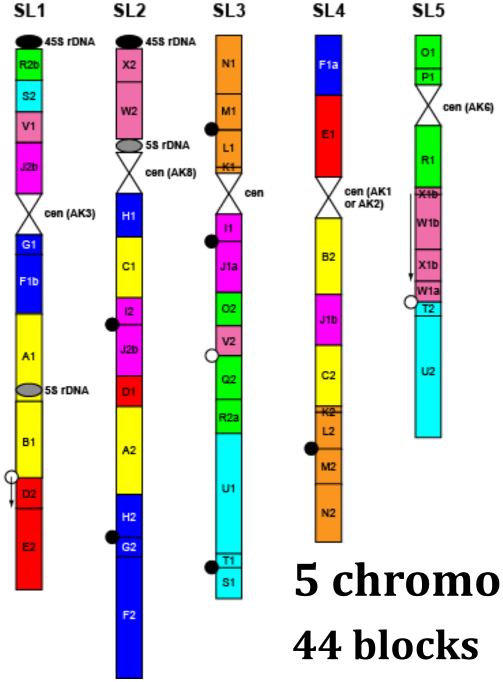
Ballantinia antipoda



6 chromosomes

40 blocks

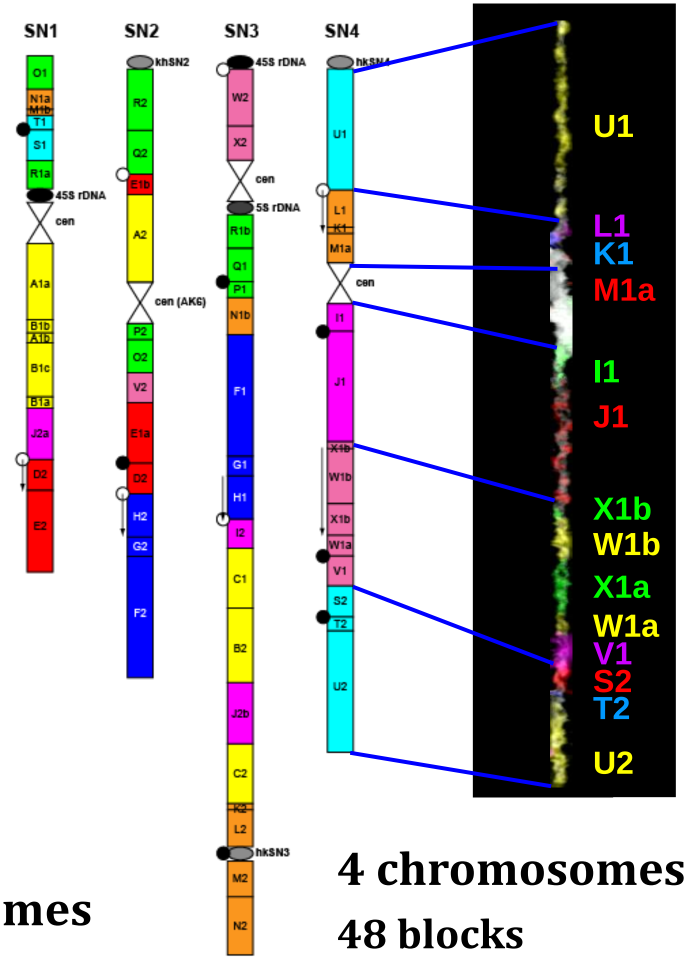
Stenopetalum lineare



5 chromosomes

44 blocks

Stenopetalum nutans



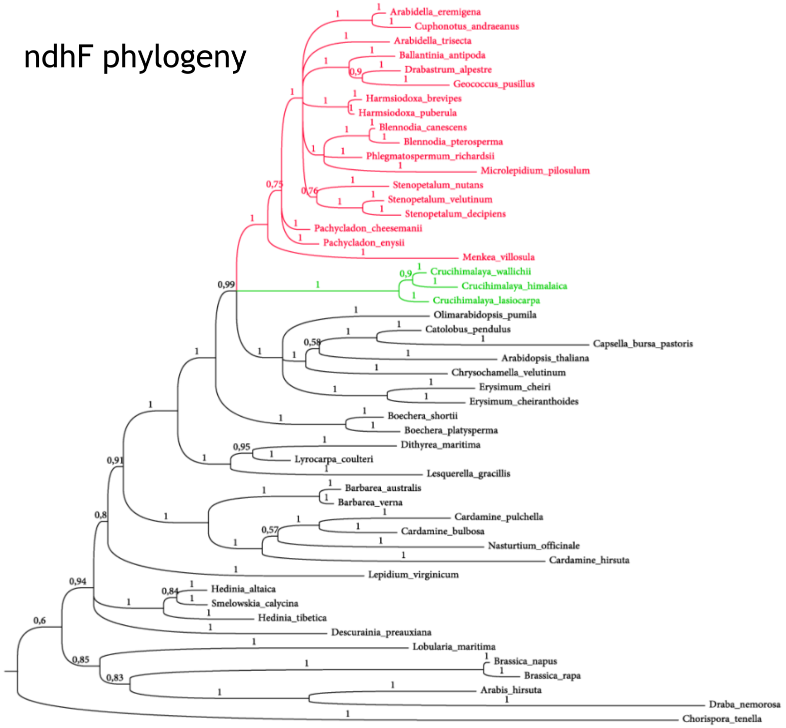
4 chromosomes

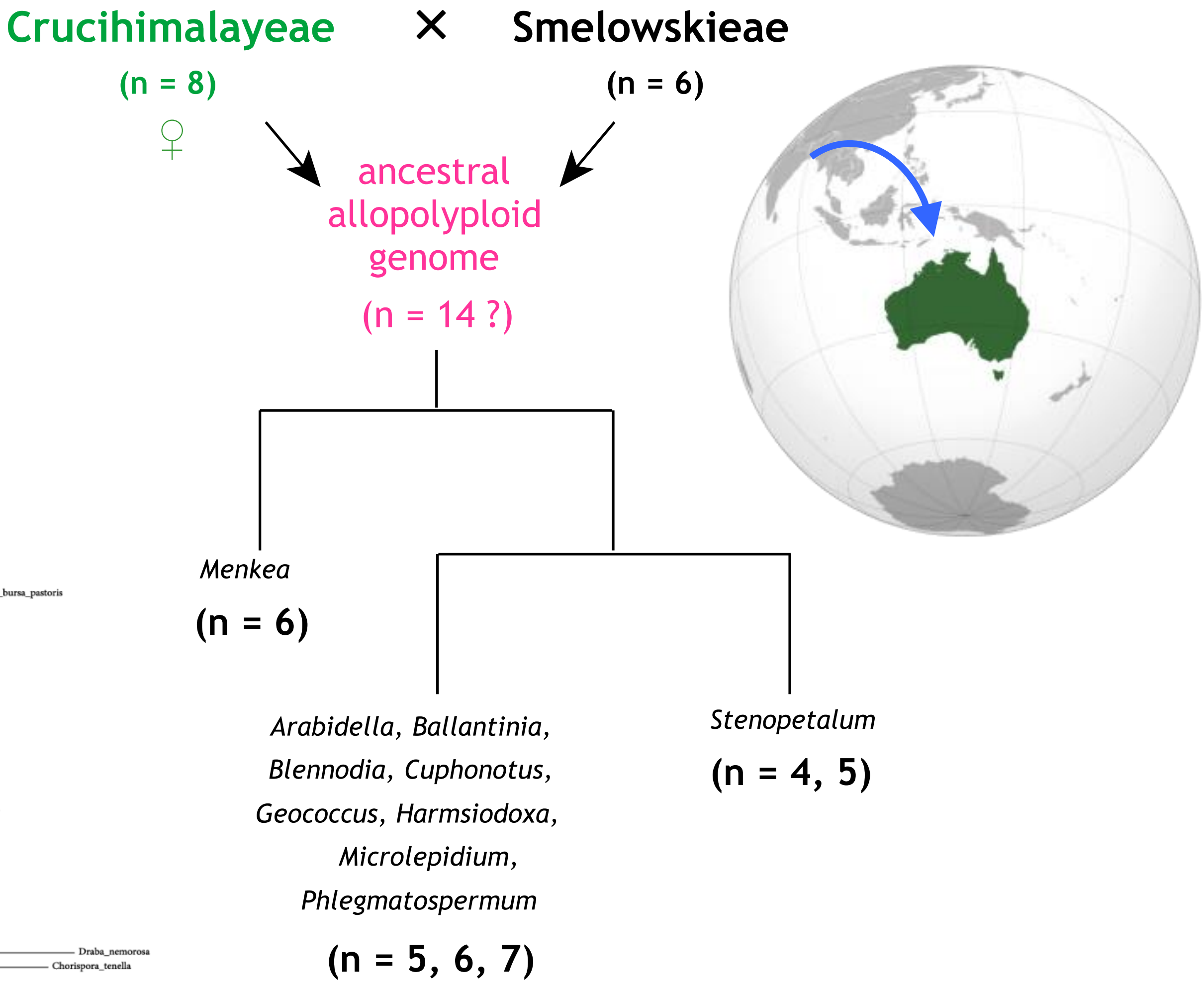
48 blocks

Mandakova et al. (2010) Plant Cell

Allopolyploid origin of the Microlepidieae

ndhF phylogeny





Crucihimalayeae × **Smelowskieae**

(n = 8) ♀

(n = 6)

ancestral allopolyploid genome

(n = 14 ?)

Menkea

(n = 6)

Arabidella, Ballantinia,
Blennodia, Cuphonotus,
Geococcus, Harmsiodoxa,
Microlepidium,
Phlegmatospermum

(n = 5, 6, 7)

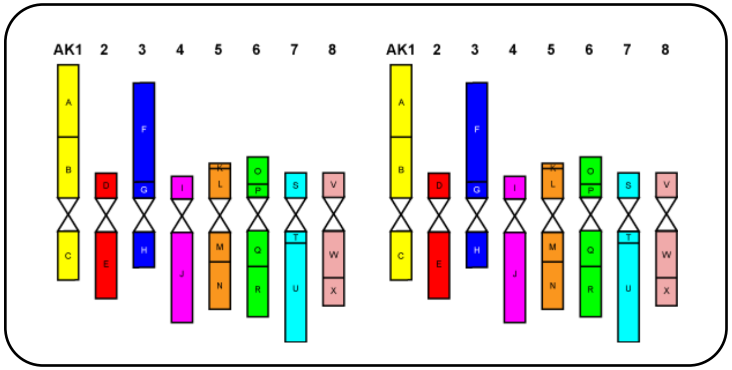
Stenopetalum

(n = 4, 5)

Mandáková et al., unpublished

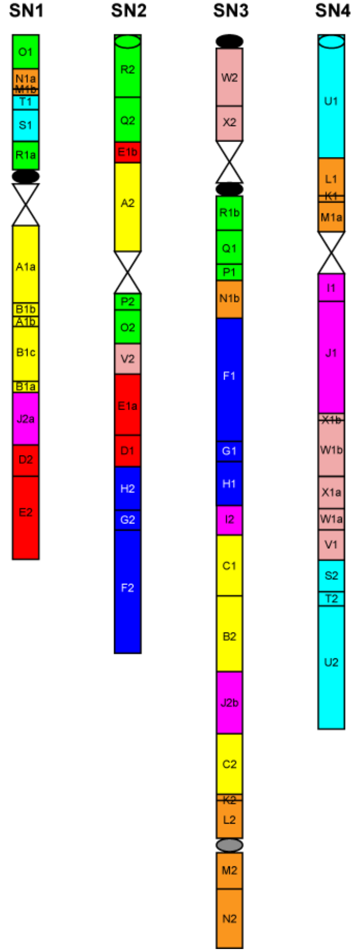
Reshuffling of 2 x 24 Genomic Blocks in Polyploids of Different Age

ACK-like allotetraploid (n ~ 16)



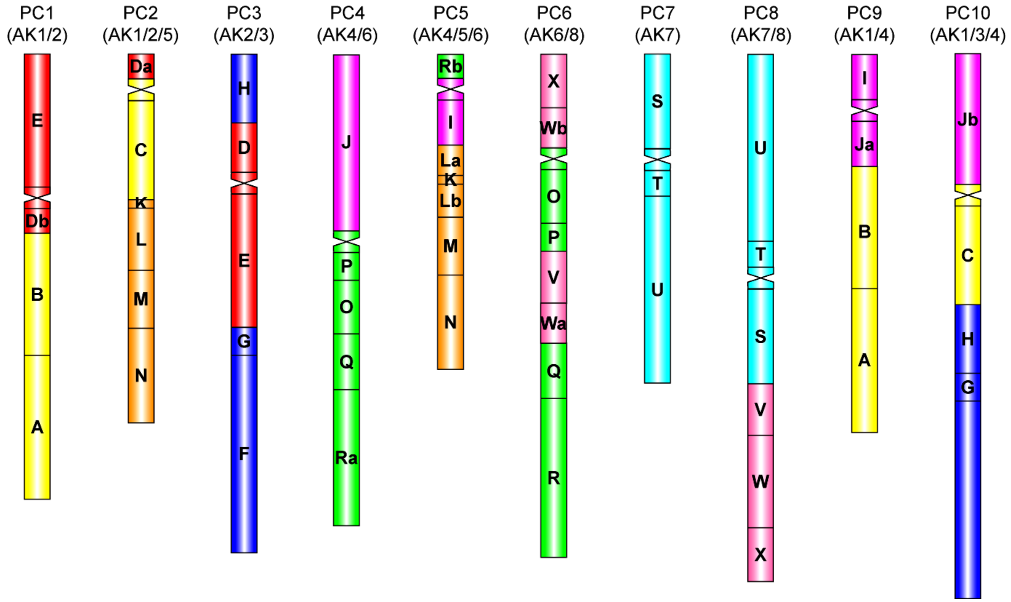
~ 6 – 9 mya

Stenopetalum (n = 4)



~ 1 – 2 mya

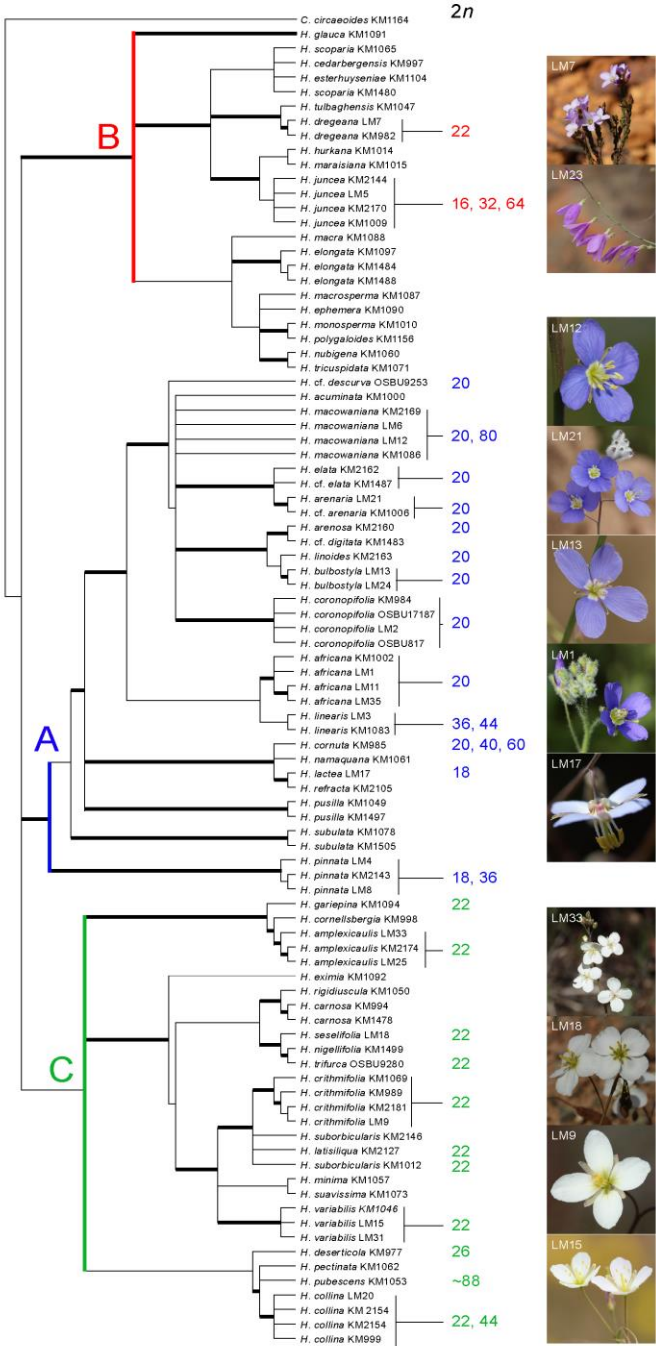
Pachycladon (n = 10)

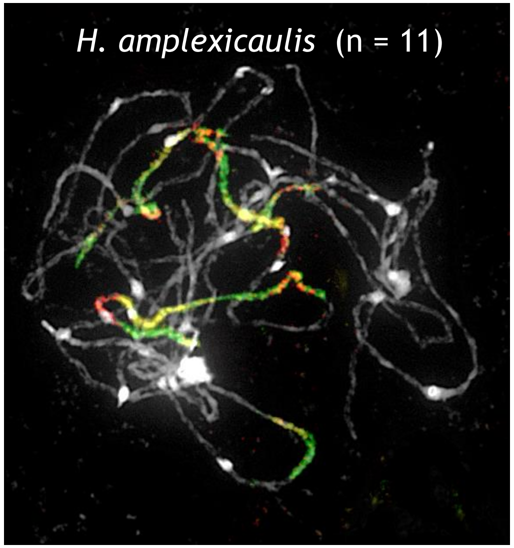


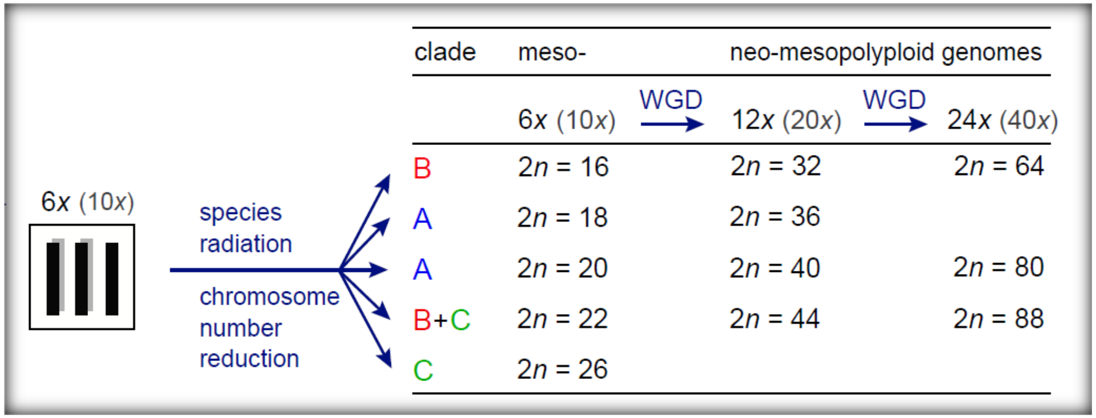
Does Ancient Polyploidy Explain the Rapid Species Radiation in *Heliophila* ?

c. 90 (-100) endemic spp.

Whole-Genome Triplication in the Southern African Tribe Heliophileae



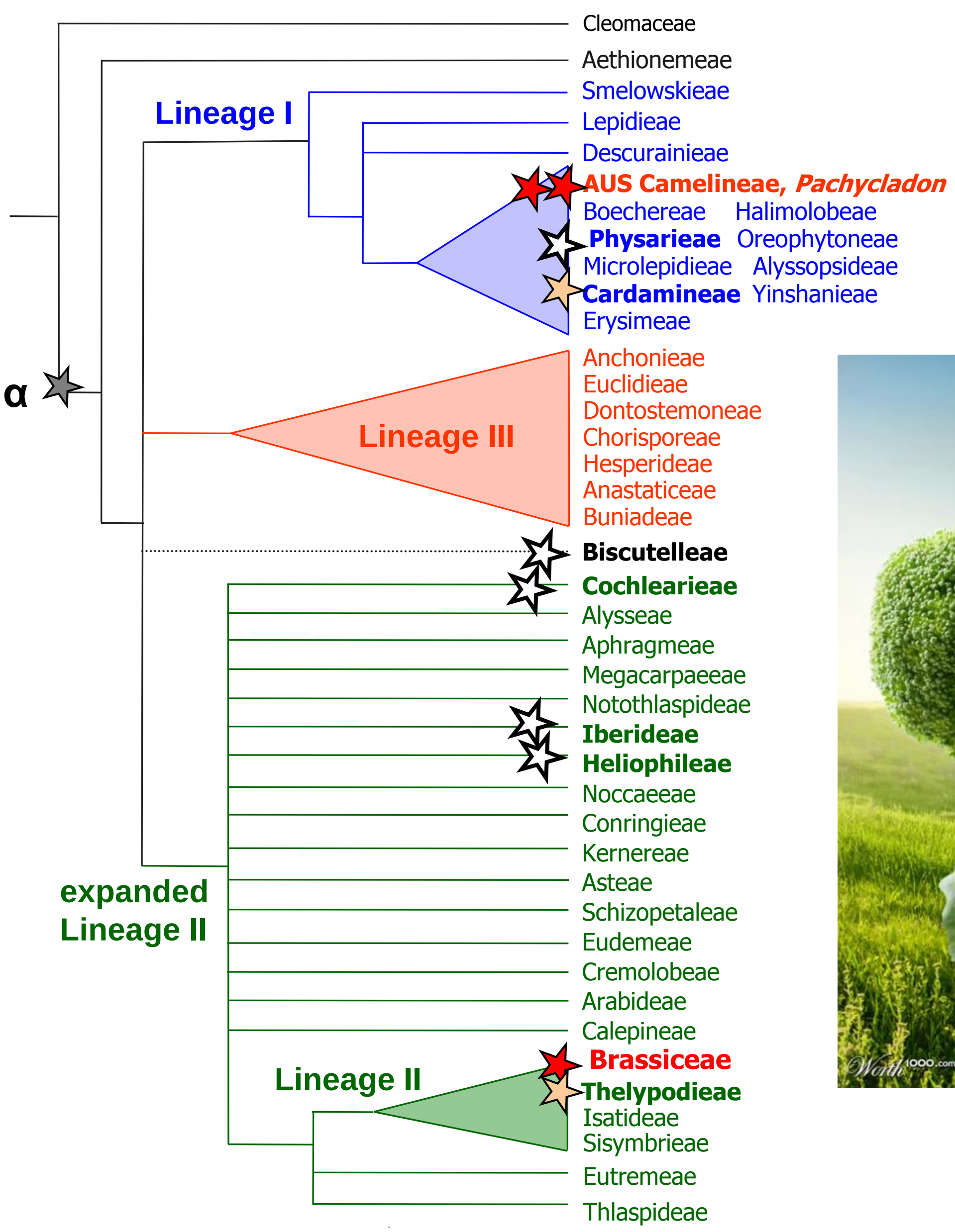




clade	meso-	neo-mesopolyploid genomes	
	6*x* (10*x*)	12*x* (20*x*)	24*x* (40*x*)
B	2*n* = 16	2*n* = 32	2*n* = 64
A	2*n* = 18	2*n* = 36	
A	2*n* = 20	2*n* = 40	2*n* = 80
B+C	2*n* = 22	2*n* = 44	2*n* = 88
C	2*n* = 26		

Mandakova et al. (2012) Taxon

Whole-Genome Duplications Drive Genome Diversification



α

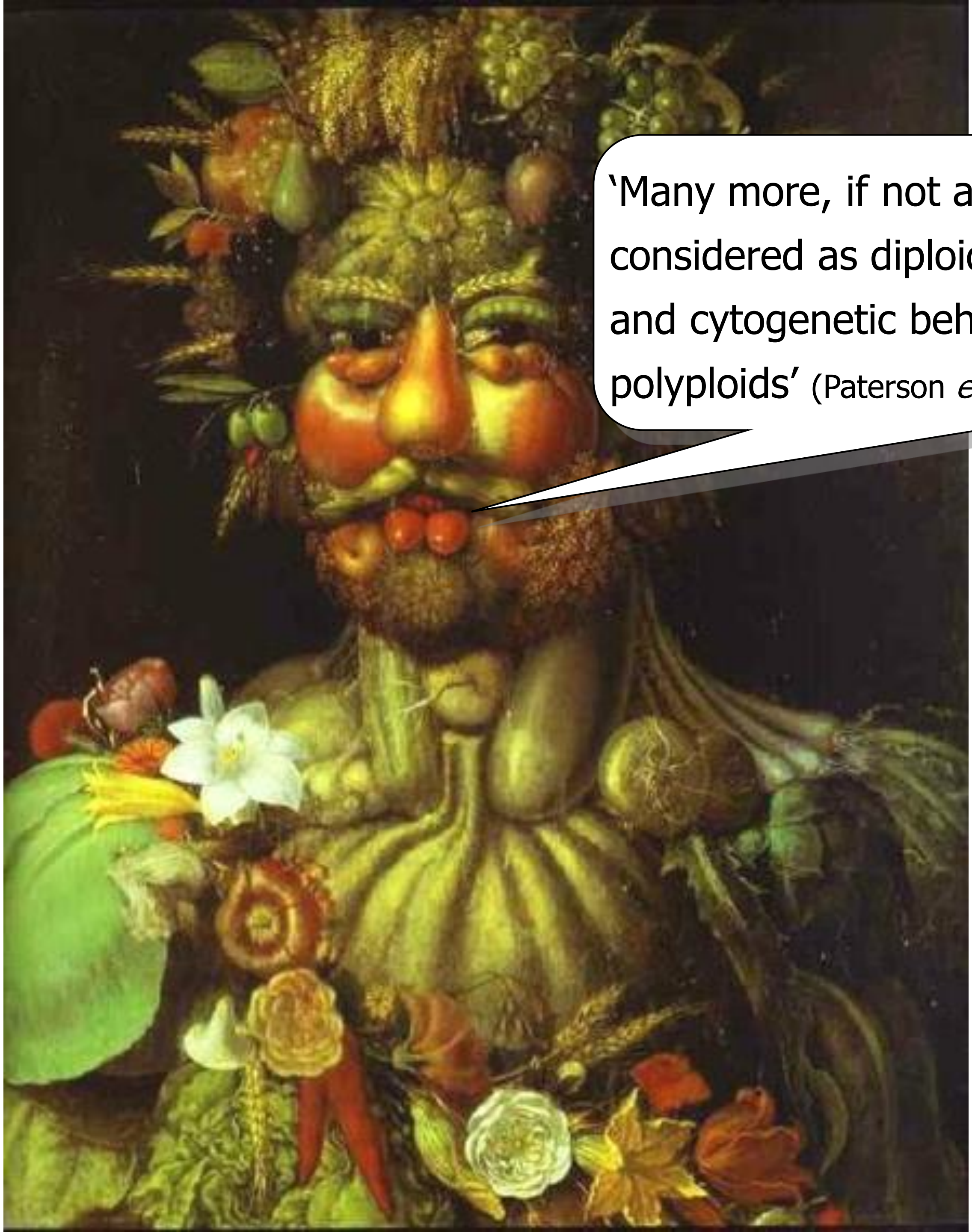
Lineage I

Lineage III

expanded Lineage II

Lineage II

Cleomaceae
Aethionemeae
Smelowskieae
Lepidieae
Descurainieae
AUS Camelineae, *Pachycladon*
Boechereae Halimolobeae
Physarieae Oreophytoneae
Microlepidieae Alyssopsideae
Cardamineae Yinshanieae
Erysimeae
Anchonieae
Euclidieae
Dontostemoneae
Chorisporeae
Hesperideae
Anastaticeae
Buniadeae
Biscutelleae
Cochlearieae
Alysseae
Aphragmeae
Megacarpaeeae
Notothlaspideae
Iberideae
Heliophileae
Noccaeeae
Conringieae
Kernereae
Asteae
Schizopetaleae
Eudemeae
Cremolobeae
Arabideae
Calepineae
Brassiceae
Thelypodieae
Isatideae
Sisymbrieae
Eutremeae
Thlaspideae



‘Many more, if not all, higher plant species, considered as diploids because of their genetic and cytogenetic behaviour, are actually ancient polyploids’ (Paterson *et al.* 2005).