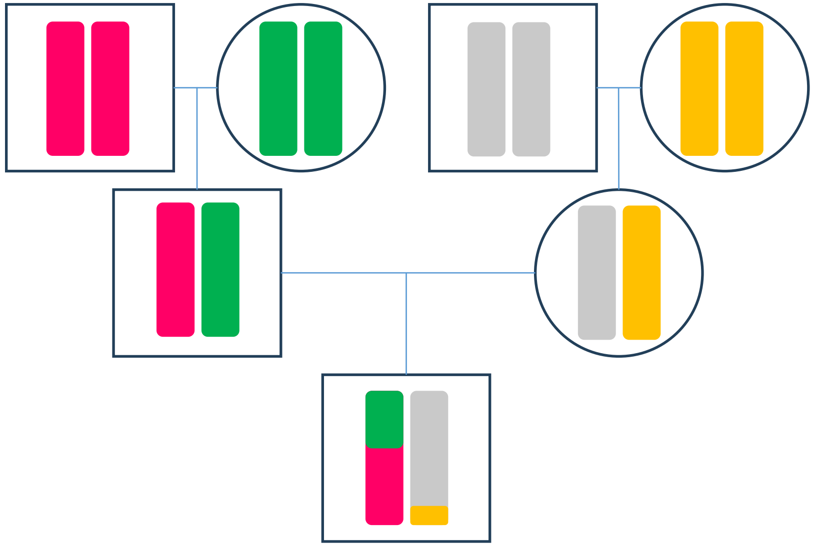
Agrobioinformatics
Giessen



HaploMAGIC

Phasing and accurately detecting recombination in multiparental populations with genotyping errors

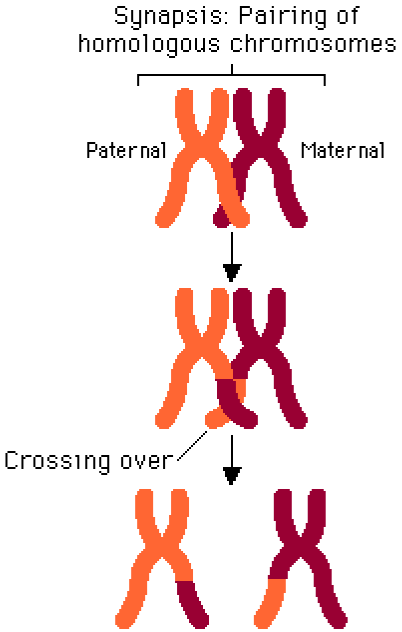
2 October, 2023

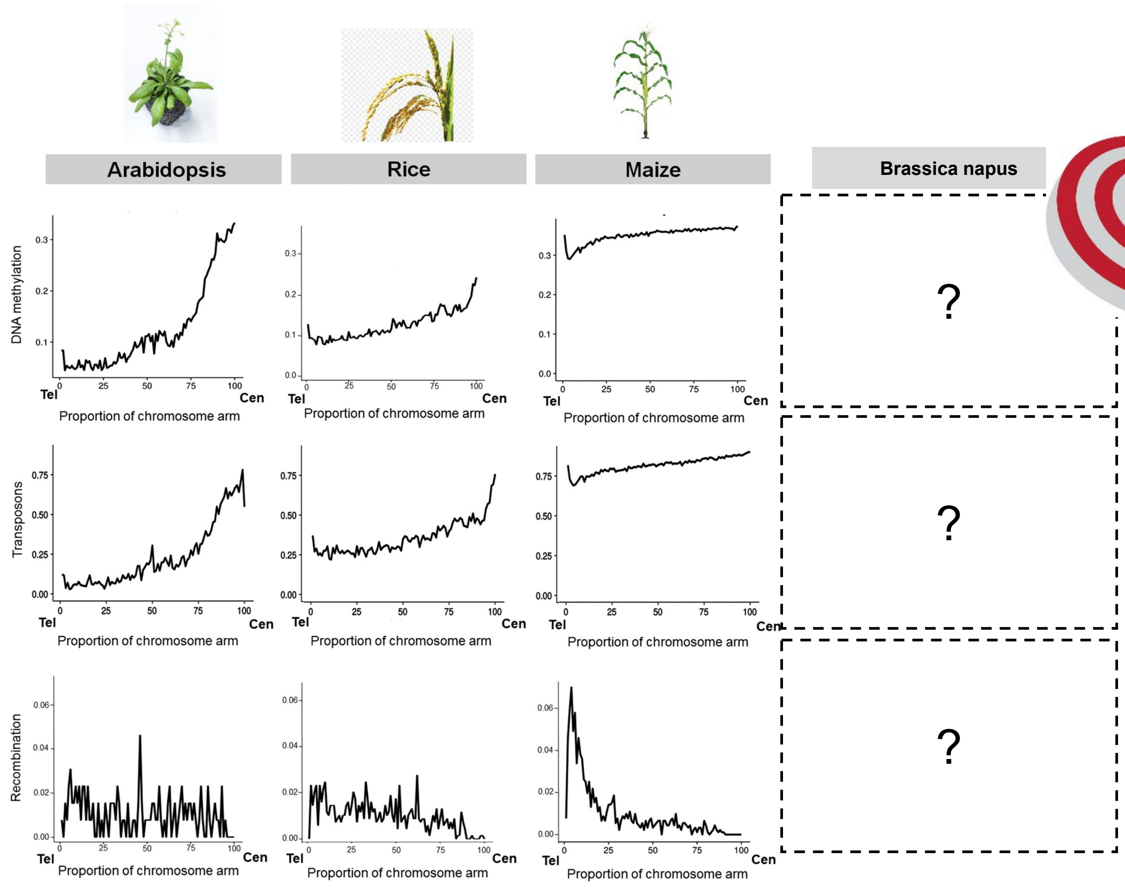
Jose Antonio Montero Tena, Justus Liebig University,
International Rapeseed Conference

Content

- Introduction
    - Meiotic recombination
    - Multiparental populations (MPPs) for detecting recombination
- Method
    - haploMAGIC. Requirements and output
    - haploMAGIC. Pipeline
- Results
    - haploMAGIC performance with simulations
    - Correlation haploMAGIC-LINKPHASE3
    - Applying haploMAGIC to real data
    - Genome-wide recombination map
- Take-home messages
- Future steps

Meiotic recombination

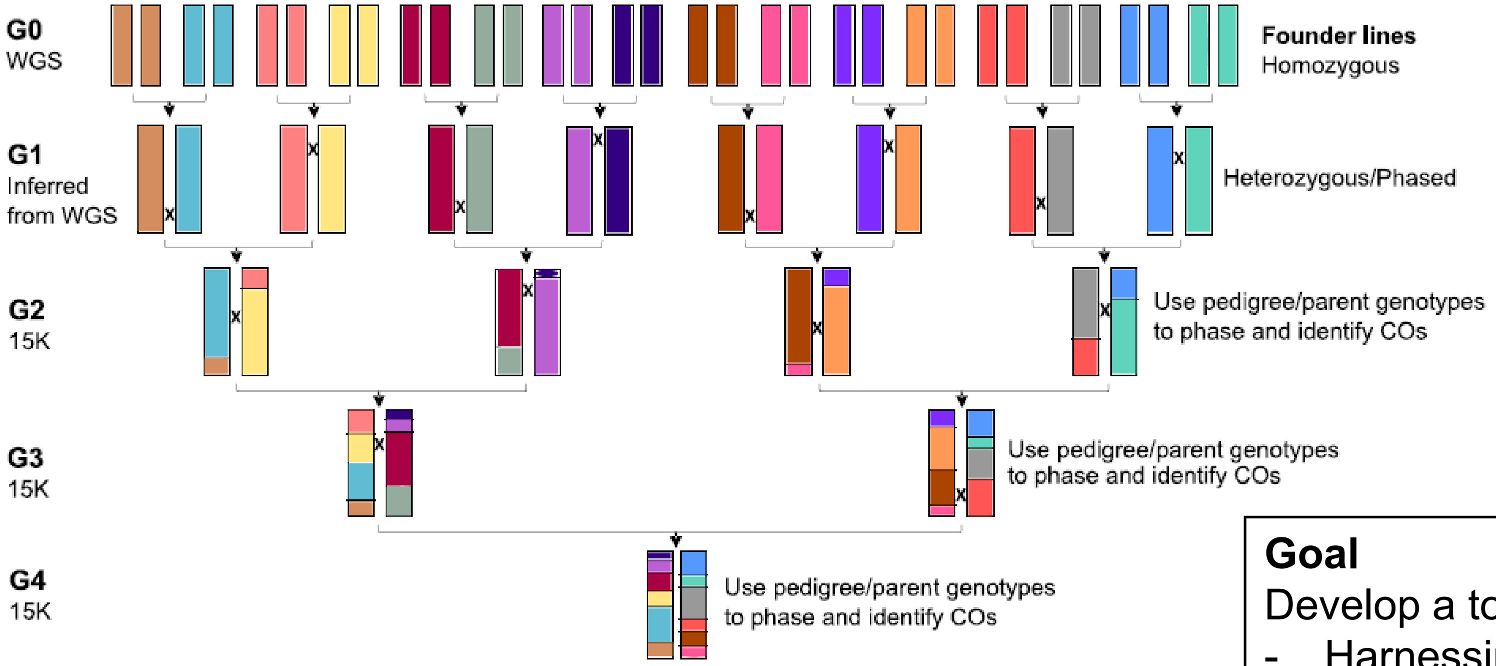




Lambing et al., 2017

**PhD project:
Defining the fine scale recombination landscape of the oilseed rape genome**

MPPs for detecting recombination



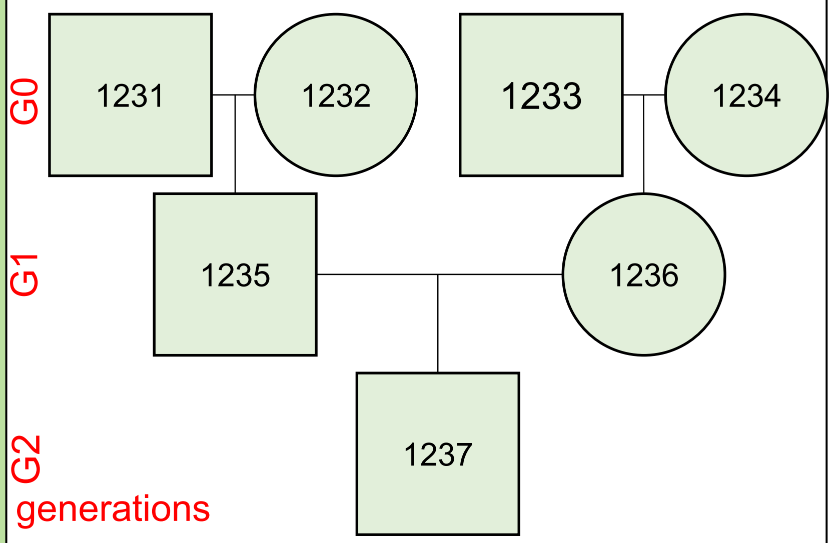
Goal

Develop a tool to detect COs by:

- Harnessing MPPs potential
- Limiting genotyping error impact

haploMAGIC. Requirements and output

FULL PEDIGREE



SNP DATA

.PED file

generation	offspring ID	father ID	mother ID	snp1	snp2	snp3	snp4	snp5	snp6	snp7	snp8	snp9	snp10
G0	ID1231	-	-	11	11	22	11	11	11	11	11	11	22
G0	ID1232	-	-	22	22	11	11	22	22	22	11	11	11
G0	ID1233	-	-	11	11	22	22	11	11	22	00	11	22
G0	ID1234	-	-	11	11	22	11	11	22	11	00	11	11
G1	ID1235	ID1231	ID1232	12	12	12	11	12	12	12	11	11	12
G1	ID1236	ID1233	ID1234	11	11	22	12	11	12	12	00	11	12
G2	ID1237	ID1235	ID1236	12	12	12	11	11	11	12	11	12	22

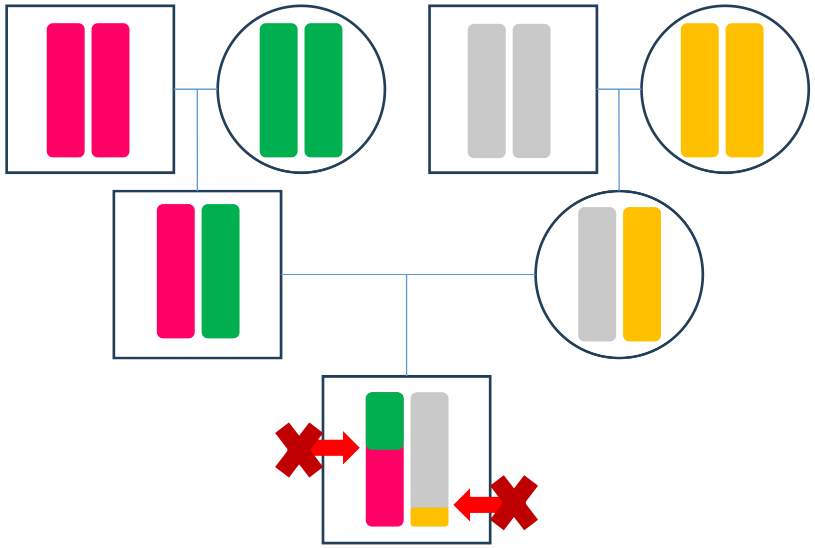
.MAP file

chr	SNP	Pos (bp)
chrC05	snp1	6008
chrC05	snp2	8322
chrC05	snp3	14390

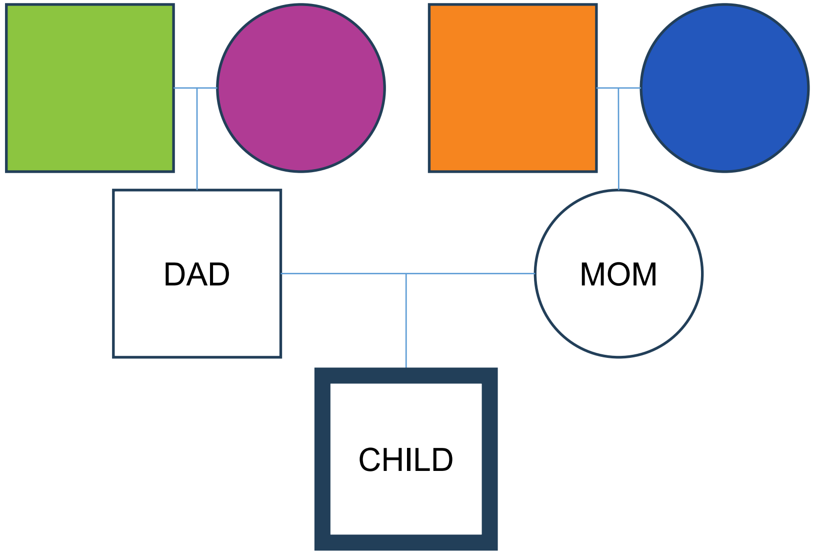
Genotypes:

0	missing allele
1	minor allele
2	major allele

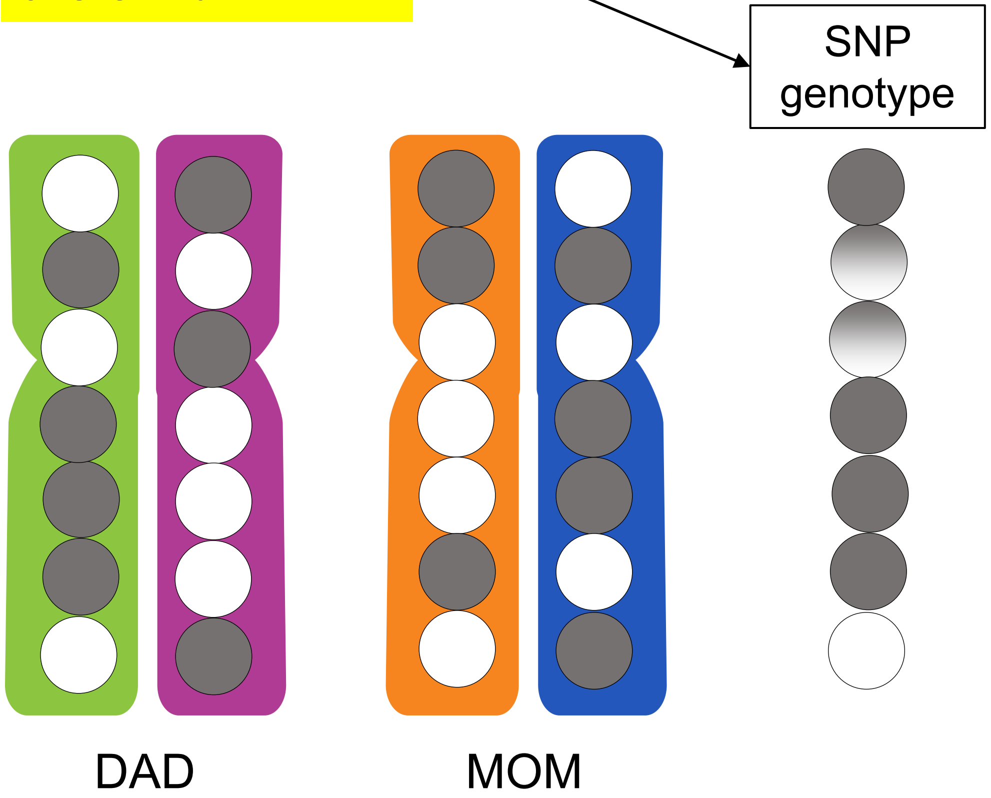
PHASE + HAPLOBLOCKS =



haploMAGIC. Pipeline

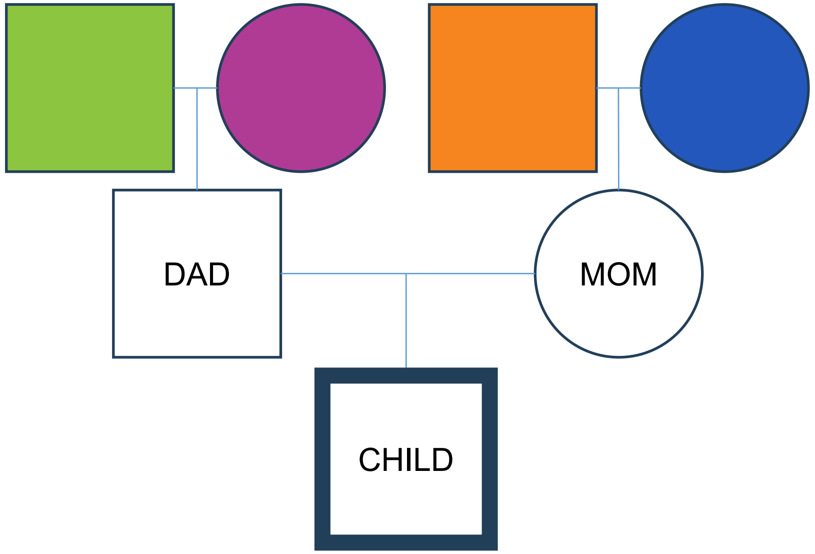


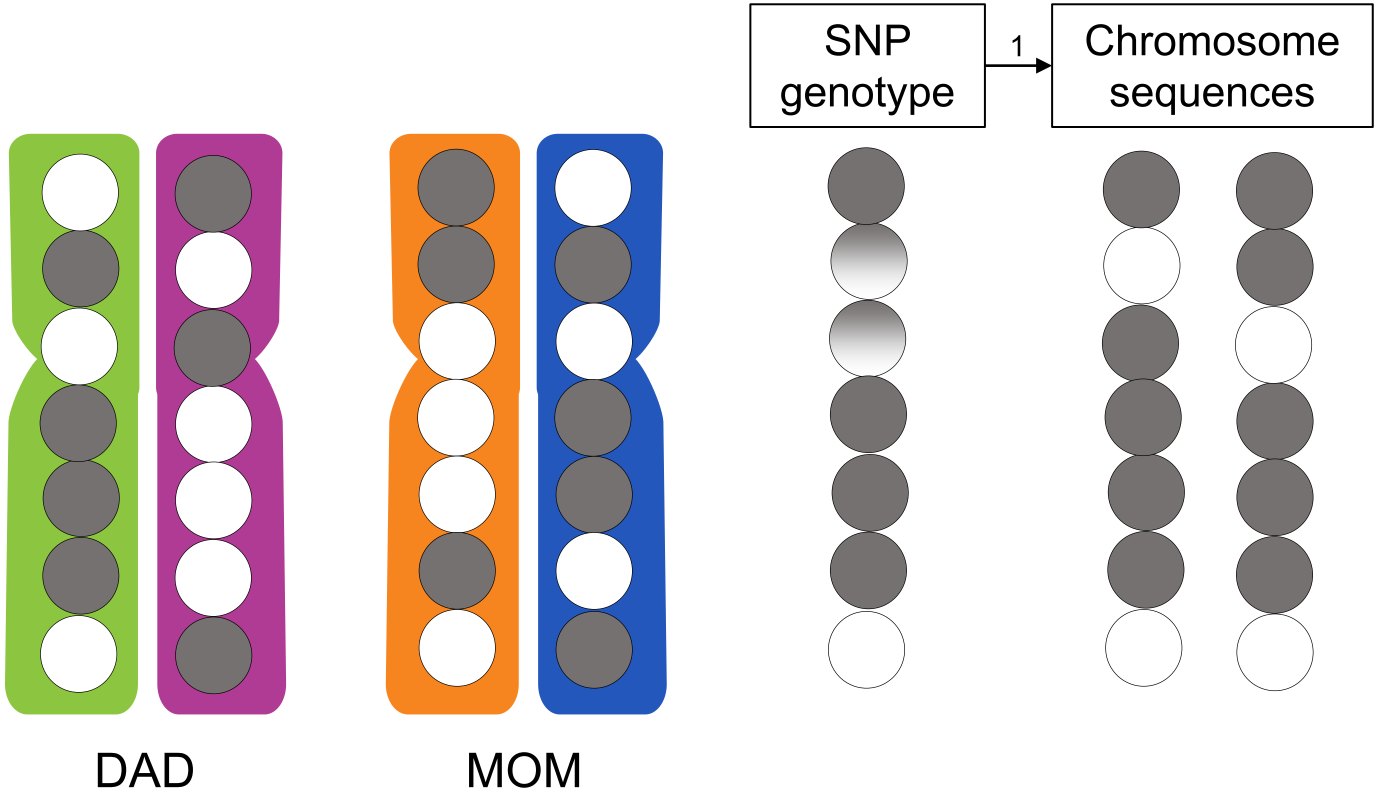
GOAL: Detect crossovers in the child



haploMAGIC. Pipeline

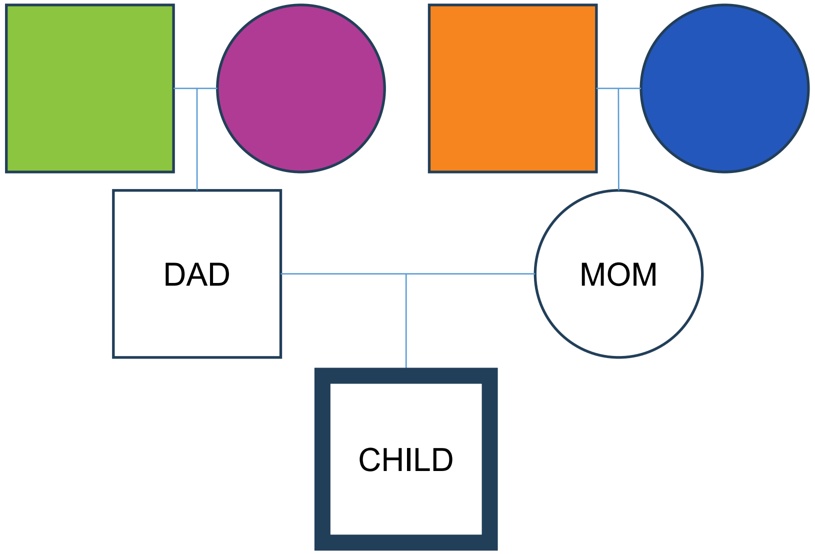
PHASING

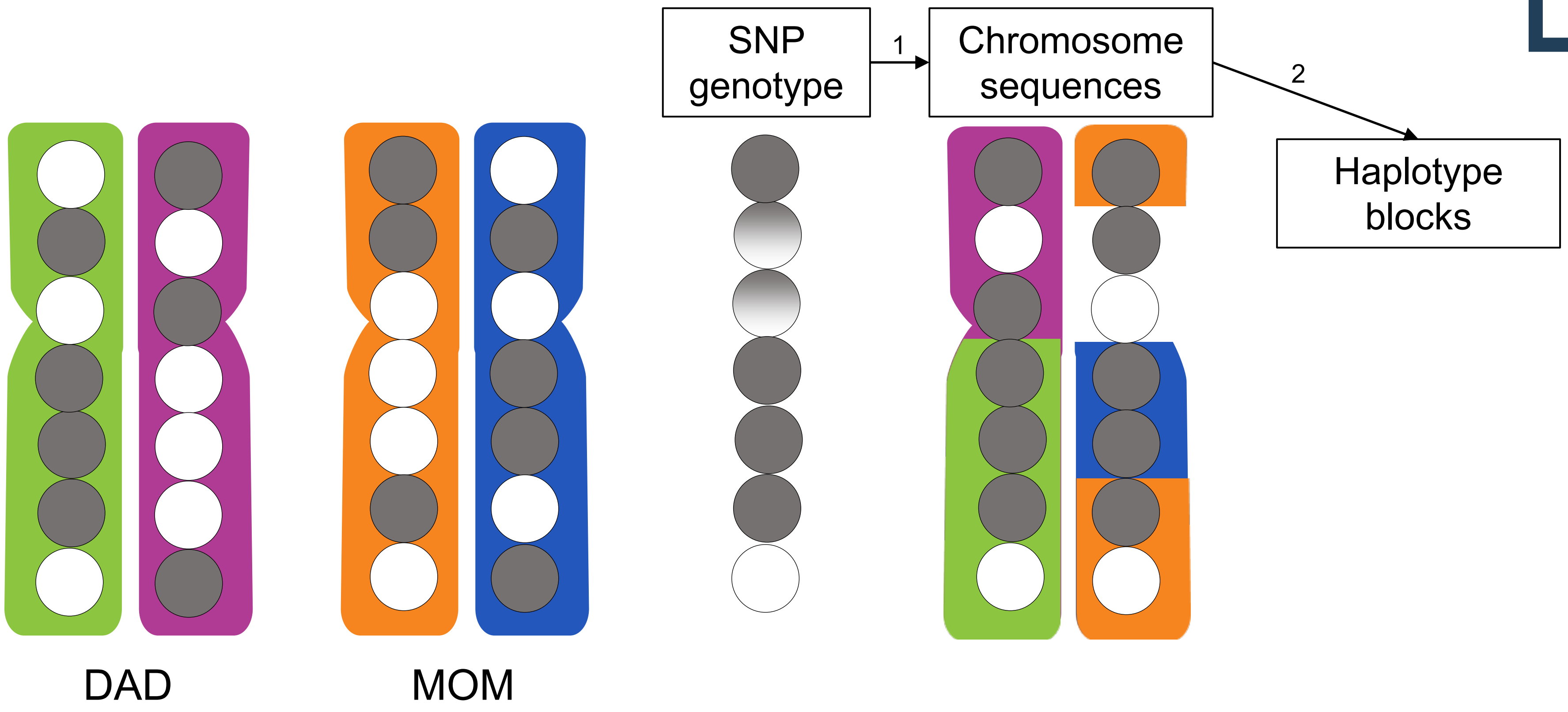




haploMAGIC. Pipeline

HAPLOTYPING

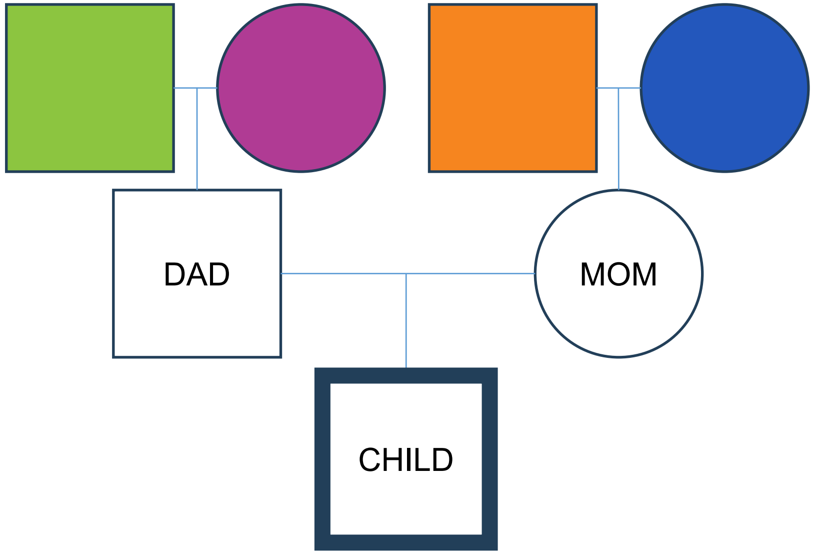



SNP genotype

Chromosome sequences

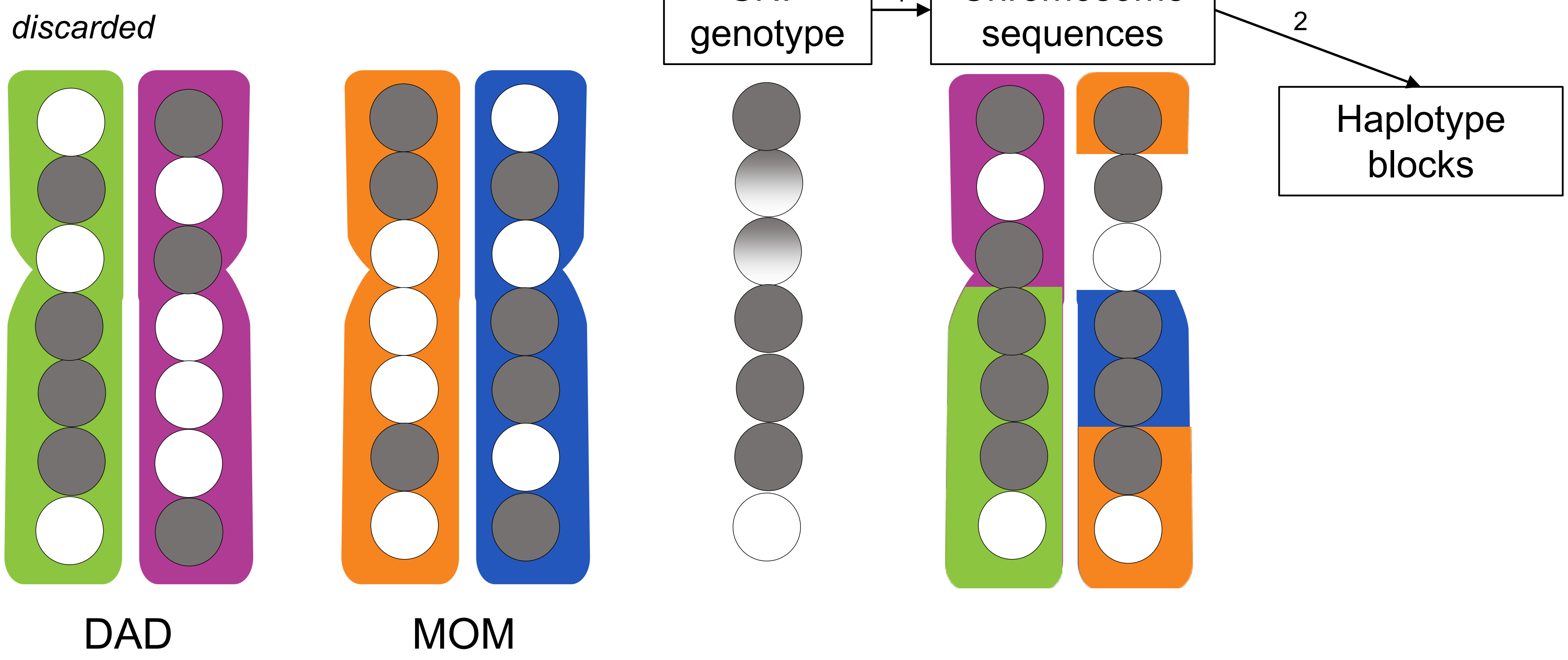
Haplotype blocks

haploMAGIC. Pipeline

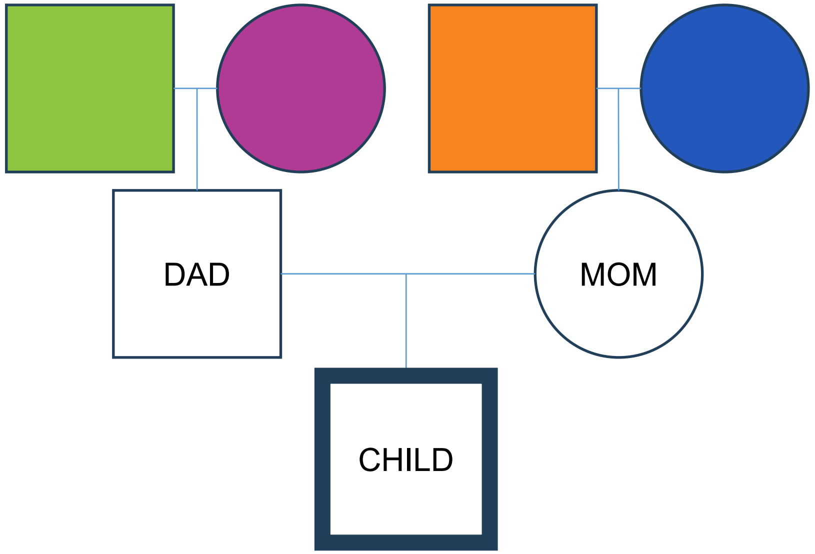


HAPLOBLOCK FILTERING

Haploblocks with ***<2*** *informative alleles are discarded*

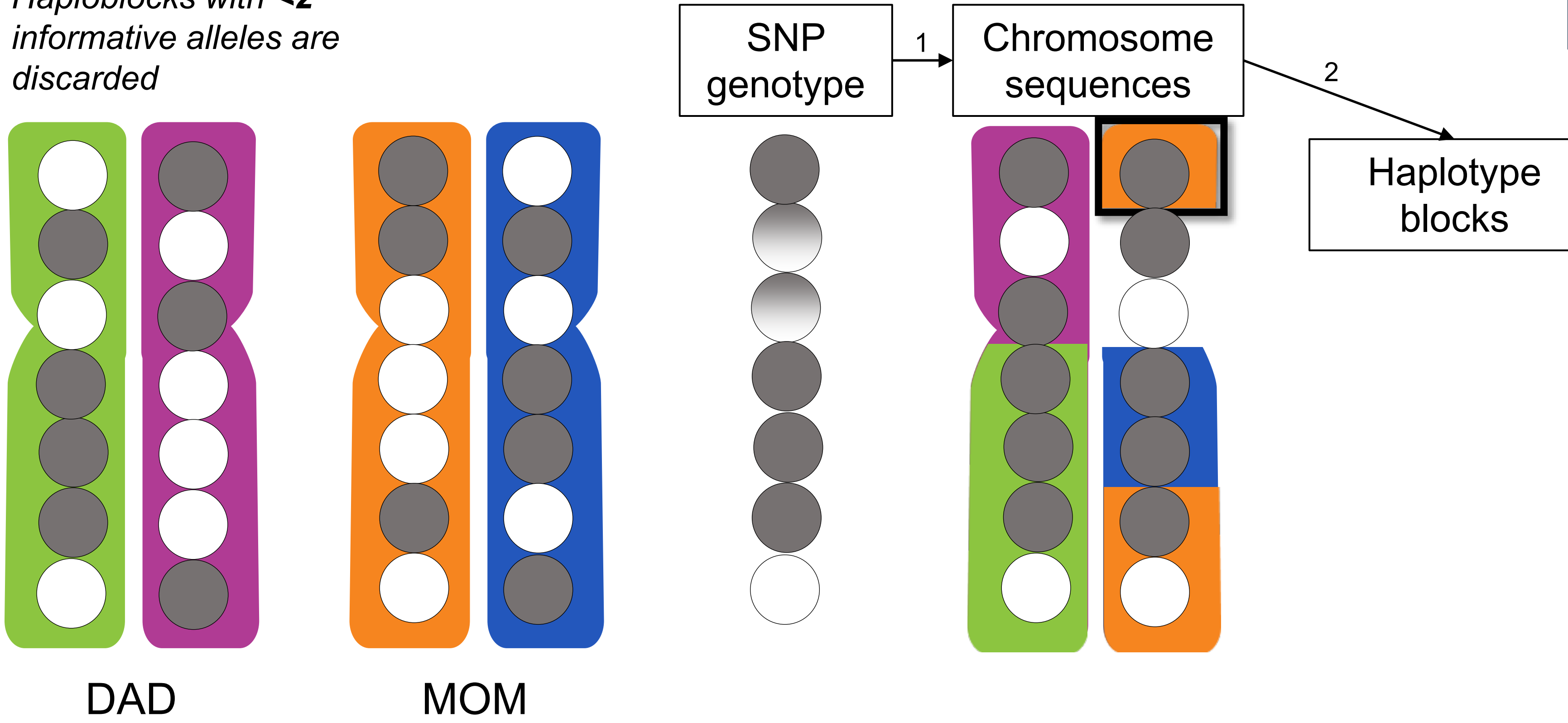


haploMAGIC. Pipeline



HAPLOBLOCK FILTERING

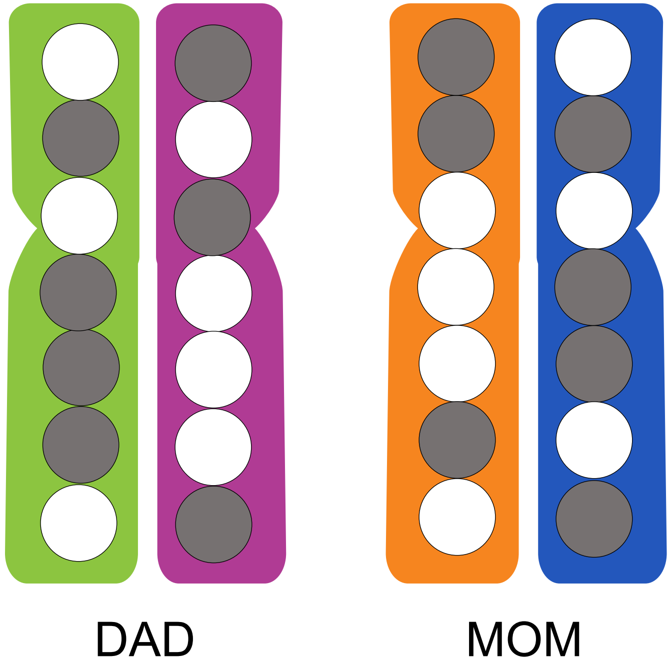
Haploblocks with ***<2*** *informative alleles are discarded*

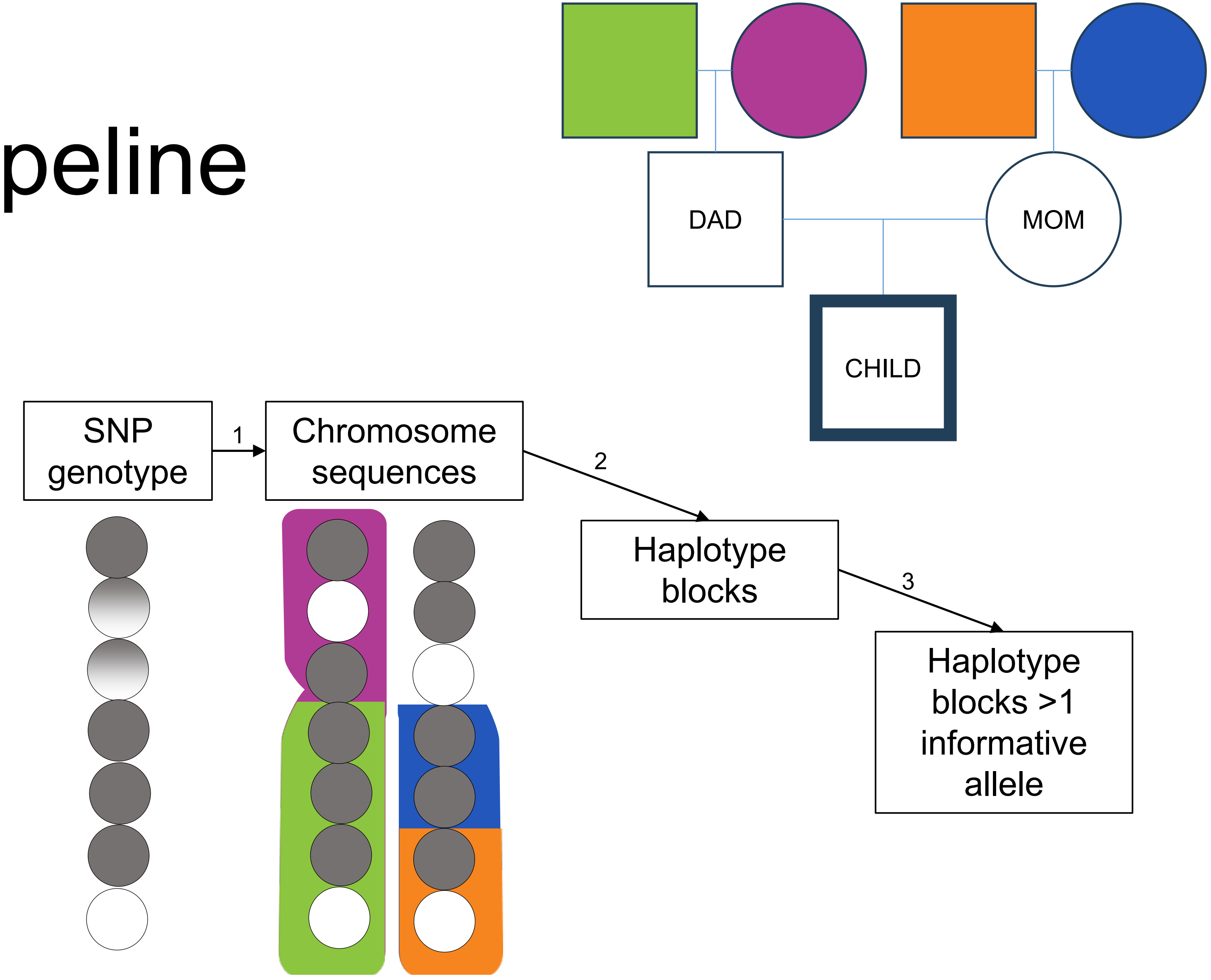


haploMAGIC. Pipeline

HAPLOBLOCK FILTERING

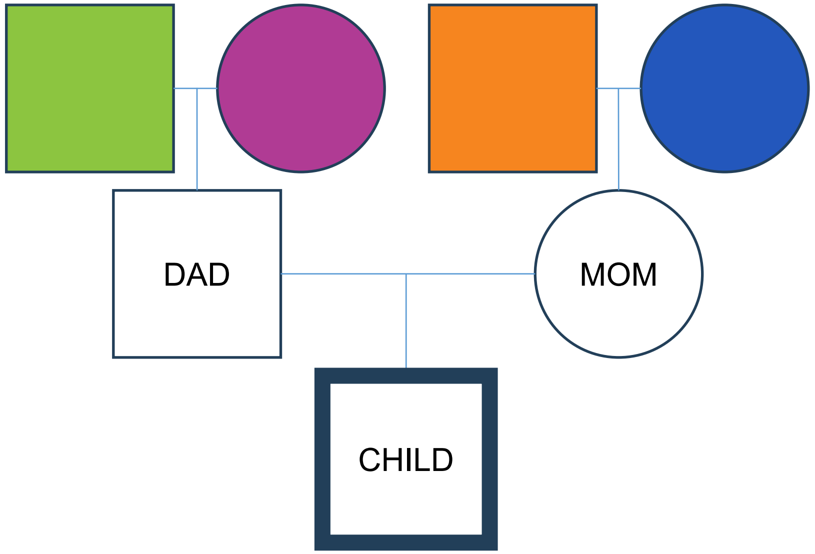
Haploblocks with ***<2*** *informative alleles are discarded*

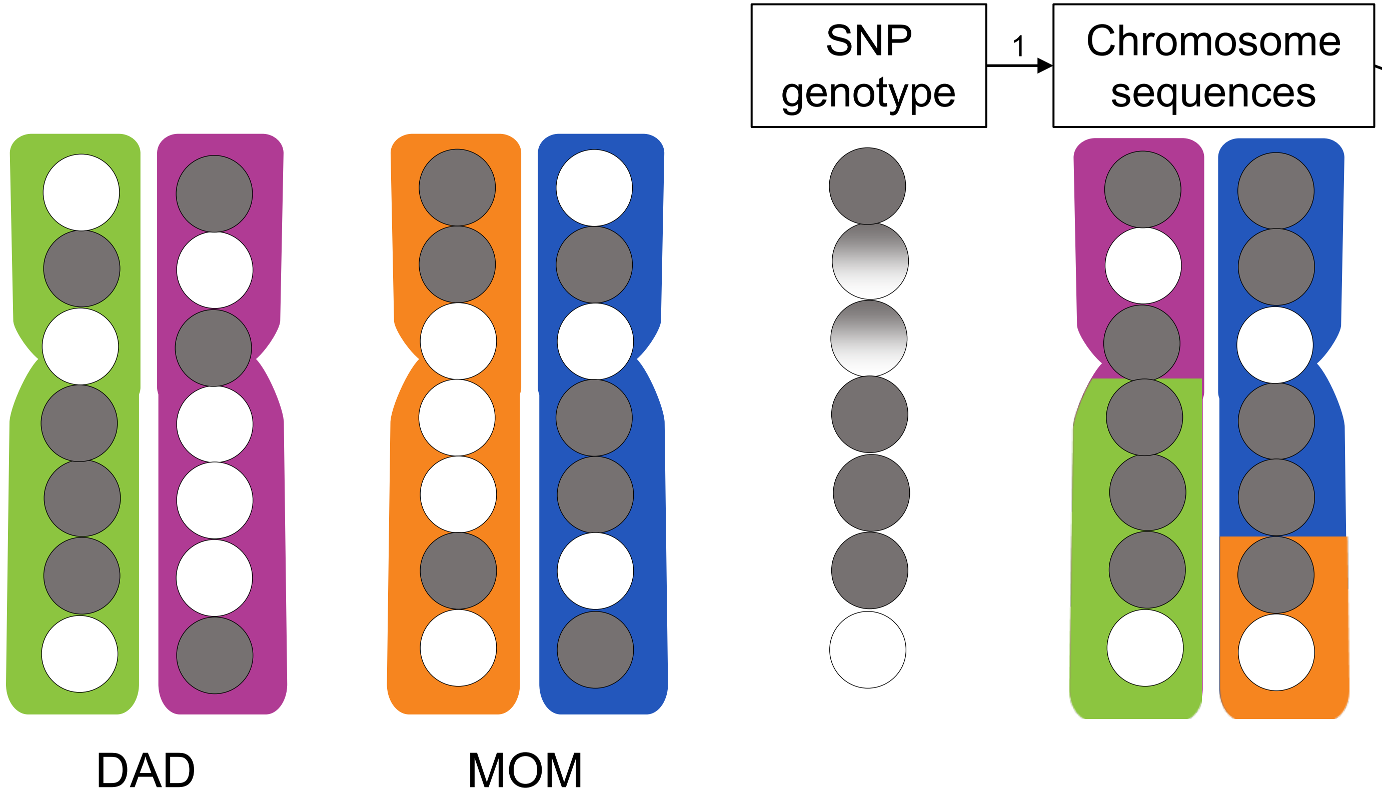


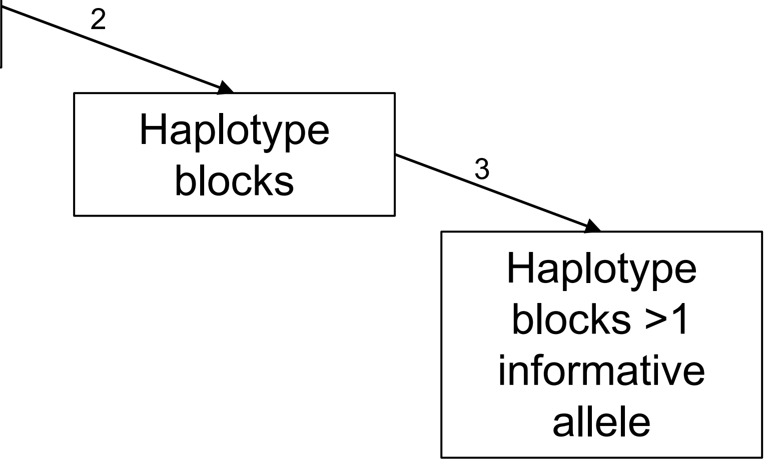


haploMAGIC. Pipeline

IMPUTATION

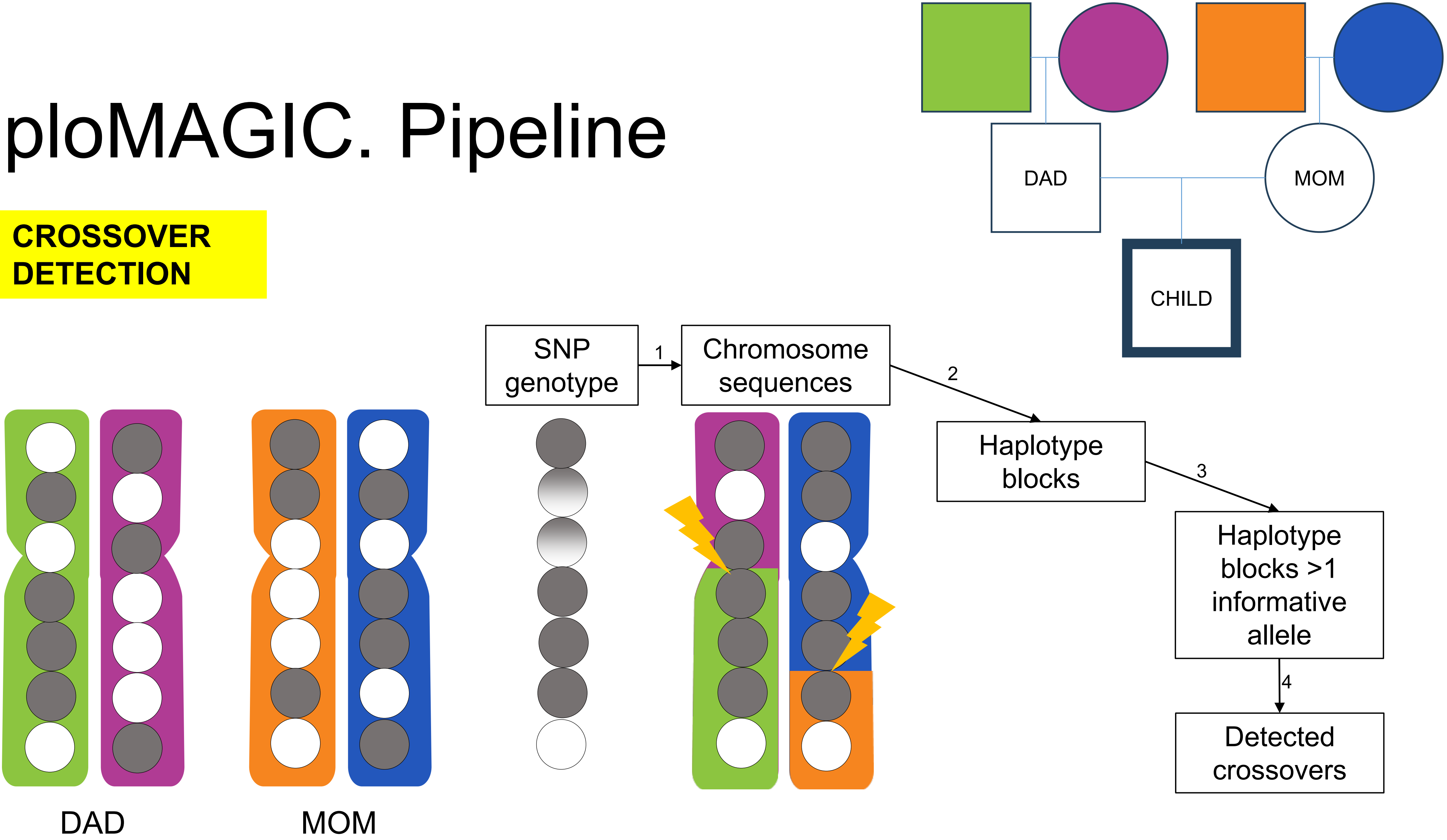






haploMAGIC. Pipeline

CROSSOVER DETECTION



SNP genotype

Chromosome sequences

Haplotype blocks

Haplotype blocks >1 informative allele

Detected crossovers

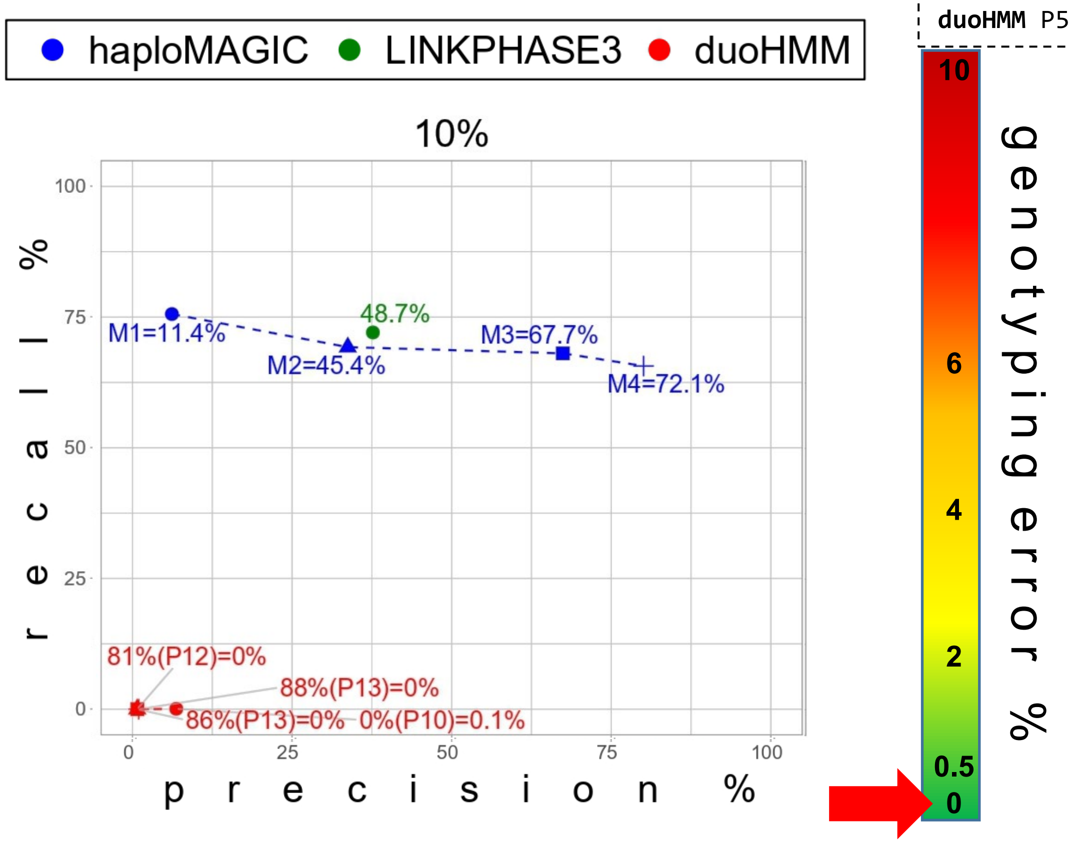
Results

haploMAGIC performance with simulations

1. **haploMAGIC**: high recall despite filtering

2. **LINKPHASE3**: no confidence levels

3. **duoHMM**: penalty on recall by filtering

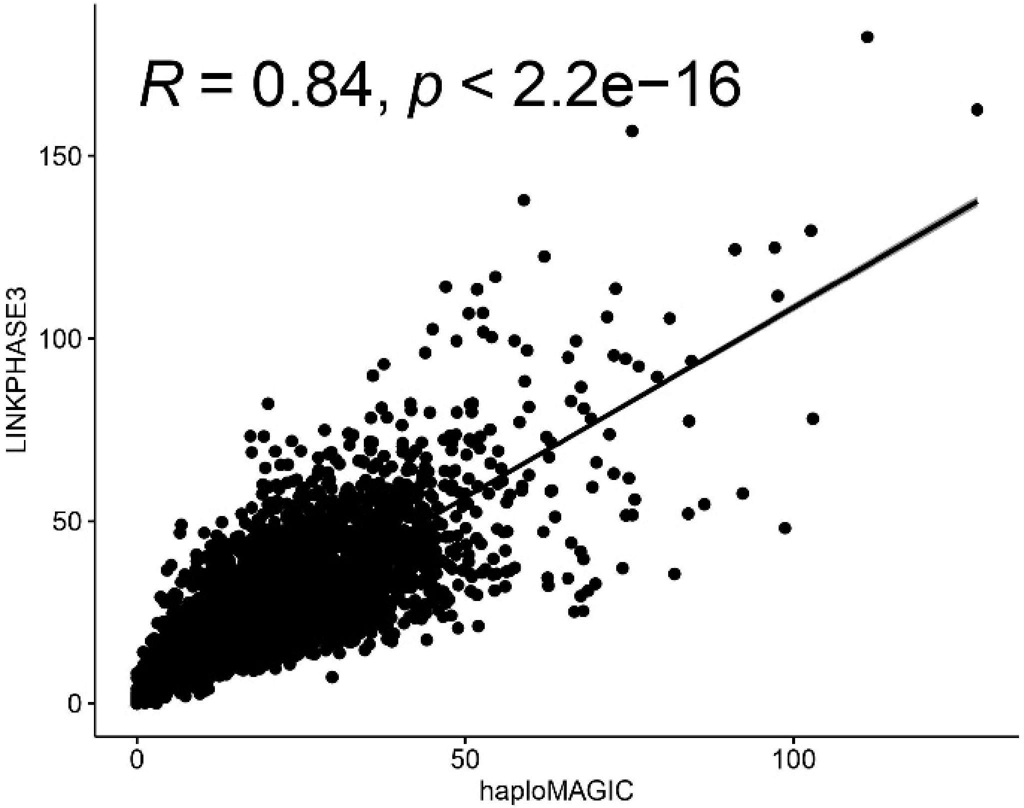


Confidence levels:

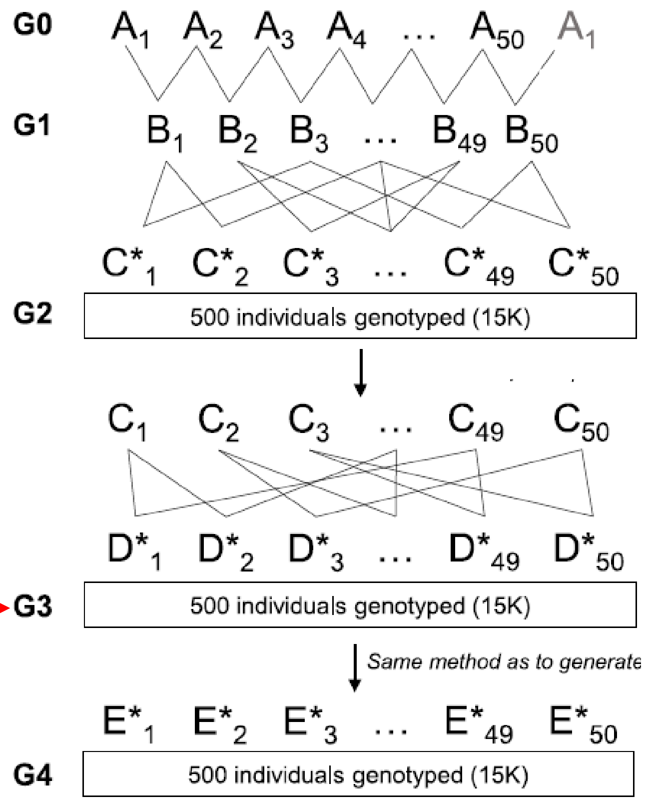
haploMAGIC M2 = “no haploblocks <2 informative alleles”

duoHMM P50 = “no crossovers <50% prob”

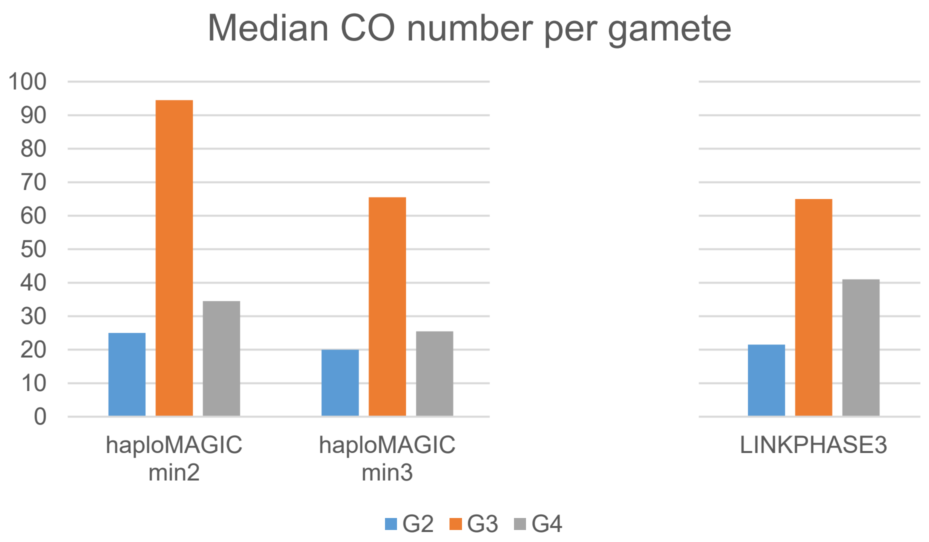
Correlation haploMAGIC-LINKPHASE3



Applying haploMAGIC to real data

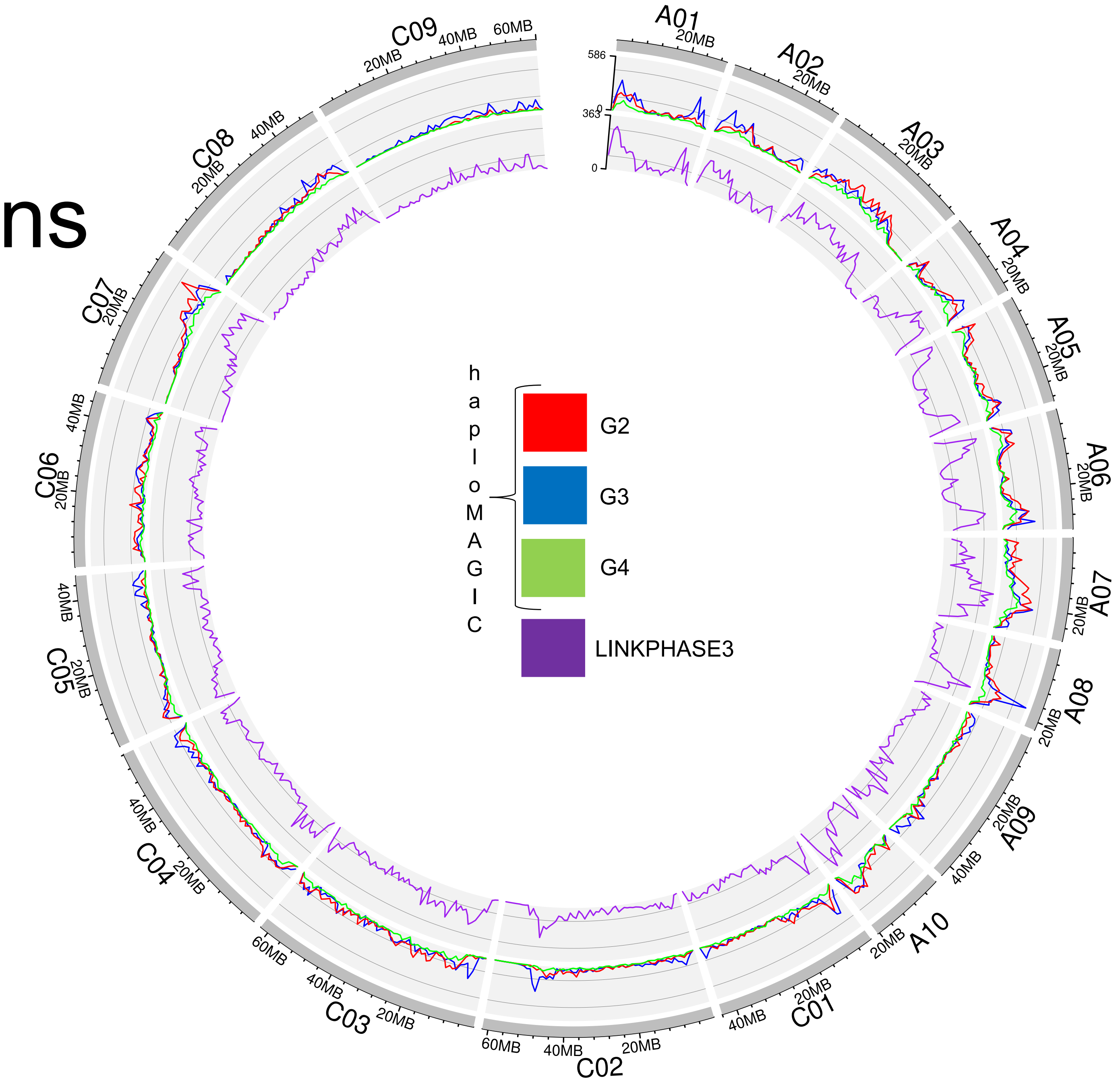


X8 genotyping error!



Filtering can be adjusted based on genotyping errors

Genome-wide recombination patterns

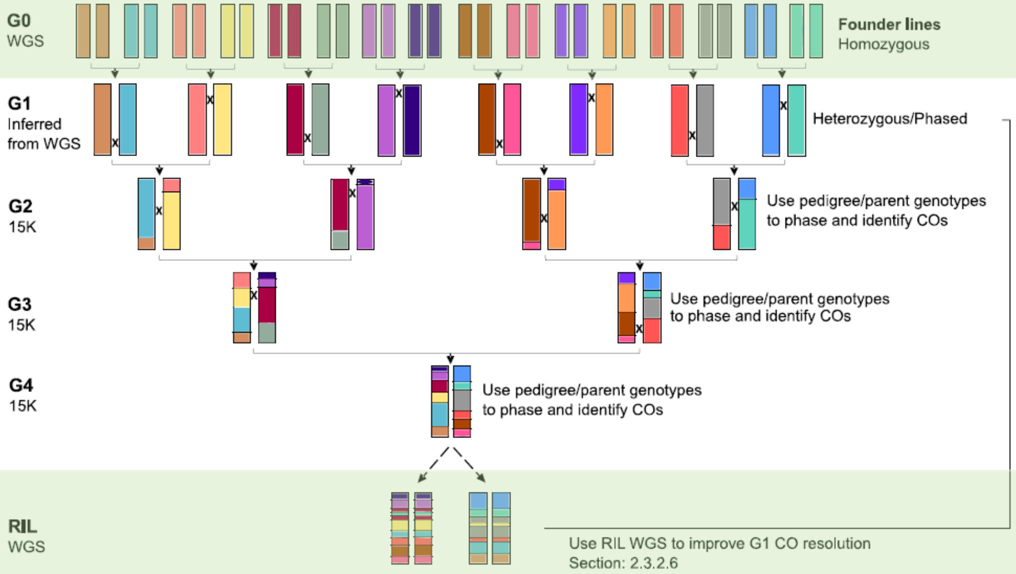


Take-home messages

- haploMAGIC is specialized in **multiparental populations** (full pedigree & SNP data) for detecting recombination events
- haploMAGIC filters haploblocks caused by **genotyping errors** by setting a minimum number of informative loci
- **Performance** based on F1-scores obtained with haploMAGIC in simulated populations was prominently better than with similar methods
- haploMAGIC will be **open-source**

Future steps

- Founder lines to increase CO resolution (coalescent method)



G0
WGS

Founder lines
Homozygous

G1
Inferred
from WGS

Heterozygous/Phased

G2
15K

Use pedigree/parent genotypes
to phase and identify COs

G3
15K

Use pedigree/parent genotypes
to phase and identify COs

G4
15K

Use pedigree/parent genotypes
to phase and identify COs

RIL
WGS

Use RIL WGS to improve G1 CO resolution
Section: 2.3.2.6

- RILs to confirm haploMAGIC haploblocks

Acknowledgments

Agrobioinformatics group (JLU)

Dr. Agnieszka Golicz[1]

Dr. Silvia Zanini[2]

Gözde Yildiz[3]

Kevin Rockenbach[4]

Kübra Arslan[5]

Venkataramana Kopalli[6]

Rishi Srivastava[7]

Plant Breeding Department (JLU)

Prof. Dr. Rod Snowdon

Dr. Nayyer Abdollahi-Sisi

Other collaborators and their location

Dr. Amine Abbadi (NPZ)

Prof. Dr. Matthias Frisch (JLU)


1


2


3


4


5


6


7

Contact:
jose.a.montero-tena@ab.uni-giessen.de
agnieszka.golicz@agrar.uni-giessen.de