

File S3

Comparison of IBD parental contribution variance with Frisch and Melchinger (2007) and simplification to IBS contribution

1 We used an algebraic formula to predict the variance of P_1 genome contribution in doubled haploid
 2 progeny derived from F1' plants. We considered two-way crosses DH-1 (called (F1)-DH) and
 3 backcrosses DH-1 (called (BC1)-DH) and compared our results with the results given by Frisch and
 4 Melchinger (2007). We considered one chromosome of 100cM for which Frisch and Melchinger (2007)
 5 derived a variance of parental contribution of 0.1419 for (F1)-DH and 0.0945 for (BC1)-DH. We varied
 6 the number of loci p used in our approach and for each, we ran ten independent samplings of loci. We
 7 observed that the results from our approach converged with increasing number of loci to the solution
 8 given by Frisch and Melchinger (2007) (Figure 1 S3).

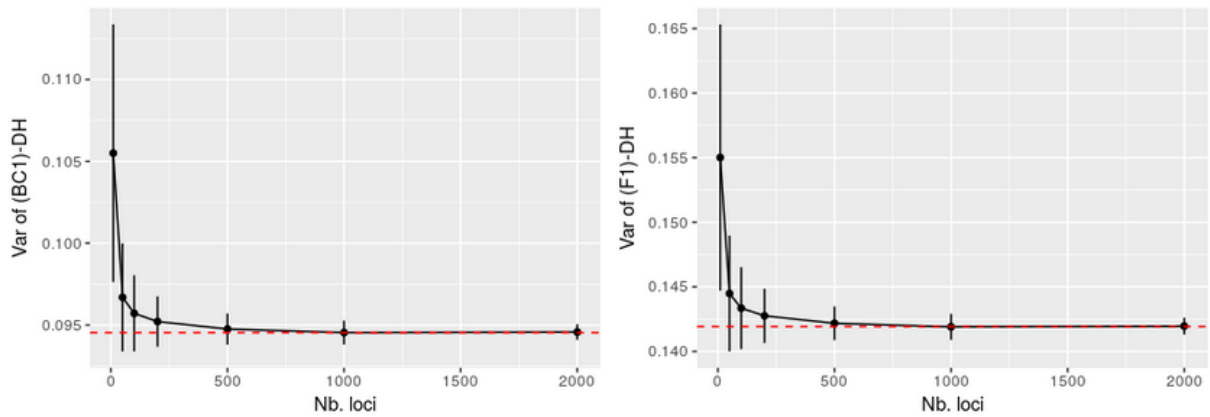


Figure 1 S3 Average parental genome contribution variance (black dots) for (BC1)-DH (left) and (F1)-DH (right) from ten simulation replications (+/- standard deviation represented by black vertical lines) with different number of considered loci. Red dotted line shows the results given by Frisch and Melchinger (2007).

9 In cases where the origin of the allele is not of interest and an identical by state (IBS) similarity between
 10 progeny and parental lines is sufficient, the multi-allelic coding can be simplified to a biallelic coding.
 11 This reduces the size of the covariance matrix from $(4p \times 4p)$ to $(p \times p)$, with p being the number of
 12 loci considered. For this, let us define the genotyping matrix of parental lines in biallelic coding:

$$13 \quad X_{IBS} = \text{diag}(X_{\text{Parental}}) = (x_1 \ x_2 \ x_3 \ x_4)'$$

14 where, \mathbf{X}_{IBS} is a $(4 \times p)$ -dimensional matrix of genotypes. The $(p \times 4)$ -dimensional matrix of global
 15 parental contribution marker effects for each of the four parents can be defined as:

$$16 \quad \boldsymbol{\beta}_{IBS} = \frac{1}{2p} \mathbf{X}_{IBS}'$$

17 where, $\forall i \in [1; 4]$ $\boldsymbol{\beta}_{IBS}(\cdot, i)$ is the p -dimensional vector of marker effect to follow the IBS
 18 contribution of parent i and p is the total number of loci considered.

19 We denote the $(N \times p)$ -dimensional genotyping matrix of N doubled haploid (DH) progeny as
 20 $\mathbf{X}_{IBS-Progeny}$ with element $\mathbf{X}_{IBS-Progeny}(j, l), \forall j \in [1, N], l \in [1, p]$ the genotype of progeny j at
 21 locus l coded as -1, 1 for the genotypes aa, AA, respectively. It results in the following $(N \times 4)$ -
 22 dimensional matrix of parental IBS contribution to progeny:

$$23 \quad \mathbf{C}_{IBS} = \mathbf{X}_{IBS-Progeny} \boldsymbol{\beta}_{IBS} + \frac{1}{2} \mathbf{1}_N \mathbf{1}_4'$$

24 where, $\forall j \in [1; N], \forall i \in [1; 4]$, $\mathbf{C}_{IBS}(j, i)$ is the parental line i contribution to progeny line j .

Literature cited

Frisch M., and A. E. Melchinger, 2007 Variance of the Parental Genome Contribution to Inbred Lines
 Derived From Biparental Crosses. Genetics 176: 477–488.