

Fractionation,
rearrangement,
consolidation,
reconstruction,
&
dissection

measures of rearrangement:

edit distances: reversals, reversals +
translocations, DCJ, etc.

breakpoints

excess adjacencies

excess adjacencies:

1 2 3 4 5 6 7 8 9

2 3 4 6 7 5 1 8 9

8

duplicate genes:

1 2 3 4 5 6 7 8 9 6 7

2 3 4 6 7 5 1 8 9

9

missing genes:

1 2 3 4 5 6 7 8 9

2 4 6 7 5 1 8 9

11

>2 genomes:

1 2 3 4 5 6 7 8 9

2 3 4 6 7 5 1 8 9

2 3 4 5 1 8 9 6 7

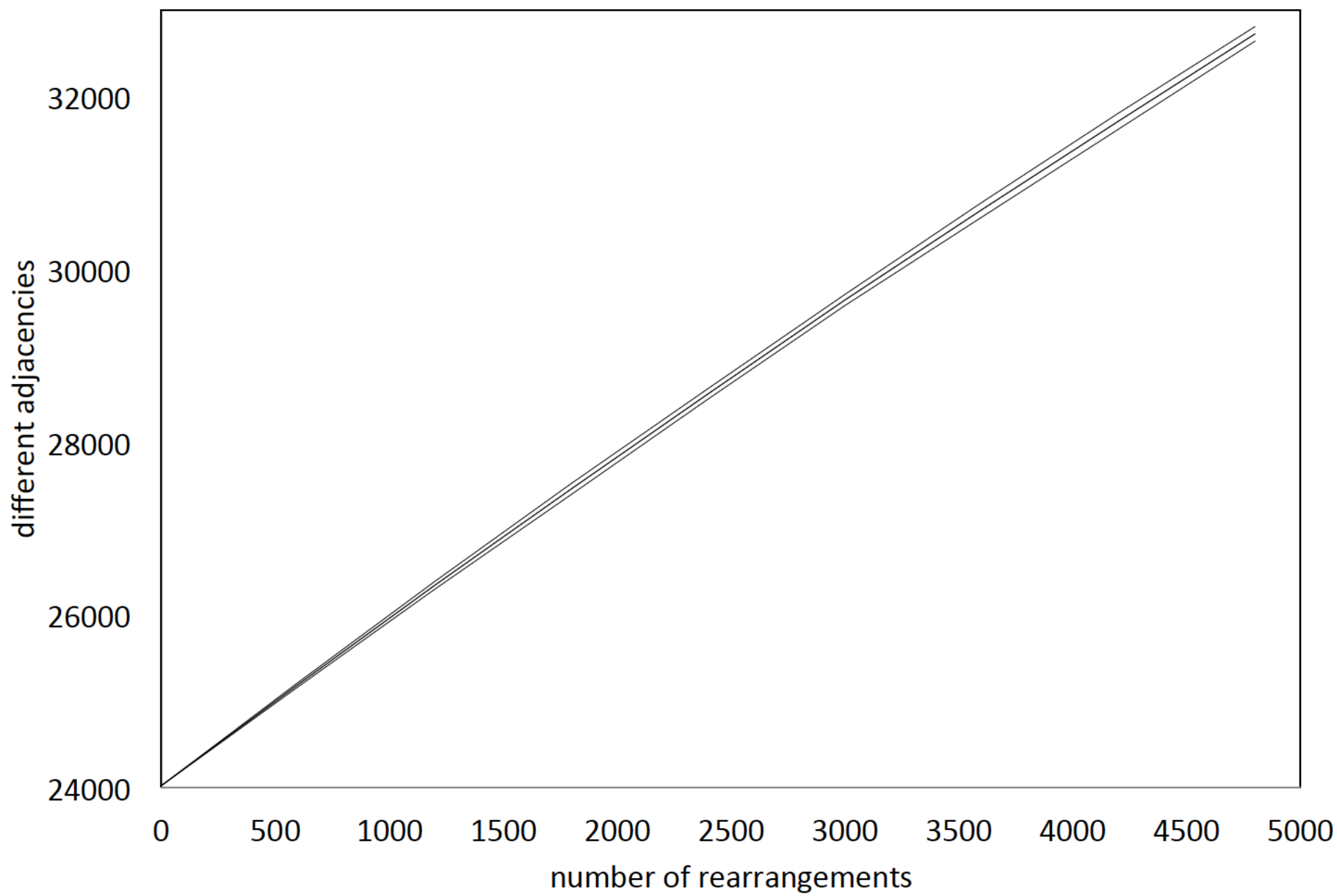
9

>1 chromosome:

1 2 3 4 5 6 7 8 9

2 3 4 6 7 5 1 8 9

9



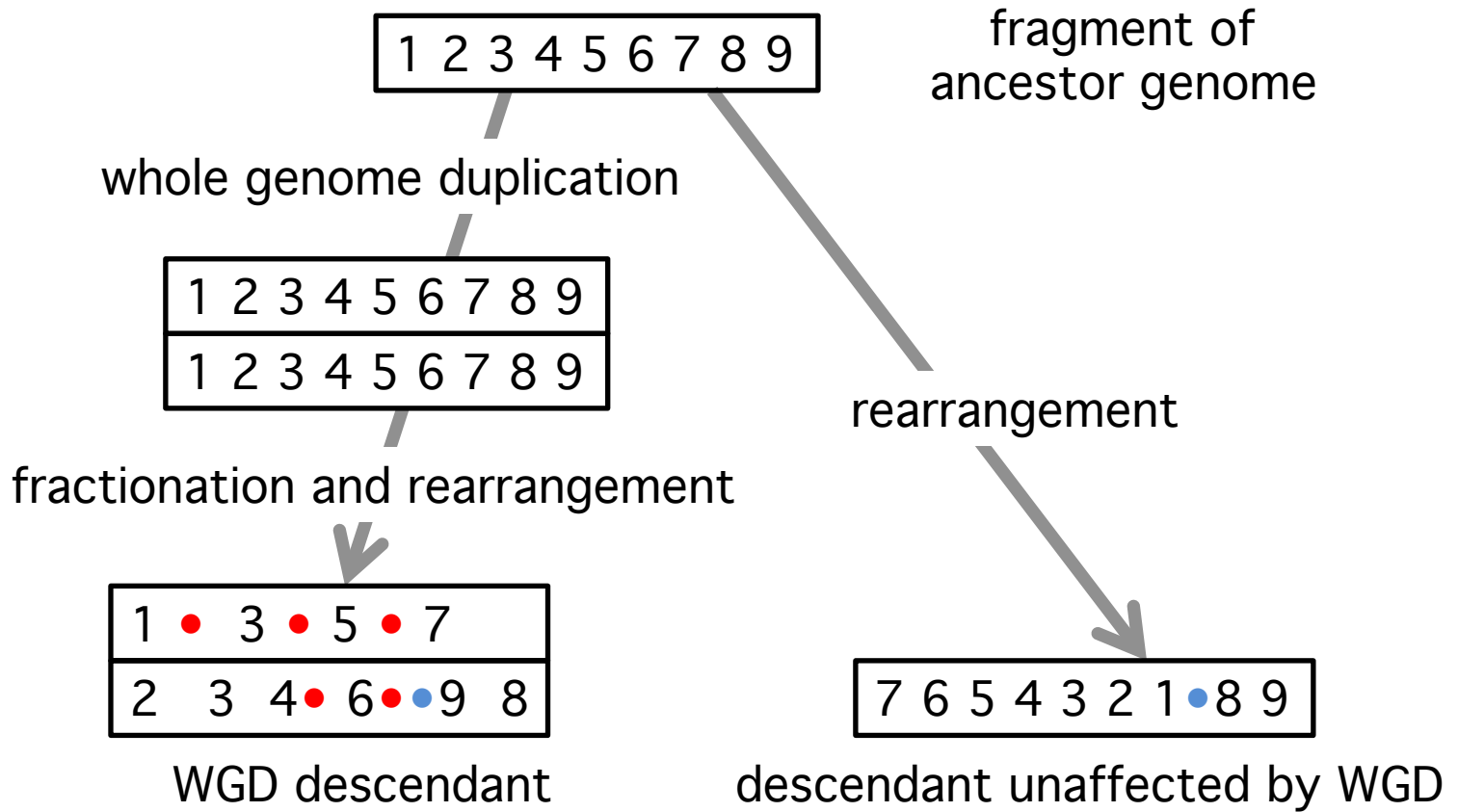
ancestral gene order reconstruction:

based on the adjacencies in the descendants
(between oriented, or signed, genes)

maximum weight matching

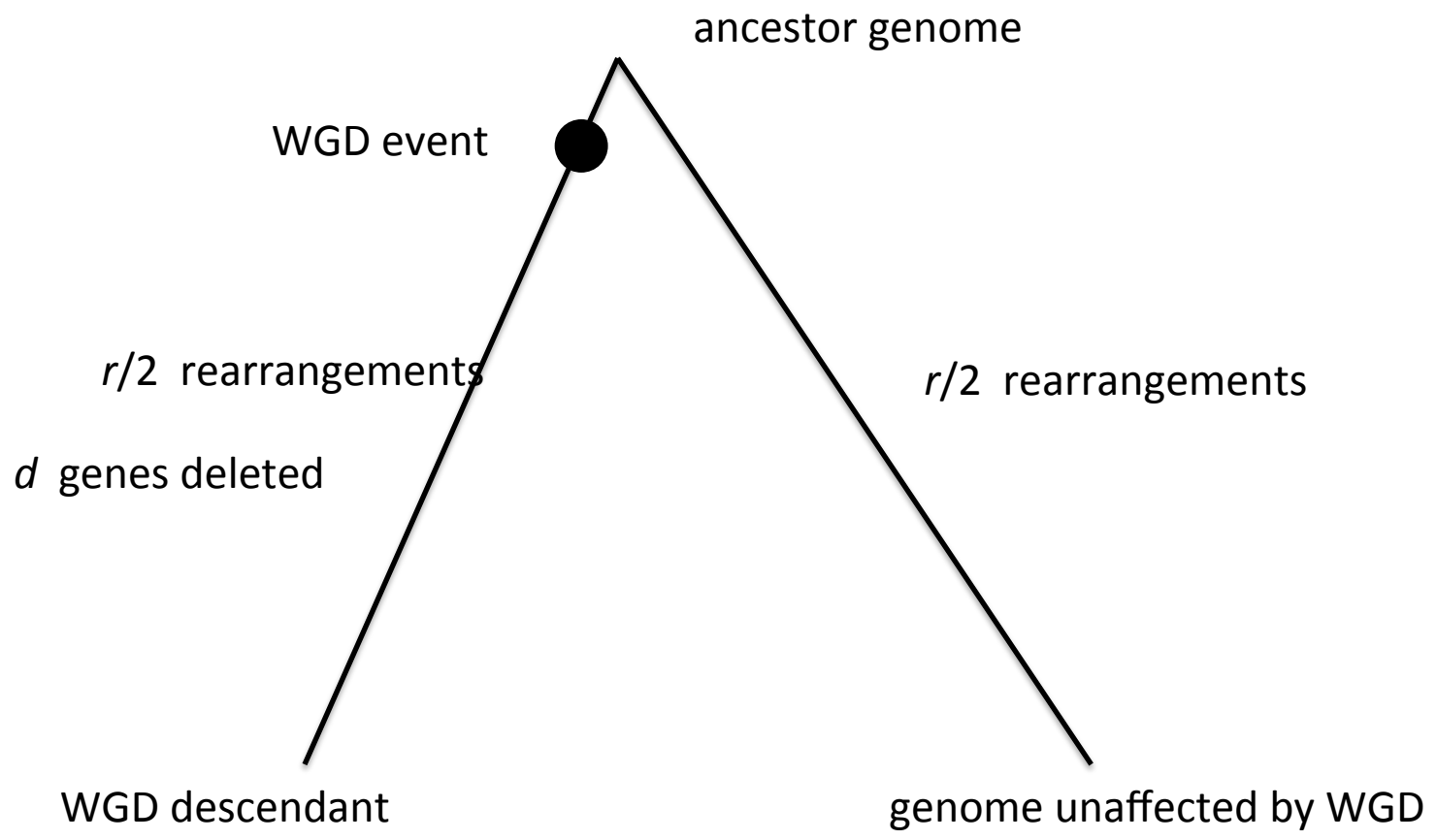
whole genome duplication (WGD)

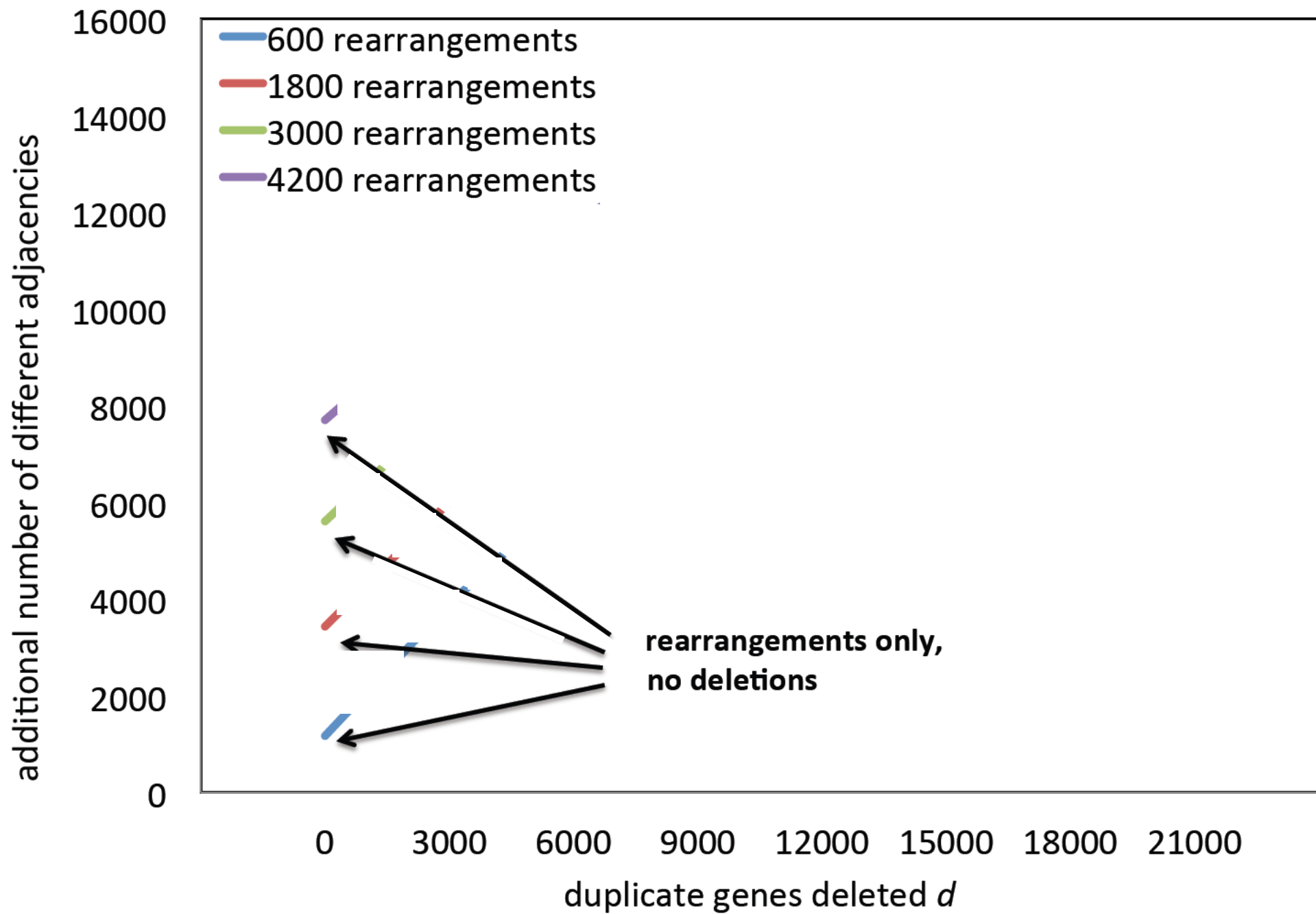
fractionation

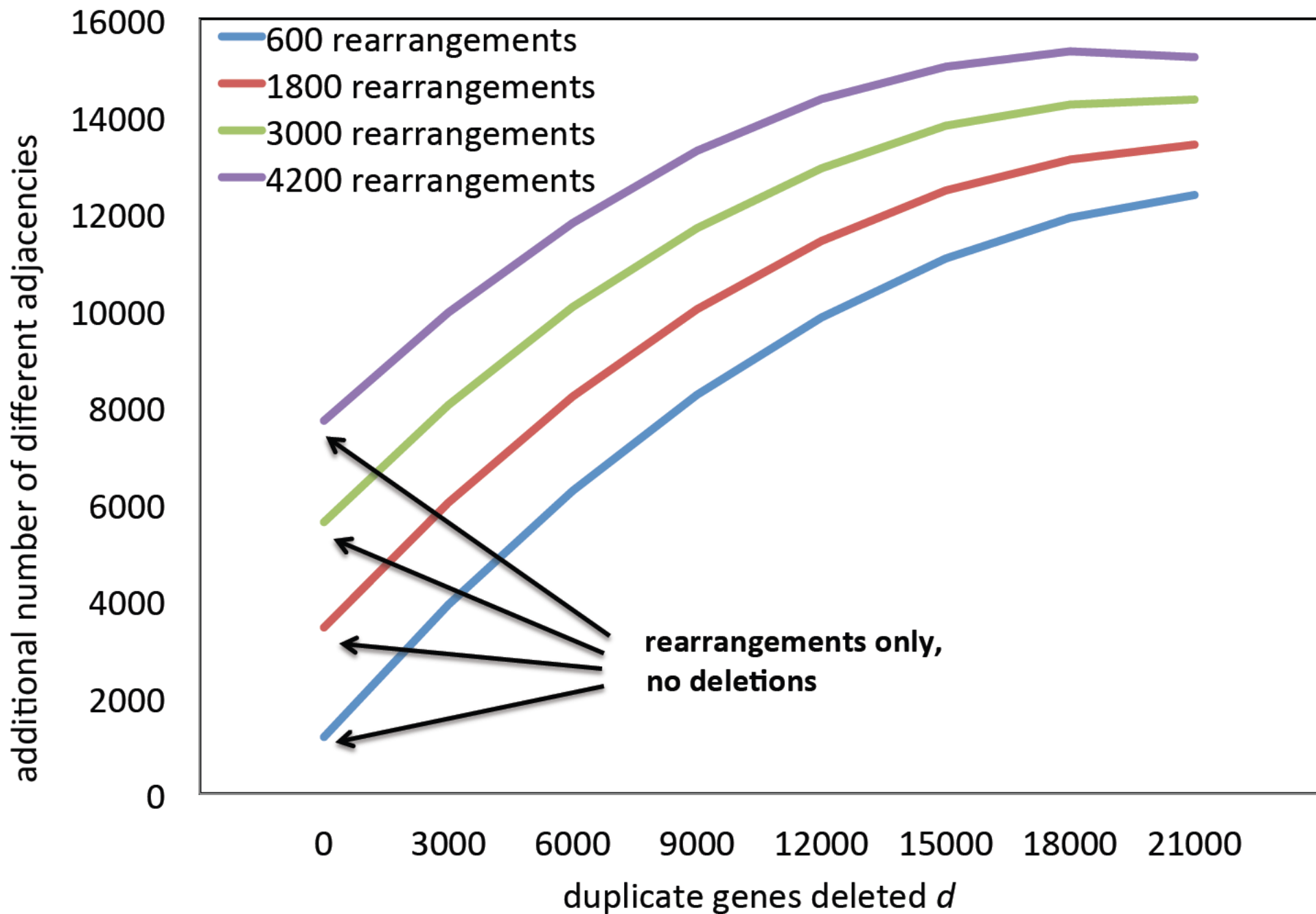


Rearrangements: **2** Fractionation: 5

Rearrangements only: **5**







Consolidation algorithm:

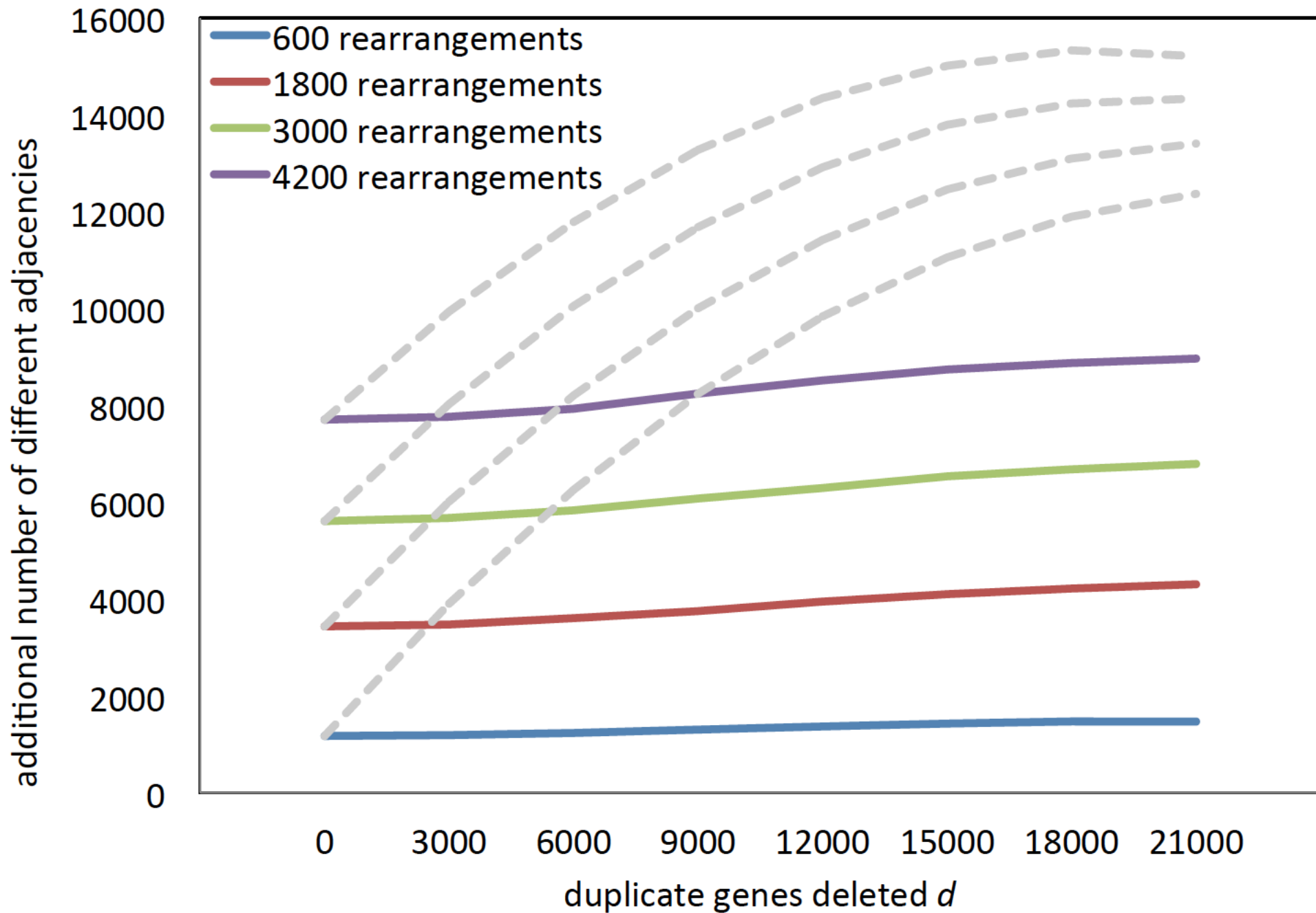
Detect triple t_i of regions: 1 2 3 4 in diploid, 1 3 and 2 4 in WGD descendant

Replace by “virtual gene”: V_i in diploid, V_i and V_i in WGD descendant

Reconstruct ancestor

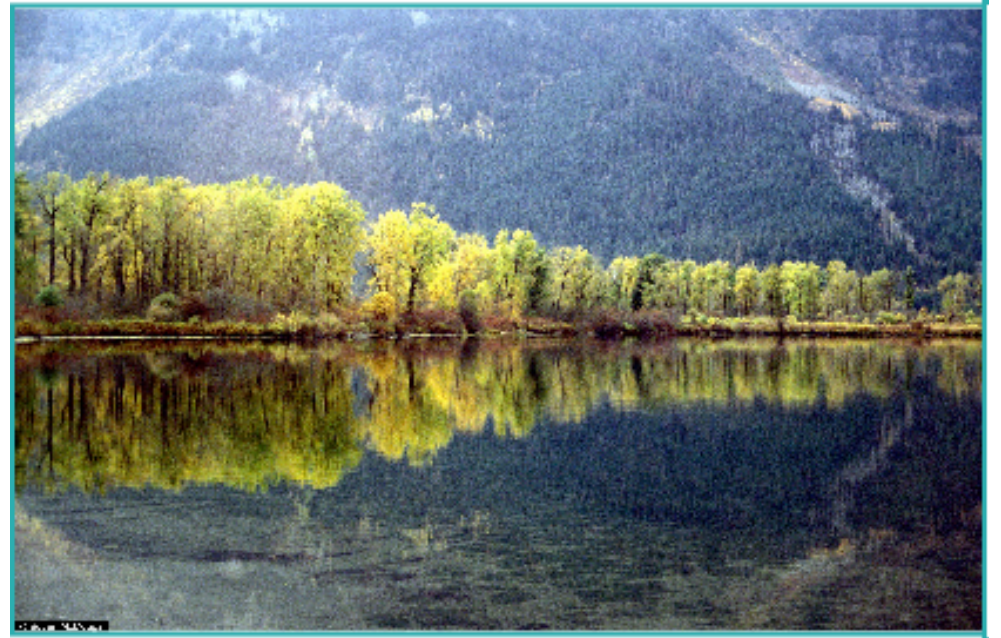
Recalculate excess adjacencies

Add in excess adjacencies within all the V_i



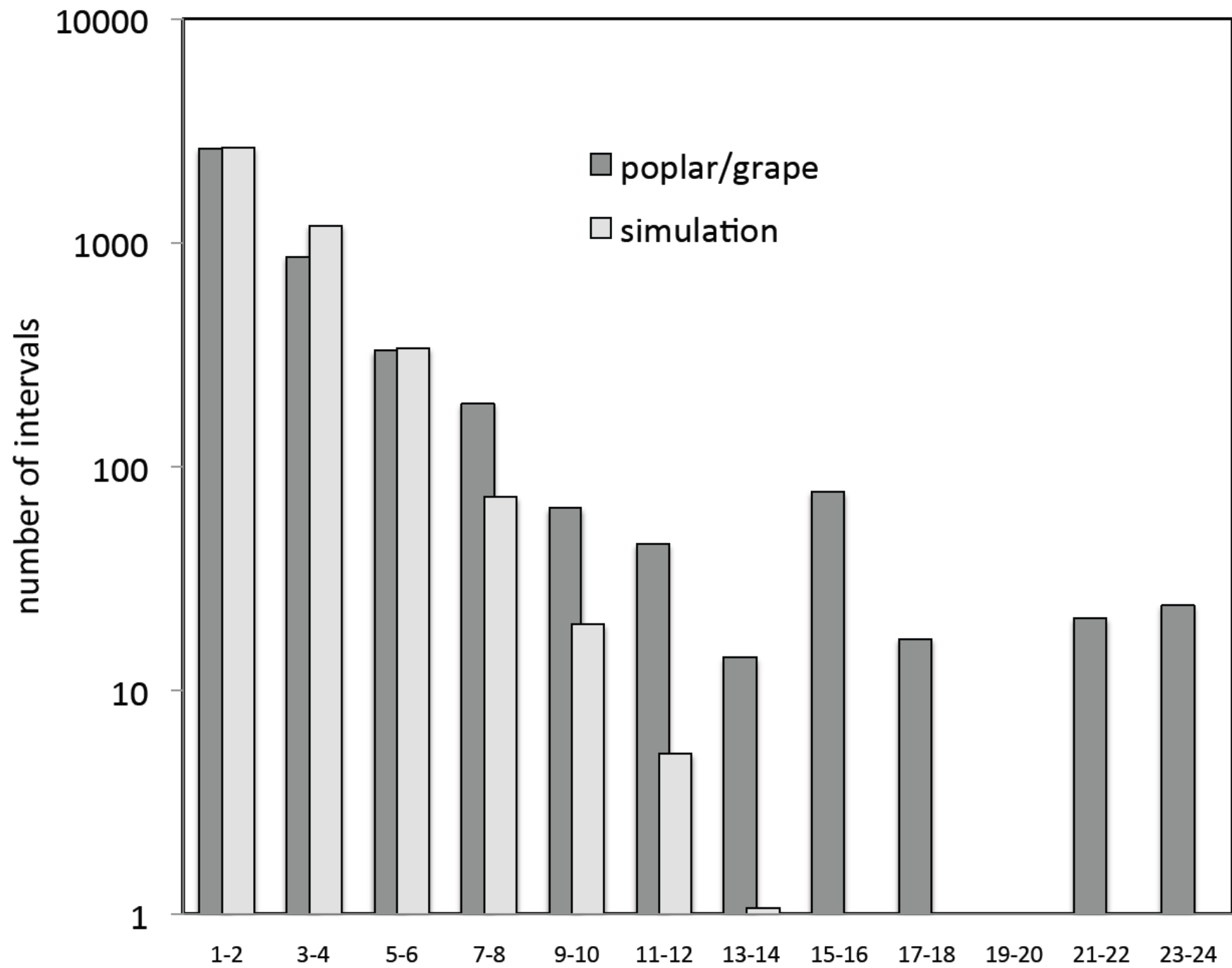
poplar (WGD descendant)

grape (unaffected by WGD)

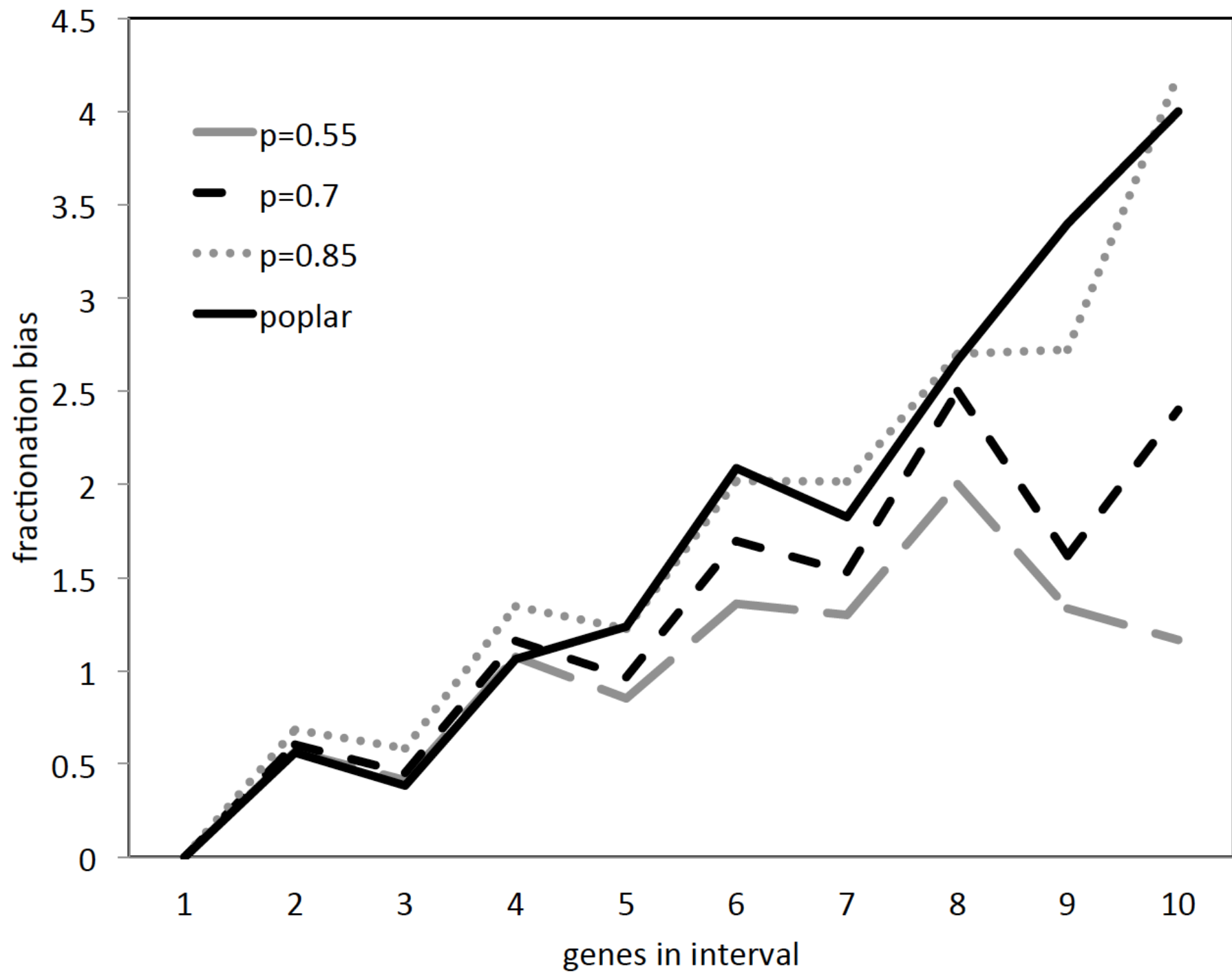


genes in comparison	grape	poplar	ancestor
single copies	12,494	4,282	8,631
in syntenic pairs	0	$2 \times 8212 = 16,424$	0
total	12,494	20,706	8,631
adjacency statistics	before fractionation analysis		
adjacencies	12,475	20,676	8,588
distinct (a)	12,475	16,165	8,588 (b)
distinct overall (c)	19,446		
excess (c-a)	6,971 (55.9%)	3,281 (20.3%)	
with ancestor (d)	9,390	11,094	total
excess (d-b)	802 (9.3%)	2,506 (29.2%)	3,308 (38.5%)
virtual genes	after fractionation analysis and consolidation		
fractionation intervals	2,462		1,888
single copies	10,674	0	8143 ¹
in syntenic pairs	0	$2 \times 10,674^2 = 21,348$	0
total	10,674	21,348	8143
adjacency statistics	after consolidation		
adjacencies	10,655	21,318	8,107
distinct (a)	10,655	13,309	8,107 (b)
distinct overall (c)	15,278		
excess (c-a)	4,623 (43.4%)	1,969 (14.8%)	
with ancestor (d)	9,079	9,844	total
excess (d-b)	972 (12.0%)	1,737 (21.4%)	2,709 (33.4%)

Delete one gene at a time or several?



fractionation: balanced or biased?

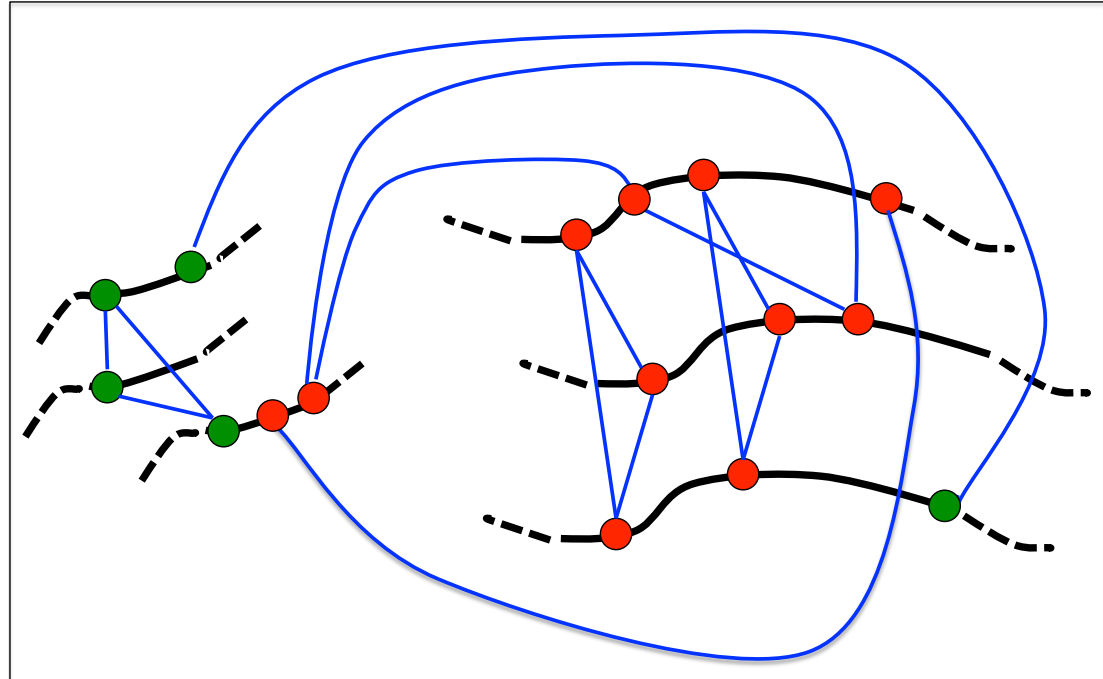
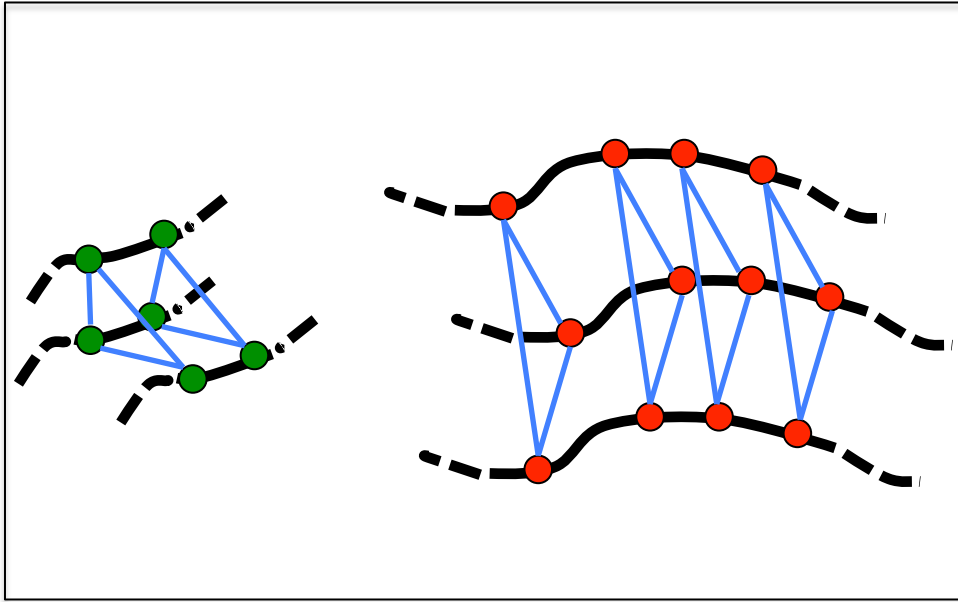


genome halving

genome aliquoting

Aliquoting: dissecting an ancient $2k$ -ploid (tetraploid, hexaploid, octoploid,...) into its k constituent subgenomes

— or whatever is left of them.



Algorithm aliquote

- **Parameters:** hypothesized ploidy parameter $k > 2$, short gap reward $r > 0$, jump $j < 0$, threshold $t \geq 0$.
- **Input:** $n > 0$ paralogy sets, each containing at most k genes. Genes distributed and ordered on C' chromosomes.
- **Output:** A number $C'' \geq 1$ of k -tuples of regions
- **Initialization:**
 - Each set of paralogs defines a k -tuple of regions, each region consisting of at most one fragment made up of one gene.
 - For all pairs of k -tuples of regions, calculate their clustering score.
- **while** there remain pairs of k -tuples of regions with positive score,
 - merge the pair of k -tuples of regions with highest score,
 - delete merged pairs and add the resulting larger k -tuple of regions,
 - calculate the clustering score of the new k -tuple of regions with all other k -tuples
- **Post-processing** If the gaps between two consecutive fragments in any region is smaller than threshold t , move the missing genes from their current location to fill in the gap. This may result in defects in the k -partition, and it is preferable to set t to as low a value as possible if this does not cause a proliferation of very small regions.



Journal content

→ [Journal home](#)

→ [Advance online publication](#)

→ [Current issue](#)

→ [Nature News](#)

→ [Archive](#)

→ [Supplements](#)

→ [Web focuses](#)

→ [Podcasts](#)

→ [Videos](#)

→ [News Specials](#)

Journal information

→ [About the journal](#)

→ [For authors](#)

→ [Online submission](#)

→ [Nature Awards](#)

→ [Nature history](#)

NPG services

→ [Advertising](#)

→ [work@npg](#)

→ [Reprints & permissions](#)

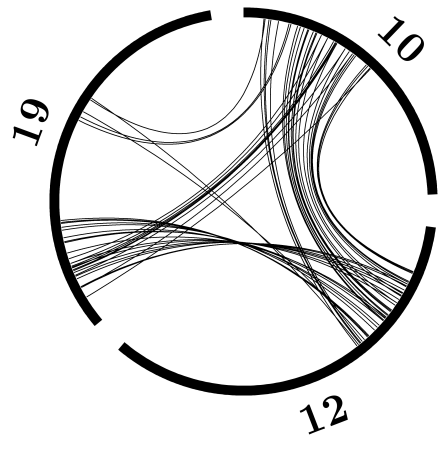
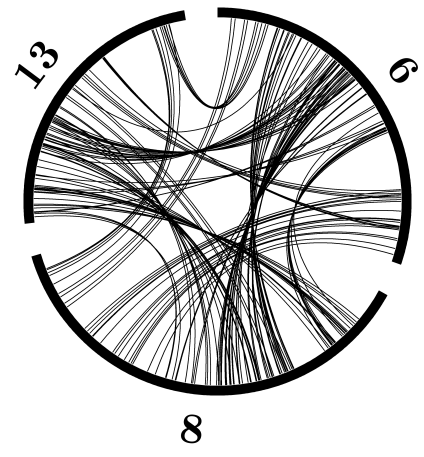
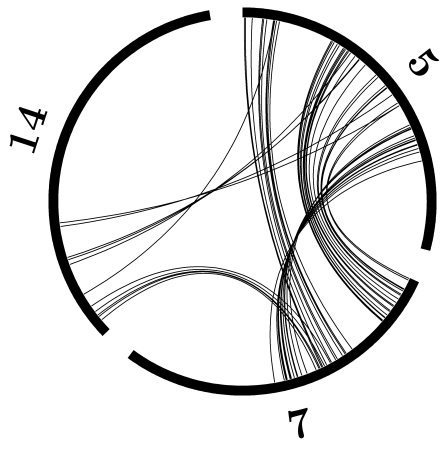
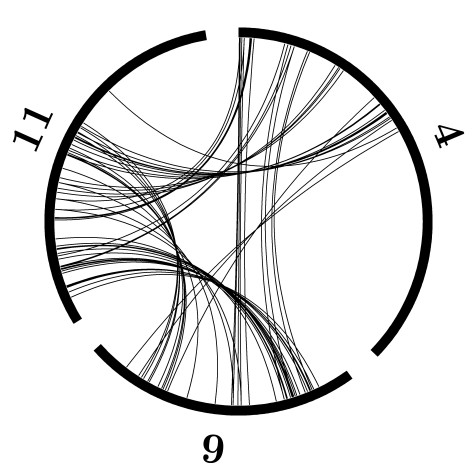
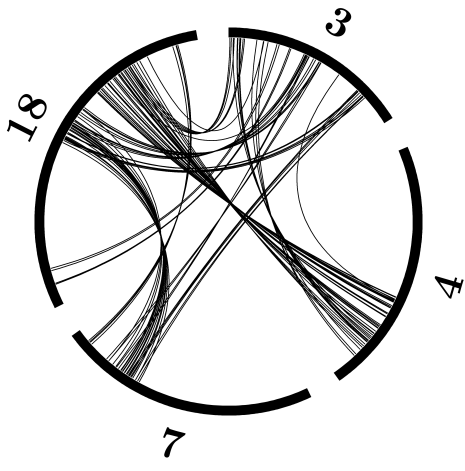
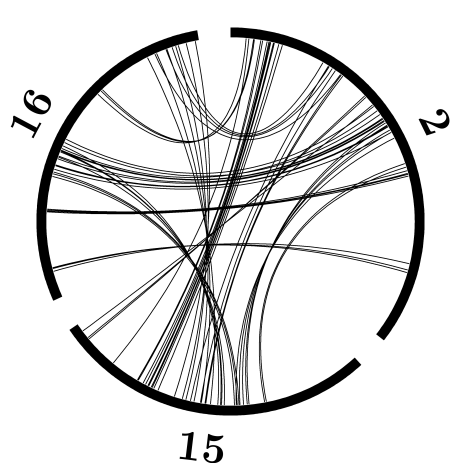
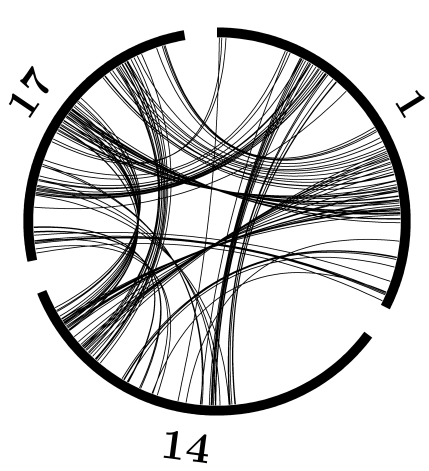
Letter

Nature **449**, 463–467 (27 September 2007) | doi:10.1038/nature06148; Received 5 April 2007; Accepted 7 August 2007; Published online 26 August 2007

The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla

Olivier Jaillon^{1,18}, Jean-Marc Aury^{1,18}, Benjamin Noel¹, Alberto Policriti^{2,3}, Christian Clepet⁴, Alberto Casagrande^{2,5}, Nathalie Choisne^{1,4}, Sébastien Aubourg⁴, Nicola Vitulo^{6,15}, Claire Jubin¹, Alessandro Vezi^{6,15}, Fabrice Legeai⁷, Philippe Huguency⁸, Corinne Dasilva¹, David Horner^{9,15}, Erica Mica^{9,15}, Delphine Jublot⁴, Julie Poulain¹, Clémence Bruyère⁴, Alain Billault¹, Béatrice Segurens¹, Michel Gouyvenoux¹, Edgardo Ugarte¹, Federica Cattonaro², Véronique Anthouard¹, Virginie Vico¹, Cristian Del Fabbro^{2,3}, Michaël Alaux⁷, Gabriele Di Gasparo^{2,5}, Vincent Dumas⁸, Nicoletta Felice^{2,5}, Sophie Paillard⁴, Irena Juman^{2,5}, Marco Moroldo⁴, Simone Scalabrin^{2,3}, Aurélie Canaguier⁴, Isabelle Le Clainche⁴, Giorgio Malacrida^{6,15}, Eléonore Durand⁷, Graziano Pesole^{10,11,15}, Valérie Laucou¹², Philippe Chatelet¹³, Didier Merdinoglu⁸, Massimo Delledonne^{14,15}, Mario Pezzotti^{15,16}, Alain Lecharny⁴, Claude Scarpelli¹, François Artiguenave¹, M. Enrico Pè^{9,15}, Giorgio Valle^{6,15}, Michele Morgante^{2,15}, Michel Caboche⁴, Anne-Françoise Adam-Blondon⁴, Jean Weissenbach¹, Francis Quétier¹ & Patrick Wincker¹ for The French–Italian Public Consortium for Grapevine Genome Characterization ¹⁷

1. Genoscope (CEA) and UMR 8030 CNRS-Genoscope-Université d'Evry, 2 rue Gaston Crémieux, BP5706, 91057 Evry, France.
2. Istituto di Genomica Applicata, Parco Scientifico e Tecnologico di Udine, Via Linussio 51, 33100 Udine, Italy.
3. Dipartimento di Matematica ed Informatica, Università degli Studi di Udine, via delle Scienze 208, 33100 Udine, Italy.
4. URGV, UMR INRA 1165, CNRS-Université d'Evry Genomique Végétale, 2 rue Gaston Crémieux, BP5708, 91057 Evry cedex, France.
5. Dipartimento di Scienze Agrarie ed Ambientali, Università degli Studi di Udine, via delle Scienze 208, 33100 Udine, Italy.
6. CRIBI, Università degli Studi di Padova, viale G. Colombo 3, 35121 Padova, Italy.
7. URGI, UR1164 Génomique Info, 523, Place des Terrasses, 91034 Evry Cedex, France.
8. UMR INRA 1131, Université de Strasbourg, Santé de la Vigne et Qualité du Vin, 28 rue de Herrlisheim, BP20507, 68021 Colmar, France.
9. Dipartimento di Scienze Biomolecolari e Biotecnologie, Università degli Studi di Milano, via Celoria 26, 20133 Milano, Italy.





[» Take the Nature Publishing Group survey for the chance to win a MacBook Air](#)

NATURE | LETTER OPEN

[日本語要約](#)

The tomato genome sequence provides insights into fleshy fruit evolution

[The Tomato Genome Consortium](#)[Affiliations](#) | [Contributions](#) | [Corresponding authors](#)*Nature* **485**, 635–641 (31 May 2012) | doi:10.1038/nature11119

Received 17 August 2011 | Accepted 03 April 2012 | Published online 30 May 2012



PDF



Citation



Reprints



Rights & permissions



Article metrics

Tomato (*Solanum lycopersicum*) is a major crop plant and a model system for fruit development. *Solanum* is one of the largest angiosperm genera¹ and includes annual and perennial plants from diverse habitats. Here we present a high-quality genome sequence of domesticated tomato, a draft sequence of its closest wild relative, *Solanum pimpinellifolium*², and compare them to each other and to the potato genome (*Solanum tuberosum*). The two tomato genomes show only 0.6% nucleotide divergence and signs of recent admixture, but show more than 8% divergence from potato, with nine large and several smaller inversions. In contrast to *Arabidopsis*, but similar to soybean, tomato and potato small RNAs map predominantly to gene-rich chromosomal regions, including gene promoters. The *Solanum* lineage has experienced two consecutive genome triplications: one that is ancient and shared with rosids, and a more recent one. These triplications set

S

C

D

O

S

P

F

T

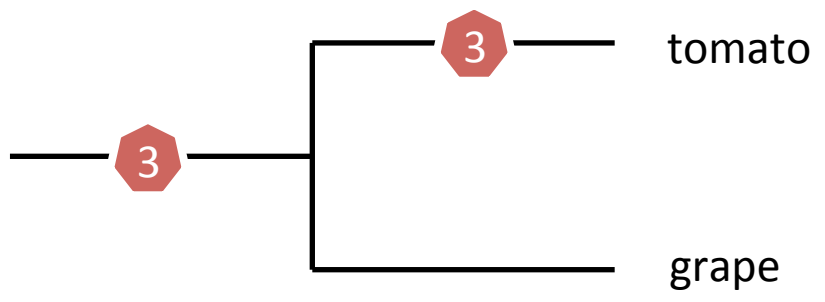
S

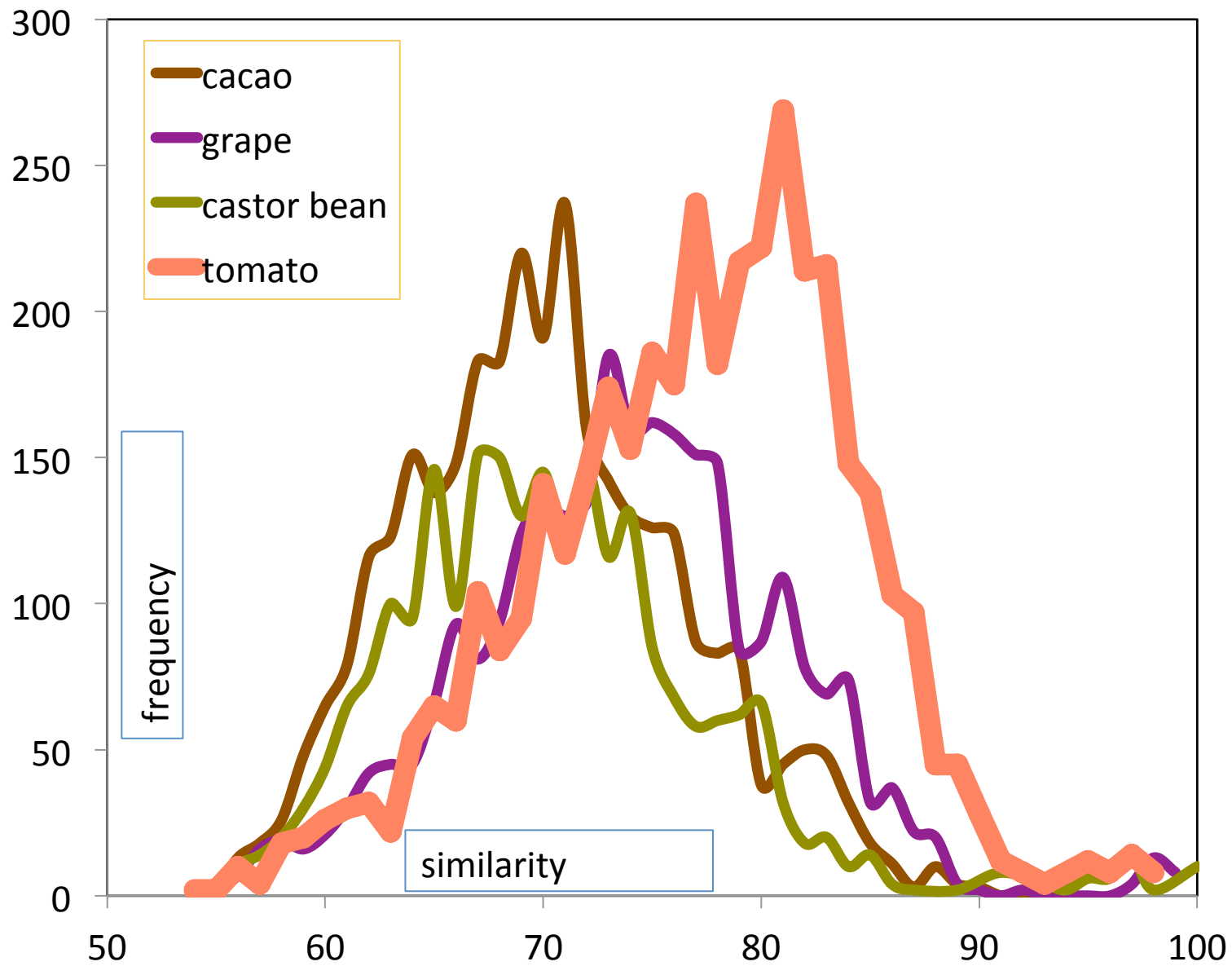
C

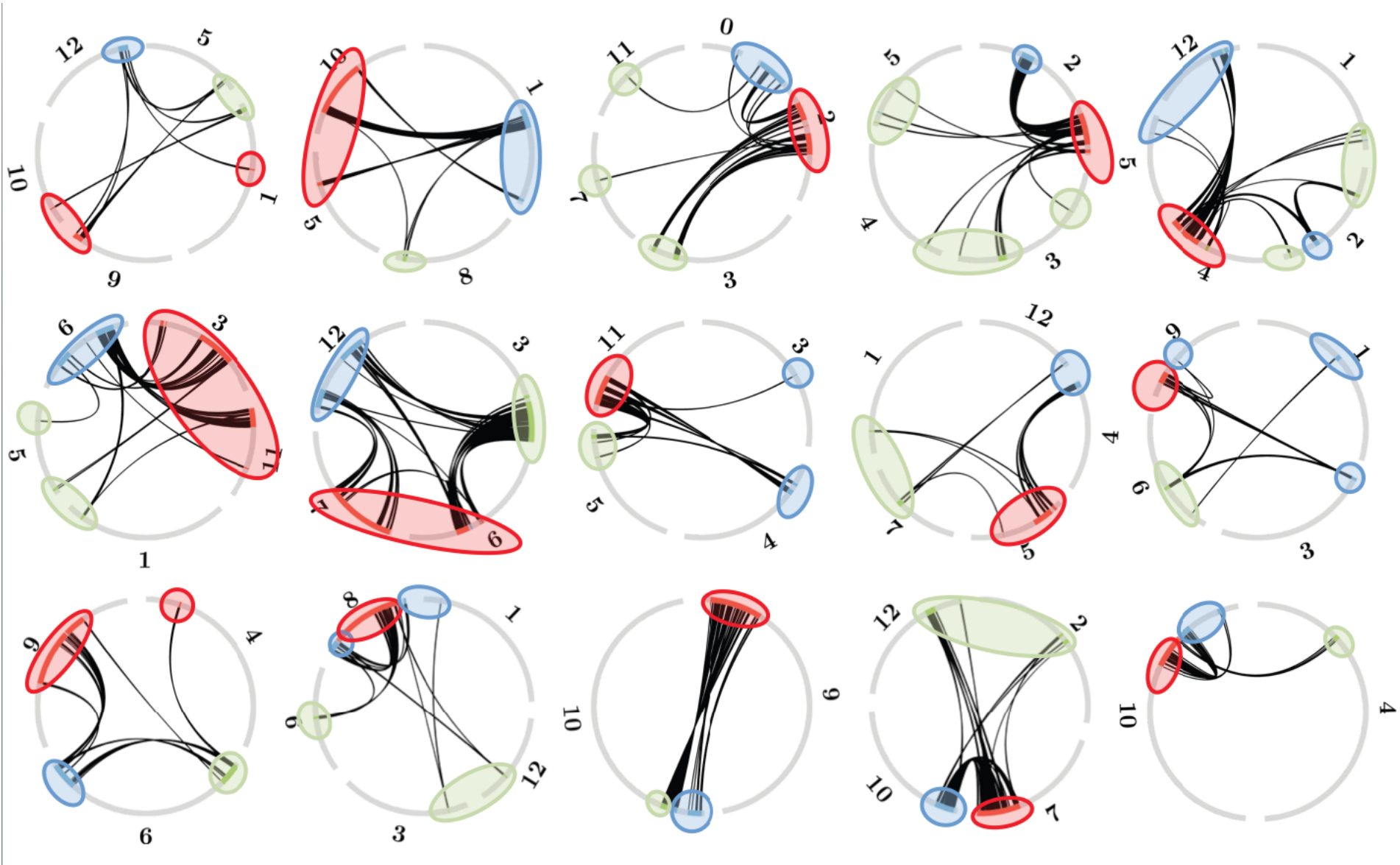
Ir

J

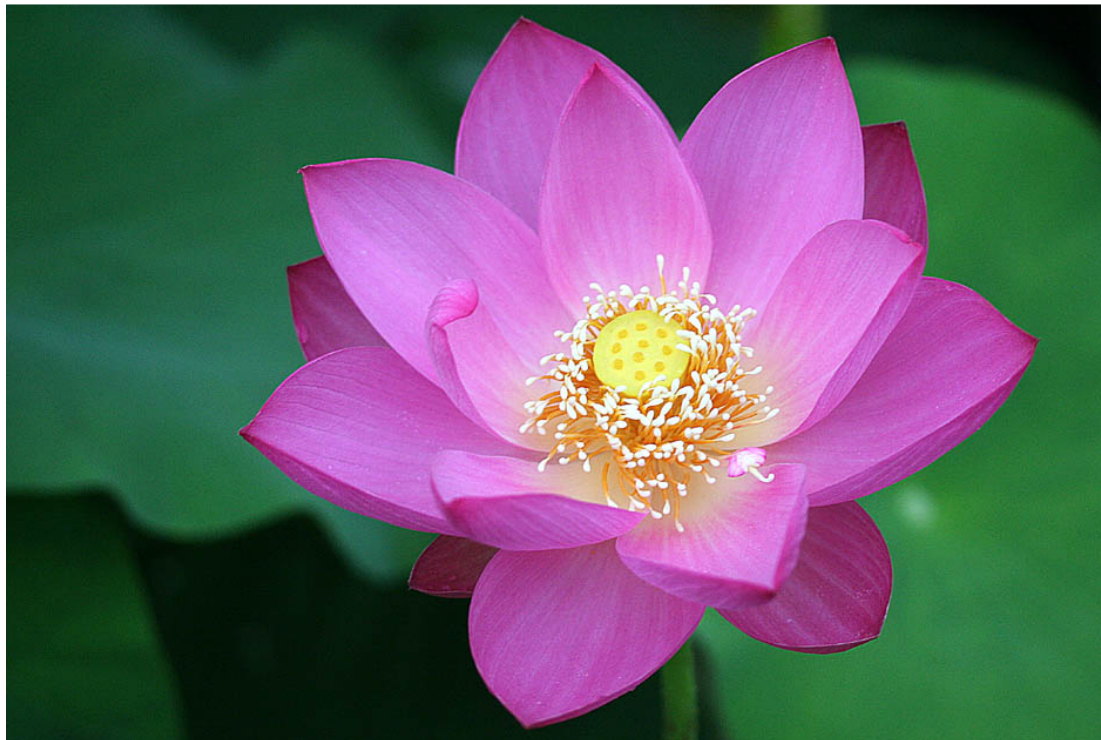
P







Ming R et al: Genome of the long-living sacred lotus (*Nelumbo nucifera* Gaertn.) *Genome Biology* 2013, **14**: R41







wiseGEEK



Aliquoting: dissecting an ancient $2k$ -ploid (tetraploid, hexaploid, octoploid,...) into its k constituent subgenomes

— or whatever is left of them.

$k = 2$: halving

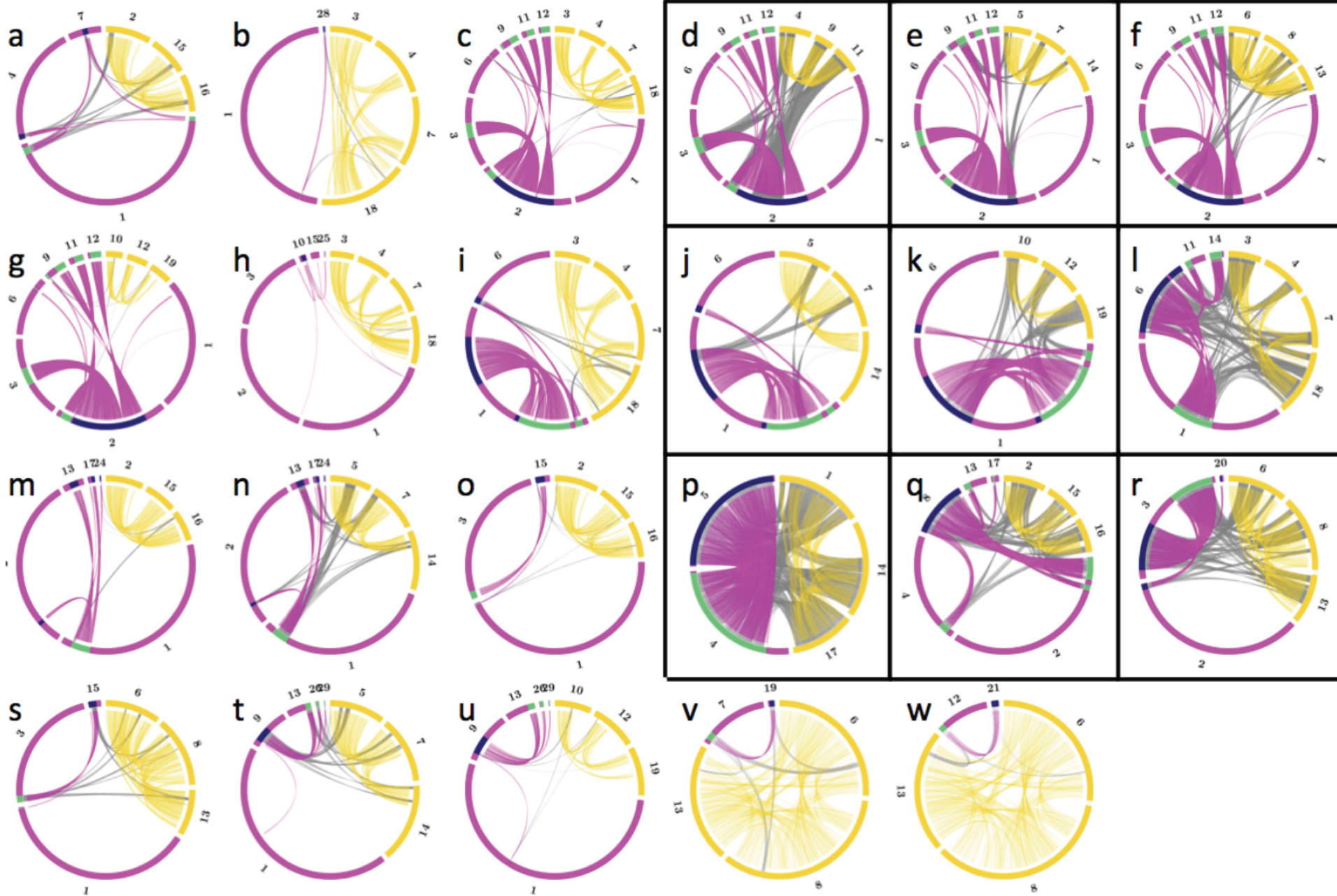
“practical halving”

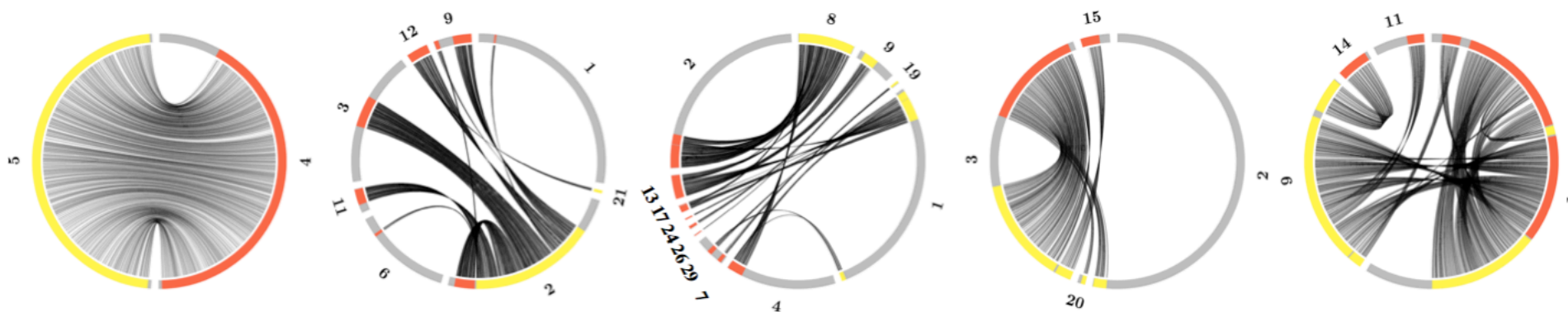
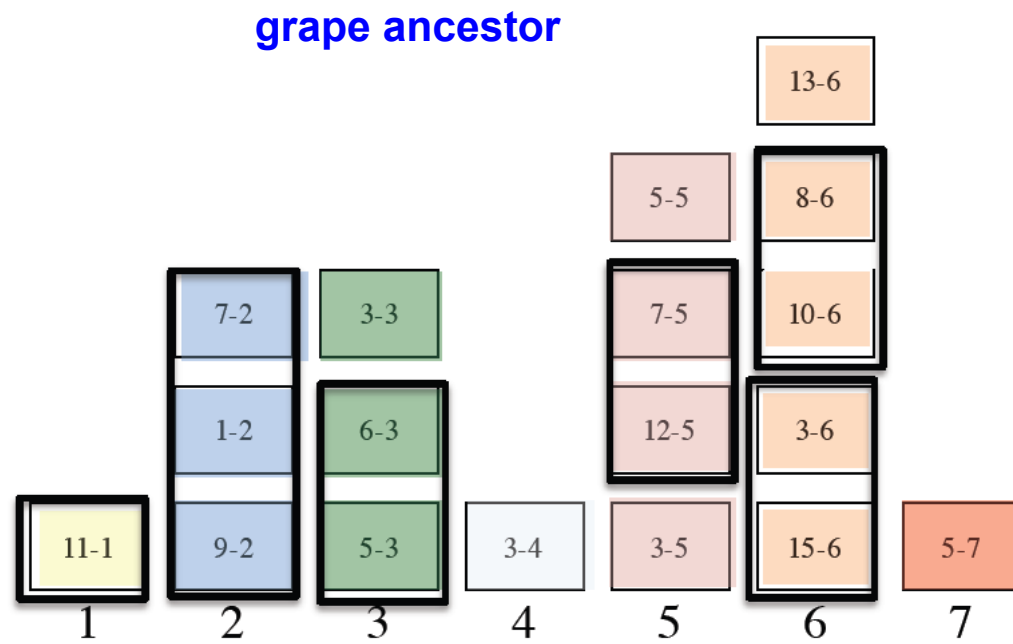
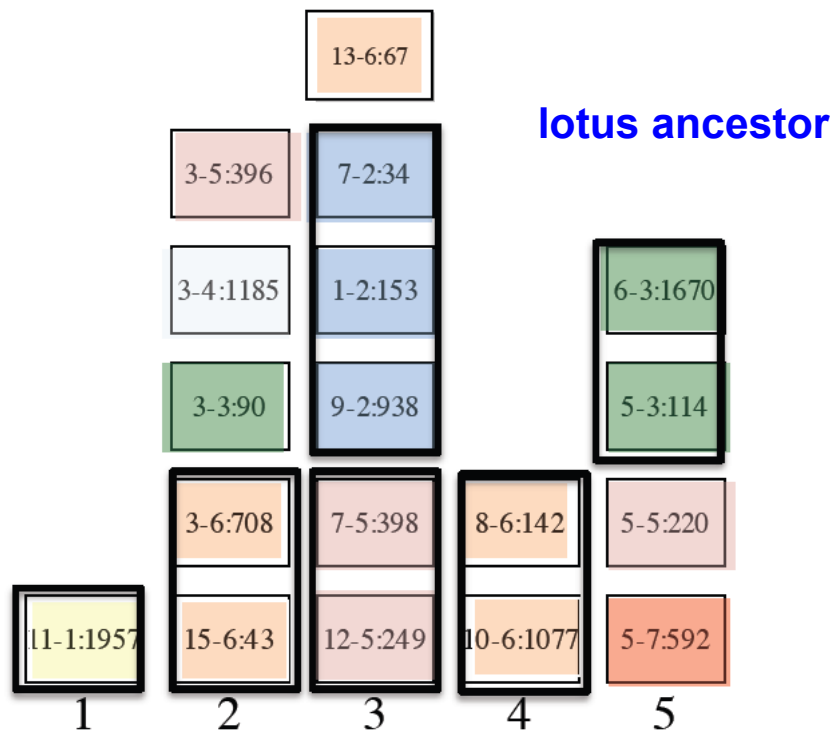
Number of genes in first estimate of chromosome pairs.

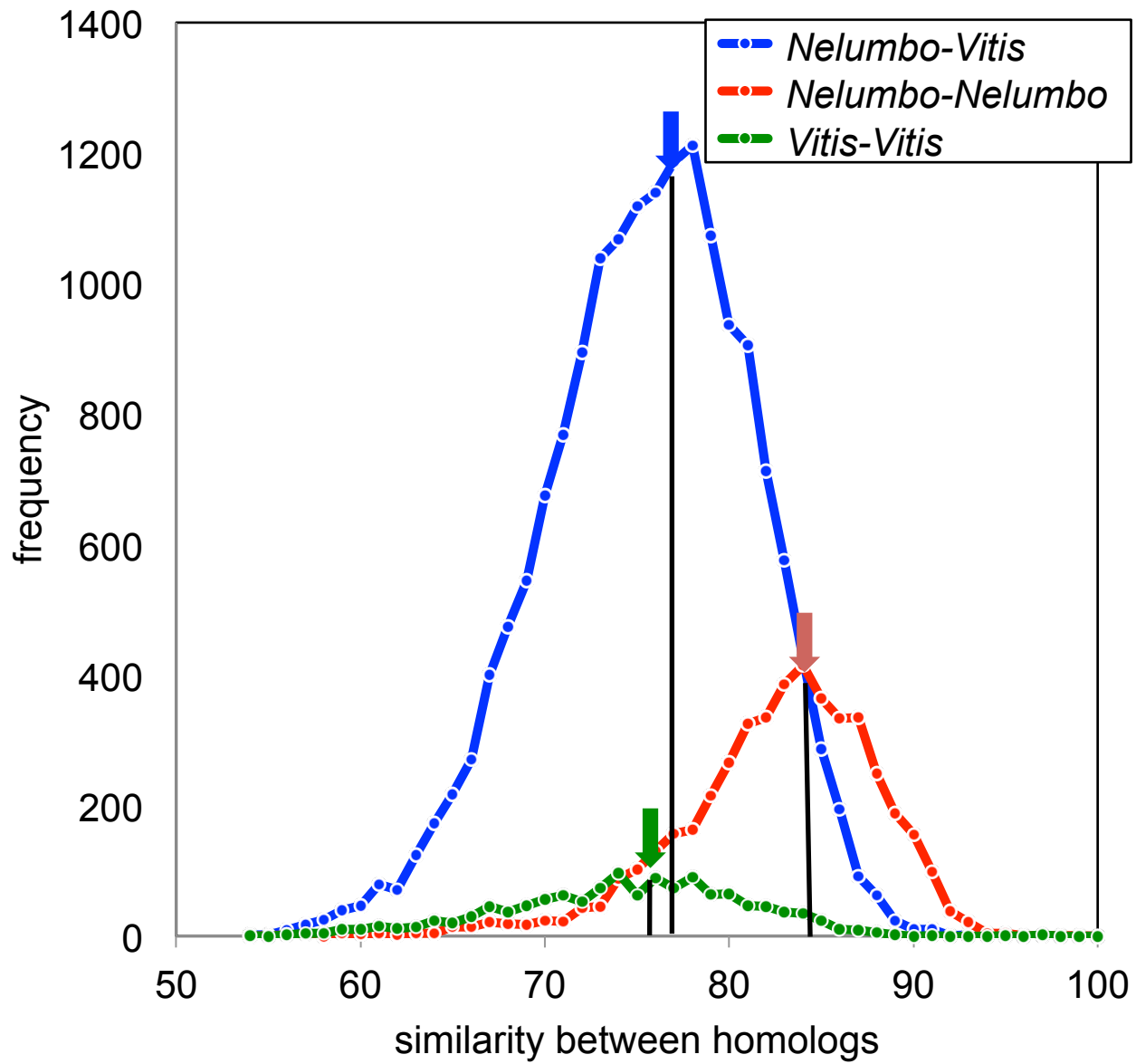
chromosome															
pair	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
genes	234	29	2717	124	1469	1839	462	195	1152	1134	2114	271	67	17	65

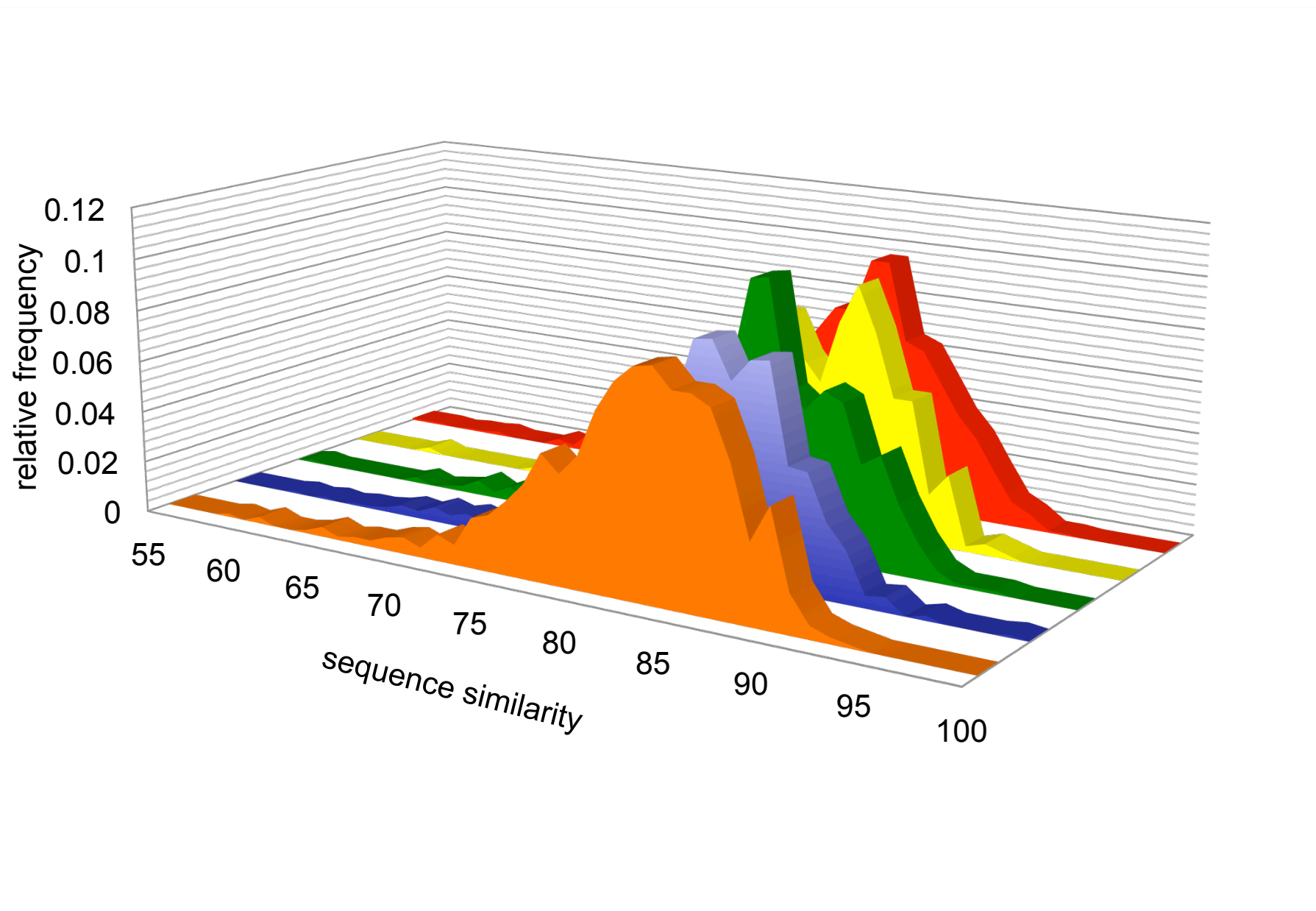
15 candidate ancestral chromosome pairs x 7 grape ancestral chromosomes

= 105 combinations



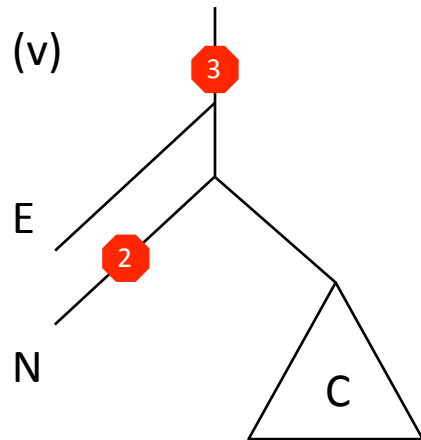
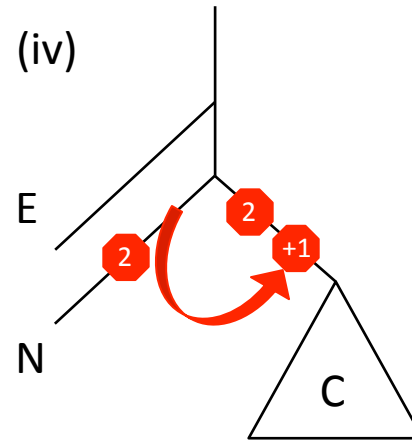
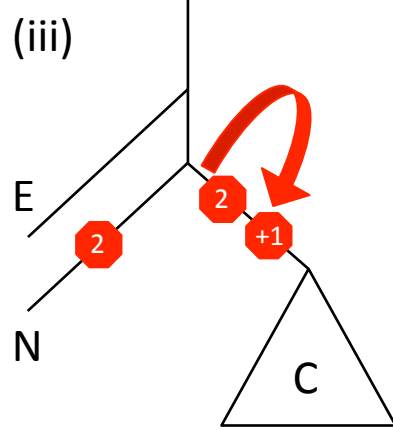
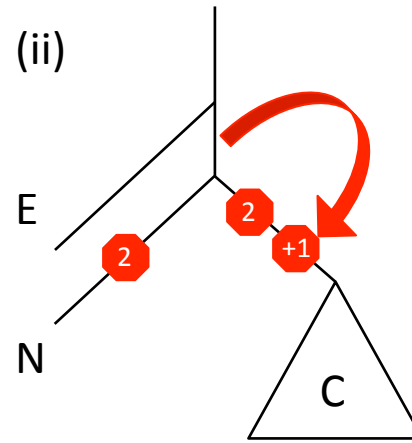
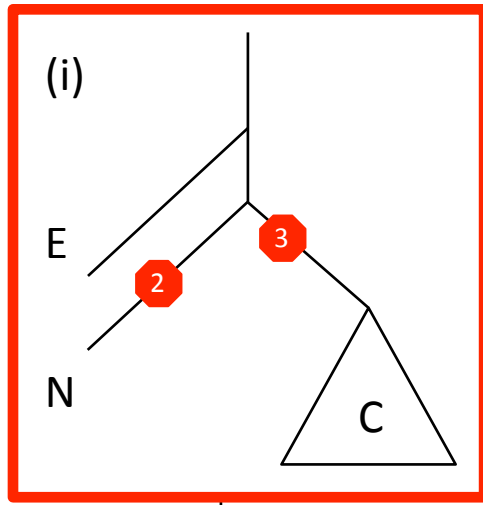






Similarities between *Nelumbo-Vitis* orthologs and between *Vitis* paralogs. Numbers in blue indicate the larger of two similarities. There is no tendency for the non-dominant genomes to be more similar, and there are almost no statistically significant differences (t-test) in any case. The more significant comparison is listed in red.

similarity of genes in <i>Vitis</i> subgenomes to <i>Nelumbo</i> orthologs				similarity of genes in pairs of <i>Vitis</i> subgenomes								
	two subgenomes being compared											
	small-medium		medium-large		small-large							
triple "colour"	720	795	795	1086	720	1086	156	172	172	186	156	186
yellow	75.47	75.63	75.47	75.62	75.63	75.62	73.45	74.2	74.2	74.28	73.45	74.28
	p<.60		p<.57		p<.97		p<.32		p<.91		p<.27	
dark blue	393	525	525	569	393	569	66	94	94	132	66	132
	75.43	74.83	74.83	75.86	75.43	75.86	72	73.13	73.13	73.64	72	73.64
	p<.11		p<.0023		p<.22		p<.21		p<.53		p<.09	
green	561	563	563	1218	561	1218	74	114	114	190	74	190
	75.46	75.35	75.35	75.45	75.46	75.45	76.11	73.82	73.82	73.75	76.11	73.75
	p<.75		p<.72		p<.99		p<.0156		p<.92		p<.002	
light blue	482	502	502	617	482	617	74	122	122	128	74	128
	75.38	75.41	75.41	75.93	75.38	75.93	73.97	76.12	76.12	75.03	73.97	75.03
	p<.93		p<.11		p<.11		p<.0317		p<.16		p<.29	
purple	315	596	596	714	315	714	20	30	30	161	20	161
	75.90	75.48	75.48	75.75	75.90	75.75	74.2	75.2	75.2	74.16	74.2	74.16
	p<.29		p<.38		p<.70		p<.66		p<.40		p<.98	
orange	730	967	967	1091	730	1091	174	197	197	258	174	258
	75.48	75.69	75.69	75.56	75.48	75.56	75.20	75.27	75.27	75.55	75.20	75.55
	p<.45		p<.59		p<.77		p<.92		p<.63		p<.58	
red	197	233	233	354	197	354	50	53	53	94	50	94
	74.88	75.70	75.70	75.39	74.88	75.39	74.08	73.91	73.91	75.34	74.08	75.34
	p<.15		p<.51		p<.31		p<.90		p<.15		p<.26	



Team:

Chunfang Zheng

Katharina Jahn (Bielefeld)