

## ORIGINAL ARTICLE

# QTL detection power of multi-parental RIL populations in *Arabidopsis thaliana*

JR Klasen<sup>1</sup>, H-P Piepho<sup>2</sup> and B Stich<sup>1</sup>

A major goal of today's biology is to understand the genetic basis of quantitative traits. This can be achieved by statistical methods that evaluate the association between molecular marker variation and phenotypic variation in different types of mapping populations. The objective of this work was to evaluate the statistical power of quantitative trait loci (QTL) detection of various multi-parental mating designs, as well as to assess the reasons for the observed differences. Our study was based on an empirical data of 20 *Arabidopsis thaliana* accessions, which have been selected to capture the maximum genetic diversity. The examined mating designs differed strongly with respect to the statistical power to detect QTL. We observed the highest power to detect QTL for the diallel cross with random mating design. The results of our study suggested that performing sibling mating within subpopulations of joint-linkage mapping populations has the potential to considerably increase the power for QTL detection. Our results, however, revealed that using designs in which more than two parental alleles segregate in each subpopulation increases the power even more.

*Heredity* (2012) **108**, 626–632; doi:10.1038/hdy.2011.133; published online 15 February 2012

**Keywords:** *Arabidopsis thaliana*; statistical power; QTL detection; multi-parental RILs

## INTRODUCTION

Quantitative traits, which include most fitness and agronomic traits, show a continuous distribution of phenotypic values as they are influenced by many genes, epistatic interactions and the environment (Lynch and Walsh, 1998). A major goal of today's biology is to understand the genetic basis of such traits. This can be achieved by means of statistical methods that evaluate the association between molecular marker variation and phenotypic variation in different types of mapping populations (for review, see Doerge, 2002; Sneller *et al.*, 2009).

For the development of biparental linkage mapping populations, two accessions are used, which differ with respect to the trait of interest. From a cross of these accessions, a segregating population is derived. The genomes of the individuals of this population are mosaics of the genomes of the parental genotypes due to the occurred recombination events (Mackay *et al.*, 2009). Many quantitative trait loci (QTL) have been detected for different quantitative traits using such biparental linkage mapping populations. With a few exceptions, however, most of these QTL have not been successfully validated in other populations (Bernardo, 2008). To overcome this problem, the detection of QTLs using a set of genotypes with unknown ancestry, which is called association mapping, has become popular (for review, see Zhu *et al.*, 2008).

The use of association-mapping populations allows the evaluation of a high number of alleles in multiple genetic backgrounds. The mapping resolution of the association-mapping populations compared with the biparental populations is high, as the former allow the utilization of historical recombination events (Mackay *et al.*, 2009). A problem of association-mapping populations, however, is that some individuals might be more related to each other than

individuals are related on average, and this leads to false-positive associations between the pheno- and genotypes (Brescaghiello and Sorrells, 2006; Sneller *et al.*, 2009). This problem cannot be completely prevented even by considering the population structure in the statistical analysis (Brachi *et al.*, 2010). Furthermore, the loci that explain the difference between subpopulations cannot be detected with such approaches. Therefore, the concept for mapping in multi-parental linkage mapping population was developed, which minimizes the effect of population structure by crossing diverse individuals, but still providing a high mapping resolution (Stich *et al.*, 2007).

Rebaï and Goffinet (1993) proposed the extension of the biparental population to a four-parental population, in which the parents were crossed in a half diallel. A method combining the strengths of linkage mapping and association mapping was proposed in the field of animal genetics (Mott *et al.*, 2000; Churchill *et al.*, 2004). In addition, statistical methods for the analysis of multi-parental populations were developed (Xu, 1998; Rebaï and Goffinet, 2000; Jannink and Wu, 2003). Subsequently, different mating designs were recommended and used for the QTL detection in a plant genetics context (Blanc *et al.*, 2006; Paulo *et al.*, 2008; Yu *et al.*, 2008; Buckler *et al.*, 2009; Kover *et al.*, 2009; Stich, 2009). These designs differ with respect to their strategy, as well as the complexity of the required crosses. The mating design underlying the nested association-mapping strategy (Yu *et al.*, 2008) is based on crosses between one parental inbred (PI) line with all other PIs. In contrast, crosses between all PIs are required for the diallel cross (Rebaï and Goffinet, 1993). In the first step of the *Arabidopsis* multi-parental recombinant inbred line (RIL) mating design (Paulo *et al.*, 2008), hybrid crosses between pairs of the PIs were performed. The second step was a diallel cross between the F<sub>1</sub> individuals. The multi-parent, advanced-generation inter-cross design

<sup>1</sup>Max Planck Institute for Plant Breeding Research, Cologne, Germany and <sup>2</sup>Bioinformatics Unit, Institute of Crop Science, University of Hohenheim, Stuttgart, Germany  
Correspondence: Dr B Stich, Max Planck Institute for Plant Breeding Research, Carl-von-Linné-Weg 10, Cologne 50829, Germany.  
E-mail: stich@mpipz.mpg.de

Received 22 July 2011; revised 27 October 2011; accepted 31 October 2011; published online 15 February 2012

(Kover *et al.*, 2009) is based on a diallel cross of all PIs, followed by four generations of random mating. Furthermore, sibling mating within the biparental populations has proven to increase the mapping resolution (Lee *et al.*, 2002). The different approaches result in mapping populations, which differ with respect to the number of combined parental genomes per individual, the number of recombination breakpoints, and the allele frequencies. This in turn is expected to influence the power to detect QTL. To the best of our knowledge, however, the relative contribution of the individual factors to increasing the power is unknown.

The objectives of this work were to evaluate the power of QTL detection of various multi-parental mating designs for *Arabidopsis thaliana* on the basis of different scenarios, as well as to assess the reasons for the observed differences.

## MATERIALS AND METHODS

Our study was based on an empirical data of 20 *A. thaliana* accessions, namely Bay-0, Bor-4, Br-0, Bur-0, C24, Col-0, Cvi-0, Est-1, Fei-0, Got-7, Ler-1, Lov-5, Nfa-8, Rrs-7, Rrs-10, Sha, Tamm-2, Ts-1, Tsu-1 and Van-0 (Clark *et al.*, 2007). These inbreds were selected on the basis of polymorphisms in 876 genome-wide distributed fragments from a sample of 96 *A. thaliana* genotypes to capture the maximum genetic diversity (Nordborg *et al.*, 2005). A total of 648 570 non-redundant single-nucleotide polymorphisms (SNPs) was available for these inbreds (Clark *et al.*, 2007). For this study, 653 sets of markers, each comprising five closely linked SNPs, were selected from the total number of SNPs (Supplementary File S1). The five SNPs of a haplomarker were located within a physical map distance of 300–3000 bp. Each set of five SNPs was considered to be one multi-allelic marker locus called the haplomarker. The 653 haplomarkers were evenly distributed throughout the physical map of *A. thaliana*. Genetic map positions for the haplomarkers were lacking. Therefore, the physical map position of the middle SNP of each haplomarker was linearly projected on to the genetic map (Singer *et al.*, 2006), resulting in an average genetic map distance of  $\sim 0.7$  cM. The number of haplotypes per haplomarker ranged from two to nine with an average of five.

### Mating designs

The 20 *A. thaliana* accessions were used to examine eight different mating designs using computer simulations.

In the first design, here referred to as the reference (REF) design, the PI line Col-0 was crossed with the other 19 PI lines (Supplementary Figure S1). Each hybrid was selfed for four generations to create a set of  $N$  RILs (Supplementary Figure S2).

For the REF design with sibling mating (REFS), sibling mating was performed for three generations among the progenies of each of the 19  $F_1$  hybrids, which were designated in our study as subpopulations. Each of the  $S=19$  sibling mating subpopulations consisted of five individuals. The 950 individuals of the third sibling-mating generation were selfed for four generations to create a set of  $N$  RILs (Supplementary Figure S2).

For the diallel cross (DC) design, each PI line was crossed with the other 19 PI lines, resulting in a total of 190 different  $F_1$  hybrids (Supplementary Figure S3). Each hybrid was selfed for four generations to create a set of  $N$  RILs.

For the DC with sibling-mating design (DCS), sibling mating was performed for three generations among the progenies of each of the 190  $F_1$  hybrids, which were designated in our study as subpopulations. Each of the  $S=190$  sibling-mating subpopulations consisted of five individuals. The 950 individuals of the third sibling-mating generation were selfed for four generations to create a set of  $N$  RILs.

For the DC with random mating design (DCR), random mating was performed for three generations among the progenies of all the 190  $F_1$  hybrids from the DC design. The 950 individuals of the third random mating generation were selfed for four generations to create a set of  $N$  RILs (Supplementary Figure S2).

For the four-way hybrids cross design (FHC), the 20 PI lines were crossed in pairs to create 10  $F_1$  hybrids. The 10  $F_1$  hybrids were further crossed in pairs to establish  $S=5$  subpopulations with a total of  $N$  four-way hybrids (Supplemen-

tary Figure S4). Each of the  $N$  four-way hybrids was selfed four times to generate  $N$  RILs.

For the two-way hybrids DC design (THDC), the 20 PI lines were crossed in pairs to create 10  $F_1$  hybrids. The 10  $F_1$  hybrids were crossed in a half diallel to establish  $S=45$  subpopulations with a total of  $N$  four-way hybrids (Supplementary Figure S5).  $N$  RILs were created by selfing these individuals for four generations.

The four-way hybrids DC design (FHDC) was examined in two scenarios. For the FHDC10 design, 20 PI lines were crossed in pairs to create 10  $F_1$  hybrids. These 10  $F_1$  hybrids were crossed in pairs to establish five subpopulations with 10 four-way hybrids per subpopulation. The four-way hybrids were crossed in a half diallel such that each four-way hybrid was crossed with one individual from the other subpopulations (Supplementary Figure S6) to establish the  $S=10$  subpopulations. With this procedure, a total of  $N F_3$  individuals was generated from which  $N$  RILs were obtained by four generations of selfing. The FHDC100 design differed from the FHDC10 design by involving 100 instead of 10 four-way hybrids per subpopulation.

The number of individuals per subpopulation  $S$  was calculated in a two-step procedure. First, the minimum number of individuals per subpopulation was calculated as  $\lfloor N/S \rfloor$ , which is the integer part of  $N/S$ . Second, a number of  $N - \lfloor N/S \rfloor * S$  random subpopulations was assigned one additional individual. The number of required generations, as well as the total number of individuals across all generations differed considerably among the examined designs (Table 1).

The mating designs were compared on the basis of different scenarios, which differed with respect to the population size  $N=1250, 2500, 5000$ , heritability and number of QTL. Choice of heritability and number of QTL will be described in the following.

### Genotypic and phenotypic values

A total of 50 simulation runs were performed for each of the examined mating designs. For each run, three subsets of haplomarkers  $l=25, 50, 100$  were randomly sampled without replacement from the linkage map and defined as QTL. The maximum genotypic effect per QTL  $a_k$  with  $k=1, 2, \dots, l$  was drawn randomly without replacement from the geometric progression  $a_k = a_0 q^k$ , with  $a_0 = 100(1-q)/(1-q^l)$  and  $q=0.90$  for 25 QTL,  $q=0.96$  for 50 QTL, and  $q=0.99$  for 100 QTL (Lande and Thompson, 1990). The number of alleles per QTL  $n$  was given by the number of haplotypes at the sampled haplomarker. The effect of each QTL allele at a given locus was randomly drawn without replacement from the arithmetic progression  $a_{k_m} = a_k - ((m-1)a_k/(n-1))$  with  $m=1, 2, \dots, n$ , where the effect  $a_k$  given from the geometric progression was gradually reduced to zero and the number of steps was given by the number of alleles  $n$  that are present at this locus. The genotypic value of an individual was the sum of all of its QTL effects. From the genotypic values of the set of PIs, the genotypic variance  $\sigma_g^2$  was calculated (Valdar *et al.*, 2006), which was the same for all mating designs. The phenotypic values of the RILs of each subpopulation were generated by adding a realization from a normal distribution  $N(0, (1-h^2)\sigma_g^2/h^2)$  to the genotypic values of the RILs, where  $h^2$  denotes the heritability. For our simulations,  $h^2=0.5, 0.8$  was assumed.

All simulations were performed with the software PLABSOFIT (Maurer *et al.*, 2007), which is implemented as an extension of the statistical software R (R Development Core Team, 2009). The data can be found under the following doi:10.5061/dryad.gn6hg74q.

### QTL detection method neglecting population structure

The comparison of statistical analyses concerning the power requires an equal empirical type I error rate  $\alpha^*$ . To meet this requirement, the following two-step procedure for QTL detection was applied. First, a stepwise multiple linear regression was used to select a set of cofactors, based on the Bayesian information criterion. The model was:

$$y = \mu + \sum_{c \in C} \sum_{j=2}^n b_{cj} x_{cj} + e \quad (1)$$

where  $y$  is the vector of the phenotypic values of all RILs,  $\mu$  is the intercept vector,  $C$  is the set of haplomarkers that are selected as cofactors,  $b_{cj}$  is the partial regression coefficients of the  $j$ th haplotype at the  $c$ th cofactor,  $x_{cj}$  is an incidence vector of the  $j$ th haplotype at the  $c$ th cofactor, and  $e$  is the vector of

**Table 1** Number of individuals per cross and number of crosses, and selfings

	REF	REFS	DC	DCS	DCR	FHC	THDC	FHDC10	FHDC100
P	20×	20×	20×	20×	20×	20×	20×	20×	20×
F <sub>1</sub>	19⊗	19⊗	190⊗	190⊗	190×	10×	10×	10×	10×
F <sub>2</sub>	5000⊗	950×	5000⊗	950×	950×	5000⊗	5000⊗	50×	500×
F <sub>3</sub>	5000⊗	950×	5000⊗	950×	950×	5000⊗	5000⊗	5000⊗	5000⊗
F <sub>4</sub>	5000⊗	5000⊗	5000⊗	5000⊗	5000⊗	5000⊗	5000⊗	5000⊗	5000⊗
F <sub>5</sub>	5000	5000⊗	5000	5000⊗	5000⊗	5000⊗	5000⊗	5000⊗	5000⊗
F <sub>6</sub>	—	5000⊗	—	5000⊗	5000⊗	5000	5000	5000⊗	5000⊗
F <sub>7</sub>	—	5000⊗	—	5000⊗	5000⊗	—	—	5000	5000
F <sub>8</sub>	—	5000	—	5000	5000	—	—	—	—
Sum individuals	20019	26919	20190	27090	27090	25010	25010	25060	25510
Crosses (×)	19	1919	190	2090	2869	15	55	115	1015
Selfings (⊗)	15019	20019	15190	20190	20000	20000	20000	20000	20000

Abbreviations: DC, diallel cross; DCR, DC with random mating; DCS, DC with sibling mating; FHC, four-way hybrids cross; FHDC, four-way hybrids DC with 10 or 100 individuals per F2 subpopulation; REF, reference design; REFS, REF with sibling mating; THDC, two-way hybrids DC.

residual errors. To avoid overparameterization, the first allele effect for every haplotype was set to zero. The assumption for the QTL analysis was that the number of haplotypes was so high that each QTL has a haplotype, which is in complete linkage disequilibrium with the QTL. Therefore, all haplotypes, including those treated as QTL, were included in the QTL detection procedure.

In the second step, a *P*-value for the association of each haplotype *i* with the phenotypic value was calculated. For this, an analysis of variance with a full model against a reduced model was calculated. The full model was:

$$y = \mu + \sum_{j=2}^n b_{ij}x_{ij} + \sum_{\substack{c \in C \\ c \neq i}} \sum_{j=2}^n b_{cj}x_{cj} + e \quad (2)$$

where  $b_{ij}$  is the partial regression coefficient and  $x_{ij}$  is the incidence vector of the *j*th haplotype at the *i*th haplotype. In the reduced model, the *i*th haplotype was neglected. Model (2) indicates that for the analysis of variance of a specific haplotype, only those cofactors are used which are not identical to the haplotype under consideration. This constraint is inevitable to detect also those QTLs for which a cofactor was selected in the first step.

The QTL detection was performed using the statistical software R (R Development Core Team, 2009).

### QTL detection method considering population structure

The following random model was used:

$$y = \mu + g + e \quad (3)$$

where the random individual effects  $g$  are  $N(0, K_v)$ ,  $K$  is the variance-covariance relationship matrix, and  $v_g$  is the additive genetic variance. The relationship matrix was calculated from pedigree records or based on the proportion of shared haplotype for each pair of individuals (Zhao *et al.*, 2007). The random model was fitted using the statistical software ASReml (Gilmour *et al.*, 2009) and the R package GenABEL (Aulchenko *et al.*, 2007).

For QTL detection, the above described two-step procedure was used, where instead of phenotypic values, the residuals of the random model (3) were considered as dependent variables (Aulchenko *et al.*, 2007).

### Power calculation

As the haplotypes that were considered as QTLs were known, the power to detect a QTL  $1-\beta^*$  was calculated as follows. For each scenario, the nominal  $\alpha$ -level was chosen in such a way that the empirical type I error rate  $\alpha^*$  was 0.5, 0.1, 0.01, 0.001, 0.0001 or 0.00001. The power for QTL detection  $1-\beta^*$  was calculated on the basis of these  $\alpha$  levels as proportion of correctly identified QTLs from the total number of QTLs *l* (Stich, 2009).

For each scenario, a Kruskal–Wallis test was performed on the 50 replications to examine the presence of significant differences among all mating

designs. If this test was significant, a Mann–Whitney test was performed to calculate the asymptotic *P*-values for pairwise differences. The pairwise differences (significance level  $P < 0.05$ ) were presented via letter-based comparisons (Piepho, 2004).

### Genome structure analysis

We calculated the number of recombination breakpoints as the average number of alterations between the parental genomes along the genome of one individual in the mapping population of the considered mating designs. Furthermore, we inferred the number of PIs that contributed to the genome of an individual of a mapping population. In both cases, identity by descent was considered as the reference point. For all mating designs, the average of measures across all individuals and all replications was calculated.

## RESULTS

### Mating designs

The lowest average number of recombination breakpoints per individual was nine for the REF and DC designs (Table 2). For the REFS and DCS design, sibling mating increased the number of recombination breakpoints to 12.7. The highest number of recombination breakpoints was observed with 20.2 for the DCR mating design.

The average number of combined parental genomes per individual in the mapping population was two for the REF, REFS, DC and DCS designs. The highest number of combined parent genomes was 9.4 per individual for the DCR design.

The allele frequency of 2.3 of the five alleles at an average QTL was 0.05 (Figure 1). In the REF and REFS design, the frequency of most of the alleles was 0.025, whereas for the other mating designs, values of about 0.05 were observed. For the DCS, DCR, FHDC10 and FHDC100, allele frequency changes due to genetic drift were observed (Figure 1).

### Method neglecting population structure during QTL detection

For the scenario with 5000 RILs and  $h^2=0.5$ , the power across all mating designs decreased from the variant with 25 QTLs to the variant with 100 QTLs from 0.72 to 0.27, whereas the variant with 50 QTLs had a power of 0.57 (Figure 2; Supplementary Table S1). In the scenario with  $h^2=0.8$ , the power to detect QTL was higher and ranged across all mating designs from 0.91 to 0.64 for 25 to 100 QTLs, respectively.

The reduction of the number of RILs from 5000 to 2500 and 1250 individuals led to a decrease of the power to detect a QTL for all mating designs (Table 3). The power trends observed in the

**Table 2 Mean number and s.d. of recombination breakpoints and combined parental genomes per individual**

Mating design	No. of recombination breakpoints		No. of combined parental genomes	
	Mean	s.d.	Mean	s.d.
REF	9.0	3.1	2.0	0.0
REFS	12.7	3.7	2.0	0.0
DC	9.0	3.1	2.0	0.0
DCS	12.7	3.7	2.0	0.0
DCR	20.2	4.5	9.4	1.5
FHC	13.1	3.7	4.0	0.2
THDC	13.1	3.7	4.0	0.2
FHDC10	17.2	4.2	7.4	0.7
FHDC100	17.2	4.2	7.4	0.7

Abbreviations: DC, diallel cross; DCR, DC with random mating; DCS, DC with sibling mating; FHC, four-way hybrids cross; FHDC, four-way hybrids DC with 10 or 100 individuals per F2 subpopulation; REF, reference design; REFS, REF with sibling mating; THDC, two-way hybrids DC.

scenarios with  $N=1250$  and  $N=2500$  were identical to those with  $N=5000$ , irrespective of the number of QTLs and  $h^2$  values considered.

The power decreased with the empirical  $\alpha^*$  level, but the ranking of the mating designs with respect to the power was largely unchanged (Supplementary Figures S7–S13). The ranking of the mating designs also remained constant across all examined QTL and  $h^2$  scenarios. The DCR mating design showed the highest power and the REF design showed the lowest power. The difference in power ( $\alpha^*=0.01$ ) between these designs was significant (significance level of 0.05) for all examined scenarios (Supplementary Tables S2–S9). The mating designs with sibling mating (REFS and DCS) had a significantly higher power than the same mating designs without sibling mating (REF and DC).

#### Methods considering population structure during QTL detection

All mating designs with the exception of DCR were also examined with QTL detection methods, considering the population structure based on pedigree information. For all examined mating designs, the power to detect QTL was lower for the methods considering population structure than for that neglecting population structure (Table 3). For all mating designs, the power of the analysis considering population structure calculated from haplotype information was lower than for the analysis considering pedigree population-based structure (Supplementary Figures S14, S15; Supplementary Tables S10, S11). The ranking of the mating design was not influenced by the QTL detection method.

## DISCUSSION

### Factors influencing the power to detect QTL

In our study, the power to detect QTL of the REF and DC design was considerably lower than that observed by Stich (2009). This finding can be explained by the different benchmarks of residual variance and hence, of heritability used in these two studies when simulating phenotypic values. Stich (2009) considered the genetic variance per subpopulation, whereas we used the genetic variance of the PIs as the basis for the simulation of phenotypic values. A second reason is the number of degrees of freedom required in the stepwise regression in our study due to the higher number of assumed alleles compared with

the study of Stich (2009). This leads to a decreased number of selected cofactors, which in turn reduces the power to detect QTL.

In our study, we assumed that the haplotypes were in complete linkage equilibrium with the QTL, which increases the power in comparison with experimental data where linkage is not complete. This simplification, however, is the same for all examined mating designs and thus, is expected not to influence the ranking of the mating designs.

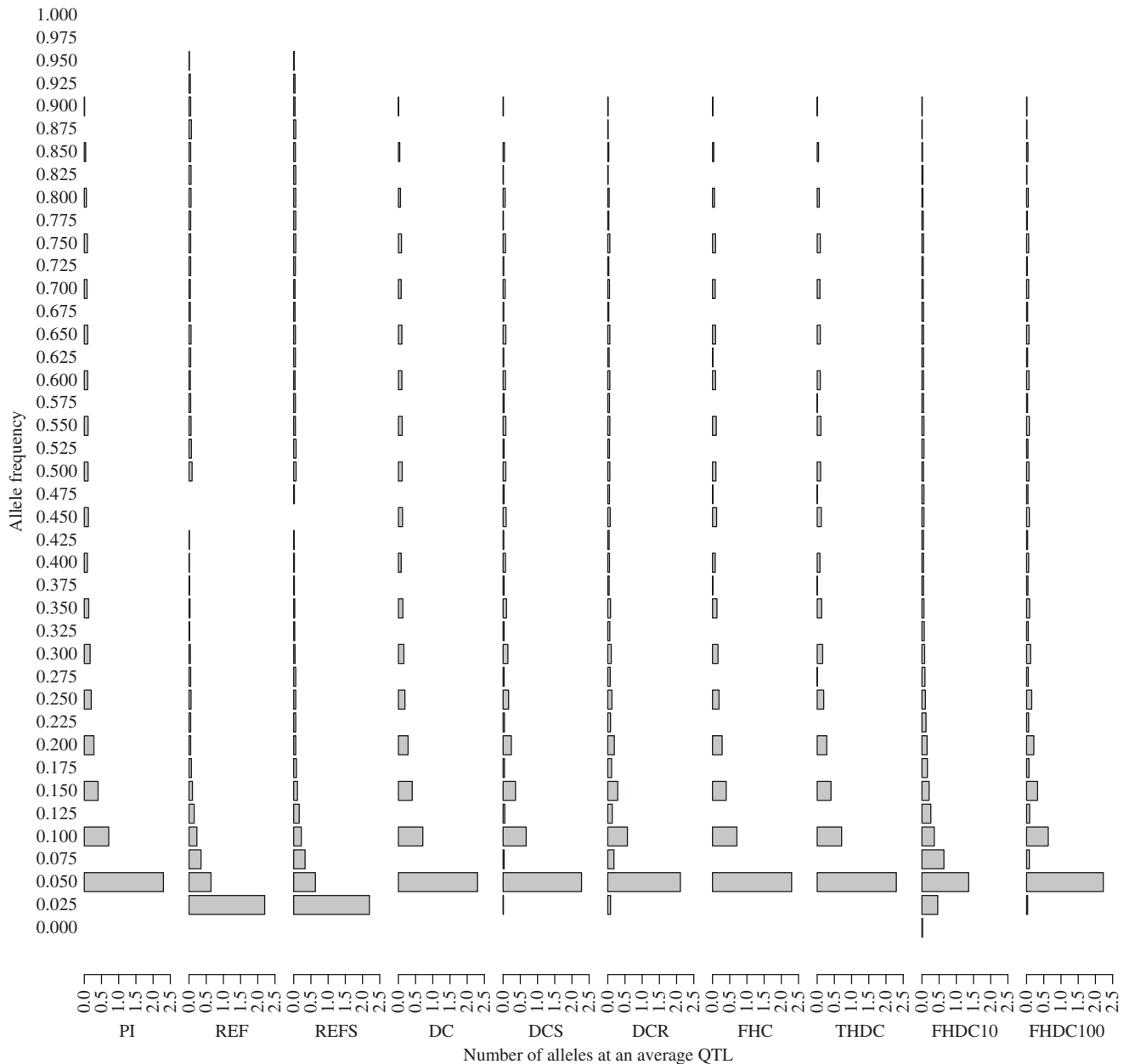
We observed a lower power to detect QTL for the approaches taking population structure into account than for the approaches neglecting this information (Table 3). This finding can be explained by the fact that association between haplotypes, which differ only in state between subpopulations, and the phenotype cannot be as simply detected when population structure is corrected for during the QTL analysis (Yu *et al.*, 2006; Sneller *et al.*, 2009; Brachi *et al.*, 2010). The analyses considering population structure calculated from haplotype information were more effective in reducing the risk of false-positive QTL than the analyses considering population structure calculated from the pedigree information. However, our strategy for calculating the significance threshold, which is described in detail in material and methods, masks this advantage. Furthermore, our results suggested that under a fixed empirical type I error rate, the former analysis leads to a lower power compared with the latter analysis.

In contrast to studies based on experimental data, the QTL underlying the phenotypic variation are known in studies using computer simulations. Therefore, in the latter case, it is possible to calculate the significance threshold in such a way that it is not influenced by false-positive associations due to population structure, as outlined in materials and methods. This, however, makes a comparison between the different QTL detection methods unfair. Nevertheless, it allows in our study to compare different designs with respect to their QTL detection power, despite their difference in the importance of population structure. When analyzing experimental data of the examined mating designs, however, population structure has to be considered to control the nominal type I error rate (data not shown).

As the ranking of the examined mating designs with respect to the power was largely constant across the studied scenarios, we discuss in the following only on the results of the scenario with  $h^2=0.5$ , 50 QTLs,  $N=5000$ ,  $\alpha^*=0.01$ , and considered the QTL detection method neglecting the population structure.

### Comparison of the examined mating designs

We examined the power of the REF design, which is similar to the design used to establish the nested association-mapping population (Yu *et al.*, 2008; McMullen *et al.*, 2009). This value was compared with that of the DC design, which corresponds to the design described by Rebaï and Goffinet (1993). Across all examined scenarios, we observed a higher power to detect QTL for the DC design than for the REF design (Table 3, Supplementary Table S1). Our observation accords with the findings of Stich (2009). This difference in power estimates between the REF, and the DC design can be explained by differences in genetic variance, which are caused by difference in allele frequencies. The allele frequency differences are due to the crossing scheme underlying the REF design, and the fact that not all parental genotypes contribute to the same extent to the segregating population. The alleles of the common parent have a high allele frequency, whereas the alleles of the other parents occur less frequently. In the DC design, however, crosses between all PIs are created and thus, the allele frequency should remain unchanged compared with that of the PIs. This explanation accords with the observed allele frequency pattern (Figure 1).



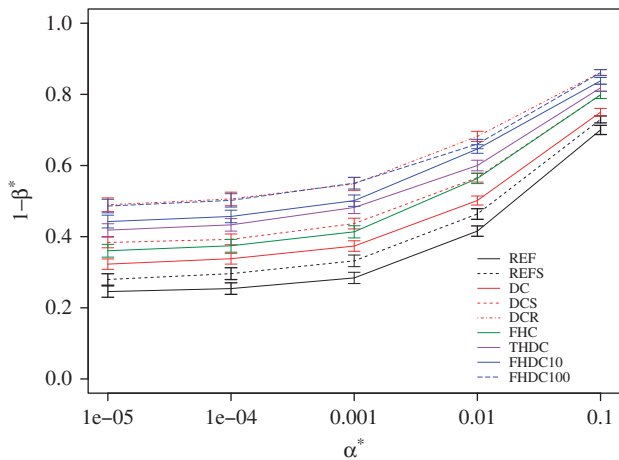
**Figure 1** Histograms of the allele frequencies at an average QTL for the following mating designs compared with PI: REF, REFS, DC, DCS, FHC, THDC and FHDC with 10 or 100 individuals per  $F_2$  subpopulation (FHDC10 or FHDC100).

Another interesting question is how the number of combined genomes per individual influences the power. Therefore, the THDC and FHC designs were examined. The THDC design is similar to the Arabidopsis multi-parental RIL design (Paulo *et al.*, 2008; Huang *et al.*, 2011), where the PIs are crossed in pairs to create two-way hybrids, which were then crossed in a diallel. Instead of a diallel cross, the FHC had a second generation of pairwise hybridisations. In all examined scenarios, the FHC design and the THDC design had a higher power to detect QTL than the DC design (Table 3, Supplementary Table S1). This difference can be explained by the higher number of combined parental genomes per individual for THDC and FHC than for the DC design (Table 2). This, in turn, results in the combination of one QTL allele with more diverse genetic backgrounds, which increases the power.

The THDC design showed a higher power than the FHC design (Figure 2). The FHC and the THDC had the same number of recombination breakpoints, as well as the same number of combined parental genomes per individual (Table 2). However, as discussed above for the REF and the DC design, the THDC design is based on the combination of all PIs, which is not the case for the FHC design. Therefore, the THDC design has a higher power than the FHC design.

The FHDC design is a combination of the Arabidopsis multi-parental RIL design and the multi-parent, advanced generation intercross design (Cavanagh *et al.*, 2008). For all examined scenarios, we observed a higher power for the FHDC10 and FHDC100 designs compared with the THDC design, despite the only marginally increased crossing effort (Table 3). The difference between the FHDC and the THDC design can be explained by a higher number





**Figure 2** Power to detect QTLs  $1 - \beta^*$  when neglecting population structure for different  $\alpha^*$  levels in a scenario with 50 QTLs, heritability  $h^2=0.5$  and population size  $N=5000$ . The following mating designs were examined: REF, REFS, DC, DCS, FHC, THDC, and FHDC with 10 or 100 individuals per  $F_2$  subpopulation (FHDC10 or FHDC100). The whiskers represent the s.e.m. across all replications.

**Table 3** Power to detect QTLs and the corresponding s.e. across replications, for different population sizes ( $N$ ), for QTL detection approaches, NPS or CPS, calculated from the P or M information

Mating design	Value	NPS			CPS-P		CPS-M
		N=1250	N=2500	N=5000	N=5000	N=5000	N=5000
REF	Power	0.22	0.31	0.42	0.34	0.30	
	s.e.	0.010	0.011	0.015	0.015	0.007	
REFS	Power	0.24	0.35	0.46	0.39	0.31	
	s.e.	0.010	0.013	0.015	0.018	0.008	
DC	Power	0.29	0.37	0.50	0.39	0.34	
	s.e.	0.012	0.013	0.013	0.016	0.008	
DCS	Power	0.30	0.43	0.57	0.45	0.38	
	s.e.	0.012	0.015	0.013	0.016	0.007	
DCR	Power	0.42	0.55	0.68	Not applicable	0.43	
	s.e.	0.013	0.017	0.014	applicable	0.008	
FHC	Power	0.36	0.44	0.56	0.53	0.37	
	s.e.	0.010	0.012	0.014	0.016	0.007	
THDC	Power	0.35	0.48	0.60	0.54	0.40	
	s.e.	0.011	0.014	0.015	0.014	0.007	
FHDC10	Power	0.37	0.51	0.65	0.62	0.40	
	s.e.	0.013	0.014	0.012	0.013	0.006	
FHDC100	Power	0.40	0.54	0.66	0.64	0.42	
	s.e.	0.013	0.012	0.013	0.015	0.007	

Abbreviations: CPS, considering population structure; DC, diallel cross; DCR, DC with random mating; DCS, DC with sibling mating; FHC, four-way hybrids cross; FHDC, four-way hybrids DC with 10 or 100 individuals per  $F_2$  subpopulation; M, marker; NPS, neglecting population structure; P, pedigree; QTL, quantitative trait loci; REF, reference design; REFS, REF with sibling mating; THDC, two-way hybrids DC. A total of 50 QTL and a heritability of 0.5 was assumed. The empirical type I error rate  $\alpha^*$  was 0.01. For the DCR design, random mating was performed across all subpopulation and thus, no P-based population structure exists.

of combined parental genomes per individual as was discussed before for THDC versus DC designs.

We observed a higher power for the FHDC100 than for the FHDC10 design (Table 3). This difference can be explained by the

reduced effect of genetic drift, that is, the random changes of the allele frequency, in the former than in the latter design. In the FHDC10 design, one of the alleles at an average QTL got lost in some of the replications (Figure 1).

For the designs with sibling mating (REFS and DCS), we observed a higher power than for the designs without sibling mating (REF and DC) in all examined scenarios (Figure 2). The increase in power by sibling mating accords with earlier results (Rockman and Kruglyak, 2008), and is due to a slower increase of homozygosity by sibling mating compared with selfing. This leads to a more genetic recombination in the segregating populations and thus, to a better resolution, but also to a higher power in the detection of QTLs (Vales *et al.*, 2005; Rockman and Kruglyak, 2008). However, the increase in power by sibling mating within subpopulations is small compared with the random crosses of the DCR design.

The DCR design is similar to the design described by Kover *et al.*, (2009), for which three generations of random crosses among all progenies followed a diallel cross. Our results indicated that this strategy has a higher power than the DCS design. This finding can be explained by the higher probability that recombination leads to new allele combinations for the DCR than for the DCS design. Our explanation is in agreement with the observation that the detected number of recombination breakpoints per individual differed considerably (Table 2). This result indicated that populations with a high number of combined parental genomes have a higher effective recombination rate, which means that recombination occurs more often between genomes of different parents. Furthermore, the finding that the DCR design requires the same effort for establishing the population as the DCS design suggests that the DCR is a very promising approach for creating multi-parental RIL populations.

## CONCLUSIONS

Our results indicate that crossing all PIs in a diallel and creating segregating populations from each  $F_1$  hybrid is a promising way of creating a multi-parental population for QTL detection. However, a diallel cross of PIs followed by hybrid crosses or random crosses among the  $F_1$  increases the number of combined parental genomes and results in an even higher power. Sibling mating increases the number of recombinations, but not the number of combined parental genomes, and is therefore less effective than the former described crossing strategies. A crossing strategy like the REF design results in populations with low power and is only useful in specific situations, for example, when the genetic diversity must be reduced to allow testing all entries in the same field experiment. The similar ranking of the examined mating designs across all studied scenarios suggests that our results are broadly applicable.

## DATA ARCHIVING

Genotype data have been submitted to Dryad: doi:10.5061/dryad.gn6hg74q.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## ACKNOWLEDGEMENTS

We thank Matthias Frisch for allowing us to use the compute cluster of the Department of Biometry and Population Genetics at the Justus Liebig University Giessen, Germany, for this study. The funding for this work was provided by the Max Planck Society. We thank the associate editor and three anonymous reviewers for their valuable suggestions.

- Aulchenko YS, de Koning DJ, Haley C (2007). Genomewide rapid association using mixed model and regression: a fast and simple method for genomewide pedigree-based quantitative trait loci association analysis. *Genetics* **177**: 577–585.
- Bernardo R (2008). Molecular markers and selection for complex traits in plants: learning from the last 20 years. *Crop Sci* **48**: 1649–1664.
- Blanc G, Charcosset A, Mangin B, Gallais A, Moreau L (2006). Connected populations for detecting quantitative trait loci and testing for epistasis: an application in maize. *Theor Appl Genet* **113**: 206–224.
- Brachi B, Faure N, Horton M, Flahauw E, Vazquez A, Nordborg M *et al.* (2010). Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature. *PLoS Genet* **6**: e1000940.
- Breseghele F, Sorrells ME (2006). Association analysis as a strategy for improvement of quantitative traits in plants. *Crop Sci* **46**: 1323–1330.
- Buckler ES, Holland JB, Bradbury PJ, Acharya CB, Brown PJ, Browne C *et al.* (2009). The genetic architecture of maize flowering time. *Science* **325**: 714–718.
- Cavanagh C, Morell M, Mackay I, Powell W (2008). From mutations to MAGIC: resources for gene discovery, validation and delivery in crop plants. *Curr Opin Plant Biol* **11**: 215–221.
- Churchill GA, Airey DC, Allayee H, Angel JM, Attie AD, Beatty J *et al.* (2004). The collaborative cross, a community resource for the genetic analysis of complex traits. *Nat Genet* **36**: 1133–1137.
- Clark RM, Schweikert G, Toomajian C, Ossowski S, Zeller G, Shinn P *et al.* (2007). Common sequence polymorphisms shaping genetic diversity in *Arabidopsis thaliana*. *Science* **317**: 338–342.
- Doerge RW (2002). Mapping and analysis of quantitative trait loci in experimental populations. *Nat Rev Genet* **3**: 43–52.
- Gilmour AR, Gogel BJ, Cullis BR, Thompson R (2009). *ASReml User Guide Release 3.0*. VSN International Ltd, Hemel Hempstead, www.vsnl.co.uk.
- Huang X, Paulo MJ, Boer M, Effgen S, Keizer P, Koornneef M *et al.* (2011). Analysis of natural allelic variation in *Arabidopsis* using a multiparent recombinant inbred line population. *Proc Natl Acad Sci USA* **108**: 1–6.
- Jannink JL, Wu XL (2003). Estimating allelic number and identity in state of QTLs in interconnected families. *Genet Res* **81**: 133–144.
- Kover PX, Valdar W, Trakalo J, Scarcelli N, Ehrenreich IM, Purugganan MD *et al.* (2009). A multiparent advanced generation inter-cross to fine-map quantitative traits in *Arabidopsis thaliana*. *PLoS Genet* **5**: e1000551.
- Lande R, Thompson R (1990). Efficiency of marker-assisted selection in the improvement of quantitative traits. *Genetics* **124**: 743–756.
- Lee M, Sharopova N, Beavis WD, Grant D, Katt M, Blair D *et al.* (2002). Expanding the genetic map of maize with the intermated B73 Mo17 (IBM) population. *Plant Mol Biol* **48**: 453–461.
- Lynch M, Walsh B (1998). *Genetics and Analysis of Quantitative Traits*. Sinauer Associates, Inc.: Sunderland.
- Mackay TFC, Stone EA, Ayroles JF (2009). The genetics of quantitative traits: challenges and prospects. *Nat Rev Genet* **10**: 565–577.
- Maurer HP, Melchinger AE, Frisch M (2007). Population genetic simulation and data analysis with Plabsoft. *Euphytica* **161**: 133–139.
- McMullen MD, Kresovich S, Villeda HS, Bradbury P, Li H, Sun Q *et al.* (2009). Genetic properties of the maize nested association mapping population. *Science* **325**: 737–740.
- Mott R, Talbot CJ, Turri MG, Collins AC, Flint J (2000). A method for fine mapping quantitative trait loci in outbred animal stocks. *Proc Natl Acad Sci USA* **97**: 12649–12654.
- Nordborg M, Hu TT, Ishino Y, Jhaveri J, Toomajian C, Zheng H *et al.* (2005). The pattern of polymorphism in *Arabidopsis thaliana*. *PLoS Biol* **3**: e196.
- Paulo MJ, Boer M, Huang X, Koornneef M, van Eeuwijk FA (2008). A mixed model QTL analysis for a complex cross population consisting of a half diallel of two-way hybrids in *Arabidopsis thaliana*: analysis of simulated data. *Euphytica* **161**: 107–114.
- Piepho HP (2004). An algorithm for a letter-based representation of all-pairwise comparisons. *J Comput Graph Statist* **13**: 456–466.
- R Development Core Team (2009). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing: Vienna, Austria.
- Rebai A, Goffinet B (1993). Power of tests for QTL detection using replicated progenies derived from a diallel cross. *Theor Appl Genet* **86**: 1014–1022.
- Rebai A, Goffinet B (2000). More about quantitative trait locus mapping with diallel designs. *Genet Res* **75**: 243–247.
- Rockman MV, Kruglyak L (2008). Breeding designs for recombinant inbred advanced intercross lines. *Genetics* **179**: 1069–1078.
- Singer T, Fan Y, Chang HS, Zhu T, Hazen SP, Briggs SP (2006). A high-resolution map of *Arabidopsis* recombinant inbred lines by whole-genome exon array hybridization. *PLoS Genet* **2**: e144.
- Sneller CH, Mather DE, Crepeux S (2009). Analytical approaches and population types for finding and utilizing QTL in complex plant populations. *Crop Sci* **49**: 363–380.
- Stich B (2009). Comparison of mating designs for establishing nested association mapping populations in maize and *Arabidopsis thaliana*. *Genetics* **183**: 1525–1534.
- Stich B, Melchinger AE, Piepho HP, Hamrit S, Schipprack W, Maurer HP *et al.* (2007). Potential causes of linkage disequilibrium in a European maize breeding program investigated with computer simulations. *Theor Appl Genet* **115**: 529–536.
- Valdar W, Flint J, Mott R (2006). Simulating the collaborative cross: power of quantitative trait loci detection and mapping resolution in large sets of recombinant inbred strains of mice. *Genetics* **172**: 1783–1797.
- Vales MI, Schön CC, Capettini F, Chen XM, Corey AE, Mather DE *et al.* (2005). Effect of population size on the estimation of QTL: a test using resistance to barley stripe rust. *Theor Appl Genet* **111**: 1260–1270.
- Xu S (1998). Mapping quantitative trait loci using multiple families of line crosses. *Genetics* **148**: 517–524.
- Yu J, Holland JB, McMullen MD, Buckler ES (2008). Genetic design and statistical power of nested association mapping in maize. *Genetics* **178**: 539–551.
- Yu J, Pressoir G, Briggs WH, Vroh Bi I, Yamasaki M, Doebley JF *et al.* (2006). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat Genet* **38**: 203–208.
- Zhao K, Aranzana MJ, Kim S, Lister C, Shindo C, Tang C *et al.* (2007). An *Arabidopsis* example of association mapping in structured samples. *PLoS Genet* **3**: e4.
- Zhu C, Gore M, Buckler ES, Yu J (2008). Status and prospects of association mapping in plants. *Plant Genome* **1**: 5–20.

Supplementary Information accompanies the paper on Heredity website (<http://www.nature.com/hdy>)