Thesis report

Inheritance for a MAGIC population of Peanut: A simulation study

Moscoso, Diego
27/03/2024

Table of Contents

Inheritance for a MAGIC population of Peanut: a simulation study

1. INTRODUCTION

Peanut (*Arachis hypogaea L.*) is one of the important oil crops in the world. In 2020, peanut world production reached 53.8 million tons. China is the biggest producer with 18 million tons, followed by India with 9.95 million tons and Nigeria with 4.5 million tons ("Peanut (groundnuts excluding shelled) production in 2020". FAOSTAT 2020).

1.1 Allopolyploids and autopolyploids

Cultivated peanut is an allotetraploid species (2n = 4x = 40). Allopolyploids usually result from hybridization of two related species involving unreduced gametes (Harlan & deWet, 1975). The two species that hybridized and donated their genomes to cultivated peanut are *Arachis duranensis* (genome A) and *Arachis. ipaënsis* (genome B) (Seijo et al., 2007). Allopolyploids usually exhibit disomic inheritance, meaning that closely related chromosome copies (i.e., two chromatids from genome A of peanuts, called homologous) may pair and recombine during meiosis. On the other hand, less-related chromosome copies (i.e., one chromatid from genome A and one from genome B of peanuts, called homoeologous) may not pair or recombine (Bourke et al., 2018).

There is another group of polyploids called autopolyploids. They result from genomic duplication within a single species involving unreduced gametes (Harlan & deWet, 1975). Autopolyploids display polysomic inheritance, meaning that all the chromatids are homologous, therefore, when meiosis occurs, pairing and recombination can occur between all chromatids that form a chromosome (Bourke et al., 2018).

In wild peanuts, sub-genomes A and B presumably started divergence around 3.5 million years ago (Moretzsohn et al., 2013; Nielen et al., 2012), considerably less than the sub-genomes of soybean and cotton with 13 and 6.7 million years of divergence, respectively (Schmutz et al., 2010; Senchina et al., 2003). Therefore, cultivated peanut usually has low polymorphisms and low diversity (Bertioli et al., 2014). Mechanisms like homoeologous recombination, deletions and mobile-element activity gave peanuts some level of diversity and genetic evolution (Bertioli et al., 2019). Allotetraploidization of peanut made it more attractive for cultivation thanks to changes like plant size, transpiration characteristics, and higher photosynthetic capacity. (S. C. M. Leal-Bertioli et al., 2017).

1.2 Segmental allopolyploid behavior in peanuts

Segmental allopolyploid organisms show disomic and tetrasomic behavior in different sections of their genome. The allotetraploid configuration of peanut suggests a disomic inheritance, with recombination mostly within homologous chromatids. Inheritance studies using F2-F3 progeny and Chi-square test report markers that fit disomic inheritance of loci for many traits (Balaiah et al., 1977; Bauer et al., 1973; W. D. Branch, 2011; Tripp, 1968). Nevertheless, other studies report different ratios of segregation of a trait depending on the parents (Balaiah et al., 1977; Essomba et al., 1993; Tripp, 1968) or segregation ratios that did not always fit as expected in a specific generation due to active cytoplasmic and/or nuclear modifying factors (William D. Branch, 2008; Pattanashetti et al., 2008).

There is also evidence of homoeologous recombination events in the progeny of *A. hypogaea* and the synthetic allopolyploid (*Arachis duranensis* × *Arachis batizocoi*). An experiment involving *A. hypogaea* and newly synthetic peanut allopolyploids suggested some degree of tetrasomic behavior in peanuts (Nguepjop et al., 2016). Another study developed Recombinant Inbred Lines (RIL) from the colchicine-induced allotetraploid (*Arachis ipaënsis* X Arachis. duranensis) and *A. hypogaea*. The linkage map construction showed recombination of blocks of markers between subgenome A and genome B (S. Leal-Bertioli et al., 2015).

1.3 Bi-parental populations

Several inheritance studies use RIL populations for the construction of genetic maps (Cobos et al., 2006; Eskandari et al., 2013; Li et al., 2013; Paul et al., 2018; Shiringani et al., 2010). Just as other types of populations such as BC1 of F2, RIL populations allow the analysis of alleles differing between two lines. This can be used to construct linkage maps and find QTLs (Pascual et al., 2015), moreover, RIL have the advantage of assessing the phenotype in a plot of plants of the same genotype instead of an only individual plant (Broman, 2005). RIL analysis relies especially on the recombination events during the intercross to obtain F1 where only two parents are involved. The selfing of the progeny in subsequent generations (F2 to Fn) still produces recombination in heterozygous sections of the genome. The heterozygous sections of the genome decrease in number the further the generation of selfing. The advantages of this setup in most cases are: (1) Simple construction, just 2 generations are needed to obtain an F2 and only six further generations of selfed plants are needed to fix the genome in each inbred line. (2) All allele frequencies expected to be close to the optimal value of 50%, which gives a high power to detect QTL. (3) Low rate of linkage disequilibrium (LD) decay with only 1 or 2 recombinants per arm (Scott et al., 2020).

Nevertheless, there are disadvantages of bi-parental populations: (1) Allelic variation is small making the number of QTLs detected smaller. (2) Limited effective recombination that results in genetic maps with poor resolutions of tens of centiMorgans (cM). Hundreds or thousands of genes are within a few tens of cM, therefore, the identification of candidate genes for QTLs can be a time-consuming process (Hall et al., 2010; Scott et al., 2020).

The resolution of a genetic map can be increased 3 times-fold compared to RIL when using AIL (Advanced Intercross lines). This method consists of an F2 derived from two inbred lines, but the later generations are randomly intercrossed instead of selfed. This increases the recombination events, increasing the resolution of the genetic map (Darvasi & Soller, 1995). On the other hand, this method doesn't overcome the small variation of alleles.

The low genetic diversity of cultivated peanut can be a challenge for high-resolution genetic maps development. A bi-parent method such as AIL can be extended to the MAGIC (Multi parent Advanced Generation Inter Cross) approach.

1.4 MAGIC multi-parental populations

Various inheritance studies use MAGIC populations for genetic analysis in allopolyploid and diploid species (Bandillo et al., 2013; Dell'Acqua et al., 2015; Bevan E. Huang et al., 2012; Huynh et al., 2018; Sallam & Martsch, 2015; Sannemann et al., 2015; Yan et al., 2020). This type of population (Fig. 1) is obtained from intercrossing 4, 8, or 16 divergent parents denominated as founders. They are intercrossed through various generations in a balanced way. Finally, the population is selfed to develop inbred lines. The chromosomes in the lines are mosaics from all the founder haplotypes. Both increased variation and recombination are present in the lines which improve the detection of QTL (Scott et al., 2020).

MAGIC populations have advantages: (1) MAGIC populations incorporate great diversity into the population that recombines many times, resulting in a high mapping power and resolution compared to bi-parental populations (C. Cavanagh et al., 2008; Dell'Acqua et al., 2015; B. Emma Huang et al., 2015; Scott et al., 2020). (2) The genomes of the population are fine-scale mosaics from all the founders (C. R. Cavanagh et al., 2013) (3) MAGIC populations can also provide excellent materials for plant breeding due to their features of high recombination and the resulting diverse phenotypic diversity (Yan et al., 2020) (4) Based on the selection of founders more targeted traits can be analyzed compared to a bi-parental population (Bandillo et al., 2013). (5) Tetrasomic behavior in segmental allotetraploids can depend on the parental genotypes (Clevenger et al., 2017; Nguepjop et al., 2016). The identification

of abnormal tetrasomic behavior in a MAGIC population is much more likely to be detected compared to a bi-parental population.

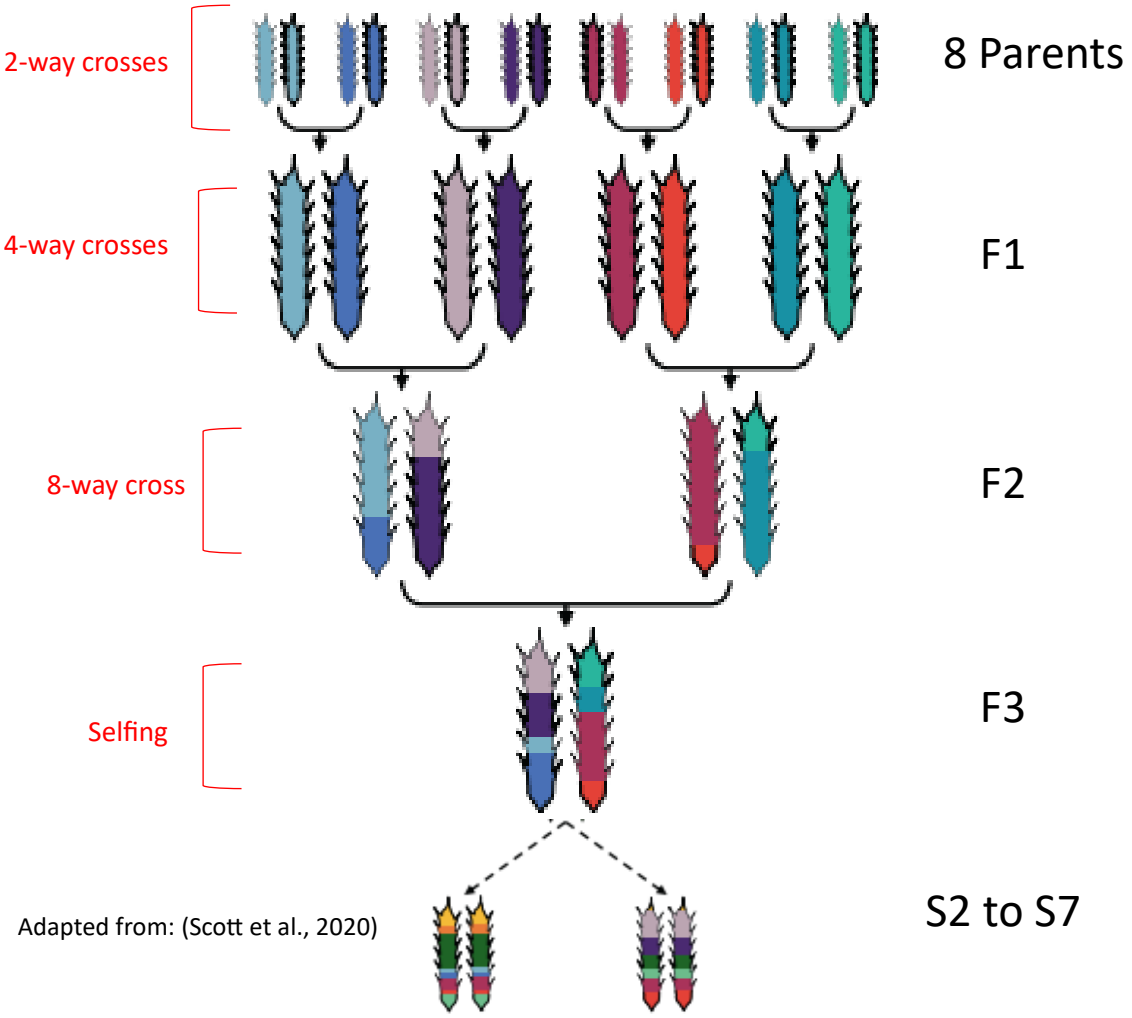

Fig 1. Equally balanced crosses in a Magic population. Mosaics from all the founder genomes are present in the population. Every generation is a product of balanced intercrossing.The F3 generation is selfed from to find segregation rates of alleles of interest.

1.5 Bi-allelic markers

Genetic maps in allotetraploid species have been developed using bi-parental populations and bi-allelic markers such as SNP (single nucleotide polymorphism) (Byers et al., 2012; Cai et al., 2014; Clarke et al., 2016). In recent years, genetic maps in autotetraploid species have been created also using bi-parental populations and SNP marker data (Bourke et al., 2016; Hackett et al., 2013). Methods such as

TetraOrigin (Zheng et al., 2016) have been developed to reconstruct haplotypes in autotetraploid mapping populations.

In genetic analysis of bi-parental progenies, SNP markers are a powerful tool. Compared to other markers the advantages include (1) the cost per data point is lower in large-scale assays (Bertioli et al., 2014); (2) The relative abundance of SNP markers in a genome is useful when the species is allopolyploid because the coupling-phase recombination frequency is identical to recombination frequencies of diploid species (Bertioli et al., 2014; Bourke et al., 2015).

1.6 Multi-allelic markers

Microsatelite (SSR) markers fall in the multi-allelic category. However, they have some disadvantages. (1) The development of an assay for SSR markers is expensive (Clevenger et al., 2017), (2) SSR markers are limited for their use in genome-wide distribution (Landergott et al., 2006; Xie et al., 2015) and (3) they are often scored as dominant (L. H. Zhang et al., 2006), meaning that the allelic diversity of SSR markers is not well exploited.

Information provided by bi-allelic SNPs can be transformed into multi-allelic markers, transferring all the advantages of SNPs to a multi-allelic information format. The resulting markers are denominated haploblocks (Sabeti et al. 2002). They consist of closely linked SNP markers in coupling phase which are inherited with rare recombination events, therefore, the whole block of markers is inherited. Each different combination of markers that occurs within that block is a haplotype that can be passed on to progeny individuals as a multi-SNP allele. A haploblock holds more information than separate SNP data, because the possible number of alleles is more than two (Voorrips & Tumino, 2022). On the other hand, the amount of haploblock markers obtained usually isn't as high as when considering all the SNP markers (Pook et al., 2019; Voorrips & Tumino, 2022). Haplotypes can be constructed from bi-allelic SNP markers using the software PolyHaplotyper (Voorrips & Tumino, 2022) and BEAGLE (Browning et al., 2021).

2. METHODS

2.1 Notation used for alleles.

2.1.1 Notation for disomic inheritance

We represented the genotypes for disomic inheritance with 2 subgenomes, denoted subgenome 1 and subgenome 2. The subgenomes are separated by an underscore '_'. Each subgenome has 2 alleles with 2 possible notations for each allele, "A" or "a" for subgenome 1 and "B" or "b" for sub genome 2 (Fig. 2).

Genotype notation	**All possible genotypes**		
Subgenome 1 ↓ AA_BB ↑ Subgenome 2	AA_BB AA_bb aa_BB aa_bb	AA_Bb aa_Bb Aa_BB Aa_bb	Aa_Bb

Fig 2. Notation of genotypes for disomic inheritance

We considered genotypes to be homozygous for disomic inheritance when the two homologs within each subgenome carry the same alleles. We also considered all the parental genotypes to be homozygous. Therefore only 4 out of 9 possible genotypes were used in each parental configuration: AA_BB, AA_bb, aa_BB and aa_bb.

2.1.2 Notation for tetrasomic inheritance

Each of the 9 possible genotypes (Fig 2) under disomic inheritance was transformed into a form with tetrasomic inheritance. We represented the genotypes for tetrasomic inheritance with 4 alleles and 2 possible notations for each allele, "A" or "a" (Fig. 3).

Genotype notation	**All possible genotypes**		
No Subgenomes AAAA	AAAA Aaaa	AAAa aaaa	AAaa

Fig 3. Notation of genotypes for tetrasomic inheritance

The tetrasomic inheritance form of each genotype was defined by the by the dosage of the two alleles. For example, the disomic genotypes AA_bb, aa_BB and Aa_Bb are equivalent to the tetrasomic genotype AAaa, two copies of each allele and no distinction according to subgenome for the tetrasomic case; the disomic genotypes AA_BB are equivalent to the tetrasomic genotype AAAA and the disomic genotype aa_bb is equivalent to the tetrasomic genotype aaaa.

2.2 Structure of the study

All analyses were conducted in R (R version 4.0.1, http://www.r-project.org). This study had 3 components to determine the different markers in a MAGIC population with 8 parents in peanut. Frist, we established the number of parental configurations under disomic inheritance and calculated the number of markers we should have. We characterized each one of the markers by giving them a nomenclature and assigned a marker type to each configuration of parents. Secondly, we calculated the exact segregation ratio in each parental configuration for selfing generations 2 and 7 (S2 and S7) and verified if the parental configurations with the same marker nomenclature had the same segregation frequency in S2 and S7. Thirdly, we transformed all the parental configurations to a tetrasomic form and compared the exact segregation ratio for each disomic parental configuration to the segregation ratio for its tetrasomic form (Fig. 4).

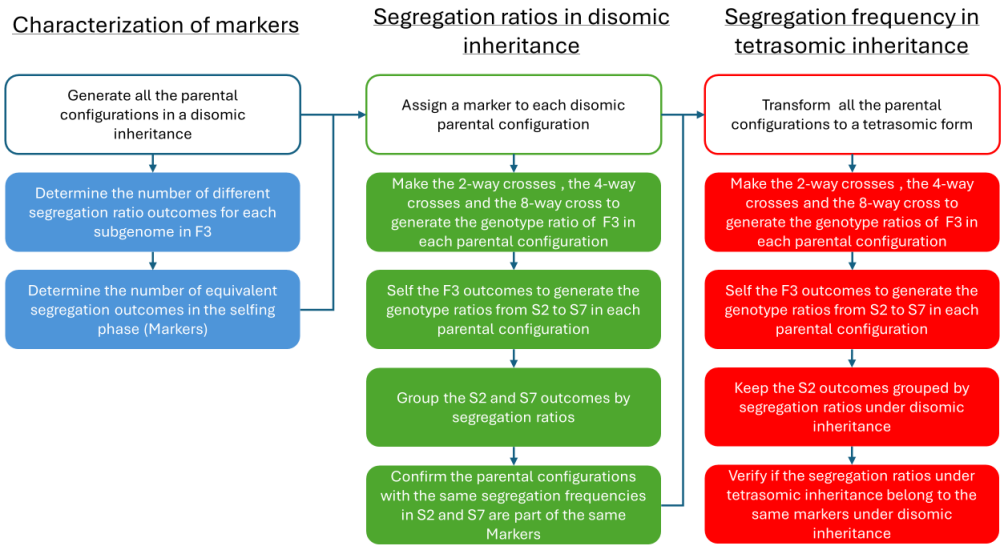

Fig 4. Flowchart of the study or genotype frequencies and ratios for parental configurations a disomic inheritance and their tetrasomic inheritance form

2.3 Characterization of markers

There are some packages in R and software used to generate linkage maps for MAGIC populations of diploid organisms like R/mpMap (B. Emma Huang & George, 2011) and GAPL (L. Zhang et al., 2019), where markers are used regardless of the order of the 2-way crosses.

In the first instances of the study, we considered that the order of the 2-way crosses in a tetraploid population under disomic inheritance didn't influence the segregation of the population, but the markers we developed were inconsistent. We changed our approach and started considering the order of the 2-way crosses. Fig. 5 shows 2 parental configurations that have the same 2-way crosses but in different order. The 2-way cross between Parent 3 and Parent 4 was swapped with the 2-way cross

between Parent 5 and Parent 6, resulting in different segregation frequencies, therefore, they are not considered as the same configuration.

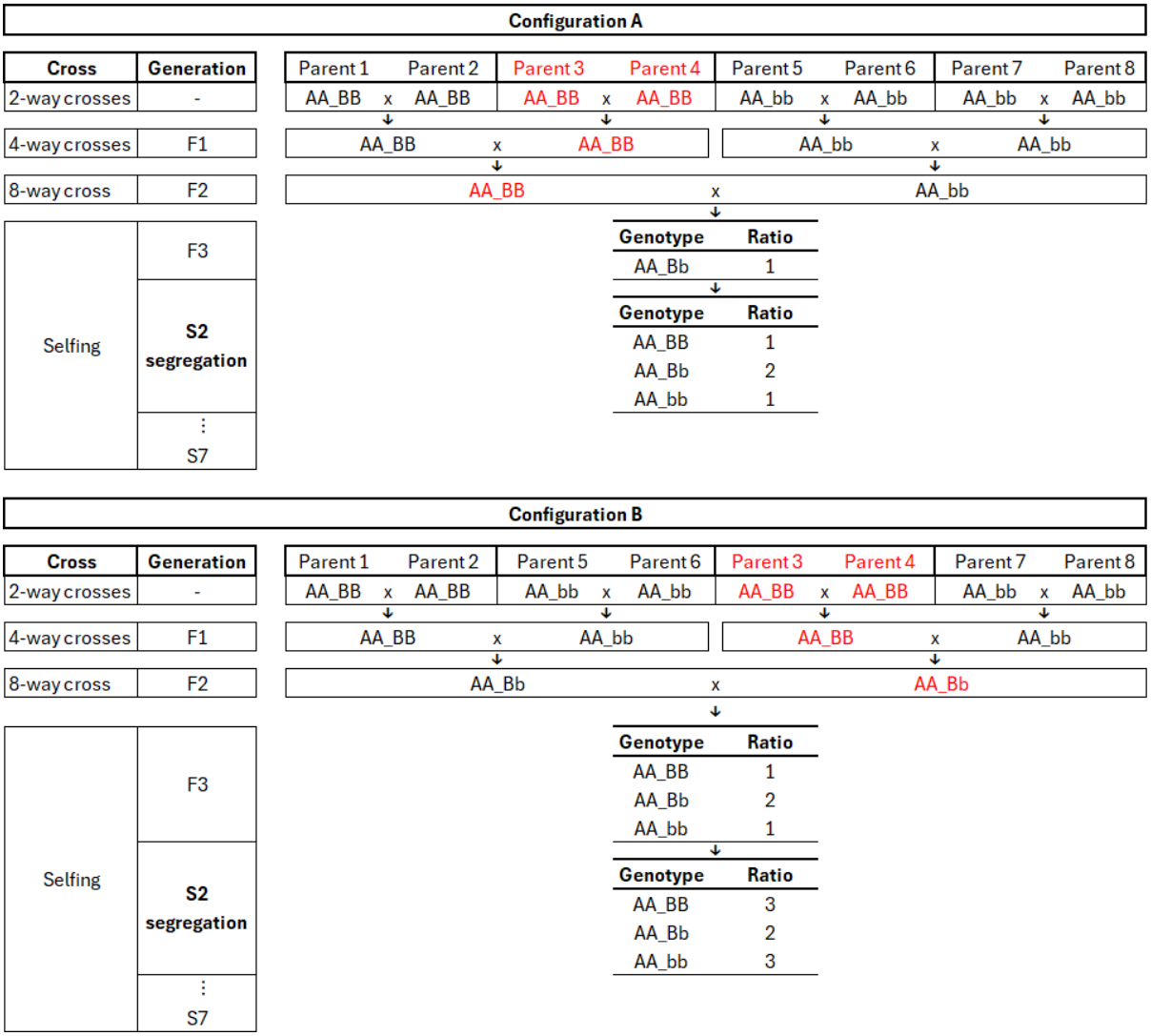

Fig. 5. Examples of parental configurations with the same parental genotypes and the same 2-way crosses but ordered differently. The segregation outcome is different between configurations.

2.3.1 Parental configurations crosses and selfings

The order of the parents, therefore, the order of the crosses was considered as influential in the segregation. We had 10 possible 2-way crosses from the parental genotypes: AA_BBxAA_BB, AA_BBxaa_BB, AA_BBxAA_bb, AA_BBxaa_bb, AA_bbxAA_bb, AA_bbxaa_BB, AA_bbxaa_bb, aa_BBxaa_BB, aa_BBxaa_bb and aa_bbxaa_bb. We generated all the possible ordered permutations of these 10 possible 2-way crosses in 4 possible results in F1. We obtained 10 000 possible parental configurations. An example of parental configurations with 4 possible results in F1 is shown in Fig. 5.

Every resulting parental configuration was crossed 3 times and then selfed 6 times. First, four ordered 2-way crosses resulted in the F1 generation (Fig. 5). Parent 1 (P1) was crossed with parent 2 (P2), parent

3 (P3) with parent 4 (P4), parent 5 (P5) with parent 6 (P6) and parent 7 (P7) with parent 8 (P8). Secondly, the resulting F1's were 4-way crossed to generate the F2 generation. The result from P1xP2 was crossed with the result from P3xP4 and the result form P5xP6 was crossed with the result from P7xP8. Thirdly, the 2 results from the 4-way cross were crossed in the 8-way cross to form the F3 that was selfed from S2 to S7.

Usually, 8-way crosses are made in a diallel of half-diallel design (Arrones et al., 2020), where all the possible 2-way, 4-way and 8-way crosses between parents are covered, meaning that 28 two-way, 14 four-way, and seven eight-way crosses could be performed per parental configuration to develop the MAGIC population (Fig. 6b).

Our study is using the funnel configuration (Arrones et al., 2020)(Fig. 6a) because we considered the order of the crosses as relevant for the segregation ratio, meaning that different order of the 2-way crosses could result in different markers.

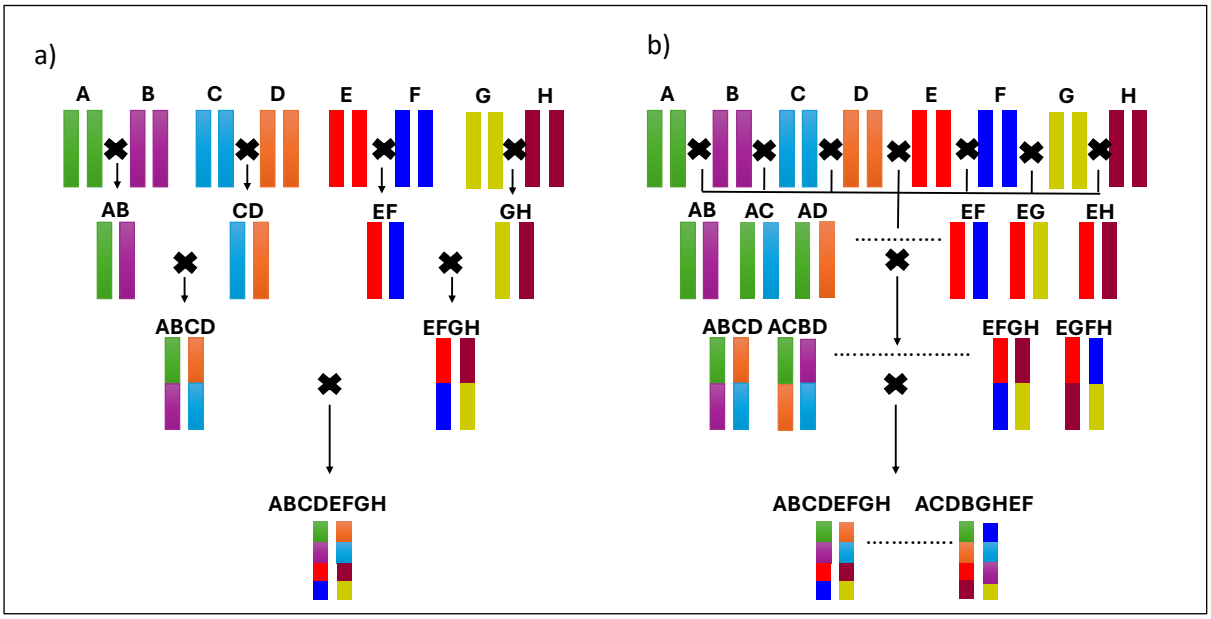

Fig. 6. Design examples for an 8-way cross MAGIC population. Parents are denominated A to H. (a) funnel design. (b) diallel or half diallel design

2.3.2 Calculations to find the total number of markers.

We calculated the total number of markers by treating each subgenome under disomic inheritance separately to join the results of both subgenomes at the end. We considered the order of the parents and the crosses for the calculation process, and we did the calculations considering disomic inheritance.

a) Number of different segregations per subgenome in F3

We wanted to know the number of different outputs each subgenome could generate in F3, just before the selfing phase, so we used the combination with repetitions calculation in the 2-way cross, the 4-way cross and the 8-way cross. The combinations with repetition calculation shown below, also named multiset coefficient occur in the negative binomial distribution.

$$C'(n,r) = \frac{(r+n-1)!}{r!\,(n-1)!}$$

Where: r is the sample size and n is the total number of objects.

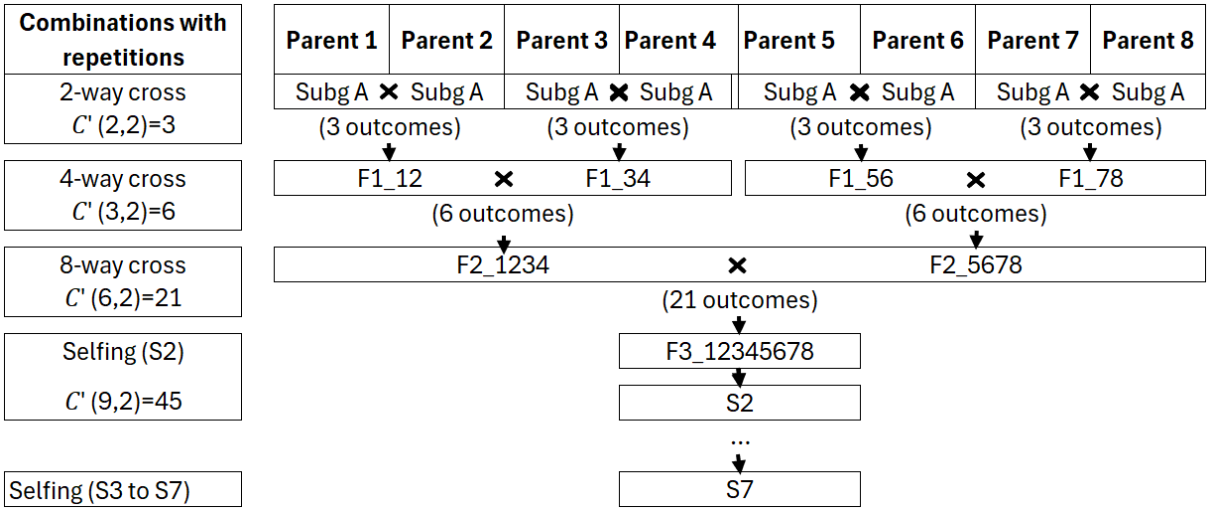

Fig. 6. Possible outcomes for each cross using the Combinations with repetition calculation.

For illustration we will make examples with subgenome 1. The results and nomenclature applied here are also applicable to subgenome 2. We grouped the crosses as in figure 6. First, we crossed P1xP2; P3xP4; P5xP6 and P7xP8. Each 2-way cross had 2 possible parental genotypes (AA, aa) and 2 parents. Therefore, $n = 2$ and $r = 2$. We had 3 possible outcomes for each pair of parents named S, s and d ("S" and "s" for same alleles and "d" for different alleles) as shown in table 1.

Table 1. Possible outcomes for F1 issued from any 2-way cross.

Possible outcomes in F1				
No.	**Subgenome crosses**	**Result subgenome**	**Ratio**	**Result Nomenclature**
1	AAxAA	AA	1	S
2	aaxaa	aa	1	s
3	AAxaa	Aa	1	d

The 2-way cross generated 4 F1's and we grouped them in pairs for the 4-way cross as shown in figure 6. We crossed F1_12 X F1_34 and F1_56 X F1_78. Each of the 4-way crosses had 3 inputs (S, s and d) from the first cross and 2 F1 slots, one for F1_12 X F1_34 and one for F1_56 X F1_78. Therefore, $n = 3$ and $r = 2$, so we had 6 possible outcomes or result nomenclatures for each pair of F1's (table 2).

Table 2. Possible outcomes for F2 issued from any 4-way cross.

Possible outcomes in F2					
No.	**Subgenome crosses**	**Nomenclature crosses**	**Result subgenome**	**Ratio**	**Result Nomenclature**
1	AAxAA	SxS	AA	1	S
2	aaxaa	sxs	aa	1	s
3	AAxaa	Sxs	Aa	1	d
4	AaxAA	dxS	AA:Aa	1:1	C
5	Aaxaa	dxs	aa:Aa	1:1	c
6	AaxAa	dxd	AA:Aa:aa	1:2:1	e

The 4-way cross generated 2 F2's denominated F2_1234 and F2_5678. We crossed them for the 8-way cross. In this case we had the 6 inputs from the previous cross (S, s, d, C, c and e) and 2 F2 slots. Therefore, $n = 6$ and $r = 2$, so we had 21 possible outcomes or result nomenclatures in the 8way-cross (table 3). Figure 7 shows configuration A from figure 5, but this time it shows the result nomenclature for each subgenome.

Table 3. Possible outcomes for F3 issued from the 8-way cross.

Possible outcomes in F3				
No.	**Nomenclature crosses**	**Result subgenome**	**Ratio**	**Result Nomenclature**
1	SxS	AA	1	S
2	sxs	aa	1	s
3	Sxs	Aa	1	d
4	dxS	AA:Aa	1:1	C
5	dxs	aa:Aa	1:1	c
6	dxd	AA:Aa:aa	1:2:1	e
7	dxC	AA:Aa:aa	3:4:1	F
8	dxc	AA:Aa:aa	1:4:3	f
9	SxC	AA:Aa	6:2	G
10	Sxc	AA:Aa	2:6	g
11	sxC	aa:Aa	2:6	H
12	sxc	aa:Aa	6:2	h
13	CxC	AA:Aa:aa	9:6:1	I
14	cxc	AA:Aa:aa	1:6:9	i
15	Cxc	AA:Aa:aa	3:10:3	j
16	exe	AA:Aa:aa	1:2:1	k
17	Sxe	AA:Aa	1:1	L
18	sxe	aa:Aa	1:1	l
19	dxe	AA:Aa:aa	1:2:1	m
20	Cxe	AA:Aa:aa	3:4:1	N
21	cxe	AA:Aa:aa	1:4:3	n

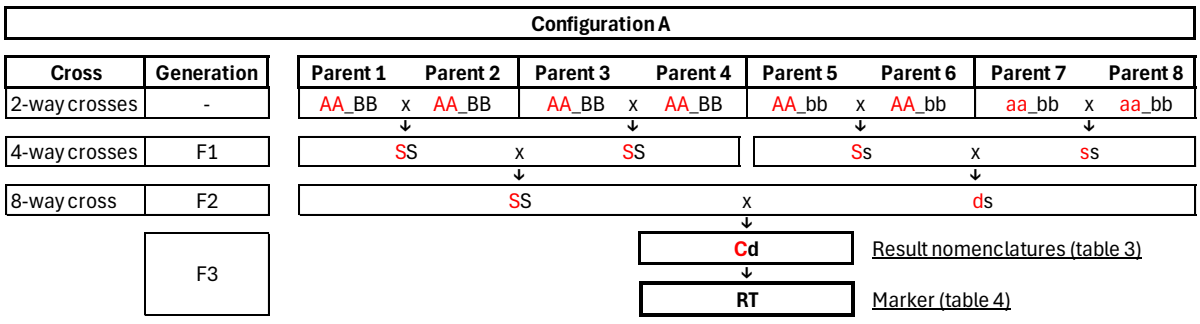

Fig. 7. Result nomenclatures for subgenome 1 (in red) and subgenome 2 (in black) and its respective marker for an example parental configuration.

b) Number of markers under disomic inheritance.

The result nomenclatures from table 3 that generated the same subgenome segregation ratio were grouped and assigned a new letter (table 4). This resulted in 9 groups, each with different subgenome ratios. The 9 resulting groups are also applicable to subgenome 2.

Table 4. Outcomes from F3 with the same expected segregation frequencies in S2 to S7

No.	Nomenclature	Result subgenome	Ratio	Group name
1	S, s	AA or aa	1	Q
2	d	Aa	1	R
3	C, c, L, l	AA:Aa	1	T
4	e, m, k	AA:Aa:aa	1:2:1	U
5	F, f, N, n	AA:Aa:aa	3:4:1 or 1:4:3	V
6	G, h	AA:Aa	6:2	W
7	g, H	aa:Aa	2:6	X
8	I, i	AA:Aa:aa	1:6:9 or 9:6:1	Y
9	j	AA:Aa:aa	3:10:3	Z

We calculated the combinations with repetitions for the 9 groups in table 4, considering both subgenomes for the selfing phase, each subgenome could have any of these 9 group values. Therefore, $n = 9$ and $r = 2$, we had 45 possible outcomes or marker types in the selfing phase of the parental configurations. Figure 7 shows the marker assigned to that configuration using the group names in table 4.

Each of the 45 markers was named after the combination of 2 group names, one for each subgenome. An example of parental configuration for each of the 45 marker types is shown in table 5. The full list of the configurations that belong to the same marker type is shown in supplementary table 1.

Table 5. Examples of parental configurations for each marker type under disomic inheritance

No.	Marker type	Parent 1	Parent 2	Parent 3	Parent 4	Parent 5	Parent 6	Parent 7	Parent 8
1	Marker_QQ	AA_BB	AA_BB	AA_BB	AA_BB	AA_BB	AA_BB	AA_BB	AA_BB
2	Marker_QR	AA_BB	AA_BB	AA_BB	AA_BB	AA_bb	AA_bb	AA_bb	AA_bb
3	Marker_QT	AA_BB	AA_BB	AA_BB	AA_BB	AA_BB	AA_BB	AA_bb	AA_bb
4	Marker_QU	AA_BB	AA_BB	AA_bb	AA_bb	AA_BB	AA_BB	AA_bb	AA_bb
5	Marker_QV	AA_BB	AA_BB	AA_BB	AA_bb	AA_BB	AA_BB	AA_bb	AA_bb
6	Marker_QW	AA_BB	AA_BB	AA_BB	AA_BB	AA_BB	AA_BB	AA_BB	AA_bb
7	Marker_QX	AA_BB	AA_BB	AA_BB	AA_BB	AA_BB	AA_bb	AA_bb	AA_bb
8	Marker_QY	AA_BB	AA_BB	AA_BB	AA_bb	AA_BB	AA_BB	AA_BB	AA_bb
9	Marker_QZ	AA_BB	AA_BB	AA_BB	AA_bb	AA_BB	AA_bb	AA_bb	AA_bb
10	Marker_RR	AA_BB	AA_BB	AA_BB	AA_BB	aa_bb	aa_bb	aa_bb	aa_bb
11	Marker_RT	AA_BB	AA_BB	AA_BB	AA_BB	AA_bb	AA_bb	aa_bb	aa_bb
12	Marker_RU	AA_BB	AA_BB	AA_bb	AA_bb	aa_BB	aa_BB	aa_bb	aa_bb
13	Marker_RV	AA_BB	AA_BB	AA_BB	AA_bb	aa_BB	aa_BB	aa_bb	aa_bb
14	Marker_RW	AA_BB	AA_BB	AA_BB	AA_BB	AA_bb	AA_bb	AA_bb	aa_bb
15	Marker_RX	AA_BB	AA_BB	AA_BB	AA_BB	AA_bb	aa_bb	aa_bb	aa_bb
16	Marker_RY	AA_BB	AA_BB	AA_BB	AA_bb	aa_BB	aa_BB	aa_BB	aa_bb
17	Marker_RZ	AA_BB	AA_BB	AA_BB	AA_bb	aa_BB	aa_bb	aa_bb	aa_bb
18	Marker_TT	AA_BB	AA_BB	AA_BB	AA_BB	AA_BB	AA_bb	aa_BB	aa_bb
19	Marker_TU	AA_BB	AA_BB	AA_bb	AA_bb	AA_BB	AA_bb	aa_BB	aa_bb
20	Marker_TV	AA_BB	AA_BB	AA_BB	AA_bb	AA_BB	AA_bb	aa_BB	aa_bb
21	Marker_TW	AA_BB	AA_BB	AA_BB	AA_BB	AA_BB	AA_bb	AA_bb	aa_BB
22	Marker_TX	AA_BB	AA_BB	AA_BB	AA_BB	AA_bb	AA_bb	aa_BB	aa_bb
23	Marker_TY	AA_BB	AA_BB	AA_BB	AA_bb	AA_BB	AA_BB	aa_BB	aa_bb
24	Marker_TZ	AA_BB	AA_BB	AA_BB	AA_bb	AA_bb	AA_bb	aa_BB	aa_bb
25	Marker_UU	AA_BB	AA_BB	aa_bb	aa_bb	AA_bb	AA_bb	aa_BB	aa_BB
26	Marker_UV	AA_BB	AA_BB	AA_bb	aa_bb	AA_BB	AA_bb	aa_BB	aa_bb
27	Marker_UW	AA_BB	AA_BB	AA_bb	AA_bb	AA_BB	AA_BB	AA_bb	aa_bb
28	Marker_UX	AA_BB	AA_BB	AA_bb	AA_bb	AA_BB	aa_BB	aa_bb	aa_bb
29	Marker_UY	AA_BB	AA_BB	AA_bb	aa_bb	AA_BB	AA_bb	AA_bb	aa_BB
30	Marker_UZ	AA_BB	AA_BB	AA_bb	aa_bb	AA_bb	aa_BB	aa_BB	aa_bb
31	Marker_VV	AA_BB	AA_BB	AA_BB	aa_bb	AA_bb	AA_bb	aa_BB	aa_BB
32	Marker_VW	AA_BB	AA_BB	AA_BB	AA_bb	AA_BB	AA_bb	AA_bb	aa_BB
33	Marker_VX	AA_BB	AA_BB	AA_BB	AA_bb	AA_bb	aa_BB	aa_BB	aa_bb
34	Marker_VY	AA_BB	AA_BB	AA_BB	aa_bb	AA_BB	AA_bb	AA_bb	aa_BB
35	Marker_VZ	AA_BB	AA_BB	AA_BB	aa_bb	AA_bb	AA_bb	aa_BB	aa_bb
36	Marker_WW	AA_BB	AA_BB	AA_BB	AA_BB	AA_BB	AA_BB	AA_bb	aa_BB
37	Marker_WX	AA_BB	AA_BB	AA_BB	AA_BB	AA_bb	AA_bb	AA_bb	aa_BB
38	Marker_WY	AA_BB	AA_BB	AA_BB	AA_bb	AA_BB	AA_BB	AA_bb	aa_BB
39	Marker_WZ	AA_BB	AA_BB	AA_BB	AA_bb	AA_BB	AA_bb	AA_bb	aa_bb
40	Marker_XX	AA_BB	AA_BB	AA_BB	AA_BB	AA_bb	aa_BB	aa_bb	aa_bb
41	Marker_XY	AA_BB	AA_BB	AA_BB	AA_bb	AA_BB	aa_BB	aa_BB	aa_bb
42	Marker_XZ	AA_BB	AA_BB	AA_BB	AA_bb	AA_bb	aa_BB	aa_bb	aa_bb
43	Marker_YY	AA_BB	AA_BB	AA_BB	aa_bb	AA_BB	AA_BB	AA_bb	aa_BB
44	Marker_YZ	AA_BB	AA_BB	AA_BB	aa_bb	AA_bb	AA_bb	AA_bb	aa_BB
45	Marker_ZZ	AA_BB	AA_BB	AA_bb	aa_BB	AA_bb	aa_BB	aa_bb	aa_bb

2.4 Calculation of exact segregation ratios for all the parental configurations under disomic inheritance

We calculated the segregation ratios in S2 and S7 for the 10000 parental configurations. We grouped the parental configurations that generated the same segregation ratios in the selfing phase and compared them to the 45 marker types generated in the previous section. We verified that parental

configurations with the same marker type corresponded to one unique segregation frequency and could be denominated equivalent configurations.

Ratios for each subgenome were calculated by generating the possible gametes in each parent. For example, the parental genotype AA_bb, generated A_b gametes with a ratio of 1 and the parental genotype aa_BB generated a_B gametes with a ratio of 1, when the gametes were crossed, we obtained genotype Aa_Bb as result. The resulting genotype would generate gametes A_B, A_b, a_B and a_b in a 1:1:1:1 ratio. When crossed for example, with genotype aa_BB, the resulting genotypes were Aa_BB, Aa_Bb, aa_BB and aa_Bb 1:1:1:1 ratio (Fig. 6).

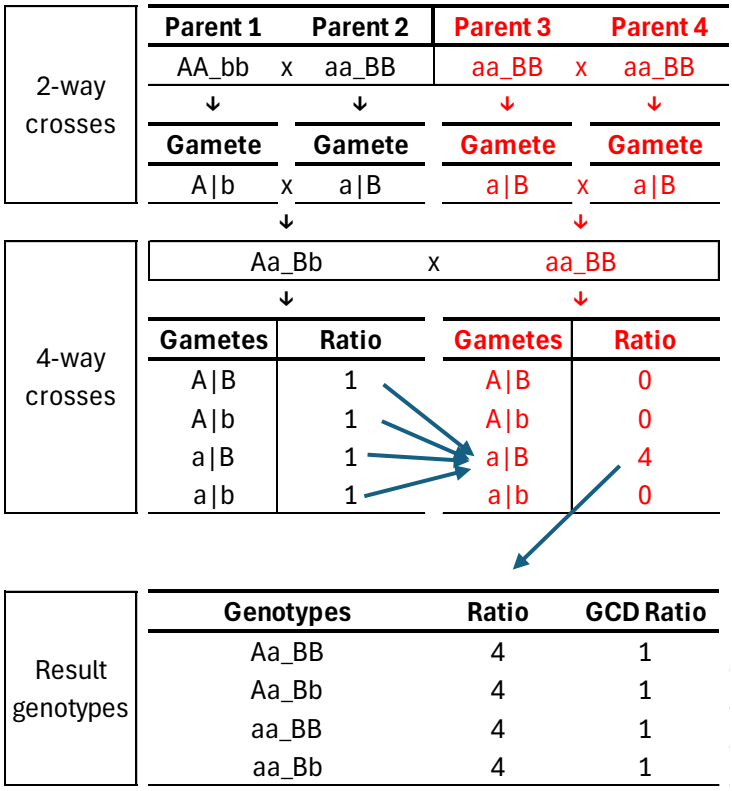

Fig. 6. Representation of the gamete and genotype ratio calculation for each cross.

2.4.1 Definition of equivalent configuration

We considered as equivalent all the configurations that generated the same segregation frequencies in heterozygous genotypes of the progeny, and the homozygous genotypes of the progeny.

For example, parental configurations A and B in figure 7 generated the same segregation frequencies in S2 although they have different parental genotypes. The heterozygous genotypes have the same frequencies between configurations, as well as the homozygous genotypes. Therefore, these two configurations are considered equivalent.

Parental configuration A

P1	P2	P3	P4	P5	P6	P7	P8
AA_BB	AA_BB	AA_BB	AA_BB	AA_bb	AA_bb	aa_bb	aa_bb

↓

S2 segregation frequencies

Genotype	Frequency	Ratio
AA_BB	0.15625	5
AA_bB	0.3125	10
AA_bb	0.15625	5
aA_BB	0.0625	2
aA_bB	0.125	4
aA_bb	0.0625	2
aa_BB	0.03125	1
aa_bB	0.0625	2
aa_bb	0.03125	1

Parental configuration B

P1	P2	P3	P4	P5	P6	P7	P8
aa_BB	aa_BB	aa_bb	aa_bb	AA_bb	AA_bb	AA_bb	AA_bb

↓

S2 segregation frequencies

Genotype	Frequency	Ratio
AA_BB	0.03125	1
AA_bB	0.0625	2
AA_bb	0.15625	5
aA_BB	0.0625	2
aA_bB	0.125	4
aA_bb	0.3125	10
aa_BB	0.03125	1
aa_bB	0.0625	2
aa_bb	0.15625	5

Parental configuration C

P1	P2	P3	P4	P5	P6	P7	P8
AA_BB	AA_BB	AA_BB	AA_bb	AA_BB	AA_bb	aa_BB	aa_bb

↓

S2 segregation frequencies

Genotype	Frequency	Ratio
AA_BB	0.3125	10
AA_bB	0.15625	5
AA_bb	0.15625	5
aA_BB	0.125	4
aA_bB	0.0625	2
aA_bb	0.0625	2
aa_BB	0.0625	2
aa_bB	0.03125	1
aa_bb	0.03125	1

Fig 7. Parental configurations and their respective segregation frequencies in S2. Homozygous genotypes from S2 are in red and heterozygous genotypes in black.

On the other hand, parental configurations A and C in figure 7 also generated the same segregation frequencies in S2, but the heterozygous genotypes have different frequencies between configurations as well as the homozygous genotypes. These two configurations are not considered equivalent because the segregation of alleles in further selfings will be different.

2.5 Markers under tetrasomic inheritance

Each configuration from the disomic scenario was transformed into its tetrasomic form using the nomenclature described. The tetrasomic forms of each parental configuration were also 8-way crossed and selfed from S2 to S7 as previously described. The segregation frequencies in the S2 and S7 from the disomic and tetrasomic form of the same configuration were compared. Table 6 shows an example of tetrasomic parental configuration per marker. Supplementary table 4 shows the full list of parental configurations under tetrasomic inheritance.

Table 6. Examples of parental configurations for each marker type under tetrasomic inheritance.

No.	Marker type	Parent 1	Parent 2	Parent 3	Parent 4	Parent 5	Parent 6	Parent 7	Parent 8
1	Marker_QQ	AAAA	AAAA	AAAA	AAAA	AAAA	AAAA	AAAA	AAAA
2	Marker_QR	AAAA	AAAA	AAAA	AAAA	AAaa	AAaa	AAaa	AAaa
3	Marker_QT	AAAA	AAAA	AAAA	AAAA	AAAA	AAAA	AAaa	AAaa
4	Marker_QU	AAAA	AAAA	AAaa	AAaa	AAAA	AAAA	AAaa	AAaa
5	Marker_QV	AAAA	AAAA	AAAA	AAaa	AAAA	AAAA	AAaa	AAaa
6	Marker_QW	AAAA	AAAA	AAAA	AAAA	AAAA	AAAA	AAAA	AAaa
7	Marker_QX	AAAA	AAAA	AAAA	AAAA	AAAA	AAaa	AAaa	AAaa
8	Marker_QY	AAAA	AAAA	AAAA	AAaa	AAAA	AAAA	AAAA	AAaa
9	Marker_QZ	AAAA	AAAA	AAAA	AAaa	AAAA	AAaa	AAaa	AAaa
10	Marker_RR	AAAA	AAAA	AAAA	AAAA	aaaa	aaaa	aaaa	aaaa
11	Marker_RT	AAAA	AAAA	AAAA	AAAA	AAaa	AAaa	aaaa	aaaa
12	Marker_RU	AAAA	AAAA	AAaa	AAaa	AAaa	AAaa	aaaa	aaaa
13	Marker_RV	AAAA	AAAA	AAAA	AAaa	AAaa	AAaa	aaaa	aaaa
14	Marker_RW	AAAA	AAAA	AAAA	AAAA	AAaa	AAaa	AAaa	aaaa
15	Marker_RX	AAAA	AAAA	AAAA	AAAA	AAaa	aaaa	aaaa	aaaa
16	Marker_RY	AAAA	AAAA	AAAA	AAaa	AAaa	AAaa	AAaa	aaaa
17	Marker_RZ	AAAA	AAAA	AAAA	AAaa	AAaa	aaaa	aaaa	aaaa
18	Marker_TT	AAAA	AAAA	AAAA	AAAA	AAAA	AAaa	AAaa	aaaa
19	Marker_TU	AAAA	AAAA	AAaa	AAaa	AAAA	AAaa	AAaa	aaaa
20	Marker_TV	AAAA	AAAA	AAAA	AAaa	AAAA	AAaa	AAaa	aaaa
21	Marker_TW	AAAA	AAAA	AAAA	AAAA	AAAA	AAaa	AAaa	AAaa
22	Marker_TX	AAAA	AAAA	AAAA	AAAA	AAaa	AAaa	AAaa	aaaa
23	Marker_TY	AAAA	AAAA	AAAA	AAaa	AAAA	AAAA	AAaa	aaaa
24	Marker_TZ	AAAA	AAAA	AAAA	AAaa	AAaa	AAaa	AAaa	aaaa
25	Marker_UU	AAAA	AAAA	aaaa	aaaa	AAaa	AAaa	AAaa	AAaa
26	Marker_UV	AAAA	AAAA	AAaa	aaaa	AAAA	AAaa	AAaa	aaaa
27	Marker_UW	AAAA	AAAA	AAaa	AAaa	AAAA	AAAA	AAaa	aaaa
28	Marker_UX	AAAA	AAAA	AAaa	AAaa	AAAA	AAaa	aaaa	aaaa
29	Marker_UY	AAAA	AAAA	AAaa	aaaa	AAAA	AAaa	AAaa	AAaa
30	Marker_UZ	AAAA	AAAA	AAaa	aaaa	AAaa	AAaa	AAaa	aaaa
31	Marker_VV	AAAA	AAAA	AAAA	aaaa	AAaa	AAaa	AAaa	AAaa
32	Marker_VW	AAAA	AAAA	AAAA	AAaa	AAAA	AAaa	AAaa	AAaa
33	Marker_VX	AAAA	AAAA	AAAA	AAaa	AAaa	AAaa	AAaa	aaaa
34	Marker_VY	AAAA	AAAA	AAAA	aaaa	AAAA	AAaa	AAaa	AAaa
35	Marker_VZ	AAAA	AAAA	AAAA	aaaa	AAaa	AAaa	AAaa	aaaa
36	Marker_WW	AAAA	AAAA	AAAA	AAAA	AAAA	AAAA	AAaa	AAaa
37	Marker_WX	AAAA	AAAA	AAAA	AAAA	AAaa	AAaa	AAaa	AAaa
38	Marker_WY	AAAA	AAAA	AAAA	AAaa	AAAA	AAAA	AAaa	AAaa
39	Marker_WZ	AAAA	AAAA	AAAA	AAaa	AAAA	AAaa	AAaa	aaaa
40	Marker_XX	AAAA	AAAA	AAAA	AAAA	AAaa	AAaa	aaaa	aaaa
41	Marker_XY	AAAA	AAAA	AAAA	AAaa	AAAA	AAaa	AAaa	aaaa
42	Marker_XZ	AAAA	AAAA	AAAA	AAaa	AAaa	AAaa	aaaa	aaaa
43	Marker_YY	AAAA	AAAA	AAAA	aaaa	AAAA	AAAA	AAaa	AAaa
44	Marker_YZ	AAAA	AAAA	AAAA	aaaa	AAaa	AAaa	AAaa	AAaa
45	Marker_ZZ	AAAA	AAAA	AAaa	AAaa	AAaa	AAaa	aaaa	aaaa

3. RESULTS

3.1 Calculation of segregation frequencies for all the parental configurations under disomic inheritance

We obtained 36 different segregation ratios in the S2 under disomic inheritance instead of 45. The apparent difference between the expected and obtained segregation ratios in the S2 is due to a coincidence of segregation when selfing the R and the V group, where both generate the same ratios. However, on further inspection, these groups don't have the same segregation ratios in the heterozygous genotypes, 2 examples are shown in figure 8. The groups of markers with their respective segregation ratios under disomic inheritance are shown in the supplementary table 2.

When the selfing continues from S3 to S7 the segregation ratios are 45, 1 for each marker type.

Parental configuration 35

P1	P2	P3	P4	P5	P6	P7	P8
AA_BB	AA_BB	AA_BB	AA_BB	aa_bb	aa_bb	aa_bb	aa_bb

↓

Result subgenomes: dd

Marker: RR

S2 segregation frequencies		
Genotype	Frequency	Ratio
AA_BB	0.0625	1
AA_bB	0.125	2
AA_bb	0.0625	1
aA_BB	0.125	2
aA_bB	0.25	4
aA_bb	0.125	2
aa_BB	0.0625	1
aa_bB	0.125	2
aa_bb	0.0625	1

Parental configuration 68

P1	P2	P3	P4	P5	P6	P7	P8
AA_BB	AA_BB	AA_BB	AA_bb	aa_BB	aa_BB	aa_bb	aa_bb

↓

Result subgenomes: dF

Marker: RV

S2 segregation frequencies		
Genotype	Frequency	Ratio
AA_BB	0.125	2
AA_bB	0.0625	1
AA_bb	0.0625	1
aA_BB	0.25	4
aA_bB	0.125	2
aA_bb	0.125	2
aa_BB	0.125	2
aa_bB	0.0625	1
aa_bb	0.0625	1

Fig 8. Groups of markers with apparent same segregation in S2. The ratios are mixed between homozygotes (red) and heterozygotes (black).

3.2 Calculation of segregation frequencies for all the parental configurations under tetrasomic inheritance

We transformed every configuration under disomic inheritance to its tetrasomic form and calculated the expected segregation frequencies for the S2 and S7 for each configuration. We found that configurations that belong to the same marker type under disomic inheritance can show different

segregations under tetrasomic inheritance in all the selfing generations as shown in table 7. The complete set of tetrasomic markers with their segregation frequencies is shown in supplementary table 3.

Table 7. Segregation ratios for Marker QQ under a disomic inheritance and under tetrasomic inheritance. Note that under tetrasomic inheritance, the marker segregates in some cases and under a disomic inheritance it doesn't segregate at all.

Inheritance scenario	Genotype	S2				S7			
		Segregation 1		Segregation 2		Segregation 1		Segregation 2	
		Frequency	Ratio	Frequency	Ratio	Frequency	Ratio	Frequency	Ratio
Disomic inheritance	AA_BB	1	1	-	-	1	1	-	-
Tetrasomic inheritance	aaaa	-	-	0.13105853	2293	-	-	0.35167329	47844763
	aaaA	-	-	0.21033379	3680	-	-	0.08475806	11531240
	aaAA	-	-	0.31721536	5550	-	-	0.12713731	17296890
	aAAA	-	-	0.21033379	3680	-	-	0.08475806	11531240
	AAAA	1	1	0.13105853	2293	1	1	0.35167329	47844763

3.3 Usefulness of markers

From a practical perspective, only homozygous genotypes (AA_BB, aa_bb, AA_bb and aa_BB) can be identified accurately under disomic inheritance. The other genotypes are mixed because we don't know which alleles correspond to subgenome 1 and subgenome 2. For instance, AA_Bb, Aa_BB, aa_Bb and Aa_bb will be detected as triplex or simple markers (AAAa or Aaaa if we can't recognize which allele belongs to each subgenome).

Markers that have one "Q" subgenome (QR, QT, QU, QW, QX, QY, QZ) will segregate into 3 genotypes under a disomic scenario, 2 homozygous and 1 heterozygous genotype. For example, Marker QU displays the following segregation ratios in S2 and S7.

Table 8. Marker QU segregation ratios for S2 and S7 under disomic inheritance

Marker type	Genotype	S2		S7	
		Frequency	Ratio	Frequency	Ratio
Marker_QU	AA_BB	0.375	3	0.49609375	127
	AA_bB	0.25	2	0.0078125	2
	AA_bb	0.375	3	0.49609375	127

The only marker that has two "Q" (QQ) subgenomes is displayed in table 7. This specific marker can be very useful because it doesn't segregate under disomic inheritance, but it does in some cases under tetrasomic inheritance. Some markers under disomic inheritance, when transformed to their tetrasomic form can generate 2 different segregation ratios as shown in table 7.

We classified the makrers by degree of usefulness.

(1) Marker QQ could be very useful, because it doesn't segregate under disomic inheritance, but it does segregate in some cases under tetrasomic inheritance as shown in table 7.

(2) The second group of markers are the ones that generate only 3 possible genotypes under disomic inheritance (those with one "Q" subgenome) and only have 1 case of segregation ratio under tetrasomic inheritance (table 9). This includes markers QR, QU, QZ.

Table 9. Marker QU segregation ratios under a disomic inheritance and under a tetrasomic inheritance.

Inheritance scenario	Genotype	S2				S7			
		Segregation 1		Segregation 2		Segregation 1		Segregation 2	
		Frequency	Ratio	Frequency	Ratio	Frequency	Ratio	Frequency	Ratio
Disomic inheritance	AA_BB	0.375	3			0.49609375	127		
	AA_Bb	0.25	2			0.0078125	2		
	AA_bb	0.375	3			0.49609375	127		
Tetrasomic inheritance	aaaa	0.015403521	539			0.13680074	18611590		
	aaaA	0.064929127	2272			0.06230936	8477120		
	aaAA	0.232853224	8148			0.09845224	13394319		
	aAAA	0.27789209	9724			0.06896446	9382538		
	AAAA	0.408922039	14309			0.6334732	86183329		

(3) The third group of markers by usefulness is composed by markers that generate only 3 possible genotypes under disomic inheritance (those with one "Q" subgenome) and have 2 cases of segregation ratio under tetrasomic inheritance (table 10). This includes markers QT, QV, QW, QX, QY.

Table 10. Marker QT segregation ratios under a disomic inheritance and under a tetrasomic inheritance.

Inheritance scenario	Genotype	S2				S7			
		Segregation 1		Segregation 2		Segregation 1		Segregation 2	
		Frequency	Ratio	Frequency	Ratio	Frequency	Ratio	Frequency	Ratio
Disomic inheritance	AA_BB	0.625	5			0.74609375	191		
	AA_Bb	0.25	2			0.0078125	2		
	AA_bb	0.125	1			0.24609375	63		
Tetrasomic inheritance	aaaa	0.001028807	2	0.04932556	863	0.05702206	574651	0.23220813	63183319
	aaaA	0.008230453	16	0.14066072	2461	0.03646945	367528	0.08014871	21808286
	aaAA	0.125	243	0.30521262	5340	0.05968944	601532	0.12326087	33539010
	aAAA	0.221193416	430	0.27029035	4729	0.04312454	434596	0.08419963	22910534
	AAAA	0.644547325	1253	0.23451075	4103	0.80369452	8099389	0.48018267	130656643

(4) The fourth group is composed by the markers that generate 9 possible genotypes under disomic inheritance but only 1 case of segregation ratio under tetrasomic inheritance (table 11). This includes markers RU, RY, TU, TZ, UY, UZ.

Table 11. Marker RU segregation ratios under a disomic inheritance and under a tetrasomic inheritance.

Inheritance scenario	Genotype	S2				S7			
		Segregation 1		Segregation 2		Segregation 1		Segregation 2	
		Frequency	Ratio	Frequency	Ratio	Frequency	Ratio	Frequency	Ratio
Disomic inheritance	AA_BB	0.09375	3			0.24417114	8001		
	AA_bB	0.0625	2			0.00384521	126		
	AA_bb	0.09375	3			0.24417114	8001		
	aA_BB	0.1875	6			0.00775146	254		
	aA_bB	0.125	4			0.00012207	4		
	aA_bb	0.1875	6			0.00775146	254		
	aa_BB	0.09375	3			0.24417114	8001		
	aa_bB	0.0625	2			0.00384521	126		
	aa_bb	0.09375	3			0.24417114	8001		
Tetrasomic inheritance	aaaa	0.094107225	3293			0.33629813	45752989		
	aaaA	0.226966164	7942			0.09354329	12726461		
	aaAA	0.357853224	12522			0.14031717	19089996		
	aAAA	0.226966164	7942			0.09354329	12726461		
	AAAA	0.094107225	3293			0.33629813	45752989		

(5) Finally, the least useful group generates 9 possible genotypes under disomic inheritance and 2 cases segregation ratio under tetrasomic inheritance (table 12). The markers that are not listed above are part of this group.

Table 12. Marker RR segregation ratios under a disomic inheritance and under a tetrasomic inheritance.

Inheritance scenario	Genotype	S2				S7			
		Segregation 1		Segregation 2		Segregation 1		Segregation 2	
		Frequency	Ratio	Frequency	Ratio	Frequency	Ratio	Frequency	Ratio
Disomic inheritance	AA_BB	0.0625	1			0.24224854	3969		
	AA_bB	0.125	2			0.00769043	126		
	AA_bb	0.0625	1			0.24224854	3969		
	aA_BB	0.125	2			0.00769043	126		
	aA_bB	0.25	4			0.00024414	4		
	aA_bb	0.125	2			0.00769043	126		
	aa_BB	0.0625	1			0.24224854	3969		
	aa_bB	0.125	2			0.00769043	126		
	aa_bb	0.0625	1			0.24224854	3969		
Tetrasomic inheritance	aaaa	0.027777778	1	0.13105853	2293	0.30464463	28427	0.35167329	47844763
	aaaA	0.222222222	8	0.21033379	3680	0.11162551	10416	0.08475806	11531240
	aaAA	0.5	18	0.31721536	5550	0.16745971	15626	0.12713731	17296890
	aAAA	0.222222222	8	0.21033379	3680	0.11162551	10416	0.08475806	11531240
	AAAA	0.027777778	1	0.13105853	2293	0.30464463	28427	0.35167329	47844763

4. DISCUSSION

R packages and software used to generate linkage maps in MAGIC populations of diploid organisms like R/mpMap (B. Emma Huang & George, 2011) and GAPL (L. Zhang et al., 2019), don't take into account the order of the 2-way crosses to define the markers of the population. In this study we demonstrated that 8-way MAGIC populations of tetraploid organisms under a disomic inheritance generate different exact segregation ratios depending on the order of the 2-way crosses (Supplementary tables 1 and 2). Therefore, parental configurations that have the same parental genotypes but have 2-way crosses organized in different order can result in different segregation ratios and must be considered as separated markers.

Tetrasomic recombination has been reported in allotetraploid *Arachis* (S. Leal-Bertioli et al., 2015; Nguepjop et al., 2016). Although these were observations in crosses between cultivated peanut *Arachis hypogaea* and synthetic allotetraploids containing the same subgenomes as cultivated peanut, tetrasomic recombination is likely to happen in crosses of cultivated peanut. Preliminary results of segregation in S2 of an 8-way cross MAGIC population of cultivated peanut show tetrasomic recombination in entire regions of chromosomes 12 and 14 of peanut (Liao, unpublished data). Two example locus with tetrasomic inheritance are shown in figure 9, one for chromosome 12 and one 14.

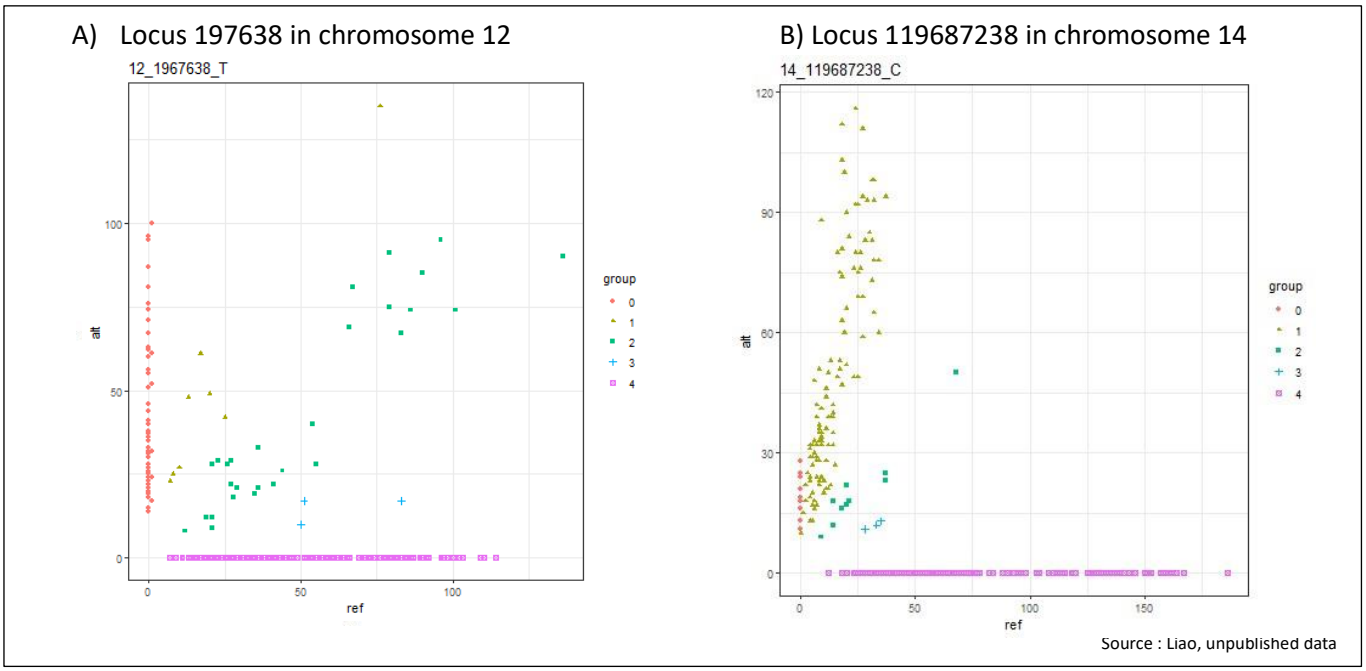

Fig. 9. S2 genotype data from chromosome 12 and 14 of cultivated peanut that corresponds to a tetrasomic segregation.

Although the ratios generated for each marker in this study are exact, they can be useful as a guide for the expected ratios from an 8-way cross MAGIC population of peanuts. The exact segregation ratio for each of the 45 markers can be validated with simulated data from software like PedigreeSim (Voorrips & Maliepaard, 2012) which can simulate meiosis of tetraploids in bivalent formation (disomic inheritance) and quadrivalent formation (tetrasomic inheritance). Experimental data could also be used to validate the exact segregation ratios if available.

Every marker has at most 3 possible exact segregation ratios; 1 under disomic inheritance and 1 to 2 under tetrasomic inheritance (tables 7 to 12). The simulated or experimental data would display segregation ratios at each marker locus that can be compared to the exact segregation ratios from this study. The goodness of fit can be calculated to identify if the locus followed a tetrasomic inheritance or a disomic inheritance. Loci that display segregation ratios that cannot be related to any of the exact segregation ratios can be labeled as unidentified.

Identity by descent (IBD) could also be inferred from the comparison between simulated or experimental data and the exact segregation ratios generated in this study. This could be useful to calculate the preferential pairing parameter (Bourke et al., 2017) in each chromosome to determine its level of disomic or tetrasomic inheritance. The results can be compared to the regions displaying tetrasomic inheritance in crosses between *Arachis hypogaea* and synthetic allotetraploids described in other studies (S. Leal-Bertioli et al., 2015; Nguepjop et al., 2016).

The identification of the preferential pairing parameter in each chromosome or the segregation behavior in each locus could be useful to build a linkage map in peanut where tetrasomic and disomic regions or chromosomes are identified. This would improve the accuracy of genetic prediction by identifying if the markers used in a linkage map are in a mostly disomic region or tetrasomic region.

The most useful marker found in this study is the Marker QQ in the specific parental configurations shown in table 13. These configurations do not segregate under disomic inheritance but do segregate under tetrasomic inheritance. Finding these marker with this particular configurations in a locus could determine very easily if it behaves in a tetrasomic or disomic inheritance.

Table 13. Parental configurations of Marker QQ that don't segregate under disomic inheritance but segregate under tetrasomic inheritance.

Inheritance	Parent_1	Parent_2	Parent_3	Parent_4	Parent_5	Parent_6	Parent_7	Parent_8
Disomic	AA_bb aa_BB	AA_bb aa_BB	AA_bb aa_BB	AA_bb aa_BB	AA_bb aa_BB	AA_bb aa_BB	AA_bb aa_BB	AA_bb aa_BB
Tetrasomic	AAaa	AAaa	AAaa	AAaa	AAaa	AAaa	AAaa	AAaa

5. CONCLUSIONS

In this study we showed that the order of the 2-way crosses influences the segregation of the progeny in an 8-way cross MAGIC population of an allotetraploid organism with disomic inheritance.

Equivalent parental configurations for an 8-way cross under disomic inheritance are not necessarily equivalent under tetrasomic inheritance.

There are different degrees in usefulness of the 45 markers, depending on their segregation behavior under disomic and tetrasomic inheritance.

(1) 1 out of the 45 markers doesn't segregate under a disomic inheritance but it segregates in half of the parental configurations under tetrasomic inheritance.
(2) 3 out of the 45 markers are the simplest to use, because they display a segregation ratio under disomic inheritance with only 3 genotypes and 1 case of segregation ratio under tetrasomic inheritance.
(3) 5 out of the 45 markers also display a segregation ratio under disomic inheritance with only 3 genotypes but 2 cases of segregation ratio under tetrasomic inheritance, making these markers more complex to use.
(4) 6 out of the 45 markers display a segregation ratio under disomic inheritance with 9 genotypes, where the heterozygotes AA_Bb, Aa_BB, aa_Bb and Aa_bb cannot display specific ratio segregations in a practical situation, because we are not able to know which alleles are part of subgenome 1 or subgenome 2. They display 1 case of segregation ratio under tetrasomic inheritance.
(5) 31 out of 45 markers display a segregation ratio under disomic inheritance with the same characteristics in (3), but they display 2 cases of segregation ratio under tetrasomic inheritance.

References

Arrones, A., Vilanova, S., Plazas, M., Mangino, G., Pascual, L., Díez, M. J., Prohens, J., & Gramazio, P. (2020). The dawn of the age of multi-parent magic populations in plant breeding: Novel powerful next-generation resources for genetic analysis and selection of recombinant elite material. *Biology*, *9*(8), 1–25. https://doi.org/10.3390/biology9080229

Balaiah, C., Reddy, P. S., & Reddi, M. V. (1977). Genic analysis in groundnut. *Proceedings / Indian Academy of Sciences*, *85*(5), 340–350. https://doi.org/10.1007/bf03052386

Bandillo, N., Raghavan, C., Muyco, P. A., Sevilla, M. A. L., Lobina, I. T., Dilla-Ermita, C. J., Tung, C. W., McCouch, S., Thomson, M., Mauleon, R., Singh, R. K., Gregorio, G., Redoña, E., & Leung, H. (2013). Multi-parent advanced generation inter-cross (MAGIC) populations in rice: Progress and potential for genetics research and breeding. *Rice*, *6*(1), 1–15. https://doi.org/10.1186/1939-8433-6-1

Bauer, K., Sy, J., & Lipmann, F. (1973). *Breeding Peanuts (Arachis hypogaea L.) for resistance to Verticillium wiltBREEDING PEANUTS (ARACHIS HYPOGAEA L.) FOR RESISTANCE TO VERTICILLIUM WILT*. *21*(July), 1154–1157.

Bertioli, D. J., Jenkins, J., Clevenger, J., Dudchenko, O., Gao, D., Seijo, G., Leal-Bertioli, S. C. M., Ren, L., Farmer, A. D., Pandey, M. K., Samoluk, S. S., Abernathy, B., Agarwal, G., Ballén-Taborda, C., Cameron, C., Campbell, J., Chavarro, C., Chitikineni, A., Chu, Y., … Schmutz, J. (2019). The genome sequence of segmental allotetraploid peanut Arachis hypogaea. *Nature Genetics*, *51*(5), 877–884. https://doi.org/10.1038/s41588-019-0405-z

Bertioli, D. J., Ozias-Akins, P., Chu, Y., Dantas, K. M., Santos, S. P., Gouvea, E., Guimarães, P. M., Leal-Bertioli, S. C. M., Knapp, S. J., & Moretzsohn, M. C. (2014). The use of SNP markers for linkage mapping in diploid and tetraploid peanuts. *G3: Genes, Genomes, Genetics*, *4*(1), 89–96. https://doi.org/10.1534/g3.113.007617

Bourke, P. M., Arens, P., Voorrips, R. E., Esselink, G. D., Koning-Boucoiran, C. F. S., van't Westende, W. P. C., Santos Leonardo, T., Wissink, P., Zheng, C., van Geest, G., Visser, R. G. F., Krens, F. A., Smulders, M. J. M., & Maliepaard, C. (2017). Partial preferential chromosome pairing is genotype dependent in tetraploid rose. *Plant Journal*, *90*(2), 330–343. https://doi.org/10.1111/tpj.13496

Bourke, P. M., Voorrips, R. E., Kranenburg, T., Jansen, J., Visser, R. G. F., & Maliepaard, C. (2016). Integrating haplotype-specific linkage maps in tetraploid species using SNP markers. *Theoretical and Applied Genetics*, *129*(11), 2211–2226. https://doi.org/10.1007/s00122-016-2768-1

Bourke, P. M., Voorrips, R. E., Visser, R. G. F., & Maliepaard, C. (2015). The double-reduction landscape in tetraploid potato as revealed by a high-density linkage map. *Genetics*, *201*(3), 853–863. https://doi.org/10.1534/genetics.115.181008

Bourke, P. M., Voorrips, R. E., Visser, R. G. F., & Maliepaard, C. (2018). Tools for genetic studies in experimental populations of polyploids. *Frontiers in Plant Science*, *9*(April), 1–17. https://doi.org/10.3389/fpls.2018.00513

Branch, W. D. (2011). First 100 years - inheritance of testa color in peanut (Arachis hypogaea L.). *Crop Science*, *51*(1), 1–4. https://doi.org/10.2135/cropsci2010.06.0349

Branch, William D. (2008). Inheritance of a one-seeded pod trait in peanut. *Journal of Heredity*, *99*(2), 221–222. https://doi.org/10.1093/jhered/esm121

Broman, K. W. (2005). The genomes of recombinant inbred lines. *Genetics*, *169*(2), 1133–1146.

https://doi.org/10.1534/genetics.104.035212

Browning, B. L., Tian, X., Zhou, Y., & Browning, S. R. (2021). Fast two-stage phasing of large-scale sequence data. *American Journal of Human Genetics*, *108*(10), 1880–1890. https://doi.org/10.1016/j.ajhg.2021.08.005

Byers, R. L., Harker, D. B., Yourstone, S. M., Maughan, P. J., & Udall, J. A. (2012). Development and mapping of SNP assays in allotetraploid cotton. *Theoretical and Applied Genetics*, *124*(7), 1201–1214. https://doi.org/10.1007/s00122-011-1780-8

Cai, G., Yang, Q., Yi, B., Fan, C., Edwards, D., Batley, J., & Zhou, Y. (2014). A complex recombination pattern in the genome of allotetraploid Brassica napusas revealed by a high-density genetic map. *PLoS ONE*, *9*(10). https://doi.org/10.1371/journal.pone.0109910

Cavanagh, C., Morell, M., Mackay, I., & Powell, W. (2008). From mutations to MAGIC: resources for gene discovery, validation and delivery in crop plants. *Current Opinion in Plant Biology*, *11*(2), 215–221. https://doi.org/10.1016/j.pbi.2008.01.002

Cavanagh, C. R., Chao, S., Wang, S., Huang, B. E., Stephen, S., Kiani, S., Forrest, K., Saintenac, C., Brown-Guedira, G. L., Akhunova, A., See, D., Bai, G., Pumphrey, M., Tomar, L., Wong, D., Kong, S., Reynolds, M., Da Silva, M. L., Bockelman, H., … Akhunov, E. (2013). Genome-wide comparative diversity uncovers multiple targets of selection for improvement in hexaploid wheat landraces and cultivars. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(20), 8057–8062. https://doi.org/10.1073/pnas.1217133110

Clarke, W. E., Higgins, E. E., Plieske, J., Wieseke, R., Sidebottom, C., Khedikar, Y., Batley, J., Edwards, D., Meng, J., Li, R., Lawley, C. T., Pauquet, J., Laga, B., Cheung, W., Iniguez-Luy, F., Dyrszka, E., Rae, S., Stich, B., Snowdon, R. J., … Parkin, I. A. P. (2016). A high-density SNP genotyping array for Brassica napus and its ancestral diploid species based on optimised selection of single-locus markers in the allotetraploid genome. *Theoretical and Applied Genetics*, *129*(10), 1887–1899. https://doi.org/10.1007/s00122-016-2746-7

Clevenger, J., Chu, Y., Chavarro, C., Agarwal, G., Bertioli, D. J., Leal-Bertioli, S. C. M., Pandey, M. K., Vaughn, J., Abernathy, B., Barkley, N. A., Hovav, R., Burow, M., Nayak, S. N., Chitikineni, A., Isleib, T. G., Holbrook, C. C., Jackson, S. A., Varshney, R. K., & Ozias-Akins, P. (2017). Genome-wide SNP Genotyping Resolves Signatures of Selection and Tetrasomic Recombination in Peanut. *Molecular Plant*, *10*(2), 309–322. https://doi.org/10.1016/j.molp.2016.11.015

Cobos, M. J., Rubio, J., Strange, R. N., Moreno, M. T., Gil, J., & Millan, T. (2006). A new QTL for

Ascochyta blight resistance in an RIL population derived from an interspecific cross in chickpea. *Euphytica*, *149*(1–2), 105–111. https://doi.org/10.1007/s10681-005-9058-3

Darvasi, A., & Soller, M. (1995). Advanced intercross lines, an experimental population for fine genetic mapping. *Genetics*, *141*(3), 1199–1207. https://doi.org/10.1093/genetics/141.3.1199

Dell'Acqua, M., Gatti, D. M., Pea, G., Cattonaro, F., Coppens, F., Magris, G., Hlaing, A. L., Aung, H. H., Nelissen, H., Baute, J., Frascaroli, E., Churchill, G. A., Inzé, D., Morgante, M., & Pè, M. E. (2015). Genetic properties of the MAGIC maize population: A new platform for high definition QTL mapping in Zea mays. *Genome Biology*, *16*(1), 1–23. https://doi.org/10.1186/s13059-015-0716-z

Eskandari, M., Cober, E. R., & Rajcan, I. (2013). Genetic control of soybean seed oil: I. QTL and genes associated with seed oil concentration in RIL populations derived from crossing moderately high-oil parents. *Theoretical and Applied Genetics*, *126*(2), 483–495. https://doi.org/10.1007/s00122-012-1995-3

Essomba, N. B., Coffelt, T. A., Branch, W. D., & Van Scoyoc, S. W. (1993). Inheritance of Leaflet Size in Peanut (Arachis hypogaea L.)1 . *Peanut Science*, *20*(2), 90–93. https://doi.org/10.3146/i0095-3679-20-2-6

Hackett, C. A., McLean, K., & Bryan, G. J. (2013). Linkage Analysis and QTL Mapping Using SNP Dosage Data in a Tetraploid Potato Mapping Population. *PLoS ONE*, *8*(5). https://doi.org/10.1371/journal.pone.0063939

Hall, D., Tegström, C., & Ingvarsson, P. K. (2010). Using association mapping to dissect the genetic basis of complex traits in plants. *Briefings in Functional Genomics and Proteomics*, *9*(2), 157–165. https://doi.org/10.1093/bfgp/elp048

Harlan, J. R., & deWet, J. M. (1975). On Ö. Winge and a prayer: The origins of polyploidy. In *The Botanical Review* (Vol. 41, Issue 4, pp. 361–390).

Huang, B. Emma, & George, A. W. (2011). R/mpMap: A computational platform for the genetic analysis of multiparent recombinant inbred lines. *Bioinformatics*, *27*(5), 727–729. https://doi.org/10.1093/bioinformatics/btq719

Huang, B. Emma, Verbyla, K. L., Verbyla, A. P., Raghavan, C., Singh, V. K., Gaur, P., Leung, H., Varshney, R. K., & Cavanagh, C. R. (2015). MAGIC populations in crops: current status and future prospects. *Theoretical and Applied Genetics*, *128*(6), 999–1017. https://doi.org/10.1007/s00122-015-2506-0

Huang, Bevan E., George, A. W., Forrest, K. L., Kilian, A., Hayden, M. J., Morell, M. K., & Cavanagh, C. R. (2012). A multiparent advanced generation inter-cross population for genetic analysis in wheat. *Plant Biotechnology Journal*, *10*(7), 826–839. https://doi.org/10.1111/j.1467-7652.2012.00702.x

Huynh, B. L., Ehlers, J. D., Huang, B. E., Muñoz-Amatriaín, M., Lonardi, S., Santos, J. R. P., Ndeve, A., Batieno, B. J., Boukar, O., Cisse, N., Drabo, I., Fatokun, C., Kusi, F., Agyare, R. Y., Guo, Y. N., Herniter, I., Lo, S., Wanamaker, S. I., Xu, S., … Roberts, P. A. (2018). A multi-parent advanced generation inter-cross (MAGIC) population for genetic analysis and improvement of cowpea (Vigna unguiculata L. Walp.). *Plant Journal*, *93*(6), 1129–1142. https://doi.org/10.1111/tpj.13827

Landergott, U., Naciri, Y., Schneller, J. J., & Holderegger, R. (2006). *Allelic configuration and polysomic inheritance of highly variable microsatellites in tetraploid gynodioecious Thymus praecox agg . 113*, 453–465. https://doi.org/10.1007/s00122-006-0310-6

Leal-Bertioli, S. C. M., Moretzsohn, M. C., Santos, S. P., Brasileiro, A. C. M., Guimarães, P. M., Bertioli, D. J., & Araujo, A. C. G. (2017). Phenotypic effects of allotetraploidization of wild arachis and their implications for peanut domestication. *American Journal of Botany*, *104*(3), 379–388. https://doi.org/10.3732/ajb.1600402

Leal-Bertioli, S., Shirasawa, K., Abernathy, B., Moretzsohn, M., Chavarro, C., Clevenger, J., Ozias-Akins, P., Jackson, S., & Bertioli, D. (2015). Tetrasomic recombination is surprisingly frequent in allotetraploid Arachis. *Genetics*, *199*(4), 1093–1105. https://doi.org/10.1534/genetics.115.174607

Li, C., Li, Y., Sun, B., Peng, B., Liu, C., Liu, Z., Yang, Z., Li, Q., Tan, W., Zhang, Y., Wang, D., Shi, Y., Song, Y., Wang, T., & Li, Y. (2013). Quantitative trait loci mapping for yield components and kernel-related traits in multiple connected RIL populations in maize. *Euphytica*, *193*(3), 303–316. https://doi.org/10.1007/s10681-013-0901-7

Moretzsohn, M. C., Gouvea, E. G., Inglis, P. W., Leal-Bertioli, S. C. M., Valls, J. F. M., & Bertioli, D. J. (2013). A study of the relationships of cultivated peanut (Arachis hypogaea) and its most closely related wild species using intron sequences and microsatellite markers. *Annals of Botany*, *111*(1), 113–126. https://doi.org/10.1093/aob/mcs237

Nguepjop, J. R., Tossim, H. A., Bell, J. M., Rami, J. F., Sharma, S., Courtois, B., Mallikarjuna, N., Sane, D., & Fonceka, D. (2016). Evidence of genomic exchanges between homeologous chromosomes in a cross of peanut with newly synthetized allotetraploid hybrids. *Frontiers in Plant Science*,

7(NOVEMBER2016), 1–12. https://doi.org/10.3389/fpls.2016.01635

Nielen, S., Vidigal, B. S., Leal-Bertioli, S. C. M., Ratnaparkhe, M., Paterson, A. H., Garsmeur, O., D'Hont, A., Guimarães, P. M., & Bertioli, D. J. (2012). Matita, a new retroelement from peanut: Characterization and evolutionary context in the light of the Arachis A-B genome divergence. *Molecular Genetics and Genomics*, *287*(1), 21–38. https://doi.org/10.1007/s00438-011-0656-6

Pascual, L., Desplat, N., Huang, B. E., Desgroux, A., Bruguier, L., Bouchet, J. P., Le, Q. H., Chauchard, B., Verschave, P., & Causse, M. (2015). Potential of a tomato MAGIC population to decipher the genetic control of quantitative traits and detect causal variants in the resequencing era. *Plant Biotechnology Journal*, *13*(4), 565–577. https://doi.org/10.1111/pbi.12282

Pattanashetti, S. K., Gowda, M. V. C., & Girija. (2008). Inheritance of morphological traits and pod features in groundnut (Arachis hypogaea L.). *Indian Journal of Genetics and Plant Breeding*, *68*(2), 157–162.

Paul, P. J., Samineni, S., Sajja, S. B., Rathore, A., Das, R. R., Chaturvedi, S. K., Lavanya, G. R., Varshney, R. K., & Gaur, P. M. (2018). Capturing genetic variability and selection of traits for heat tolerance in a chickpea recombinant inbred line (RIL) population under field conditions. *Euphytica*, *214*(2), 1–14. https://doi.org/10.1007/s10681-018-2112-8

Pook, T., Schlather, M., de los Campos, G., Maye, r M., Schoen, C. C., & Simianer, H. (2019). *HaploBlocker: Creation of Subgroup-Speci fi c Haplotype Blocks and Libraries*. *212*(August), 1045–1061.

Sallam, A., & Martsch, R. (2015). Association mapping for frost tolerance using multi-parent advanced generation inter-cross (MAGIC) population in faba bean (Vicia faba L.). *Genetica*, *143*(4), 501–514. https://doi.org/10.1007/s10709-015-9848-z

Sannemann, W., Huang, B. E., Mathew, B., & Léon, J. (2015). Multi-parent advanced generation inter-cross in barley: high-resolution quantitative trait locus mapping for flowering time as a proof of concept. *Molecular Breeding*, *35*(3). https://doi.org/10.1007/s11032-015-0284-7

Schmutz, J., Cannon, S. B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., Hyten, D. L., Song, Q., Thelen, J. J., Cheng, J., Xu, D., Hellsten, U., May, G. D., Yu, Y., Sakurai, T., Umezawa, T., Bhattacharyya, M. K., Sandhu, D., Valliyodan, B., … Jackson, S. A. (2010). Genome sequence of the palaeopolyploid soybean. *Nature*, *463*(7278), 178–183. https://doi.org/10.1038/nature08670

Scott, M. F., Ladejobi, O., Amer, S., Bentley, A. R., Biernaskie, J., Boden, S. A., Clark, M., Dell'Acqua, M., Dixon, L. E., Filippi, C. V., Fradgley, N., Gardner, K. A., Mackay, I. J., O'Sullivan, D., Percival-

Alwyn, L., Roorkiwal, M., Singh, R. K., Thudi, M., Varshney, R. K., … Mott, R. (2020). Multi-parent populations in crops: a toolbox integrating genomics and genetic mapping with breeding. *Heredity*, *125*(6), 396–416. https://doi.org/10.1038/s41437-020-0336-6

Seijo, G., Lavia, G. I., Fernández, A., Krapovickas, A., Ducasse, D. A., Bertioli, D. J., & Moscone, E. A. (2007). Genomic relationships between the cultivated peanut (Arachis hypogaea, Leguminosae) and its close relatives revealed by double GISH. *American Journal of Botany*, *94*(12), 1963–1971. https://doi.org/10.3732/ajb.94.12.1963

Senchina, D. S., Alvarez, I., Cronn, R. C., Liu, B., Rong, J., Noyes, R. D., Paterson, A. H., Wing, R. A., Wilkins, T. A., & Wendel, J. F. (2003). Rate variation among nuclear genes and the age of polyploidy in Gossypium. *Molecular Biology and Evolution*, *20*(4), 633–643. https://doi.org/10.1093/molbev/msg065

Shiringani, A. L., Frisch, M., & Friedt, W. (2010). Genetic mapping of QTLs for sugar-related traits in a RIL population of Sorghum bicolor L. Moench. *Theoretical and Applied Genetics*, *121*(2), 323–336. https://doi.org/10.1007/s00122-010-1312-y

Tripp, L. D. R. (1968). *Germplasm evaluation and inheritance studies in peanuts, Arachis hypogaea L.*

Voorrips, R. E., & Maliepaard, C. A. (2012). The simulation of meiosis in diploid and tetraploid organisms using various genetic models. *BMC Bioinformatics*, *13*(248). https://doi.org/10.1186/1471-2105-13-248

Voorrips, R. E., & Tumino, G. (2022). PolyHaplotyper: haplotyping in polyploids based on bi-allelic marker dosage data. *BMC Bioinformatics*, *23*(1), 1–16. https://doi.org/10.1186/s12859-022-04989-0

Xie, K. D., Xia, Q. M., Wang, X. P., Liang, W. J., Wu, X. M., Grosser, J. W., & Guo, W. W. (2015). Cytogenetic and SSR-marker evidence of mixed disomic, tetrasomic, and intermediate inheritance in a citrus allotetraploid somatic hybrid between 'Nova' tangelo and 'HB' pummelo. *Tree Genetics and Genomes*, *11*(6). https://doi.org/10.1007/s11295-015-0940-4

Yan, W., Zhao, H., Yu, K., Wang, T., Khattak, A. N., & Tian, E. (2020). Development of a multiparent advanced generation intercross (MAGIC) population for genetic exploitation of complex traits in Brassica juncea: Glucosinolate content as an example. *Plant Breeding*, *139*(4), 779–789. https://doi.org/10.1111/pbr.12820

Zhang, L. H., Byrne, D. H., Ballard, R. E., & Rajapakse, S. (2006). Microsatellite marker development in rose and its application in tetraploid mapping. *Journal of the American Society for Horticultural*

Science, *131*(3), 380–387. https://doi.org/10.21273/jashs.131.3.380

Zhang, L., Meng, L., & Wang, J. (2019). Linkage analysis and integrated software GAPL for pure-line populations derived from four-way and eight-way crosses. *Crop Journal*, *7*(3), 283–293. https://doi.org/10.1016/j.cj.2018.10.006

Zheng, C., Voorrips, R. E., Jansen, J., Hackett, C. A., Ho, J., & Bink, M. C. A. M. (2016). Probabilistic multilocus haplotype reconstruction in outcrossing tetraploids. *Genetics*, *203*(1), 119–131. https://doi.org/10.1534/genetics.115.185579