

GN/HS 757

Quantitative Genetics Theory and Methods

Fall 2022

5. Linkage, Phasing and Haplotyping in Polyploid

Marcelo Mollinari

mmollin@ncsu.edu

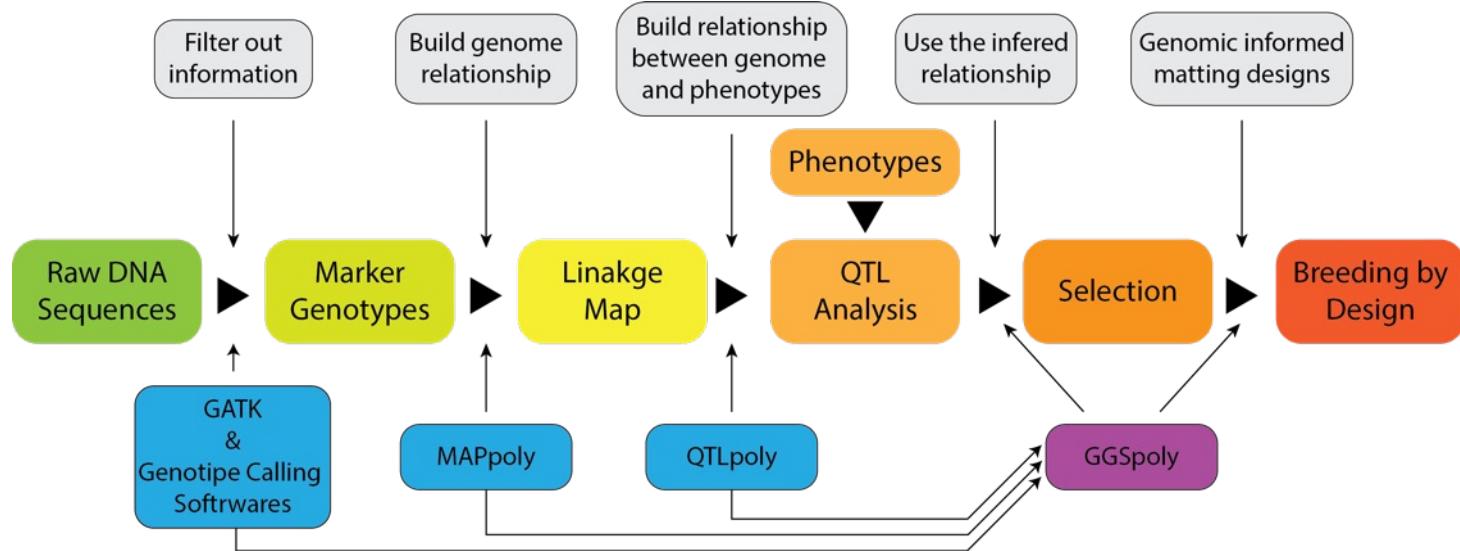
Bioinformatics Research Center

Department of Horticultural Science

North Carolina State University

Objective

- Develop computational tools for the whole pipeline data analysis in polyploids

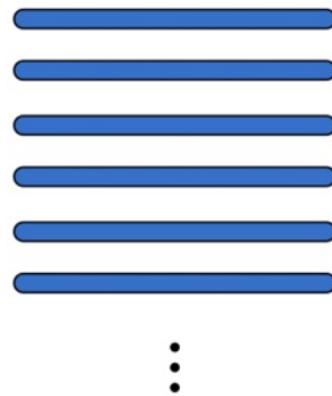


- **GATK and calling software:** Variant discovery and genotype calling
- **MAPpoly:** Genetic mapping of autopolyploid species: 4x, 6x, and 8x. Also, 2x.
- **QTLpoly:** QTL analysis in autopolyploid species: 4x, 6x, and 8x. Also, 2x.
- **GGSpoly:** a **user-friendly computational tool** to assist breeders in making long and short-term **breeding decisions** based on collected and learned **information** about their breeding populations.

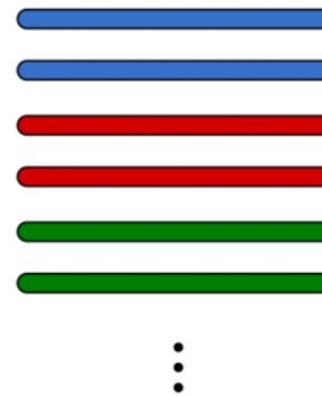
Polyplloid Species

- Organisms with more than 2 complete sets of chromosomes

Autohexaploid

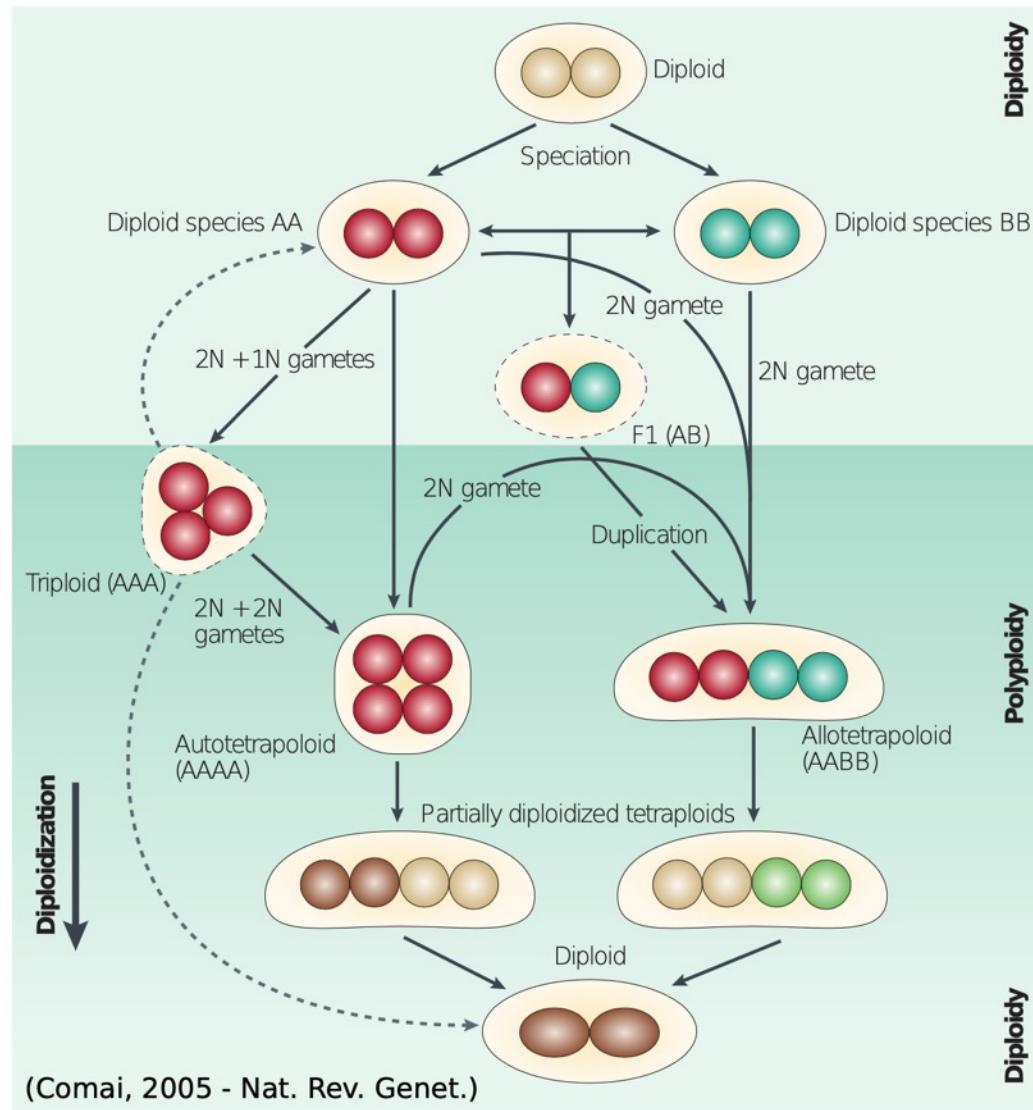


Allohexaploid



- Multiple sets have the **same** origin
- Multiple sets have **different** origins

How are polyploids formed



Biparental Population - BT

- Beauregard x Tanzania
- 315 individuals
- GBS – GBSpoly protocol (Bode Olukolu – U Tennessee)
- Two reference genomes *I. trifida* and *I. triloba* (Zhangjun Fei's group – BTI Cornell)



Beauregard



Tanzania



<https://doi.org/10.1534/g3.119.400620>



Biparental Population - BT



X

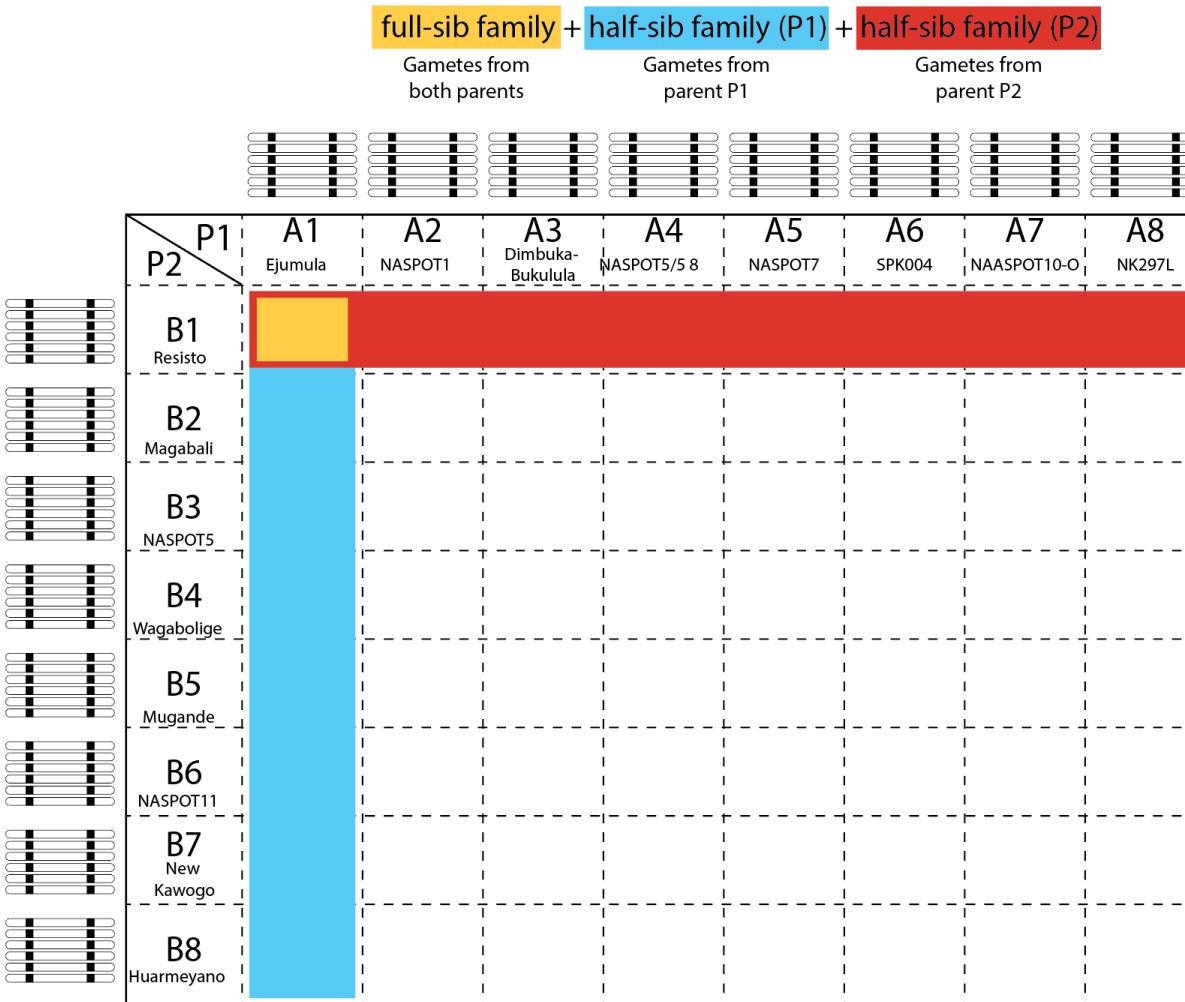


Beauregard

Tanzania

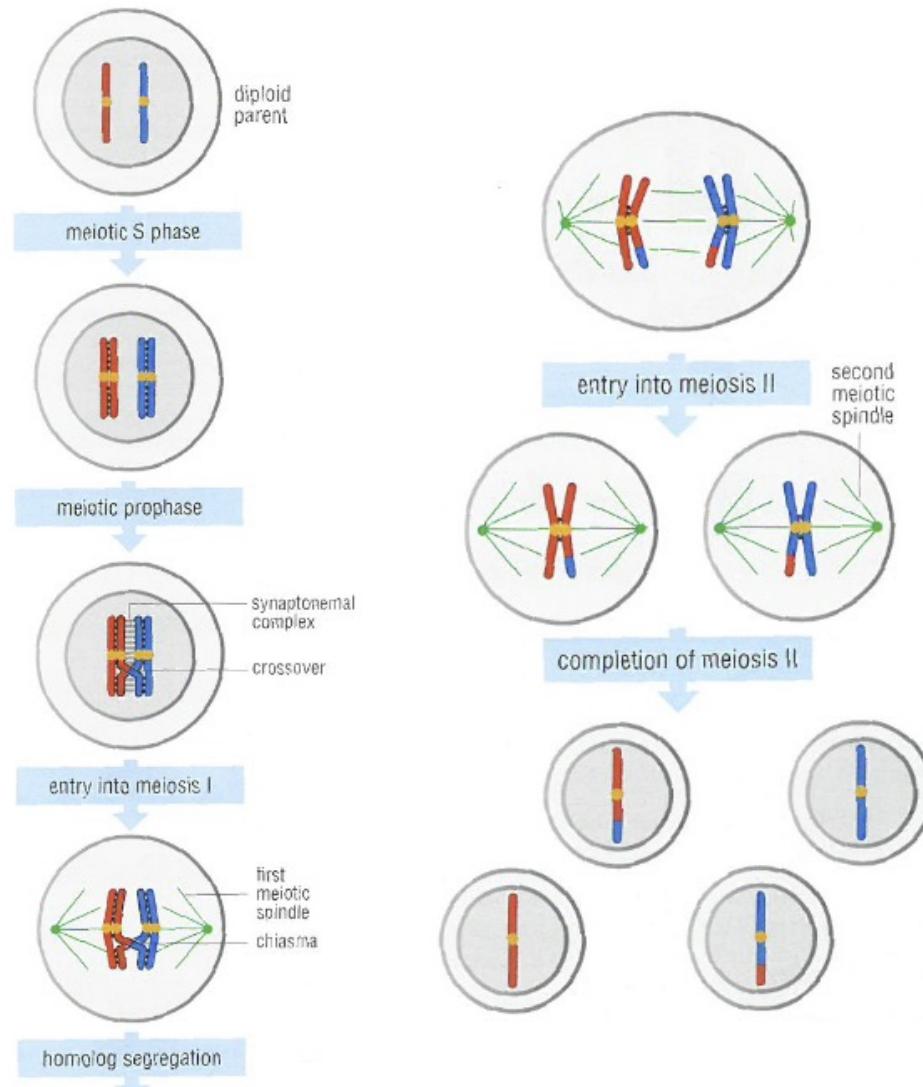


Mwanga Diversity Panel (MDP) is an 8 x 8 factorial

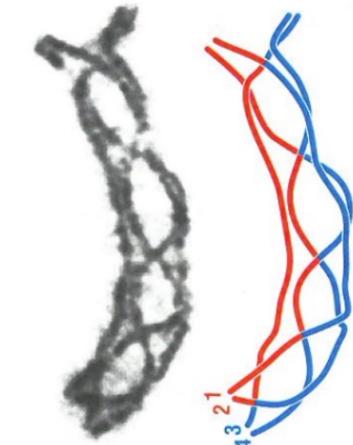
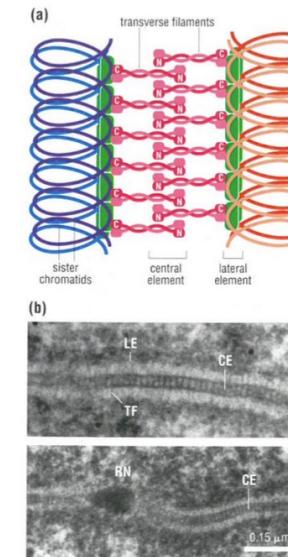


- Up to 96 different alleles at a given locus
- Full-sib families ~ 30 ind.
- Half-sib families ~ 240 ind.
- Total ~ 1800 ind.
- 200 ind. hard to keep alive
- **Recombination fraction** between markers from a joint analysis
- **Specific linkage phases** for each parent
- Reconstruct the **haplotypes of each offspring** in terms of the haplotypes of the 16 parents.

Meiosis

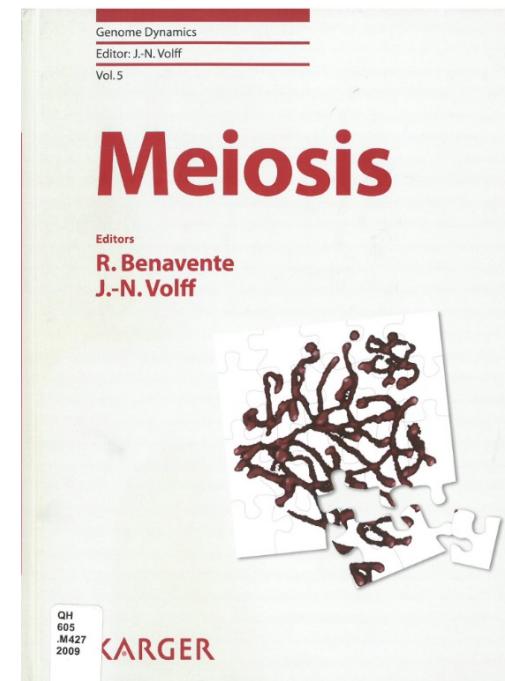
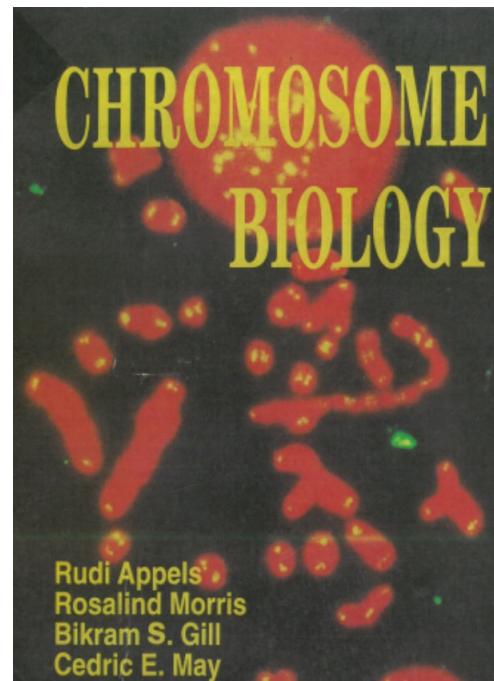
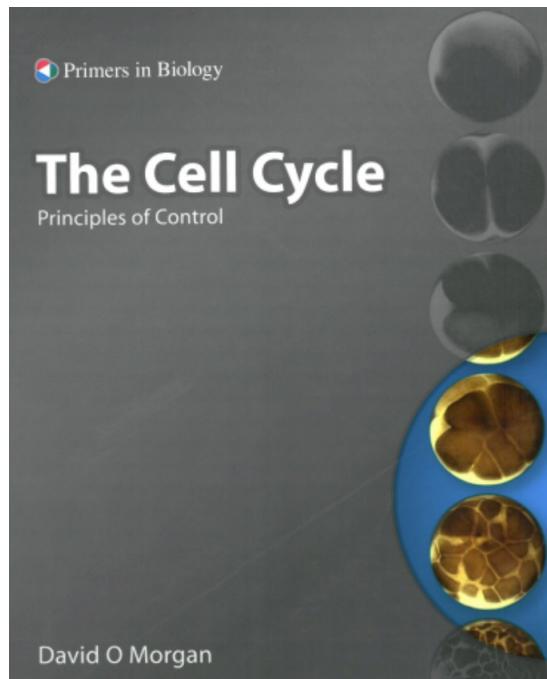


Meiosis is a specialized form of nuclear division that generates nuclei carrying **half** the normal complement of chromosomes.



Synaptonemal complex

Meiosis literature



An overview of plant recombination data

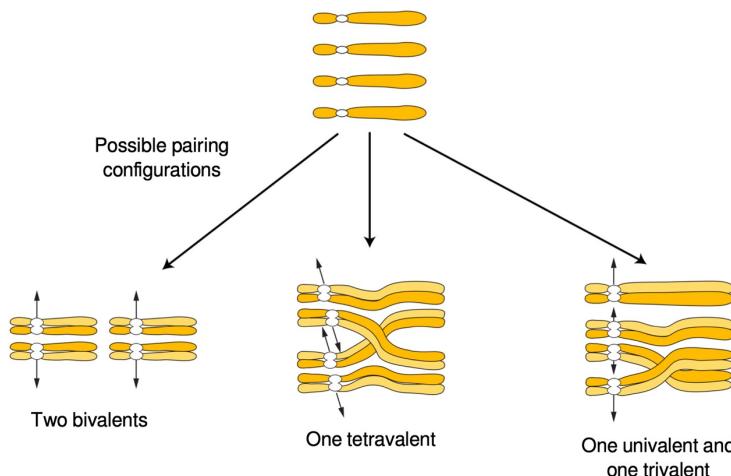
	Organisms	Genome size Mb	Haploid chr. number	Genetic size cM	cM/Mb	ENs or RAD51/DMC1 foci	LNs or chiasma number	CO/DSB ^f
Arabidopsis	<i>A. thaliana</i>	120 [74]	5	470 [75]	3.9	220 [14, 53]	9.2 [76]	24
Medicago	<i>M. truncatula</i>	475 ^a	8	1,125 [77]	2.4			
Lotus	<i>O. sativa</i>	430 [78]	12	1,530 [79]	3.55			
Rice	<i>L. japonicus</i>	475 ^a	6	500 [80]	1.05			
Populus	<i>P. trichocarpa</i>	485 [81]	19	2,500 [82]	5.2			
Tomato	<i>L. esculentum</i>	824 ^a	12	1,469 [83]	1.8	292 ^e	22 [84]	13
Palm	<i>E. guineensis</i>	1,750 [85]	16	1,743 [86]	1			
Maize	<i>Z. mays</i>	2,365 ^b	10	1,729 ^b	0.73	500 [87]	21.9 [88]	23
Rye	<i>S. cereale</i>	8,300 ^a	7	921 ^c	0.11			
Green Onion	<i>A. fistulosum</i>	9,900 ^a	8			669 [89]	15 [90]	44
Garlic	<i>A. sativa</i>	11,400 ^a	21	2,932 [91]	0.25			
Onion	<i>A. cepa</i>	15,000 ^a	8	2,000 [92]	0.13	614 [89]	19 [90]	32
Wheat	<i>T. aestivum</i>	17,000 ^d	42	3,600 ^d	0.2			
Lily	<i>L. longiflorum</i>	19,500 ^a	12			2,000 [93]	55 [94]	36
Sweetpotato	<i>I. batatas</i>	526.4	15	2,708	0.12			

^awww.rbgkew.org.uk/cval^bwww.maizegdb.org^c<http://www.ncbi.nlm.nih.gov>^dP. Sourdille, pers. com.^eL. Anderson, pers. com.

^fThe ratio CO/DSB is calculated by considering that the number of ENs or RAD51/DMC1 foci is equivalent to the number of DSB sites.

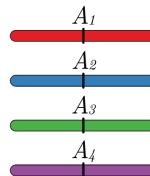


Gamete formation in polyploids*



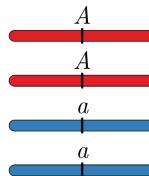
Griffiths et al. (2004)

Multiallelic



- A_1A_2
- A_1A_3
- A_1A_4
- A_2A_3
- A_2A_4
- A_3A_4

Biallelic



- | | | |
|-------|--|---|
| $A A$ | | 1 |
| $A a$ | | 4 |
| $a a$ | | 1 |

Number of possible gamtes considering
one locus with no double-reduction in
one and two parents

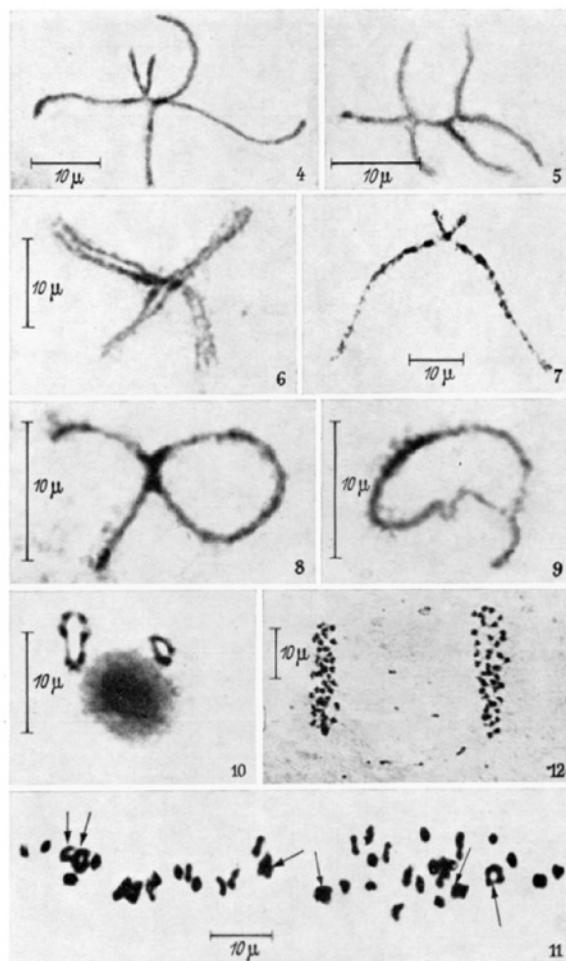
Ploidy	$\left(\frac{p}{2}\right)$	$\left(\frac{p}{2}\right)^2$
4	6	36
6	20	400
8	70	4900
10	252	63504
12	924	853776
14	3432	11778624
16	12870	165636900

*random pairing and no double reduction



Meiotic Pairing in polyploids

Sweetpotato



Hexavalents, quadrivalents and bivalents
in sweetpotato (Magoon *et al.* 1970)

Sugarcane

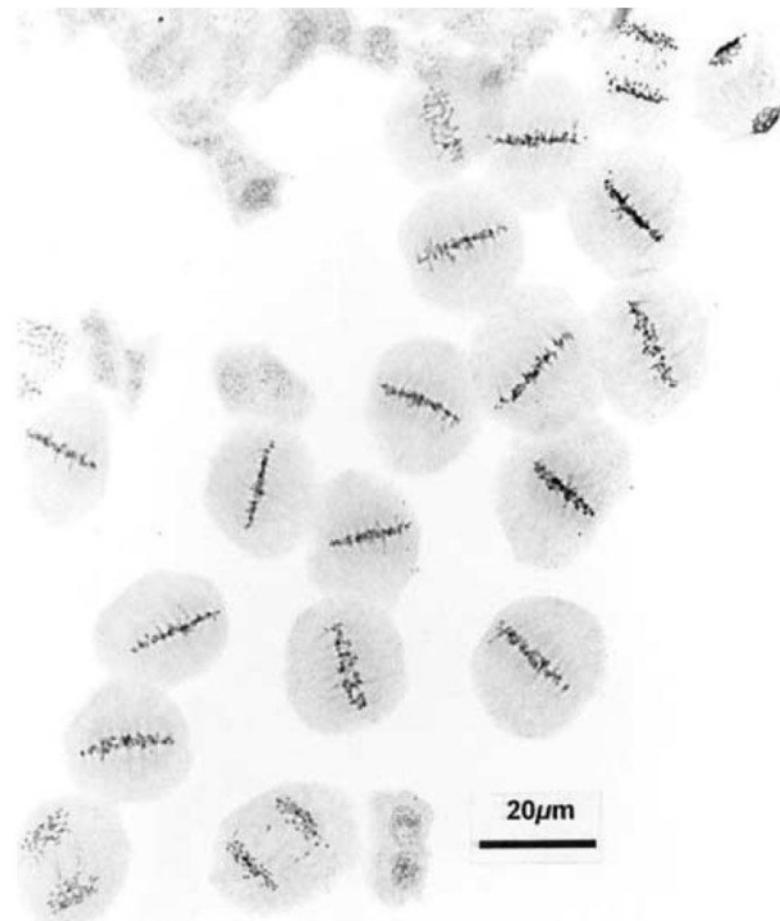
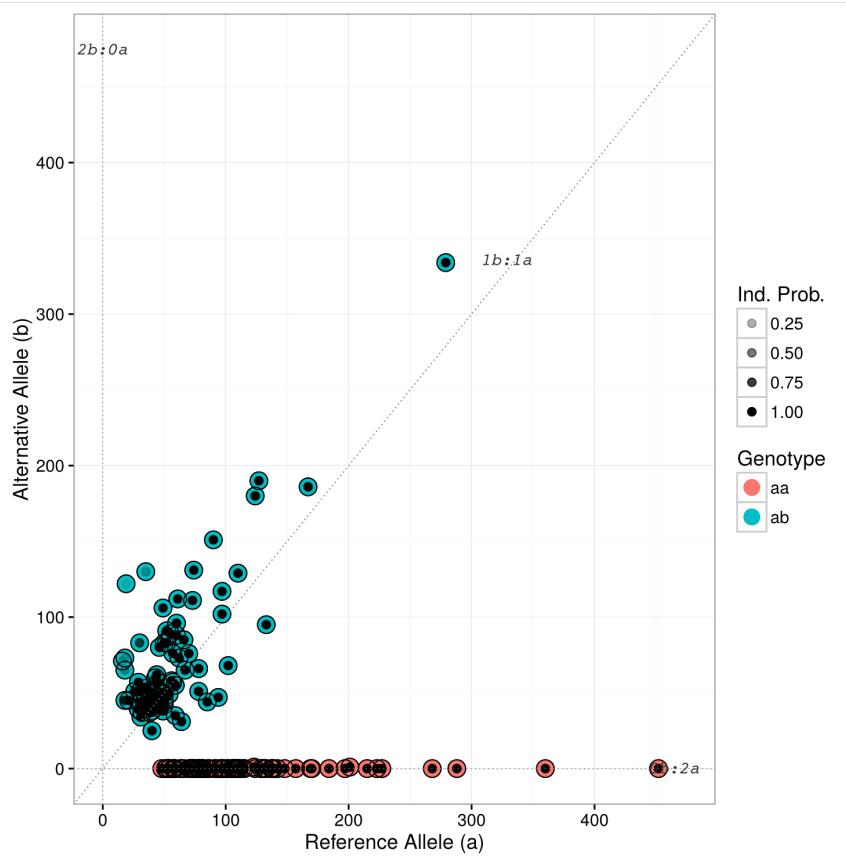


Figure 1. PMCs of *Saccharum* spp. hybrid clone 79N9059 at meiosis. As was the case in other clones, pairing was regular, bivalents generally formed. (Bielig *et al.* 2003)

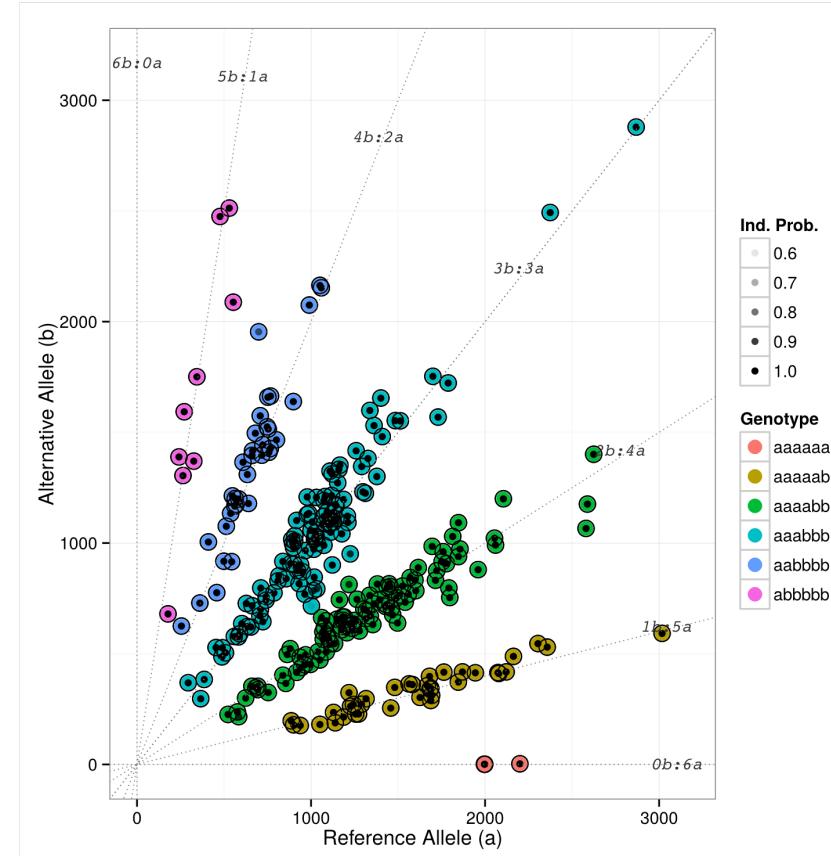
Genotyping Calling using SuperMASSA

Dosage calling Including the probability distribution of the genotypes

Diploid



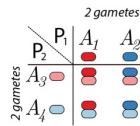
Hexaploid



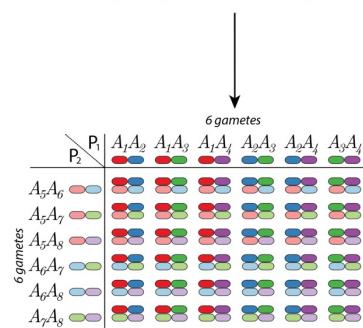
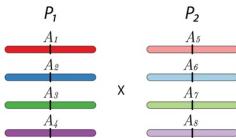
Segregation in polyploids*

Multiallelic

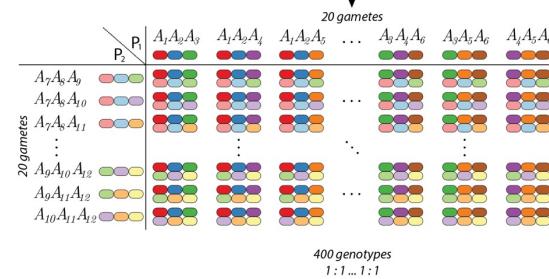
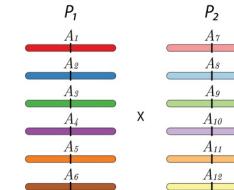
Diploid



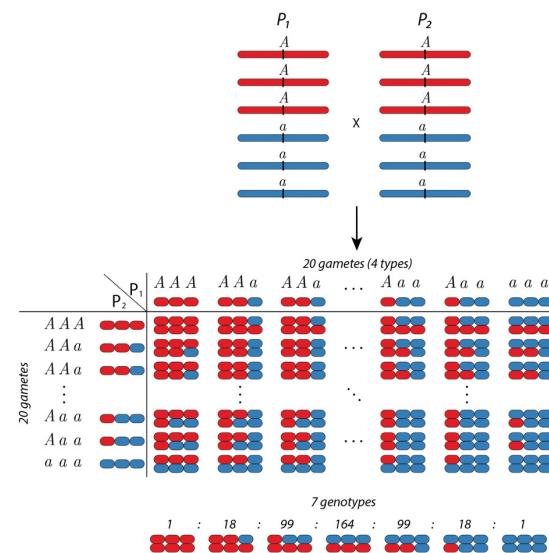
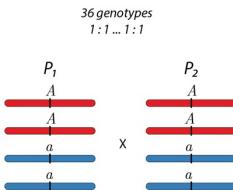
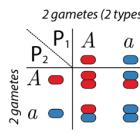
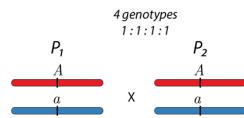
Tetraploid



Hexaploid



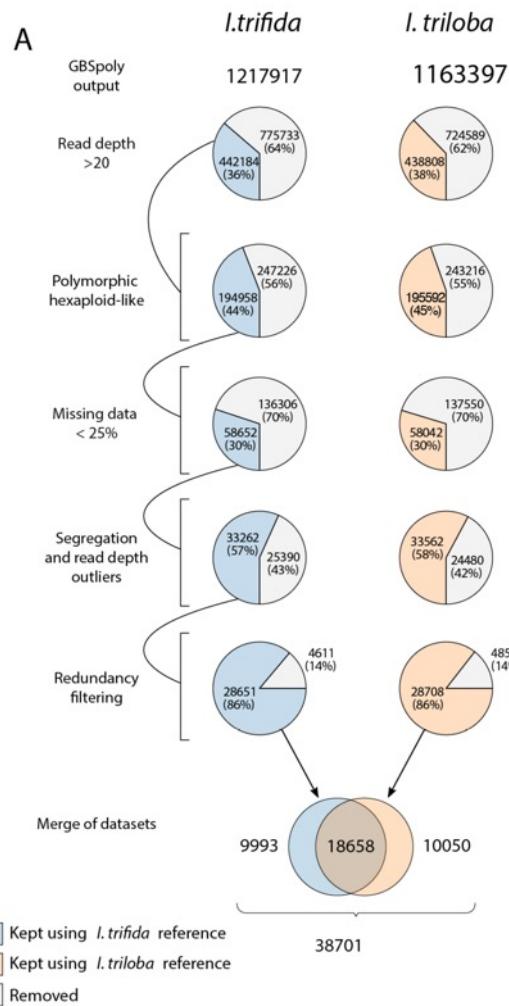
Biallelic



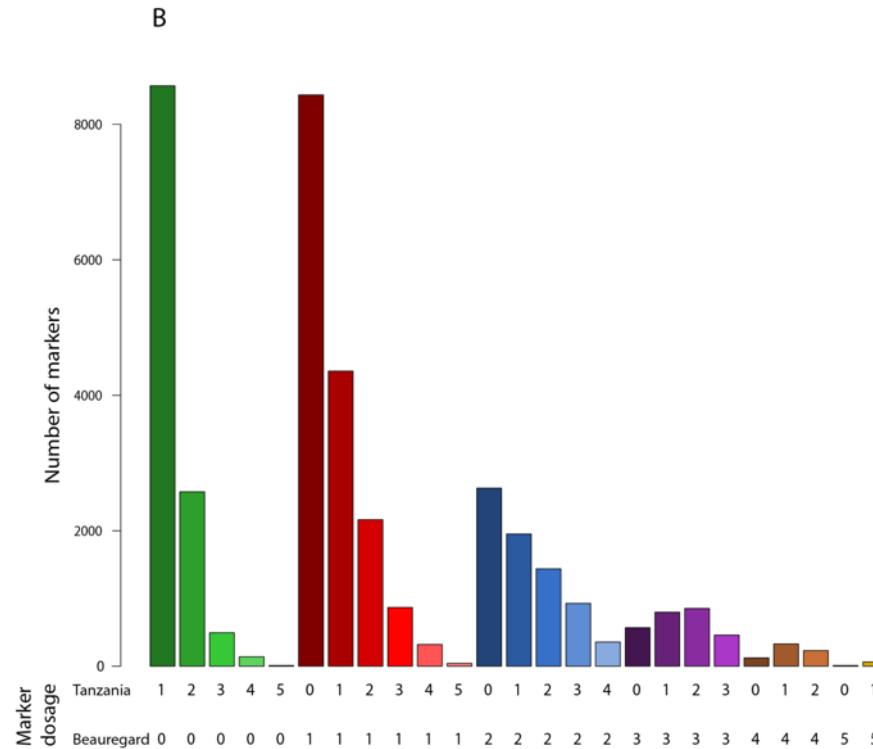
*random pairing and no double reduction

Results - Genotyping Calling - BT population

Filtering process

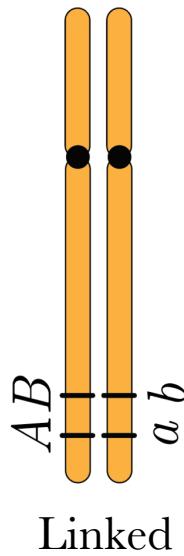


Distribution of the marker doses in *Beauregard* and *Tanzania* (38,701 SNPs).

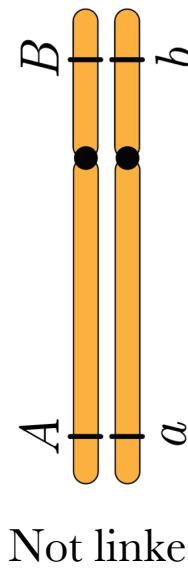


Introduction

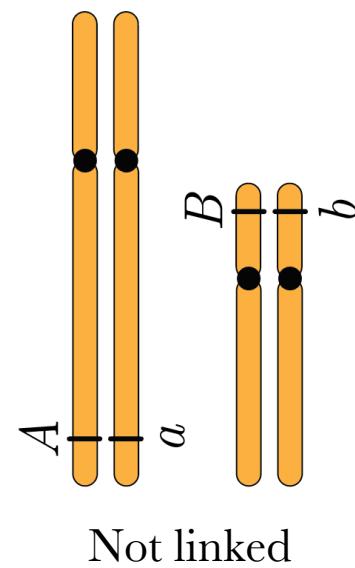
- Genetic linkage is the phenomenon where markers are likely to be inherited together.
- The closer the markers are, the lower the probability of crossing over events occur between them; consequently, the more likely they will be co-inherited.



Linked



Not linked



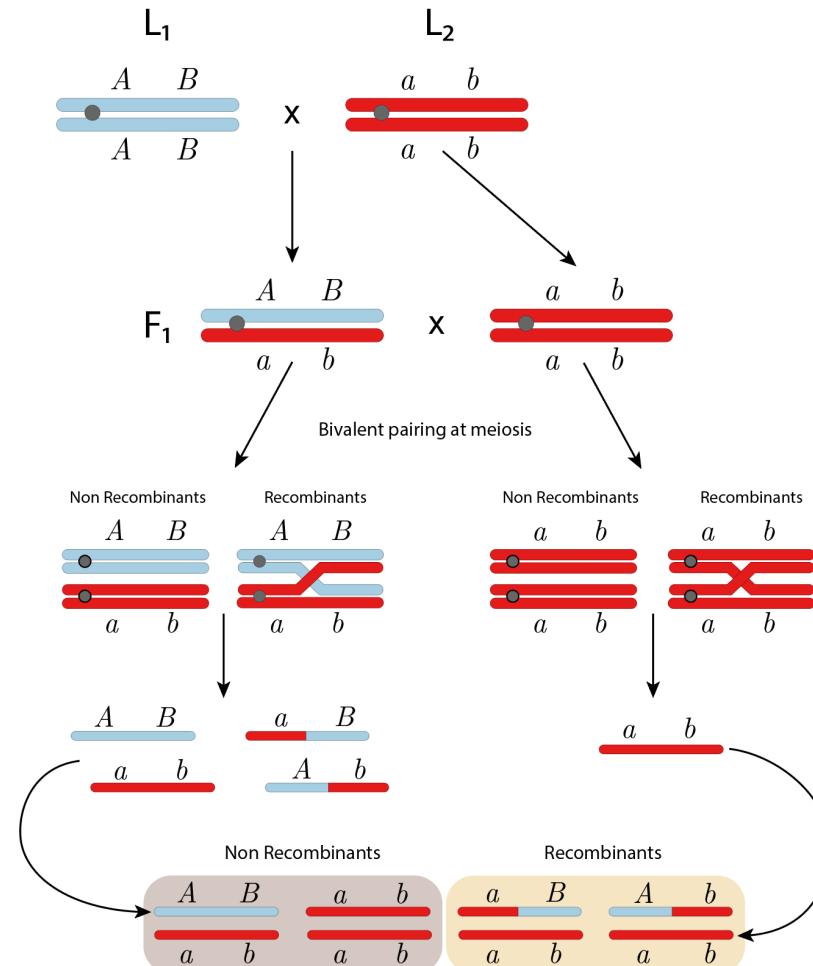
Not linked

- How can we measure how likely A and B are co-inherited?

Linkage analysis

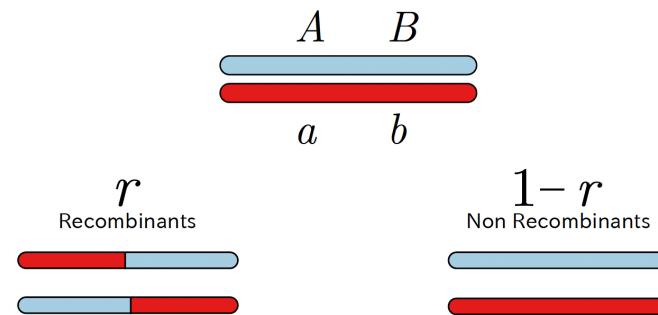
- Genetic linkage is a concept applied to *at least two loci*.
- We measure linkage using the *recombination frequency* (or fraction) in a segregation population.
- Recombination frequency is the *probability* that an odd number of crossovers occurs between the markers. Ranges from 0.0 to 0.5 (considering double reduction this number can be higher)
- We can transform these probabilistic values into distances using *mapping functions*. (Morgan, Haldane, Kosambi, etc.)
- By computing the recombination frequencies between pairs of markers and using mapping functions, we can construct *linkage maps* which show the linear order and relative distance between adjacent markers.
- First, let us address the *behavior o a single loci* when transited across generations.

Recombination fraction in diploids



$$\hat{r} = \frac{\#\text{recombinants}}{\#\text{total}}$$

Recombination fraction in diploids - Likelihood



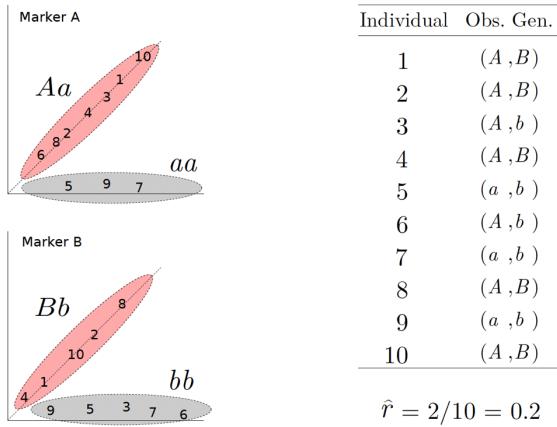
$$\mathbf{V} = \begin{bmatrix} \Pr(A, B) & \Pr(a, B) \\ \Pr(A, b) & \Pr(a, b) \end{bmatrix} = \begin{bmatrix} \frac{1-r}{2} & \frac{r}{2} \\ \frac{r}{2} & \frac{1-r}{2} \end{bmatrix}$$

$$L(r) = \prod_n \Pr(G_A, G_B \mid r)$$

where n is the number of individuals. The maximum likelihood estimator of r is

$$\hat{r} = \operatorname{argmax}_r L(r)$$

Recombination fraction in diploids

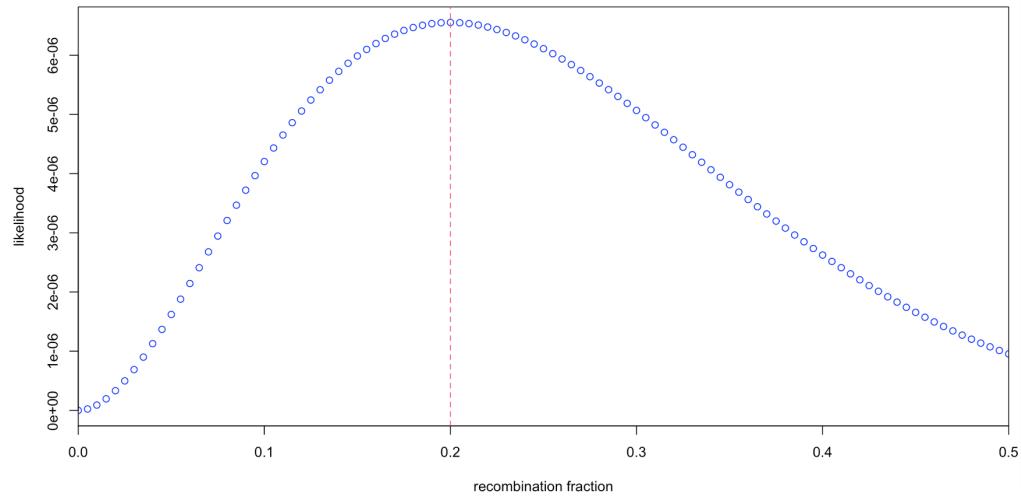


$$L = \prod_n \Pr(\text{loc}_B, \text{loc}_A \mid Data)$$

Individual	Obs. Gen.	$\Pr(\text{loc}_B, \text{loc}_A)$
1	(A ,B)	$\frac{1}{2}(1-r)$
2	(A ,B)	$\frac{1}{2}(1-r)$
3	(A ,b)	$\frac{1}{2} (r)$
4	(A ,B)	$\frac{1}{2}(1-r)$
5	(a ,b)	$\frac{1}{2}(1-r)$
6	(A ,b)	$\frac{1}{2} (r)$
7	(a ,b)	$\frac{1}{2}(1-r)$
8	(A ,B)	$\frac{1}{2}(1-r)$
9	(a ,b)	$\frac{1}{2}(1-r)$
10	(A ,B)	$\frac{1}{2}(1-r)$

$$L = \left(\frac{r}{2}\right)^2 \left(\frac{1-r}{2}\right)^8$$

Toy example



$$L = \prod_n \Pr(G_A, G_B) = \left(\frac{r}{2}\right)^2 \left(\frac{1-r}{2}\right)^8$$

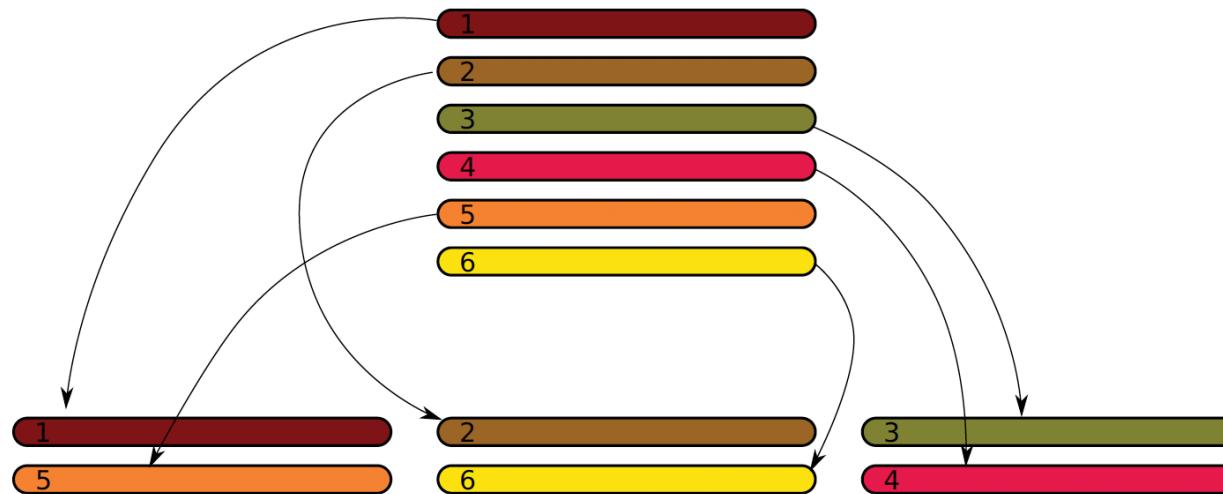
The MLE (maximum likelihood estimate) of r is $\hat{r} = 0.2$

Reference for MLE estimation in categorical data (Kalbfleisch, 1985)
<https://catalog.lib.ncsu.edu/catalog/NCSU652190>

Computing recombination frequencies in diploids using R and C++
https://github.com/mmollina/Cpp_and_R



Gamete formation in polyploids*



$$\psi_4$$

In this case: 15 possible configurations. For any ploidy level p

$$\frac{1}{\frac{p}{2}!} \prod_{i=1}^{\frac{p}{2}} \binom{2i}{2}$$

<https://doi.org/10.1534/g3.119.400378>

*no double reduction



Expected gametic frequency given a bivalent configuration

$$\begin{array}{ccccc}
 A_1 & B_1 & & A_2 & B_2 \\
 \text{---} & \text{---} & & \text{---} & \text{---} \\
 A_6 & B_6 & & A_5 & B_5 \\
 & & & & \\
 \left[\begin{array}{cc} \frac{1-r}{2} & \frac{r}{2} \\ \frac{r}{2} & \frac{1-r}{2} \end{array} \right] & \otimes & \left[\begin{array}{cc} \frac{1-r}{2} & \frac{r}{2} \\ \frac{r}{2} & \frac{1-r}{2} \end{array} \right] & \otimes & \left[\begin{array}{cc} \frac{1-r}{2} & \frac{r}{2} \\ \frac{r}{2} & \frac{1-r}{2} \end{array} \right]
 \end{array}$$

In general:

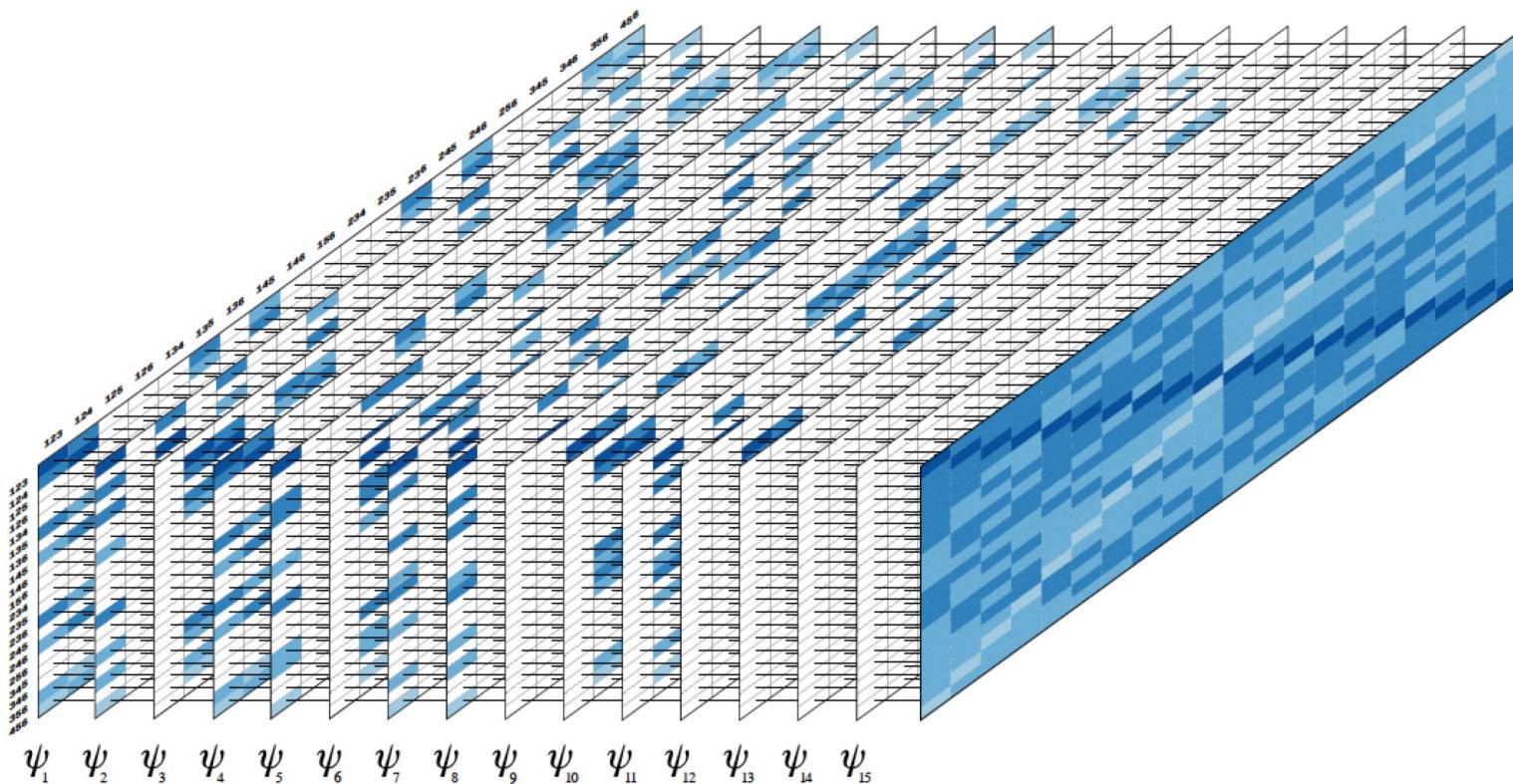
$$\mathbf{V}_1 \otimes \mathbf{V}_2 \otimes \cdots \otimes \mathbf{V}_{\frac{p}{2}}$$

All elements of this product are of the form

$$\Pr(G_A, G_B \mid \psi_j, r) = \frac{(1-r)^{(\frac{p}{2}-l)} p^l}{2^{\frac{p}{2}}}$$

l : known number of recombinant bivalents between loci A and B

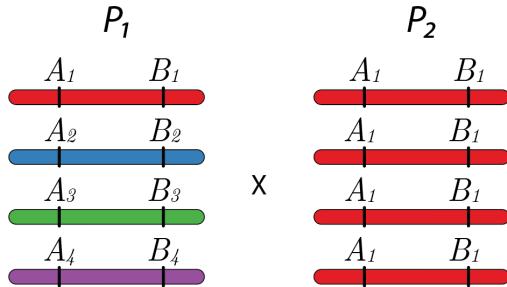
Unconditional gametic probability



$$\Pr(G_A, G_B \mid r) = \sum_j \Pr(G_A, G_B \mid \psi_j, r) \Pr(\psi_j)$$

Recombination Fraction – autotetraploid

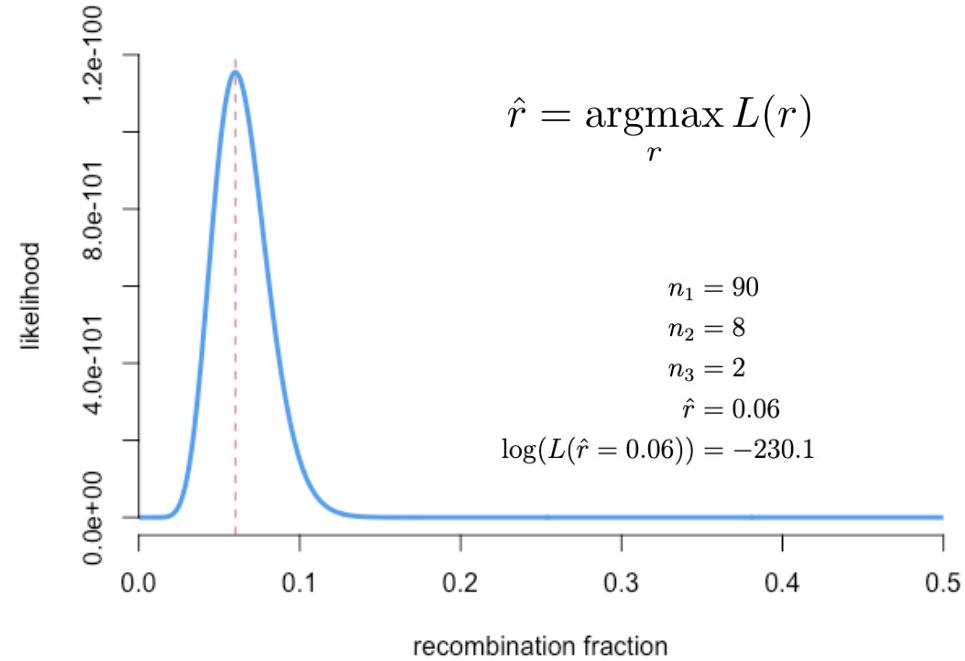
Fully informative marker



$$\Pr(G_A, G_B \mid r)$$

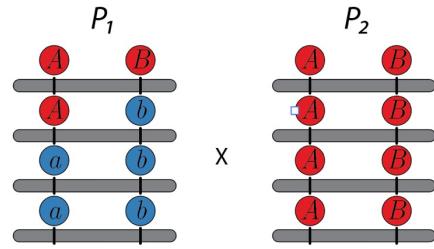
	B_1B_2	B_1B_3	B_1B_4	B_2B_3	B_2B_4	B_3B_4
A_1A_2	$\frac{(1-r)^2}{6}$	$\frac{(1-r)r}{12}$	$\frac{(1-r)r}{12}$	$\frac{(1-r)r}{12}$	$\frac{(1-r)r}{12}$	$\frac{r^2}{6}$
A_1A_3	$\frac{(1-r)r}{12}$	$\frac{(1-r)^2}{6}$	$\frac{(1-r)r}{12}$	$\frac{(1-r)r}{12}$	$\frac{r^2}{6}$	$\frac{(1-r)r}{12}$
A_1A_4	$\frac{(1-r)r}{12}$	$\frac{(1-r)r}{12}$	$\frac{(1-r)^2}{6}$	$\frac{r^2}{6}$	$\frac{(1-r)r}{12}$	$\frac{(1-r)r}{12}$
A_2A_3	$\frac{(1-r)r}{12}$	$\frac{(1-r)r}{12}$	$\frac{r^2}{6}$	$\frac{(1-r)^2}{6}$	$\frac{(1-r)r}{12}$	$\frac{(1-r)r}{12}$
A_2A_4	$\frac{(1-r)r}{12}$	$\frac{r^2}{6}$	$\frac{(1-r)r}{12}$	$\frac{(1-r)r}{12}$	$\frac{(1-r)^2}{6}$	$\frac{(1-r)r}{12}$
A_3A_4	$\frac{r^2}{6}$	$\frac{(1-r)r}{12}$	$\frac{(1-r)r}{12}$	$\frac{(1-r)r}{12}$	$\frac{(1-r)r}{12}$	$\frac{(1-r)^2}{6}$

$$L(r) = \left[\frac{(1-r)^2}{6} \right]^{n_1} \left[\frac{(1-r)r}{12} \right]^{n_2} \left[\frac{r^2}{6} \right]^{n_3}$$



Recombination Fraction – autotetraploid

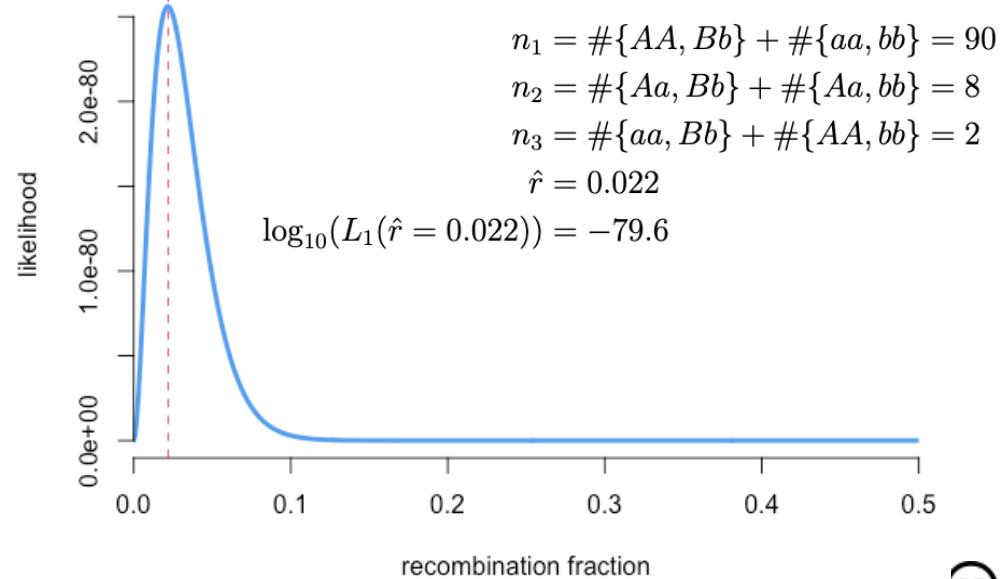
Partially informative marker – Duplex/simplex – Association



	Bb	Bb	Bb	bb	bb	bb
AA	$\frac{(1-r)^2}{6}$	$\frac{(1-r)r}{12}$	$\frac{(1-r)r}{12}$	$\frac{(1-r)r}{12}$	$\frac{(1-r)r}{12}$	$\frac{r^2}{6}$
Aa	$\frac{(1-r)r}{12}$	$\frac{(1-r)^2}{6}$	$\frac{(1-r)r}{12}$	$\frac{(1-r)r}{12}$	$\frac{r^2}{6}$	$\frac{(1-r)r}{12}$
Aa	$\frac{(1-r)r}{12}$	$\frac{(1-r)r}{12}$	$\frac{r^2}{6}$	$\frac{(1-r)^2}{6}$	$\frac{(1-r)r}{12}$	$\frac{(1-r)r}{12}$
Aa	$\frac{(1-r)r}{12}$	$\frac{r^2}{6}$	$\frac{(1-r)r}{12}$	$\frac{(1-r)r}{12}$	$\frac{(1-r)^2}{6}$	$\frac{(1-r)r}{12}$
aa	$\frac{r^2}{6}$	$\frac{(1-r)r}{12}$	$\frac{(1-r)r}{12}$	$\frac{(1-r)r}{12}$	$\frac{(1-r)r}{12}$	$\frac{(1-r)^2}{6}$

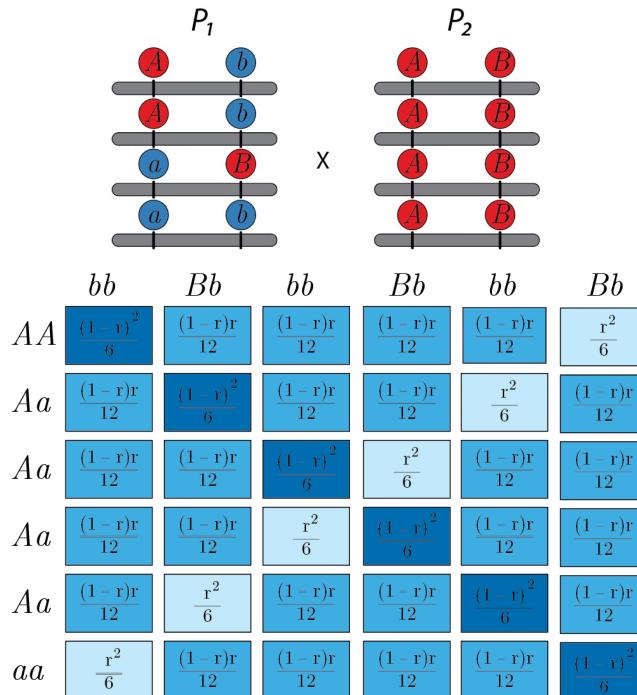
	Bb	bb
AA	$\frac{(1-r)}{6}$	$\frac{r}{6}$
Aa	$\frac{1}{3}$	$\frac{1}{3}$
aa	$\frac{r}{6}$	$\frac{(1-r)}{6}$

$$L_1(r) = \left[\frac{(1-r)}{6} \right]^{n_1} \left[\frac{1}{3} \right]^{n_2} \left[\frac{r}{6} \right]^{n_3}$$



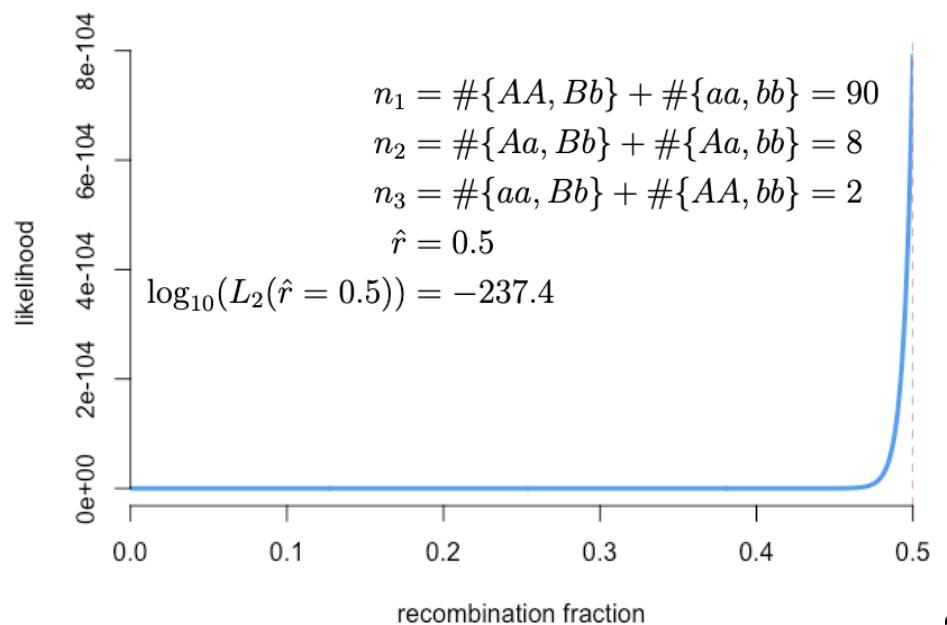
Recombination Fraction – autotetraploid

Partially informative marker – Duplex/simplex – Repulsion

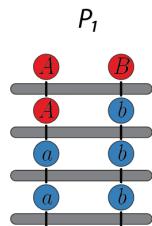


	<i>Bb</i>	<i>bb</i>
<i>AA</i>	$\frac{r}{6}$	$\frac{(1-r)}{6}$
<i>Aa</i>	$\frac{1}{3}$	$\frac{1}{3}$
<i>a a</i>	$\frac{(1-r)}{6}$	$\frac{r}{6}$

$$L_2(r) = \left[\frac{r}{6} \right]^{n_1} \left[\frac{1}{3} \right]^{n_2} \left[\frac{(1-r)}{6} \right]^{n_3}$$

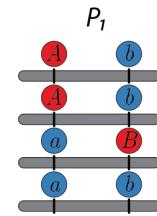


Recombination Fraction – assessing linkage phases



$$\hat{r}_1 = \arg \max_r L_1(r) \implies L_1(0.022) = -79.6$$

Compare likelihoods
choosing the **most likely**
configuration



$$\hat{r}_2 = \arg \max_r L_2(r) \implies L_2(0.5) = -237.4$$

logarithm-of-odds (LOD-score)

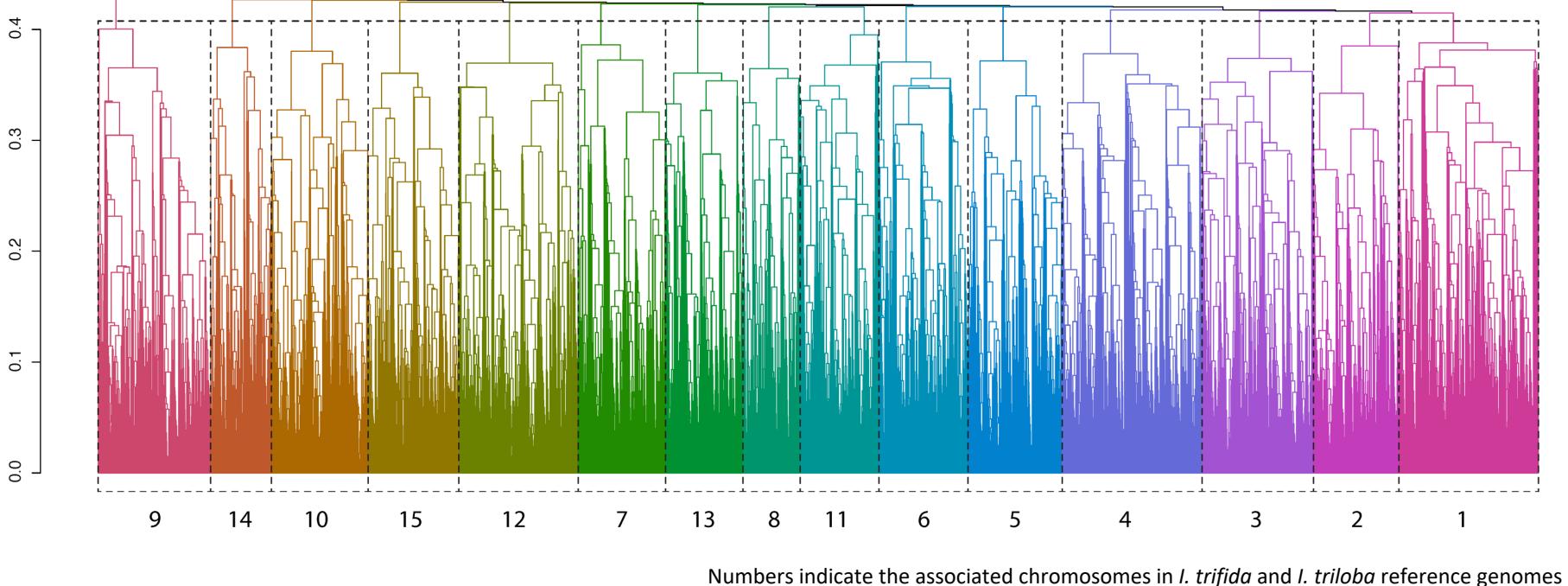
$$\begin{aligned} LOD &= \log_{10} \left(\frac{L_1}{L_2} \right) \\ &= \log_{10}(L_1) - \log_{10}(L_2) \\ &= -79.6 - (-237.4) = 157.8 \end{aligned}$$

Markers A and B are most
likely in **association**

- Pairwise MLEs of r are used to group markers into linkage groups and order markers within each linkage group using optimization algorithms such as MDS.
- Given a sequence of ordered markers, it is possible to extend the idea of comparing the likelihoods of competing linkage phases throughout multiple markers.

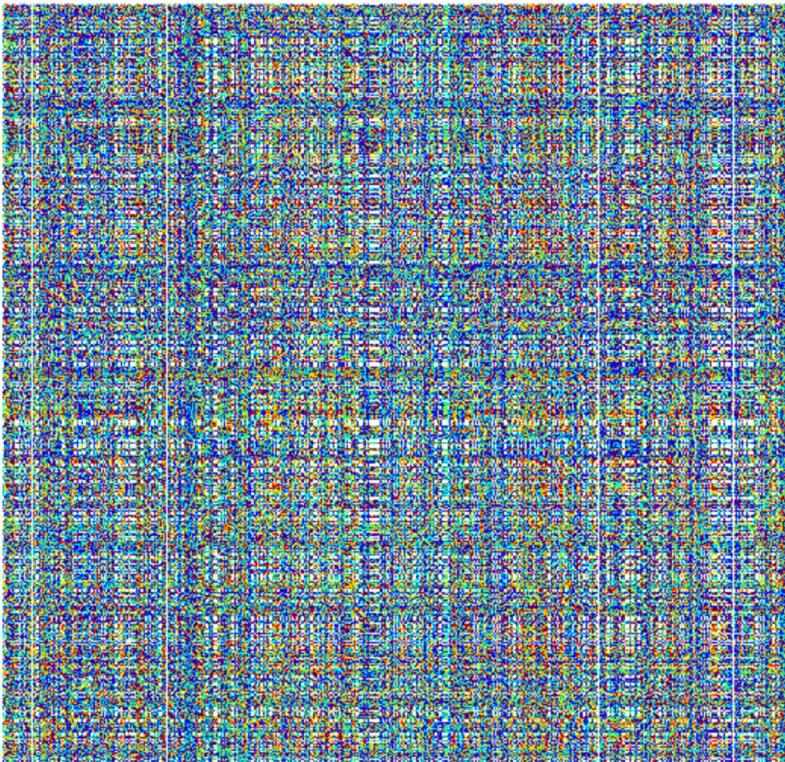
Two – point analysis and grouping

- Number of markers: 38,701
- Number of recombination fractions: ~749 million pairs

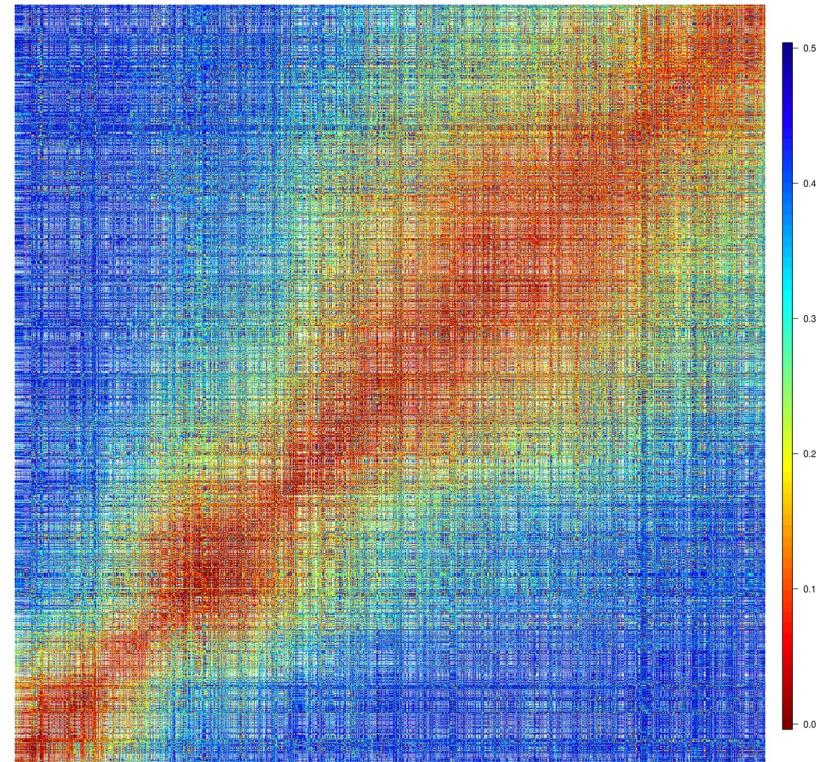
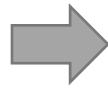


Multidimensional Scaling Algorithm (MDS)

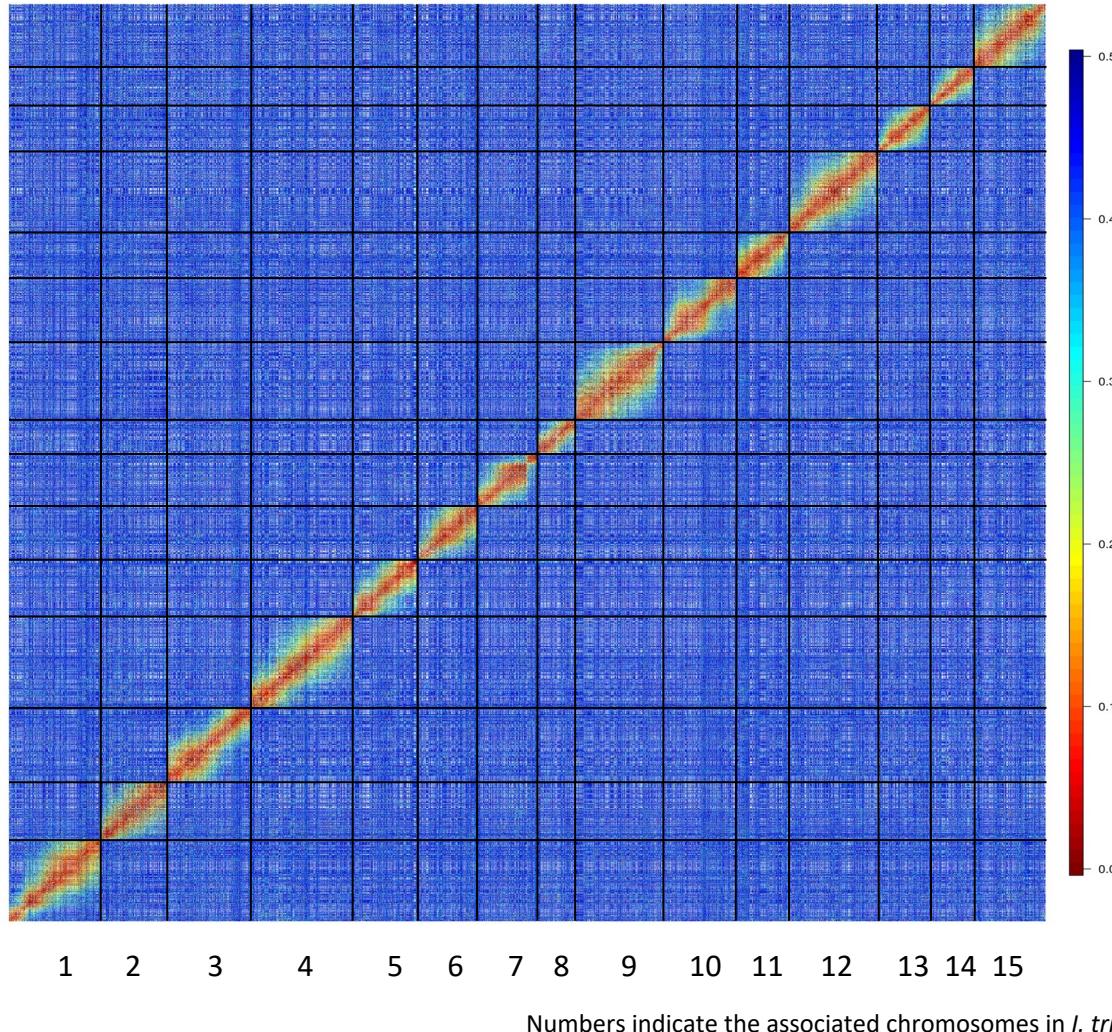
- Reduce data from many dimensions preserving the observed distances between points by minimizing a loss function L .



Sweetpotato linkage group 1: 2745 markers

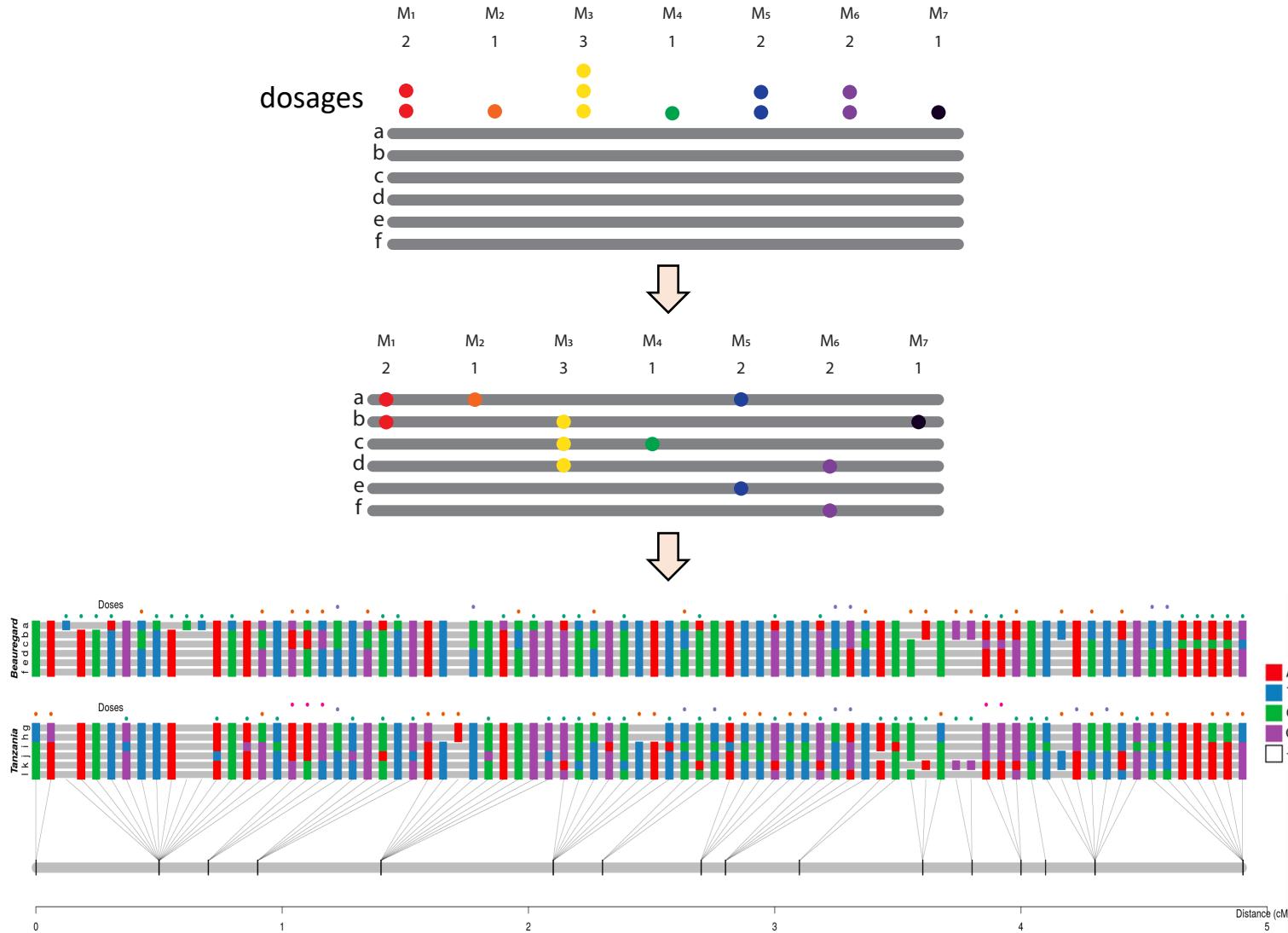


Ordering with MDS – 15 linkage groups, 30684 SNPs



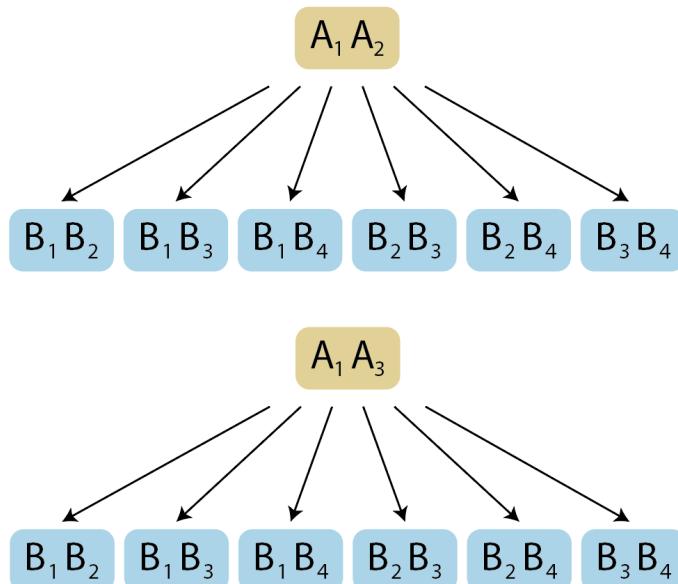
Haplotyping in polyploids

- Placement of allelic variants in the homologs in a homology group

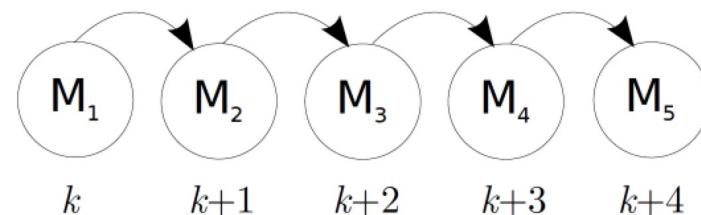


Multilocus linkage analysis in polyploids

What is the probability of observing a specific **state** at a moment (or position), **given** we observed some state in a previous moment?



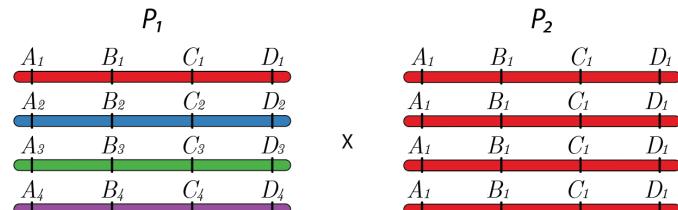
Markov Model: Conditional independence



$$\Pr(G_{k+1}|G_k) = \frac{(1 - r_k)^{\frac{p}{2} - l} (r_k)^l}{\binom{\frac{p}{2}}{l}}$$

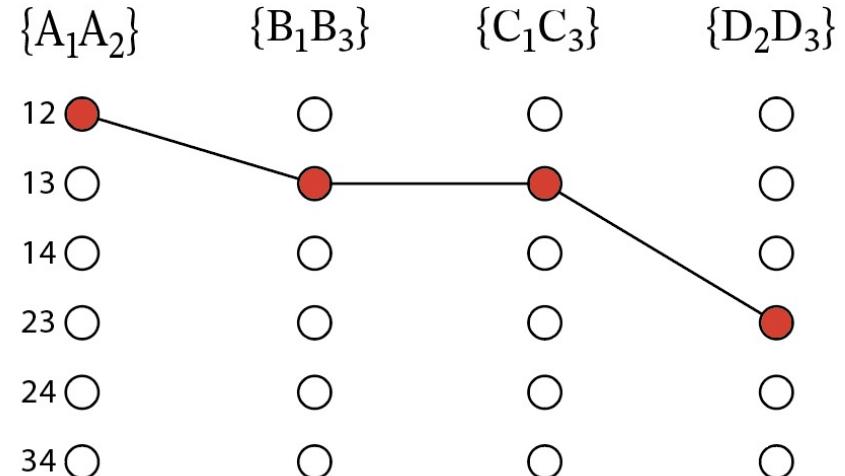
where r_k is the recombination frequency between loci k and $k+1$, p is the ploidy level and l is the number of recombinant events between k and $k+1$.

Markov model



Individual 3

	M ₁	M ₂	M ₃	M ₄
Ind ₁	{A ₁ , A ₂ }	{B ₁ , B ₃ }	{C ₃ , C ₄ }	{D ₃ , D ₄ }
Ind ₂	{A ₁ , A ₂ }	{B ₁ , B ₃ }	{C ₁ , C ₃ }	{D ₃ , D ₄ }
Ind ₃	{A ₁ , A ₂ }	{B ₁ , B ₃ }	{C ₁ , C ₃ }	{D ₂ , D ₃ }
⋮	⋮	⋮	⋮	⋮
Ind _n	{A ₂ , A ₃ }	{B ₂ , B ₃ }	{C ₂ , C ₃ }	{D ₂ , D ₃ }

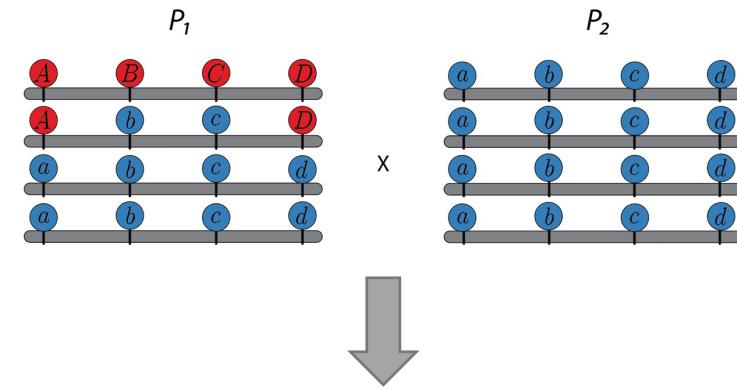


$$L(\mathbf{r}) = \prod_n \Pr(G_A, G_B, G_C, G_D \mid \mathbf{r})$$

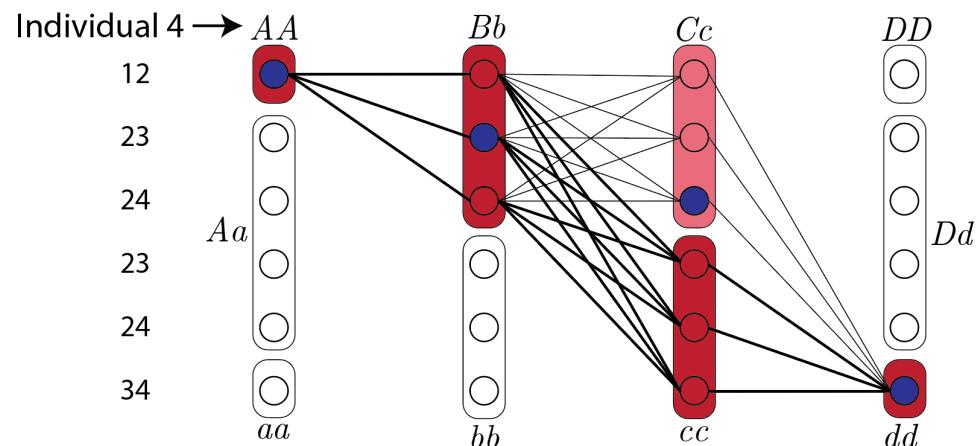
$$\hat{\mathbf{r}} = \underset{\mathbf{r}}{\operatorname{argmax}} L(\mathbf{r})$$

Hidden Markov Model - HMM

$$b_j(O) = \Pr\left(O \middle| \mathcal{G}_{k,j}^m, \varphi_P^k, \varphi_Q^k\right) = \begin{cases} 1 - \epsilon & \text{if } O = \delta(k, j) \\ \frac{\epsilon}{m} & \text{otherwise} \end{cases}$$



	M ₁	M ₂	M ₃	M ₄
Ind ₁	AA	Bb	cc	dd
Ind ₂	AA	Bb	Cc	dd
Ind ₃	AA	Bb	Cc	Dd
Ind ₄	AA	Bb	Cc/cc 0.2 0.8	Dd
⋮	⋮	⋮		
Ind _n	Aa	Bb	Cc	Dd

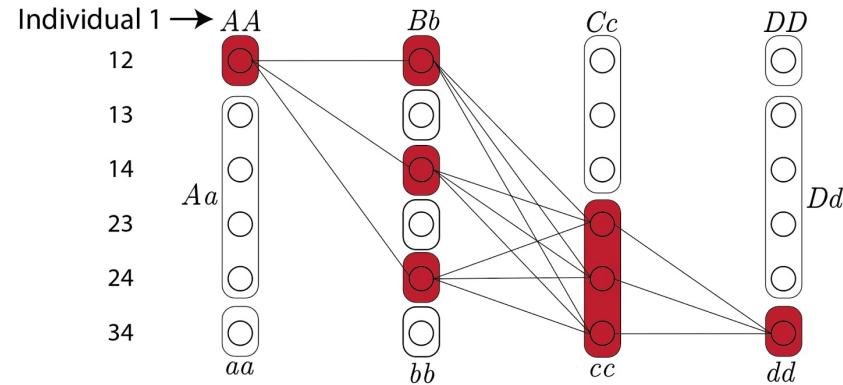
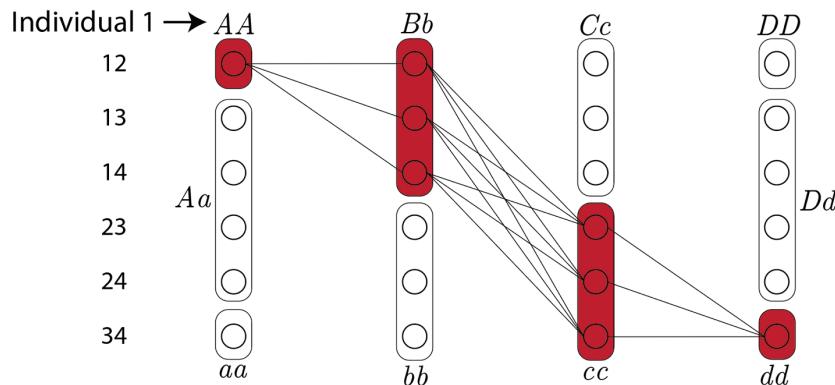
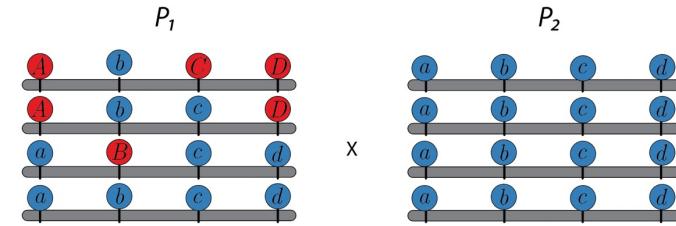
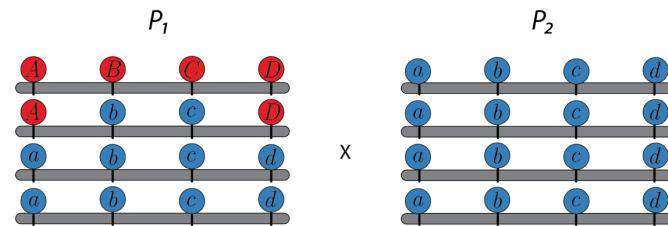


$$L(\mathbf{r}) = \prod_n \Pr(G_A, G_B, G_C, G_D \mid \mathbf{r})$$

$$\hat{\mathbf{r}} = \operatorname{argmax}_{\mathbf{r}} L(\mathbf{r})$$

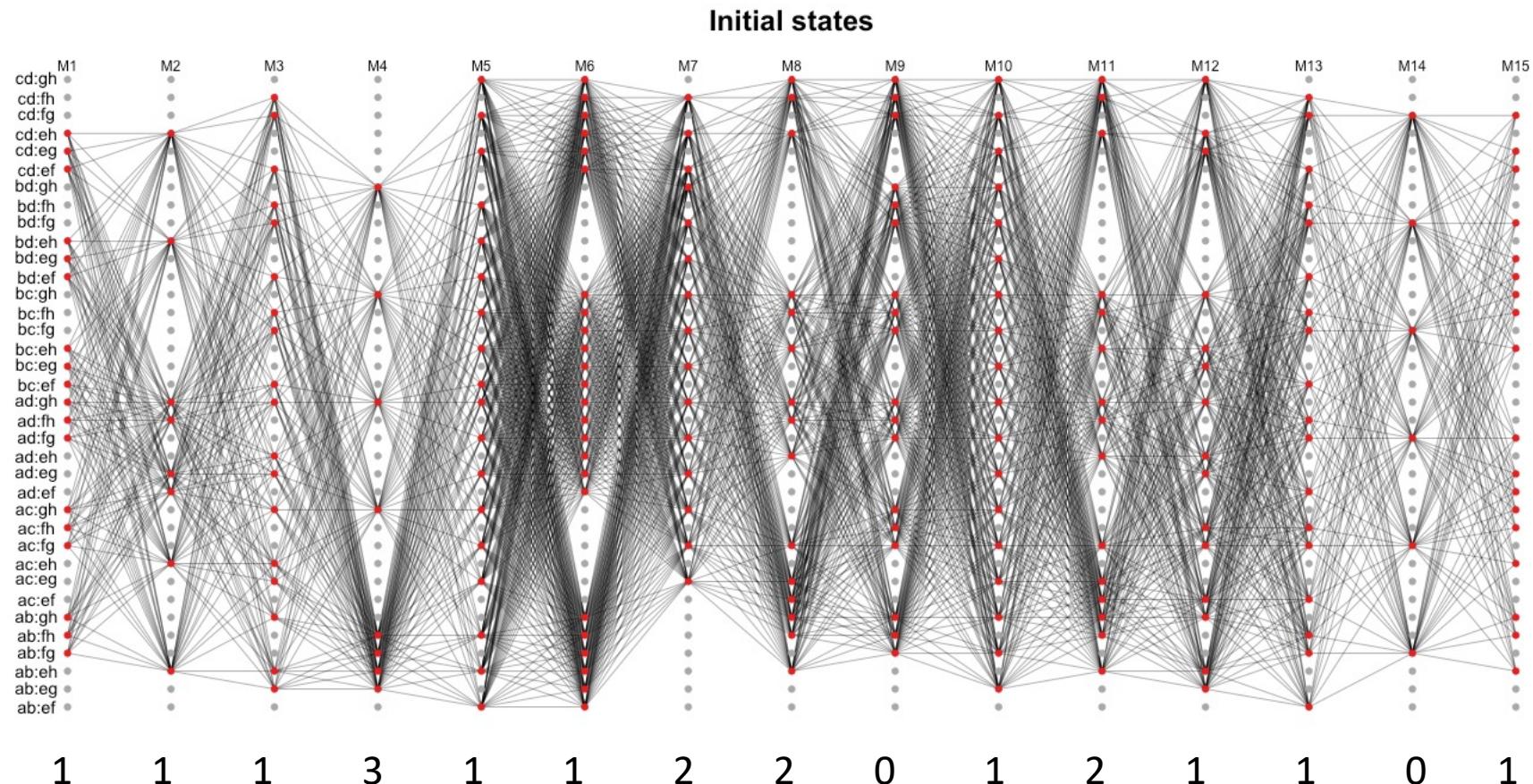
Hidden Markov Model - HMM

Assessing *different linkage phases* using multilocus analysis

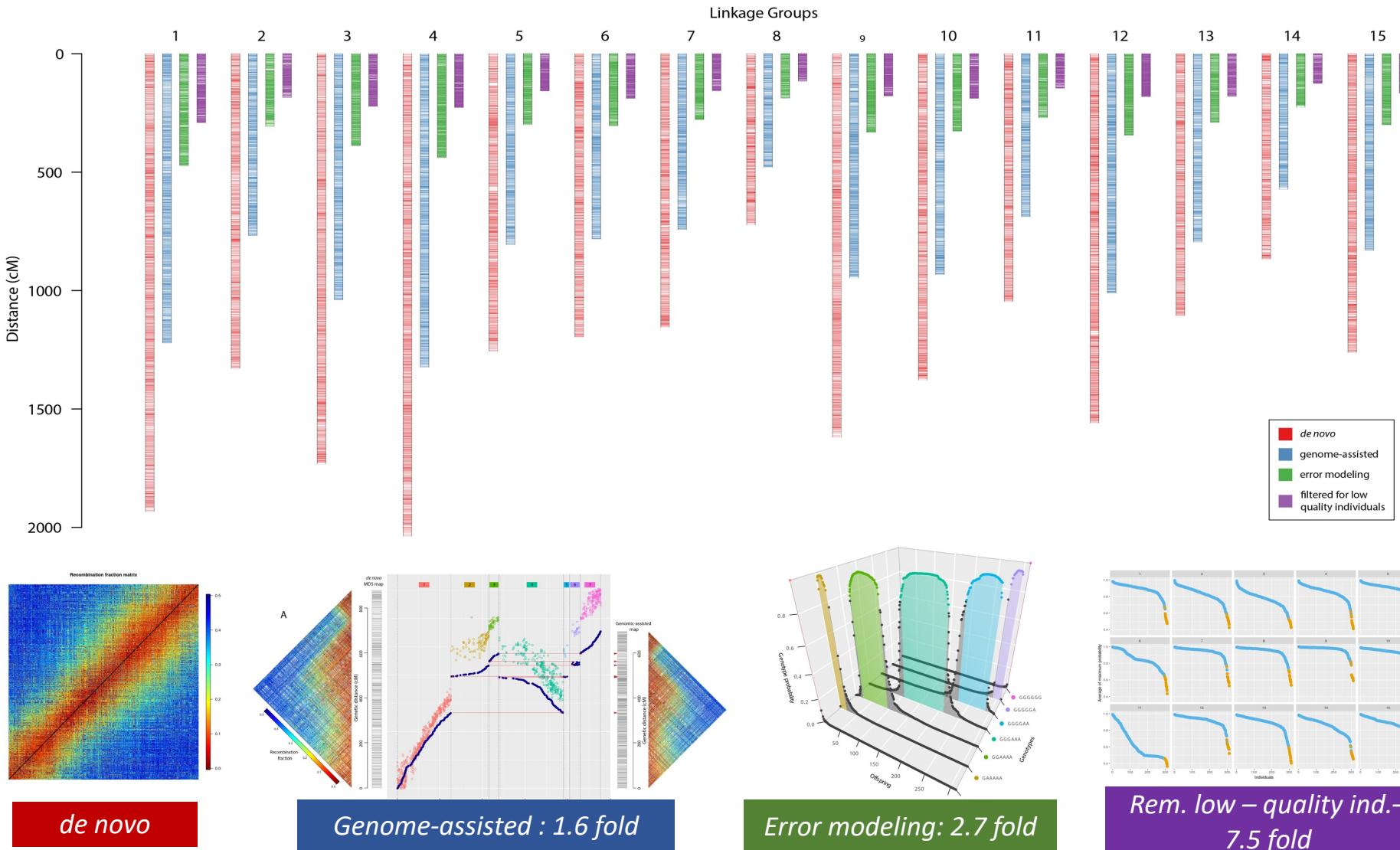


Hidden Markov Model - HMM

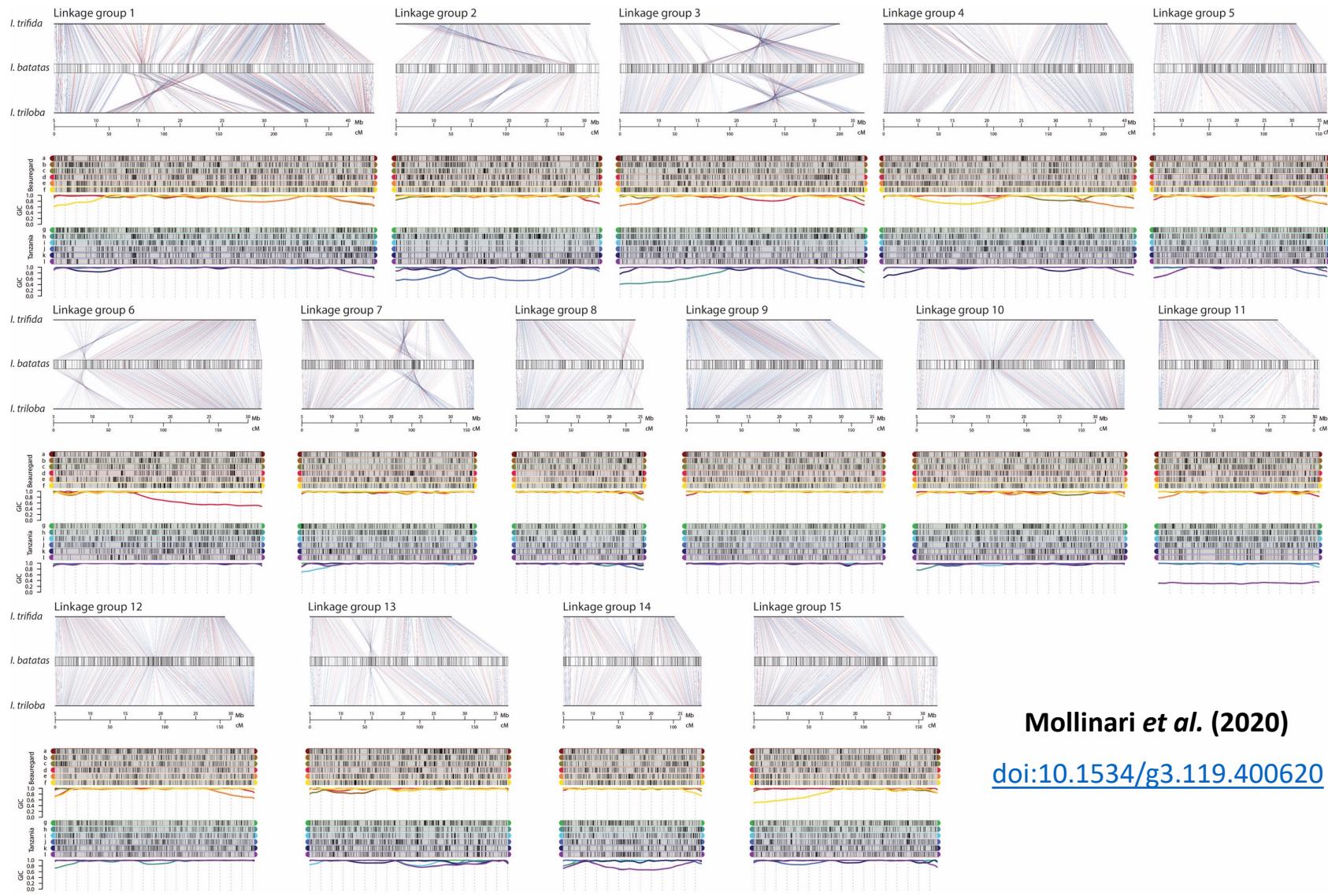
- Tetraploid example, one individual, 15 markers



Strategies to improve map quality



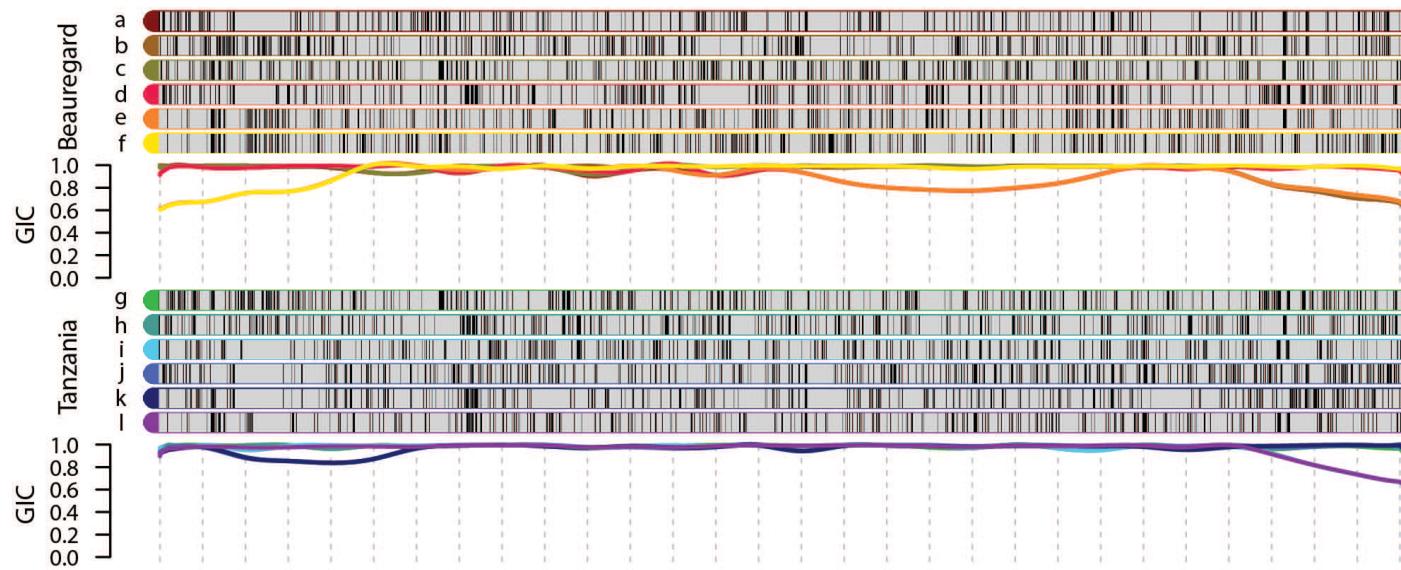
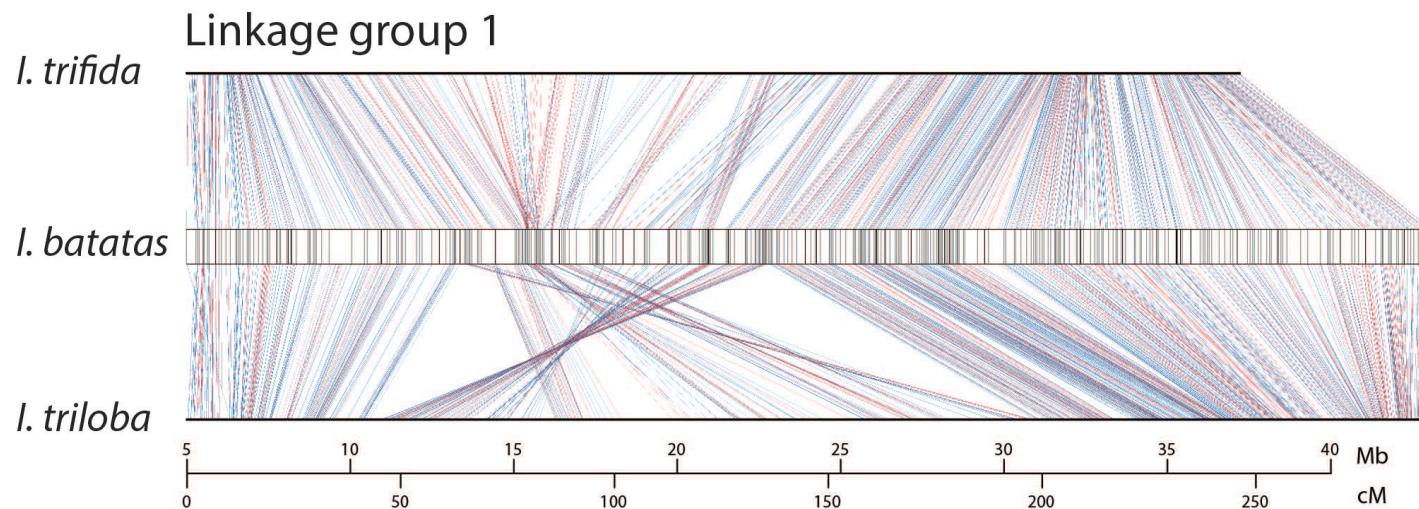
Sweetpotato genetic map



Molinari et al. (2020)

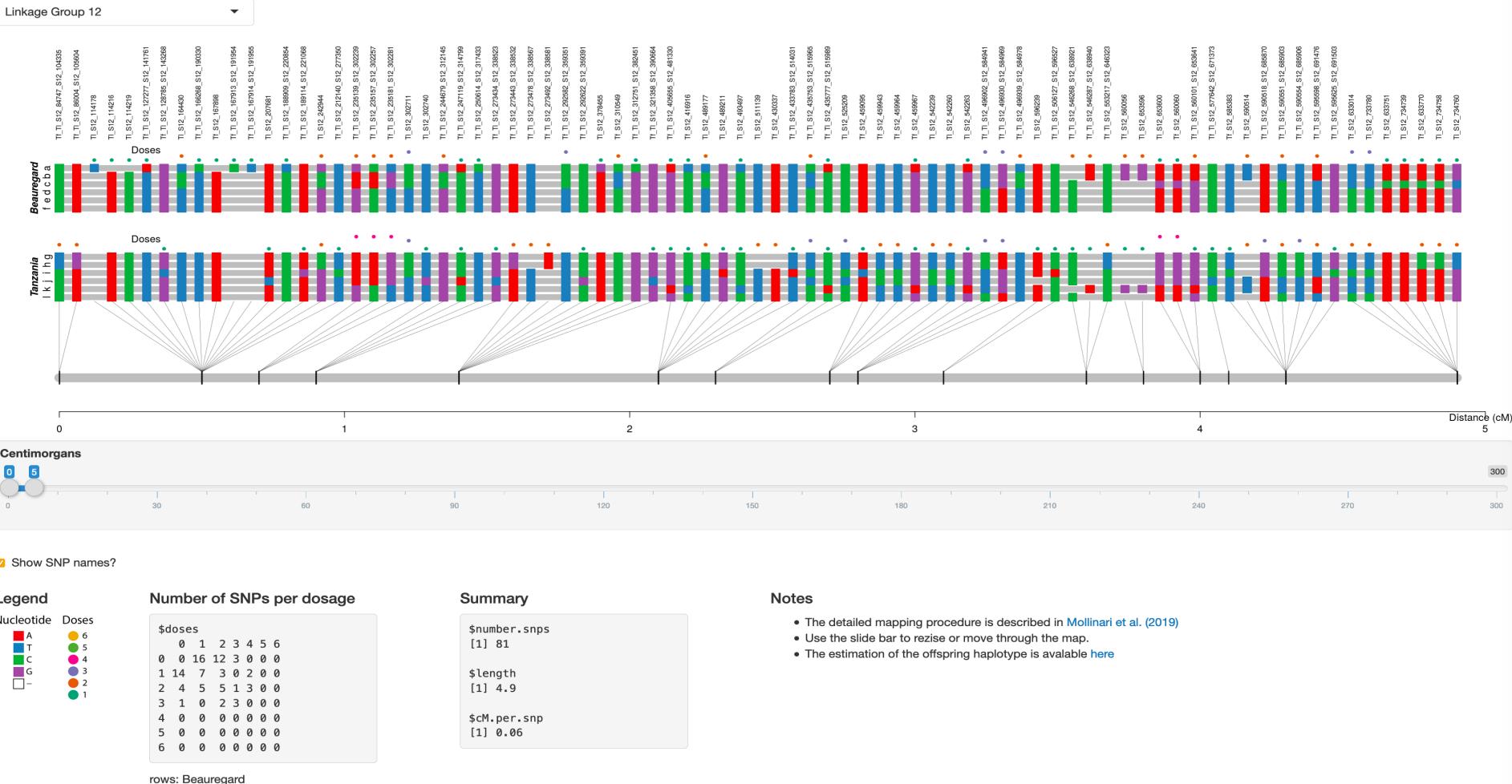
[doi:10.1534/g3.119.400620](https://doi.org/10.1534/g3.119.400620)

Sweetpotato genetic map



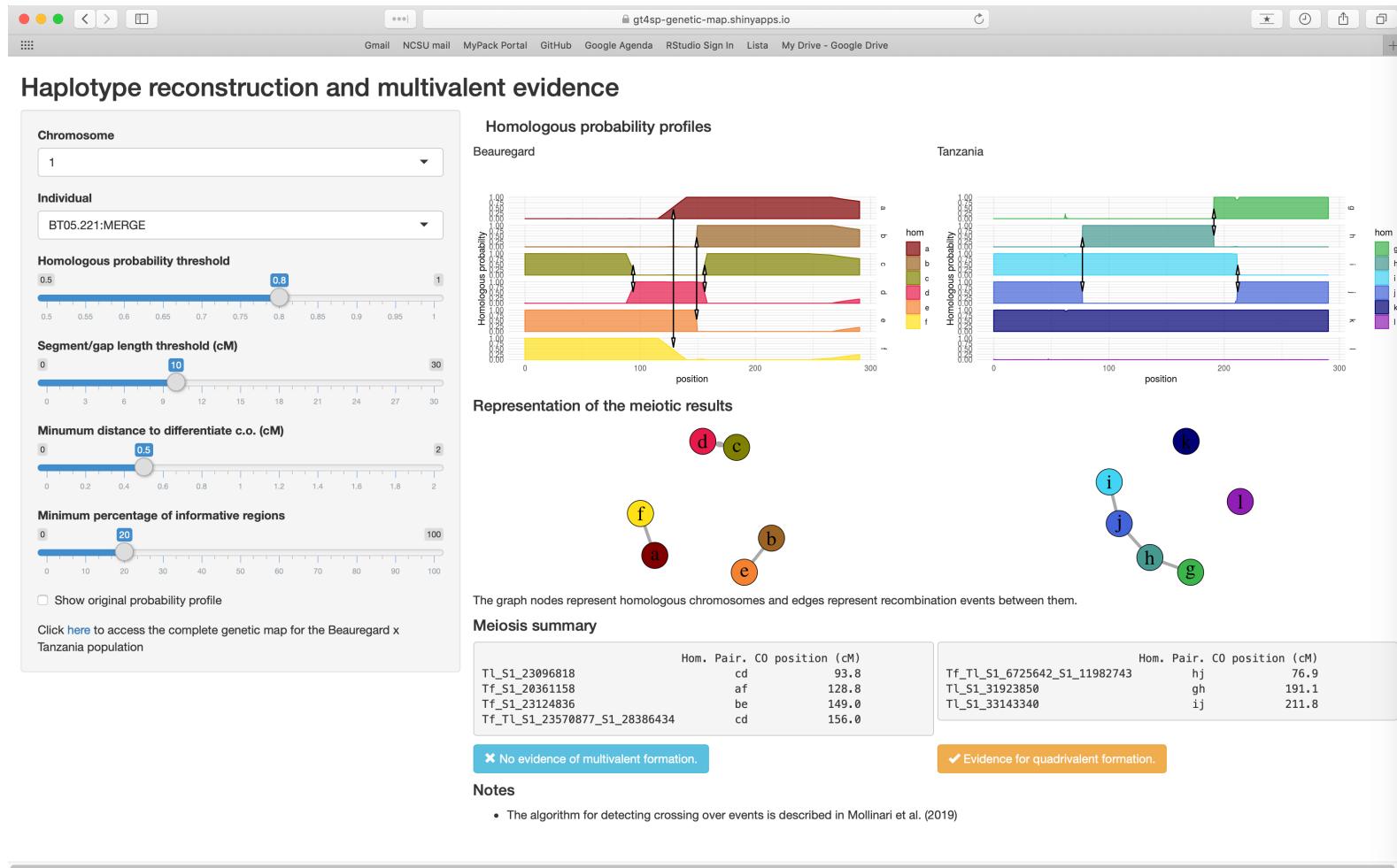
Genetic mapping – Linkage group 12 – 2661 SNPs

Sweetpotato genetic map - Beauregard x Tanzania (BT)

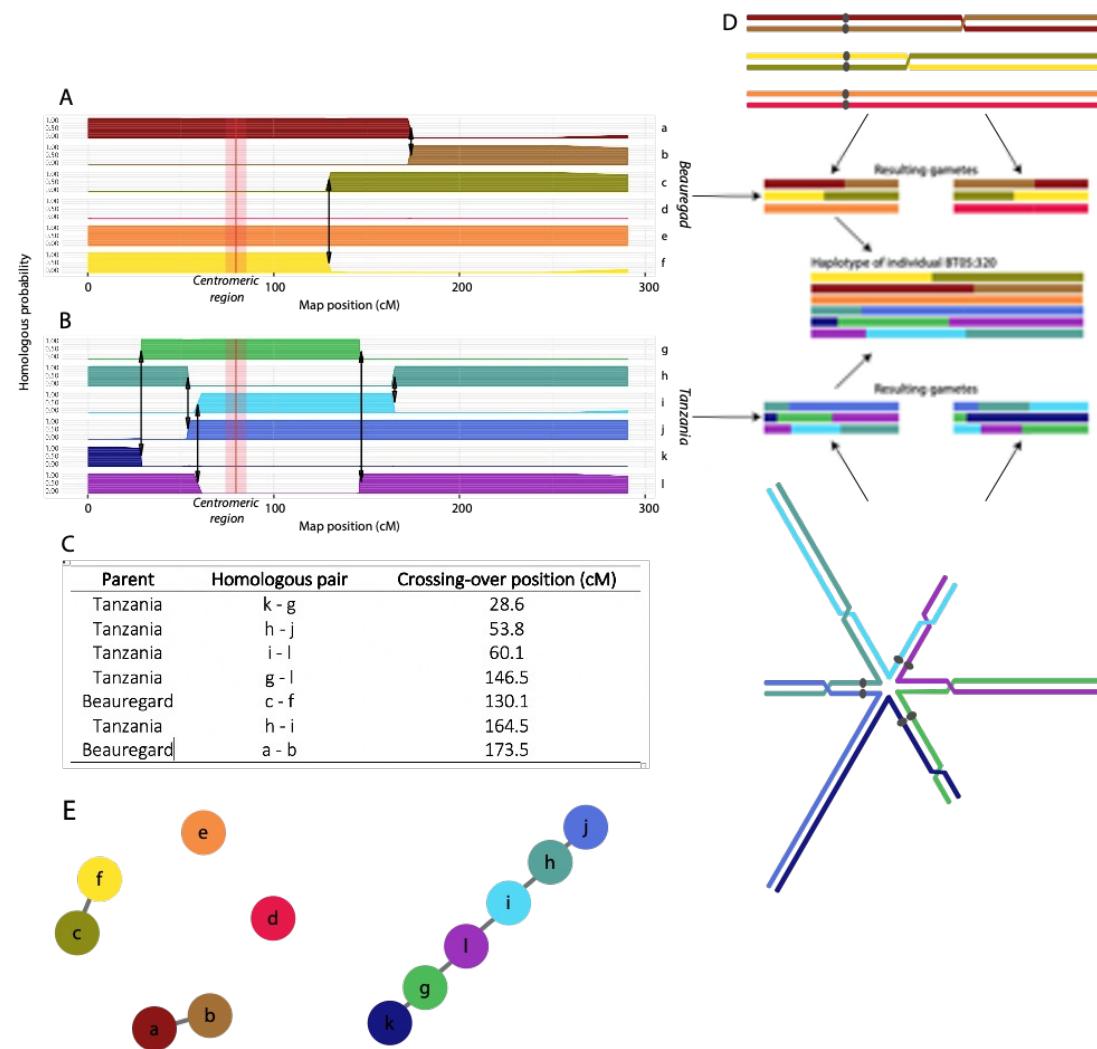


Interactive version: https://gt4sp-genetic-map.shinyapps.io/bt_map/

Haplotype reconstruction in the offspring



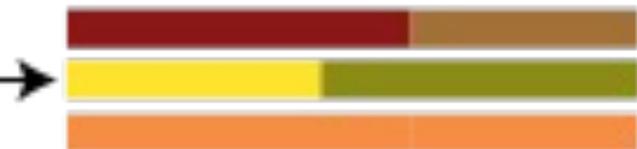
Haplotype reconstruction in the offspring



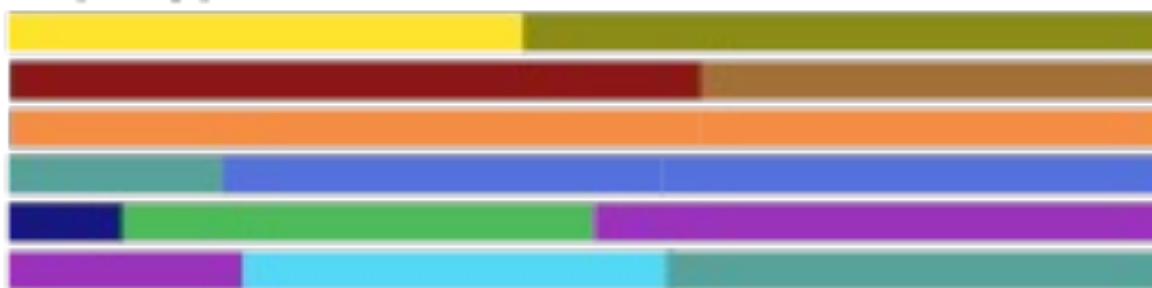
https://gt4sp-genetic-map.shinyapps.io/offspring_haplotype_BT_population/



Resulting gametes



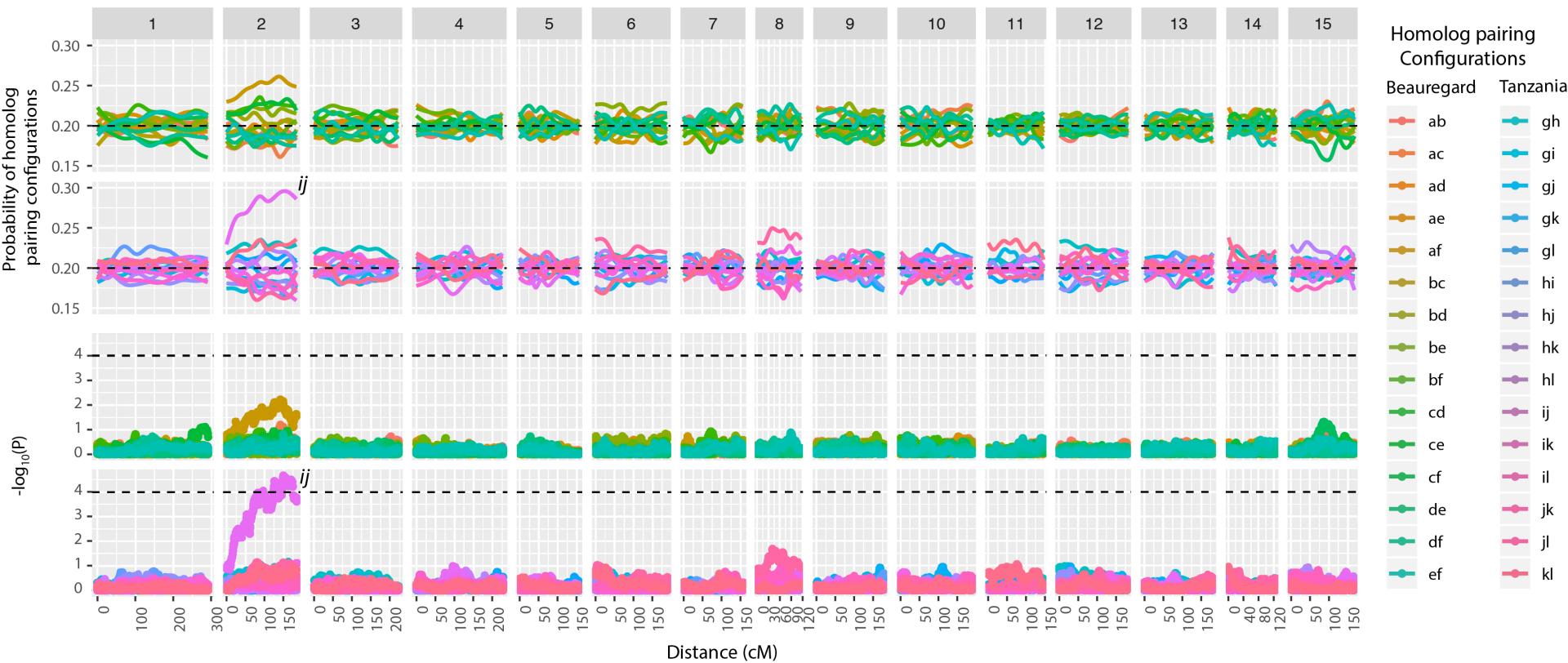
Haplotype of individual BT05:320



Resulting gametes



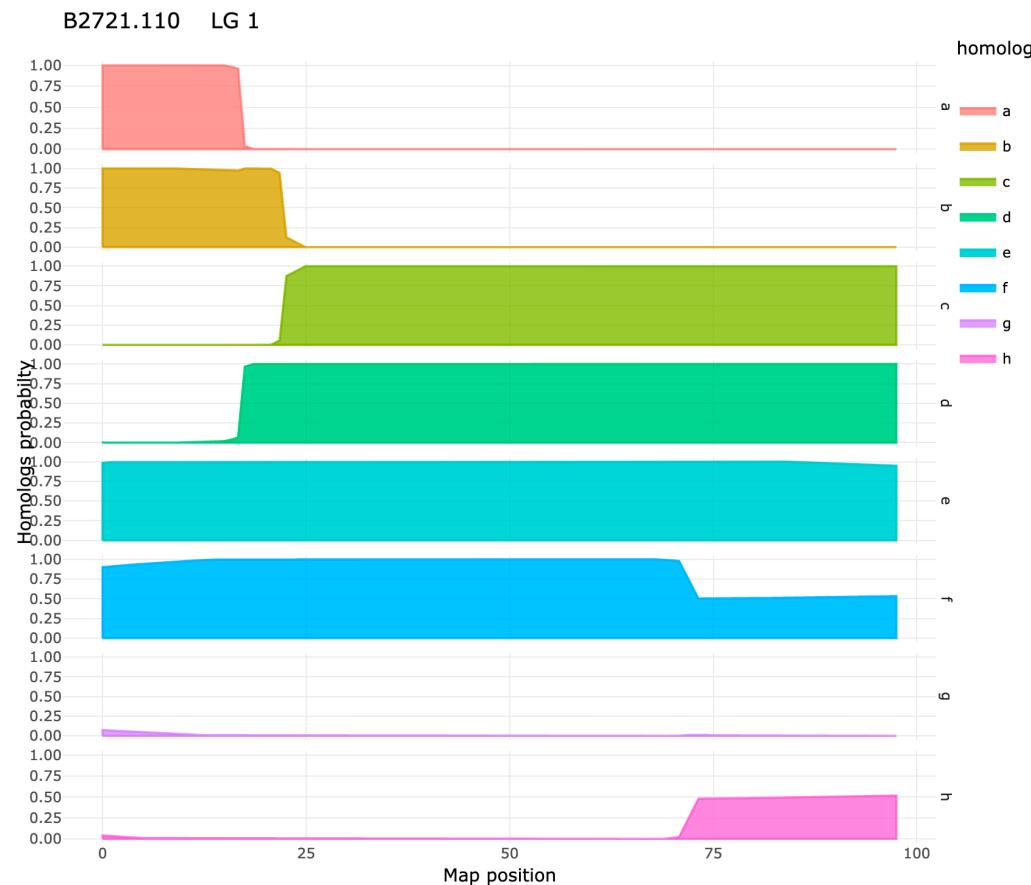
Preferential pairing profiles: Sweetpotato is vastly **auto**hexaploid



Probabilistic haplotype reconstruction

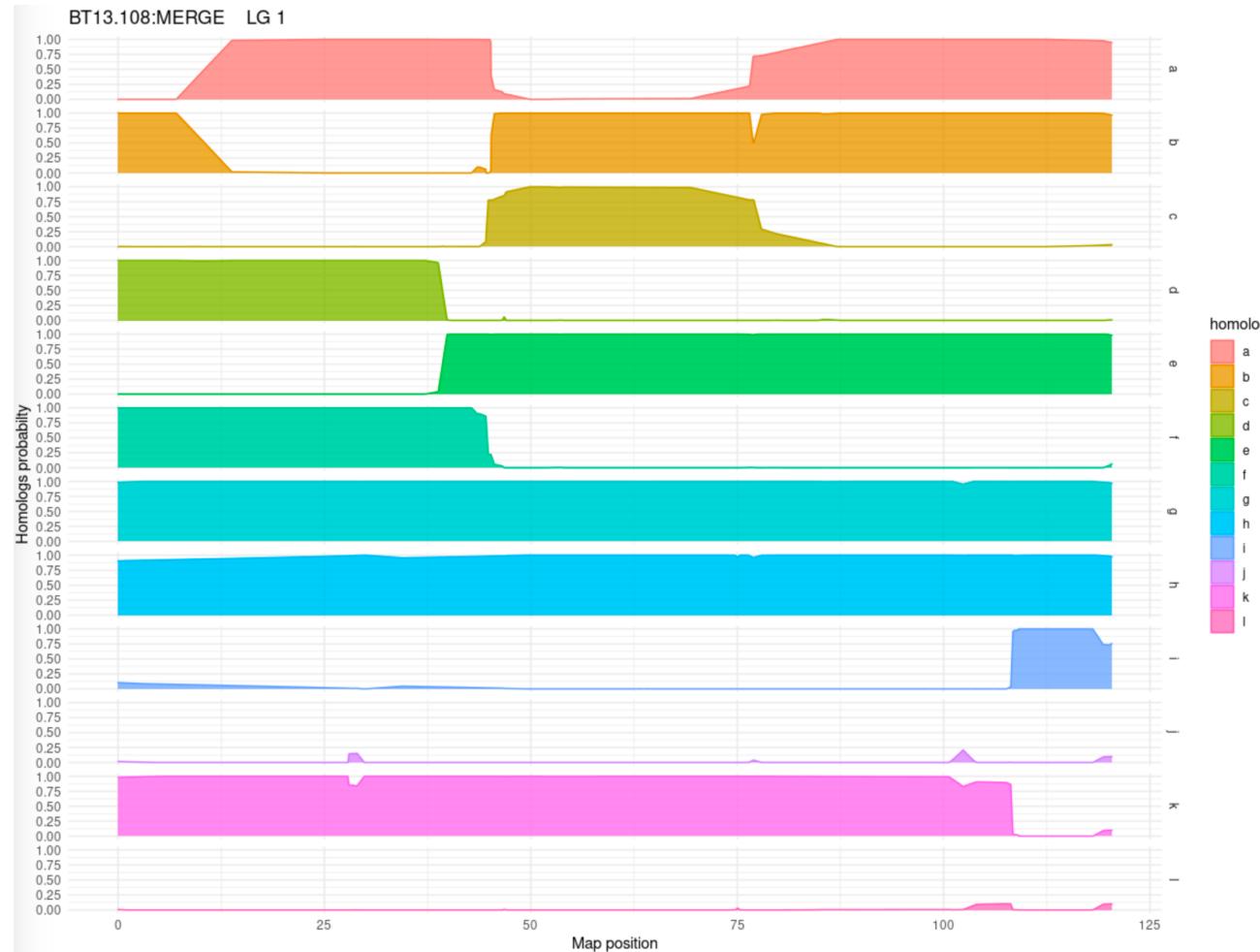
- When assuming a *prior probability* distribution of the genotypes, multilocus strategies can improve the quality of the inferred haplotypes

Tetraploid potato



Probabilistic haplotype reconstruction

Hexaploid sweetpotato

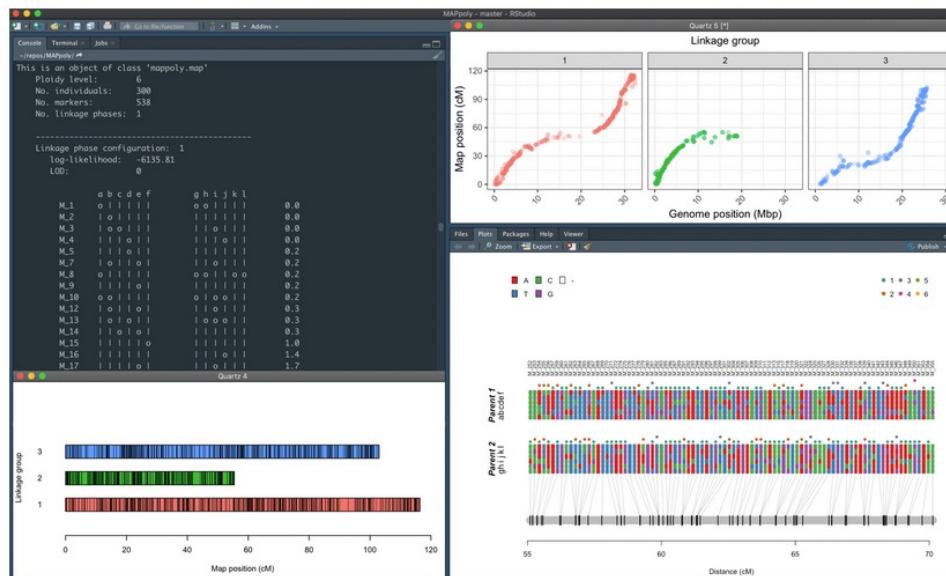
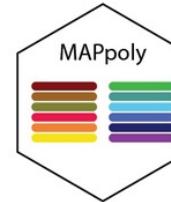


MAPpoly demonstration

R-CMD-check passing build passing development active License GPL v3 codecov 71% CRAN 0.3.1 r-universe 0.3.1
downloads 9623

MAPpoly

MAPpoly (v. 0.3.1) is an R package to construct genetic maps in autopolyploids with even ploidy levels. In its current version, MAPpoly can handle ploidy levels up to 8 when using hidden Markov models (HMM), and up to 12 when using the two-point simplification. When dealing with large numbers of markers (> 10,000), we strongly recommend using high-performance computation.



Development version and other resources: <https://github.com/mmollina/MAPpoly>
Stable version: <https://cran.r-project.org/package=mappoly>

References

- Haldane, J. Theoretical Genetics of Autopolyploids. *J. Genet.* **22**, 359–372 (1930).
- Mather, K. Reductive and equational separation of the chromosomes in bivalents and multivalents. *J. Genet.* **30**, 53–78 (1935).
- Mather, K. Segregation and linkage in autotetraploids. *J. Genet.* **32**, 287–314 (1936).
- Mather, K. *The measurement of linkage in heredity*. (1938).
- Fisher, R. A. The theory of linkage in polysomic inheritance. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **233**, 55–87 (1947).
- Ripol, M. I., Churchill, G. A., Silva, J. A. G. Da & Sorrells, M. Statistical aspects of genetic mapping in autopolyploids. *Gene* **235**, 31–41 (1999).
- Luo, Z. W., Zhang, R. M. & Kearsey, M. J. Theoretical basis for genetic linkage analysis in autotetraploid species. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 7040–7045 (2004).
- Leach, L. J., Wang, L., Kearsey, M. J. & Luo, Z. Multilocus tetrasomic linkage analysis using hidden Markov chain model. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 4270–4274 (2010).
- Hackett, C. a., McLean, K. & Bryan, G. J. Linkage Analysis and QTL Mapping Using SNP Dosage Data in a Tetraploid Potato Mapping Population. *PLoS One* **8**, (2013).
- Zheng, C. *et al.* Probabilistic Multilocus Haplotype Reconstruction in Outcrossing Tetraploids. *Genetics* **203**, 119–131 (2016).
- Bourke, P. M. Genetic mapping in polyploids. (Wageningen University, 2018).
- Mollinari, M. & Garcia, A. A. F. Linkage Analysis and Haplotype Phasing in Experimental Autopolyploid Populations with High Ploidy Level Using Hidden Markov Models. *G3; Genes/Genomes/Genetics* **9**, 3297–3314 (2019).